Thinking Modally

Thinking Modally:

English and Contrastive Studies on Modality

Edited by

Juan Rafael Zamorano-Mansilla,
Carmen Maíz, Elena Domínguez
and Mª Victoria Martín de la Rosa

Cambridge
Scholars
Publishing

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The editors of this volume would like to express their gratitude to the Spanish Ministry of Science and Innovation[1] and the department of English Studies I (Universidad Complutense of Madrid) for their financial support.

Our thanks must also be extended to all those that made possible the success of ModE4 in the academic and the human dimensions: the members of the organizing committee (Juana Marín Arrese, Marta Carretero, Jorge Arús, Elena Domínguez, Carmen Maíz, Mª Victoria Martín and Juan Rafael Zamorano-Mansilla), the BA and MA students who collaborated with this committee, and the members of the scientific committee (Francisco Alonso Almeida, Johan van der Auwera, Antonio Barcelona, Enrique Bernárdez, Marisa Blanco, Frank Brisard, Pierre Busuttil, Rosario Caballero, Bert Cornillie, Margarita Correa, Nicole Delbecque, Angela Downing, Marciano Escutia, Roberta Facchinetti, Teresa Fanego, Francisco Garrudo, Dirk Geeraerts, Gaëtanelle Gilquin, María de los Ángeles Gómez, Montserrat González, Santiago I. González Fernández-Corugedo, Pilar Guerrero, Christopher Hart, Liesbet Heyvaert, Laura Hidalgo, Iraide Ibarretxe, István Kecskes, Gitte Kristiansen, Paul Larreya, Julia Lavid, Javier Martín Arista, Mª Ángeles Martínez, Begoña Núñez, Ana Pinto, Jesús Romero, Francisco Ruiz de Mendoza, Jesús Sanchez, Augusto Soares da Silva, Anastasios Tsangalidis and Christopher Williams).

INTRODUCTION

JUAN RAFAEL ZAMORANO-MANSILLA
CARMEN MAÍZ
ELENA DOMÍNGUEZ
Mª VICTORIA MARTÍN DE LA ROSA

This volume contains a selection of the papers presented at the 4th International Conference on Modality in English, held in Madrid on 9-11 September 2010. The contributions have been divided into two parts. The first part includes papers that focus on the notions of modality, evidentiality and temporality, while the papers in the second part explore modality and its connection with stance and evaluation in specific genres and discourse domains.

The book starts with the paper by Svenja Kranich and Volker Gast, who report on an empirical investigation of small-scale diachronic changes in the domain of modality between 1961 and 1991. From a corpus-driven perspective, the authors show how this period witnessed a significant increase of modal auxiliaries at the expense of other (lexical) means of modal marking in the expression of epistemic modality. They interpret this development as a decrease of "explicitness" in modal marking, either due to a global reorganization of the English modal system or a register-specific change affecting primarily (popular) scientific texts.

In the second paper Olivier Polge carries out a diachronic analysis of the verbs *need* and *want*. Employing an extensive and varied corpus that includes a wide range of examples from American English, Renaissance drama plays or the British National Corpus; the author sets out to explain why the two constructions converge in terms of meaning and syntactic properties in Present Day English.

The next paper analyses central modals *must* and *should* and semi-modals *have to*, *have got to*, *gotta*, *need to* and *want to/wanna* in the dialogic BNC genres of Broadcast discussion, Classroom, Tutorial and Consultation. The author, Soili Nokkonen, explores these forms in connection with register variation, the sociolinguistic variable of gender

and the type of subject, revealing that there are several intertwining factors involved.

The following paper also focuses on the analysis of modal verbs; more specifically *can* and *may*. As indicated in the title, Patrick J. Duffley sets out to answer the question why epistemic *may* is rare in questions and epistemic *can* in statements. In his view, the semantics postulated for these two auxiliaries has been so impoverished that it has left practically all of the work to the pragmatic component. Hence, he considers it is crucial to strike the proper balance between semantics and pragmatics.

Adopting a functional perspective, Daisuke Suzuki's paper focuses on the analysis of the three modal adverbs *no doubt*, *doubtless* and *undoubtedly*. Using data from the British National Corpus, his analysis demonstrates that the three adverbs fulfil different functions at the discourse-pragmatic level, and that the factors influencing their usage are strongly associated with the parameters of modality and discourse.

Grégory Furmaniak and Paul Larreya, basing their analysis on the work being done in recent literature on the use of epimistic *would* in "would be" and would have been", set out to extend the study of epistemic *would* – the uncertainty bears either on the whole of the proposition or on one of its components only, broadening the concept of "epistemicity" – to cover a wider range of uses, all of them known as "conjectural would", which can be found in three different contexts. The result is that the authors provide a unified account of the three uses of *would,* where they can be seen to form a radial category, each of them sharing some semantic components with at least one of the other members of the category.

