Explicitness of epistemic modal marking: Recent changes in British and American English

Svenja Kranich (Hamburg), Volker Gast (Jena)

Abstract

This paper reports on an empirical investigation of small-scale diachronic changes in the domain of modality between 1961 and 1991 as documented in corpora from the BROWN-family (BROWN, LOB, FROWN, FLOB). This period witnesses a significant increase of modal auxiliaries at the expense of other (lexical) means of modal marking in the expression of epistemic modality. This development is interpreted as a decrease of 'explicitness' in modal marking, as lexical markers of modality tend to convey more information than grammatical markers. Two explanations for the observed changes are considered, (i) a global reorganization of the English modal system with auxiliaries shifting towards 'more grammaticalized' (epistemic) meanings, and (ii) a register-specific change affecting primarily (popular) scientific texts. While the data available to us do not allow for any definite conclusions, the second hypothesis appears more likely, as our data show a considerable register effect setting the (popular) scientific texts apart from other registers.

Key words: register, modal, epistemic, explicitness

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 recorded in 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.

---

[1] This paper is a result of a research visit by Volker Gast at the Sfb 538 'Mehrsprachigkeit' (University of Hamburg) in March 2009. Financial support from the DFG is gratefully acknowledged. Moreover, we would like to thank the audience of the conference 'Modality in English IV' (09–11 Sep, 2010), two anonymous reviewers, Martin Schweinberger, and Daniel Wiechmann for valuable comments. Any inaccuracies are our own. The association plots and mosaic plots used in this paper were generated with the open source software *R*, version 2.9.2 (© *The R Foundation for Statistical Computing*).

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 leave this as a suggestion for 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

[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).

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)
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 in terms of 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) or 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

[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).

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 either inflectionally (e.g. *may* vs. *might*) or lexically (cf. *must* vs. *have to*).

(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 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 approx. 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 indicated in 0, using the same method as used by 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.

[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.

	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 / 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 values '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 $\chi^2 = 8.6887$, df = 3, $p < .05$, AmE $\chi^2 = 10.3009$, 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 whereas modal auxiliaries gained ground. We will return to the relationship between British and American data in Section 3.3.

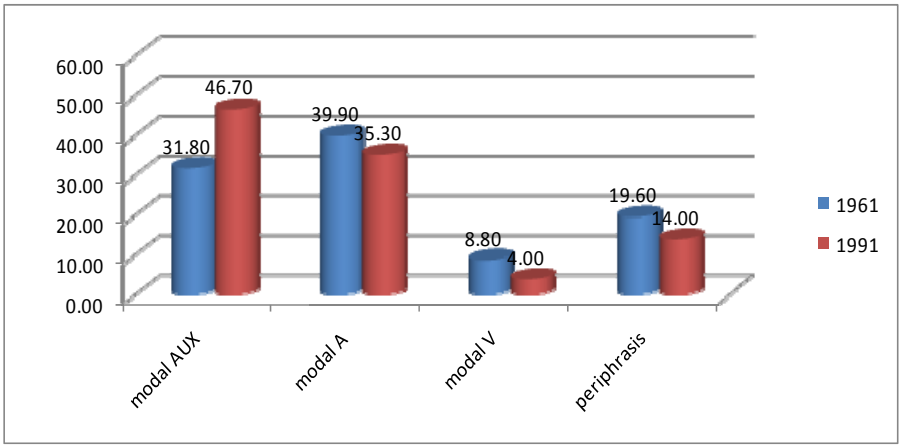


Diagram 1: Relative frequencies of epistemic modal constructions in British English (1961 and 1991)