In the light of the consideration that auxiliares (including modal ones) function as essential constitutive elements of clauses and utterances, two English auxiliaries, *would* and *used to*, are analysed in Katherine Hrisonopulo's paper. There are two guiding questions: firstly, the basis for the correlation of modal and aspectual meanings expressed by *would* and *used to* in habitual utterances; secondly, the functional load of the two auxiliaries is considered. Finally, the conclusion is reached that the choice of either modal *would* or aspectual *used to* relates to a number of indexical values such as knowledge (sub)type, viewpoint and the degree of a conceptualizer's engagement when making a clausal prediction.

In spite of the claim being made in recent literature that modality is intertwined with tense, Eleni Staraki's paper gives evidence to support the opposite claim that, in fact, modal verbs and tenses constitute separate entities and, as a result, they have a distinct semantic contribution. Besides that, she examines Greek modal verbs, *bori* and *prepi*, and contrasts their temporal interpretation with modal verbs like English may and might.

Moreover, the author claims that counterfactuality is not the result of a past tense and a future orientation but the result of nonveridicality. And finally, al alternative semantic and syntactic analysis is provided to account for the interaction of modality and tense.

Marc Fryd's paper explores the novel construction of present perfect in present-day English in conjunction with implicit or explicit definite past time reference, in a corpus of Australian police reports (where examples of this construction abound), as something may be happening to the present perfect, in such a way that the constraints we seem to find in other varieties such as Standard English do not apply to the Australian variety. Finally, the paper proposes that the motivating factor for this new use may not just be the mechanics of plot progression but speaker-based considerations of testimonial involvement.

Jelena Timotijevic's *paper explores* the relationship between modality and evidentiality, which has been addressed differently by different authors, with the result that a diversity of definitions about modals and the way they are analysed have been provided. In line with that, three main controversial issues are dealt with in this paper by applying notions from the position known as Contextualisation (it claims that a clear-cut distinction between semantics and pragmatics is not possible because what a speaker says and what (s)he implies are both pragmatic). The three main issues are: Is evidentiality a type of modality or is it a separate category?; Is evidentiality encoded in particular expressions and grammatical constructions or dependent on context?; What is the relationship between evidentiality and epistemic modality?

Laura Alba-Juez and Elena Martínez Caro's paper is a corpus-based study of the use the expression *no wonder* in English. More specifically, the authors focus on two aspects of the use of the expression: 1) its function as a marker of epistemic modality and its relation to the expression of negative evaluation or stance; and 2) its apparent gradual change into a grammaticalized expression. The authors conclude that the expression *no wonder* mainly indicates the speaker's opinion concerning the likelihood of the event described and that the expression has taken on a negative evaluative meaning. Structurally, the authors have found similarities between this expression and so-called pragmatic fragments, such as *I think, I know, I guess.*

Marta Carretero and Maite Taboada reports results of research carried out as part of the CONTRANOT project, which aims to produce and validate annotation guidelines for semantic and pragmatic meanings in English and Spanish. In this paper the authors explore the interplay between Appraisal and modality. More specifically, they investigate the

possible overlap between two of the categories proposed in the framework of Appraisal (Attitude and Engagement) and epistemic modality. Using the Simon Fraser University Review Corpus, the authors identify the areas of veracity, security, normality and epistemic and quality/capacity implicatures as the main areas of overlap.

Analysing the collection of lectures contained in the MICASE corpus, Silvia Molina attempts to identify a pattern in the use of epistemic and stance adverbs and adjectives across three subsets of the corpus. These subcorpora, established based on their content, are: biological and health science; humanities and arts; and social science and education. The pattern of use investigated here concerns the variation in the frequency of use of epistemic modality and attitudinal stance across the subcorpora, the frequency of use of epistemic adjectives and adverbials and the frequency of use of attitudinal and epistemic adverbs.

Cristina Luque Agulló's paper analyses epistemic modality within the framework of Cognitive Linguistics. Drawing on data from fiction works and dictionaries, the paper discusses the relevance of the notion of time, proactivity and Potential Reality for the conceptualisation of the experiential domain of Fear in English. The analysis offered here links the experience of the emotion to its conceptualisation and the way we talk or understand it. Adjectives such as (ANXIOUS), (APPREHENSIVE), (NERVOUS) and (WORRIED), nouns such as (PHOBIA) and (DANGER) or verbs such as (DREAD) or (THREATENED) are shown to profile the proactive aspect of the emotion, and, in particular, the ambiguity/unability on the side of the experiencer to locate/identify the type or nature of the stimulus.

Margarita Sánchez Cuervo provides an analysis of the use of stance and (inter)subjectivity in the essays of Virginia Woolf. The author sets out to classify expressions of stance and (inter)subjectivity in Virginia Woolf's essays, trying to find patterns of use both in terms of frequency and motivation. The author concludes that epistemic modality is prevalent in the essays, and establishes connections between different subtypes of stance and the situations described by Virginia Woolf in each passage.

Pilar López Sánchez focuses on Modality in English and Spanish advertising slogans, her main aim being that to analyze significant elements of modality in 200 written advertising slogans published in *Cosmopolitan*—UK, US and Spanish issues. Different modality expressions are analyzed, compared and studied in both corpora from a Systemic Functional approach.

Advertising is also the focus of the following paper. María Victoria Martín and Elena Domínguez's main intention is to test the persuasive

effect of multimodal, pictorial metaphors in advertising, particularly focusing on the analysis three advertisements taken from the publications *Newsweek* and *The Economist*. The analysis is mostly based on the theory on pictorial metaphor developed by Forceville (1996 and 2006), where a new approach to metaphor is advanced: the all-pervasive verbal realization of metaphor is just one mode of communication which lies side by side with other modes such as the visual/pictorial manifestations of metaphor.

Francisco Alonso-Almeida investigates the expression of sentential evidential and epistemic devices in a corpus of forty-six academic abstracts preceding journal articles in English and Spanish in the field of computing. Even if the author admits that evidentiality could be understood under the umbrella of epistemic modality, he opts for keeping them two as distinct categories since this disassociation may bring better results in empirical studies such as the one reported in his study. In this sense, his description of both categories draws mainly on Carretero's (2004) intersective approach, and Cornillie's (2009) disjunctive model.

Within the field of modality in academic discourse, scholar articles are the focus of the chapter by Luz Gil-Salom and Carmen Soler-Monreal. More specifically, the coding of modality in research article discussions is the topic of. These two authors follow previous studies on writer/reader interaction and textual practices and differences among disciplines and genres (Dahl, 2008; Gil-Salom and Soler-Monreal, 2009; Hyland, 1996, 1998c, 2005; Koutsantoni, 2006; Mauranen, 1997; Meyer, 1997) to analyze Research Articles Discussions in what the authors call four "hard science disciplines"―namely Computing, Nanotechnology, Robotics and Telecommunications―and to explore epistemic devices as discourse features aiding the author's persuasive goals and the reader's engagement in the construction of academic knowledge. More specifically, it concentrates on the use of markers of uncertainty that tone down propositional contents.

María Ángeles Martínez's study on Modality awareness development in narratives for young readers of English focuses on the linguistic organization of narrative discourse for young readers of English, with the aim of identifying variations connected to target age and target native/non-native addressee hence confirming a gradual exposure hypothesis. The cognitive principles of *foreground* and *background* provide the theoretical framework for the analysis of two equally proportioned sets of data. The first consists of 12 narratives for native young readers of English, and the second of 16 graded—simplified or abridged—narratives for foreign learners of the language, with a total of 9,841 clauses: 5,401 in the L1 sample and 4,440 in the EFL one.

Jorge Arús Hita, Juan Rafael Zamorano-Mansilla, and Julia Lavid López bring together two important areas of linguistic research—modality and corpus linguistics—to which they add a third component functioning as a binding element between the former two: linguistics teaching. More specifically, they look at a methodology for the teaching of modality by means of corpora. With this aim in mind, they present some corpus-based activities carried out in a contrastive English-Spanish linguistics course belonging to the English Studies curriculum. Their intention is to unveil the benefits these activities provide students with when it comes to the acquisition of modality. Following the authors' line of argumentation, the corpus is not only used for substantial addition of evidence but for checking the correct understanding of modal meanings. To this end, they draw on previous approaches to the didactical applications of annotated corpora by Sinclair (2004) or McEnery and Wilson (2004). Nevertheless, the didactic applications of manual corpus annotation remain largely unexplored to this date.

Juan Rafael Zamorano-Mansilla teaches English language and linguistics at the Facultad de Filología Inglesa, Universidad Complutense de Madrid. He obtained his PhD in 2006 with a dissertation on the expression of tense and aspect in English and Spanish. His publications are mainly on tense, aspect and modality as well as the use of new technologies in linguistics.

Carmen Maíz obtained her PhD from Universidad Complutense of Madrid, where she currently lectures in the English Department. She has published numerous book chapters and articles on intercultural pragmatics, multimodality and contrastive linguistics (English-Spanish).

Elena Domínguez lectures in the English Department of the Complutense University of Madrid (Spain). Some of her publications include a multimodal approach to literary, journalistic and advertising discourse. She is currently collaborating as an external member with the European Evidentiality and Modality (EUROEVIDMOD) Research Project team.

Victoria Martín de la Rosa works as an instructor in the UCM. She has a PhD in English Linguistics from the UCM (2002). Since then, she has been applying the tool of metaphor to the analysis of multimodal messages such as advertisements and to the analysis of the North-American education policy (Obama's educational discourse).

PART I:

STUDIES ON MODALITY, EVIDENTIALITY AND TEMPORALITY

EXPLICITNESS OF EPISTEMIC MODAL MARKING: RECENT CHANGES IN BRITISH AND AMERICAN ENGLISH

SVENJA KRANICH AND VOLKER GAST

1. Introduction[1]

A comparative study of epistemic modality in English and German popular scientific texts (Kranich 2010) has shown that there is a significant decrease of explicitness in the English texts between around 1980 and 2000. Roughly speaking, writers show a tendency towards using more grammaticalized expressions of modality, without an indication of the "source of evidence", and with less precise information on the (degree of) "modal force" (cf. Section 2). The question arises whether this change is characteristic of the register investigated by Kranich (2010) (popular scientific texts), or whether it affects the entire modal system of English. The present study addresses this question by investigating mixed-genre corpora from the BROWN-family. By considering both British and American data we are moreover in a position to determine the degree of (large-scale) regional variation. More specifically, the following questions are addressed:

(i)   To what extent did the distribution of major types of epistemic modal markers change between 1961 (as reflected in BROWN and LOB) and 1991 (as reflected in FROWN and FLOB)?

(ii)   To what extent do the observed changes (if any) differ between British and American English?

(iii) To what extent are the observed changes (if any) register-specific?

(iv)   How can the observed changes (if any) be explained?

Our results show that there is a significant increase in the use of modal auxiliaries for the expression of epistemic modality at the expense of lexical markers, and hence a decrease in explicitness, between the mixed-genre corpora from 1961 and those from 1991. However, this change is not consistent across registers and is in fact most marked in scientific and popular scientific texts. The other registers, too, show a trend towards an increasing use of modal auxiliaries, but the results obtained from the four 50,000 word samples are not significant at a five per cent level. We have not found any significant differences between British and American English.

We consider two hypotheses concerning explanations for our results: First, they may reflect a global reorganization of the mapping from form to function in the domain of modality; second, they may be symptoms of stylistic changes characteristic of specific registers. While our results suggest that the second hypothesis is more likely than the first, a more comprehensive investigation of both deontic and epistemic modality in the BROWN-corpora is required for a detailed assessment of this matter. We plan to investigate these hypotheses further in future research.

After a few remarks on markers of epistemic modality (Section 2) the results of our quantitative study are presented in Section 3. These results are interpreted and discussed in Section 4. The paper concludes with a brief summary and an outlook in Section 5.

2. Epistemic modal markers

2.1. Basic distinctions

The category of epistemic modality concerns "the degree of commitment by the speaker to what he/she is saying" (Palmer 2001: 51; cf. also van der Auwera et al. 2005: 201, Verstraete 2007: 17, Larreya 2009: 13, among others). In non-modalized declarative sentences, a speaker is fully committed to the utterance made, and someone uttering (1) would be accused of insincerity if the director is not in fact sleeping at the moment of utterance.

(1)   The director is sleeping.

If a speaker does not have sufficient evidence justifying a "bare" indicative sentence as in (1) or wishes to mitigate his/her claim for other reasons (e.g. politeness),[2] s/he can weaken the degree of commitment by "modalizing" the statement in some way, e.g. with a modal auxiliary such as *must* or *may*:

> (2) a. The director must be sleeping.
>   b. The director may be sleeping.

Both examples in (2) make a weaker claim than (1), indicating as they do that the information provided is not based on direct evidence but on beliefs or inferences. (2a) contains a "universal" modal, i.e. one expressing necessity. The modal here indicates that the sentence is believed to be true under all conceivable circumstances. (2b), which contains an "existential" modal, merely states that it is considered possible (i.e. not necessarily false) that the director is sleeping. In other words, there are conceivable circumstances under which the sentence is true (hence "existential" modal; cf. Kratzer 1991 for an explicit treatment of modality, and von Fintel 2006 for a more recent survey).[3]

We can distinguish four major (syntactic) types of epistemic modal markers in English (cf. Kranich 2010):

> (i)   modal auxiliaries (*may, might, can, could, must*)
> (ii)   (lexical) modal verbs (*seem, appear*)
> (iii) modal adjectives or adverbs (*likely, probably, perhaps,* etc.)
> (iv)   modal periphrases (*I would wager that …, I doubt if …*, etc.)

While categories (i) and (ii) represent closed classes, categories (iii) and (iv) are open. Adjectives and adverbs can be derived (e.g. *reportedly*), and modal periphrases can be formed *ad hoc*. This latter class contains two major types, i.e. superordinate predications (e.g. *I doubt if …*) and parentheses (*… – I guess – …*). Obviously, modal markers of different types can be combined in a single statement, as in (3):

> (3)   a. *I would wager* that he has *probably* never been there.
>   (periphrasis + adverb)

---

[2] It has been observed in various studies that epistemic modal markers are used for two types of reason: "content-oriented caution" (i.e. a speaker is not sure) and "addressee-oriented caution" (i.e. a speaker does not wish to appear bold or boastful, intends to leave room for other opinions, etc.; see e.g. Hyland 1994, 1996, White 2003, White & Sano 2006, Kranich 2011).

[3] Pragmatically, even a sentence as the one in (2b) is, however, taken to indicate that the speaker considers the possibility that the director is sleeping more likely than the possibility that the director is not sleeping (see Kranich 2011: 83, 87–88).

    b.   He *may possibly* like to add a few words to that.
       (modal auxiliary + adverbial)

2.2. Degrees of explicitness

The various ways of indicating epistemic modality differ in their degrees of explicitness along at least two dimensions, (i) the indication of a source of evidence, and (ii) the degree of precision in the indication of "epistemic force" (cf. Kratzer 1991, von Fintel 2006, Declerck 2009: 48). As a general tendency, modal constructions involving lexical markers (i.e. adjectives, adverbs or verbs) and modal periphrases are more explicit with respect to both dimensions than modal auxiliaries.

    The set of sources of evidence includes categories such as "direct evidence", "hearsay", "inference", "general knowledge", etc. For example, the (attested) example in (4) indicates that the statement made is a belief ("what I take to be") based on general knowledge ("the prevailing view"), while the (made-up) example in (5), where a modal auxiliary is used, merely characterizes the statement as some type of (more or less inescapable) inference.

    (4)   According, then, to what I take to be the prevailing view, these rioters were merely a handful of irresponsible, Stalinist-corrupted provocateurs. (BROWN, science)

    (5)   These rioters must have been merely a handful of irresponsible, Stalinist-corrupted provocateurs.

    Similarly, the precision with which the degree of epistemic commitment is indicated is typically higher in lexical and periphrastic modal marking than in modal auxiliary constructions. As mentioned above, we can make a rough distinction between "existential" and "universal" markers of modality (possibility vs. necessity), but there are obviously many degrees of modal force located in between (cf. Kratzer 1991). For instance, (6) (from a scientific text) is very explicit with respect to the degree of commitment made. An equivalent degree of precision could not be expressed with a modal auxiliary, as the latter are basically restricted to the binary distinction mentioned above (possibility vs. necessity), with a few more fine-grained distinctions being encoded in what were historically speaking the past and present tense forms of the modals (e.g. *may* vs. *might*).

    (6)   *It is suggested that* it *is unlikely that* the rhythmites represent annual glacial varves as interpreted by Shotton and *it is considered more probable that* they represent small scale turbidites. (FLOB, science)

The asymmetry in explicitness between modal auxiliaries on the one hand, and lexical and periphrastic markers on the other, parallels differences between more grammaticalized and less grammaticalized alternatives in other ontological categories. For example, morphosyntactically expressed tenses are typically less precise than temporal adverbs with respect to the location of an event or "Topic Time" (e.g. "past time" vs. "yesterday"; cf. Klein 1994). It will therefore be useful to make a distinction between "grammatical" and "lexical" markers of epistemic modality, with the first group corresponding to modal auxiliaries and the second group comprising expressions of types (ii) – (iv), i.e. lexical modal verbs, modal adjectives and adverbs as well as modal periphrases. It should be borne in mind that "grammatical" here means "more grammatical(ized)" than alternative ("lexical") types of expression, as the status of modal auxiliaries as grammatical markers is probably not beyond doubt.

3. The corpus study

3.1. Methodological remarks

Given that we do not have access to suitable mixed-genre corpora from the time frame investigated by Kranich (2010) for popular scientific texts (around 1980–2000), we have used corpora of the BROWN-family as an empirical basis of our study. This has allowed us to probe into the role of regional differences by comparing the two major national varieties of English, i.e. British and American English. The BROWN-family comprises two corpora from 1961 (BROWN/American English and LOB/British English) and two corpora from 1991 (FROWN/American and FLOB/British).

In order to determine the distribution of different types of modal expressions in the four corpora under consideration, parallel samples were extracted from each corpus. The samples consist of fifteen sub-samples from the registers distinguished in the BROWN-corpora, covering such varied genres as editorials in newspapers, scientific texts, romances and adventure novels. Each sub-sample contains approximately ten randomly chosen blocks of ten to fifteen sentences. Each sample thus created amounts to approximately 50,000 words (15 registers, 10 blocks, ~ 333 words per block). Altogether, 200,000 words were thus extracted from the four corpora.

Through "close reading", markers of epistemic modality were identified and classified into the four groups mentioned above, using the

same method as Kranich (2011: 86–88).[4] The results were analysed using common statistical procedures.

3.2. Differences between 1961 and 1991

The frequencies of specific types of modals in our four corpus samples are shown in Table 1 for British and American English separately.

	British English		American English	
	LOB (1961)	FLOB (1991)	BROWN (1961)	FROWN (1991)
Modal auxiliaries	47 31.8%	70 46.7%	70 38.5%	78 54.5%
Modal adjectives or adverbs	59 39.9%	53 35.3%	53 29.1%	38 26.6%
Lexical modal verbs	13 8.8%	6 4.0%	23 12.6%	11 7.7%
Modal periphrases	29 19.6%	21 14.0%	36 19.8%	16 11.2%
Σ	148	150	182	143

**Table 1. Four types of epistemic modal expressions in the four samples**

The null hypothesis is that the independent variable "year" (with the factor levels "1961" and "1991") does not have an influence on the dependent variable "type of modal expression" ("auxiliary", "verb", "adjective/adverb", "periphrasis"). Table 1 shows that it must be rejected for both varieties under consideration. Both distributions deviate significantly from chance (BrE χ^2 = 8.6887, df = 3, p < .05, AmE χ^2 = 10.3009, df = 3, p < .05). This is a first indication that something changed in the domain of modality between 1961 and 1991.

The data are visualized in Diagram 1 for British English and in Diagram 2 for American English (relative frequencies). A comparison of the two diagrams reveals an amazing parallelism between the two varieties under investigation. In both British and American English, all types of lexical epistemic markers decreased in the time frame under consideration,

---

[4]   Instead of following the form-to-function (semasiological) approach characteristic of most corpus-based studies, we thus used a function-to-form (onomasiological) approach (as called for by Nuyts 2005: 14f. for the study of modality in particular). Based on a semantic definition of epistemic modality all linguistic expressions fulfilling this definition were manually searched.

whereas modal auxiliaries gained ground. We will return to the relationship between British and American data in Section 3.3.
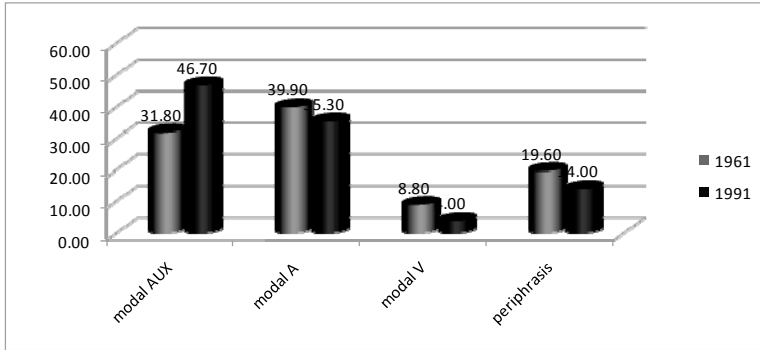
Diagram 1. Relative frequencies of epistemic modal constructions in British English (1961 and 1991)
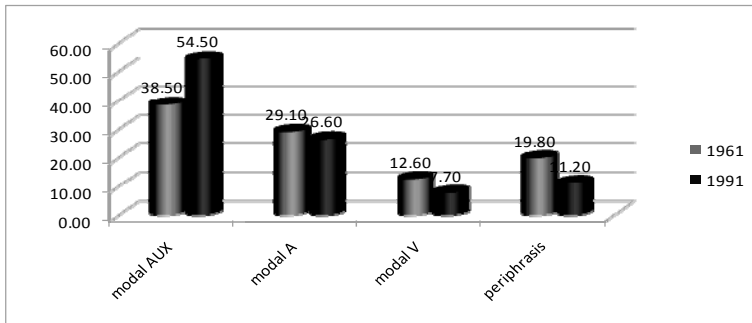
Diagram 2. Relative frequencies of epistemic modal constructions in American English (1961 and 1991)

The data displayed in Diagram 1 and Diagram 2 suggest that the major divide is between grammatical and lexical marking of epistemic modality. If we subsume categories (ii)–(iv) under the single category of "lexical modal markers", we get the distribution displayed in Table 3 (where British and American data are lumped together). The proportion of modal auxiliaries rises from 35.5% in 1961 to 50.5% in 1991. The deviation from chance is very highly significant ($\chi^2 = 17.237$, df = 1, p < .001).

	1961 (BROWN/LOB)	1991 (FROWN/FLOB)

lexical modal	213
	                                    64,5%
		145
		                                    49,5%

modal AUX	117
	                                    35,5%
		148
		                                    50,5%

Σ	330	293

Table 2. Modal auxiliaries vs. lexical modal expressions in 1961 and 1991 (British and American English)

3.3. Differences between British and American English

As mentioned above, the distributions of modal expressions in British and American English shown in Diagram 1 and Diagram 2 look rather similar. It remains to be seen whether there are any significant deviations from statistical independence. The data are reorganized in Table 4, which allows for a direct comparison of British and American English in the two time frames under consideration.

	1961		1991	
	LOB (Br)	BROWN (Am)	FLOB (Br)	FROWN (Am)
Modal auxiliaries	47 31.8%	70 38.5%	70 46.7%	78 54.5%
Modal adjectives / adverbs	59 39.9%	53 29.1%	53 35.3%	38 26.6%
Lexical modal verbs	13 8.8%	23 12.6%	6 4.0%	11 7.7%
Modal periphrases	29 19.6%	36 19.8%	21 14.0%	16 11.2%
Σ	148	182	150	143

Table 3. Comparing British and American English

The null hypothesis is that the independent variable "variety" ("British", "American") does not have an influence on the dependent variable "type of modal expression" in either 1961 or 1991. On the basis of the data in Table 2, this null hypothesis cannot be rejected for either time frame. In neither year is there a significant deviation from independence (1961: $\chi^2 = 4.9237$, df = 3, p > .05; 1991: $\chi^2 = 4.8868$, df = 3, p > .05). Given that the individual distributions (1961 and 1991) do not differ significantly in British and American English, and that the observable changes within the two

distributions are entirely parallel (cf. Diagram 1 and Diagram 2) – we conclude that the variable "variety" can be neglected, and we will aggregate British and American data in the following.

3.4. Distribution of epistemic modal expressions across registers

Having shown that the trend towards an increasing use of modal auxiliaries in the epistemic domain identified by Kranich (2011) for popular scientific texts is also evidenced in a mixed genre corpus of texts from 1961 and 1991, we will now have a closer look at the influence of register. For this purpose, we classified twelve of the fifteen BROWN-genres into four major "register types": "Popular science and science" (F, J), "Press" (A, B, C), "General fiction and biographies, essays, belles lettres" (G, K) and "Popular fiction" (L, M, N, P, R). The BROWN-registers D ("Religion"), E ("Skill and hobbies"), F ("Popular lore") and H ("Miscellaneous") were disregarded, as they could not easily be subsumed under any of the other register types and would have delivered too small numbers to allow for any quantitative inferences if considered on their own.

Table 4 shows the distribution for the 1961-corpora (BROWN and LOB) and Table 5 for the 1991-corpora (FROWN and FLOB). The percentages indicate the proportion of a given type of expression within the register in question (i.e. they represent the quotient of the observed frequency and the marginal sum on the right).

	AUX	A	V	periphr.	Σ
Popular science, science (F, J)	22 36.7%	12 20.0%	18 30.0%	8 13.3%	60
Press (A, B, C)	31 43.1%	7 9.7%	24 33.3%	10 13.9%	72
General fiction and biographies, essays, belles lettres (G, K)	11 25.6%	8 18.6%	15 34.9%	9 20.9%	43
Popular fiction (L, M, N, P, R)	37 34.6%	7 6.5%	33 30.8%	30 28.0%	107
Σ	101	34	90	57	282

Table 4. Distribution of modals in four major register types (1961, British and American English)

	AUX	A	V	periphr.	Σ
Popular science, science (F, J)	46	4	11	5	66
	69.7%	6.1%	16.7%	7.6%	
Press (A, B, C)	28	2	23	6	59
	47.5%	3.4%	39.0%	10.2%	
General fiction and biographies,	11	3	12	4	30
essays, belles lettres (G, K)	36.7%	10.0%	40.0%	13.3%	
Popular fiction	41	4	41	15	101
(L, M, N, P, R)	40.6%	4.0%	40.6%	14.9%	
Σ	126	13	87	30	256

**Table 5. Distribution of modals in four major register types
(1991, British and American English)**

The null hypothesis is that the distribution of modal auxiliaries is independent of the variable "register type". For the 1961 data, this hypothesis cannot be rejected, as the deviation from independence is not significant at a five per cent level, though it comes very close to significance (χ^2 = 16.4156, df = 9, p = .059).[5] The 1991-distribution, by contrast, is significant (χ^2 = 20.1596, df = 9, p < .05), and in fact close to being highly significant (p = 0.01695).[6]

The two distributions can be visualized with Cohen-Friendly association plots as shown in Diagram 3. These plots can be regarded as graphic versions of contingency tables. Each box in the plots corresponds to a cell in Table 4 or Table 5. For example, the black box in the bottom right corner of the left plot corresponds to the cell representing the combination of the values "periphrases" (for the variable "type of modal expression") and "popular fiction" (for the variable "register") in Table 4. The association plots show whether a given cell is larger or smaller than expected (on the assumption of statistical independence of the variables), with cells exceeding expectations being located above the baseline (in black) and cells that are lower than expected underneath (in white). The area of each cell is proportional to the difference between expected and observed frequency.[7]

[5] Fisher's exact test: p = .063.
[6] Fisher's exact test: p = .011.
[7] The height of each cell is proportional to the corresponding Pearson residual, and the width is proportional to the square root of the expected frequency.

1961   1991
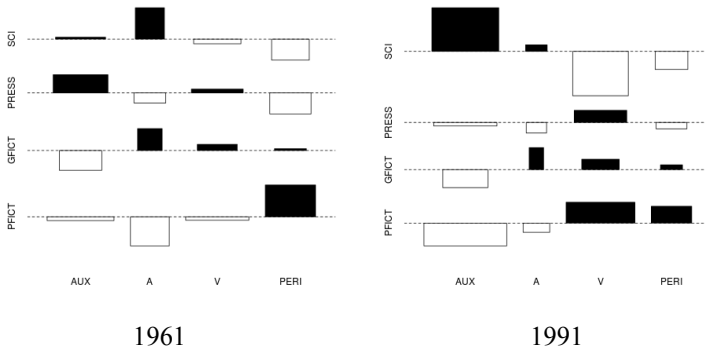
Diagram 3. Association plots for Table 4 and Table 5

A comparison of the two plots reveals a major difference in the distribution of epistemically used modal auxiliaries over the four register groups under consideration. While the 1961-plot shows only minor deviations from independence (which are jointly not significant at a five per cent level), in 1991 there is a clear positive association between the scientific registers (top row) and modal auxiliaries (first column). At the same time, modal auxiliaries are negatively associated with the three other registers, though to a minor extent (first column/rows 2–4). Considerable differences between the two distributions can also be observed in the (third) column showing the frequency of lexical modal verbs. While they are close to expectation in the 1961 samples (the boxes are very small), they are negatively associated with the genre "science" in 1991 (first row, third column), and positively associated with all other genres.

The question arises to what extent the trends shown in Diagram 4 are significant. This cannot be gathered from the association plots shown in Diagram 3, where attested frequencies are compared with expected ones without testing for significance. Significance testing can be illustrated with mosaic plots of the type shown in Diagram 4, which is based on a chi-square test. Cells with values that are significantly higher than expected on the assumption of statistical independence are blue (dark grey in black and white print), those that are significantly lower than expected are white. Diagram 4 shows that only two cells in the 1991-distribution are in themselves significantly higher or lower than expected, i.e. the cells "modal auxiliary × science" (> expectation) and "modal verb × science" (< expectation).

1961   1991
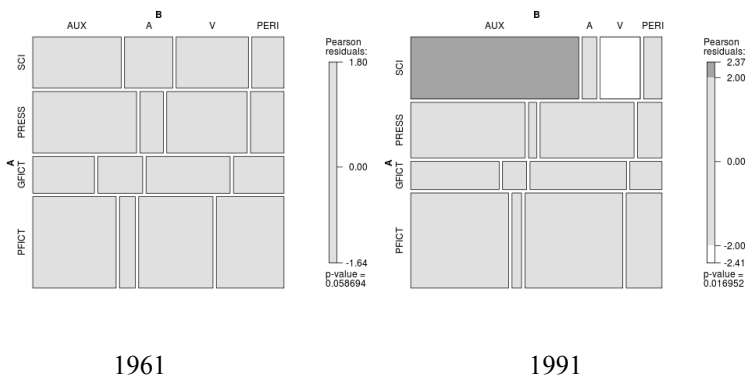
**Diagram 4. Mosaic plots for the distribution across registers
(1961 vs. 1991)**

The data considered in this section suggest that we are dealing with a register-specific effect. This impression is confirmed when we collapse the three types of expression subsumed under the category "lexical". The data are shown in Table 6 (1961) and Table 7 (1991).

	AUX	lex	Σ
Popular science, science (F, J)	22 36.7%	38 63.3%	60
Press (A, B, C)	31 43.1%	41 56.9%	72
General fiction and biographies, essays, belles lettres (G, K)	11 25.6%	32 74.4%	43
Popular fiction (L, M, N, P, R)	37 34.6%	70 65.4%	107
Σ	101	181	282

**Table 6. Distribution of modals in four major groups of registers
(1961, British and American English)**