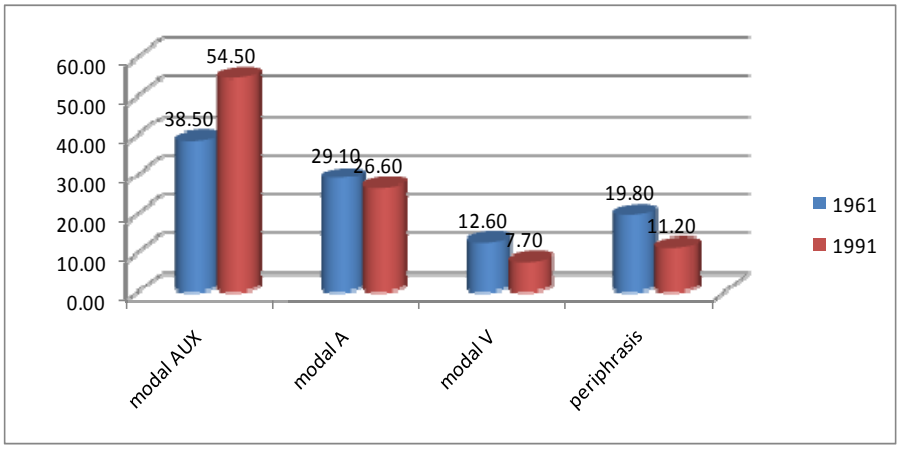


Diagram 2: Relative frequencies of epistemic modal constructions in American English (1961 and 1991)

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 departure 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: Four types of epistemic modal expressions in the four samples

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 in different corpus 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).

	modal AUX	modal A	modal V	periphrases	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)

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

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

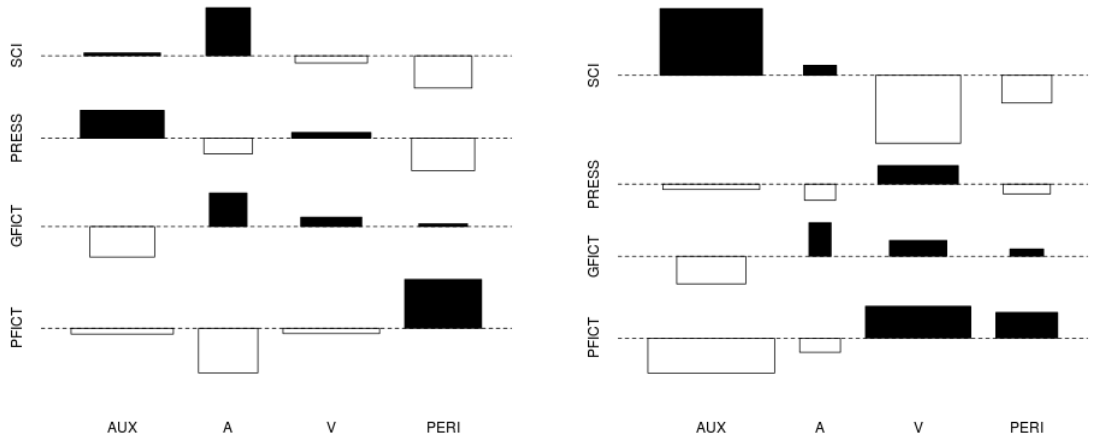
The null hypothesis is that the distribution of modal auxiliaries is independent of the variable 'register type'. For the 1961 data, this hypothesis cannot be rejected, as the deviation from independence is not significant at a five per cent level, though it comes very close to significance ($\chi^2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

[5] Fisher's exact test: p = .063.

[6] Fisher's exact test: p = .011.

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]



1961 1991

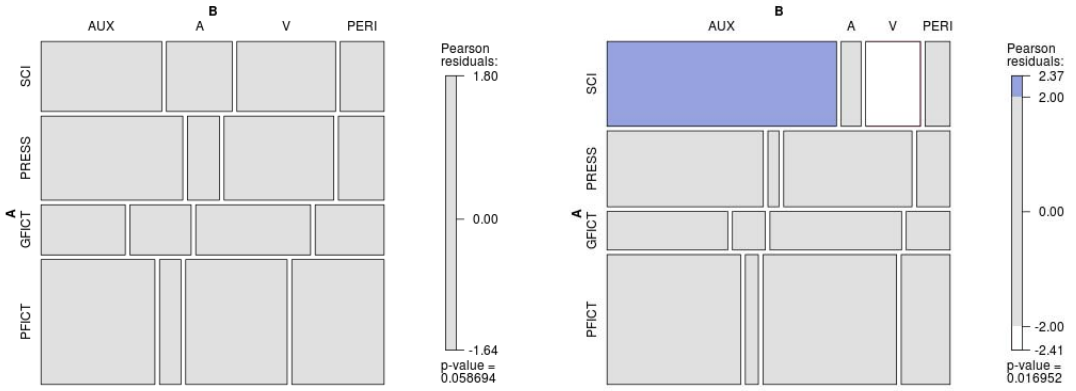
Diagram 3: Association plots for Table 4 and Table 5

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 very close to independence 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

[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. Each rectangle is thus proportional to the difference between expected and observed frequencies.

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, 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).

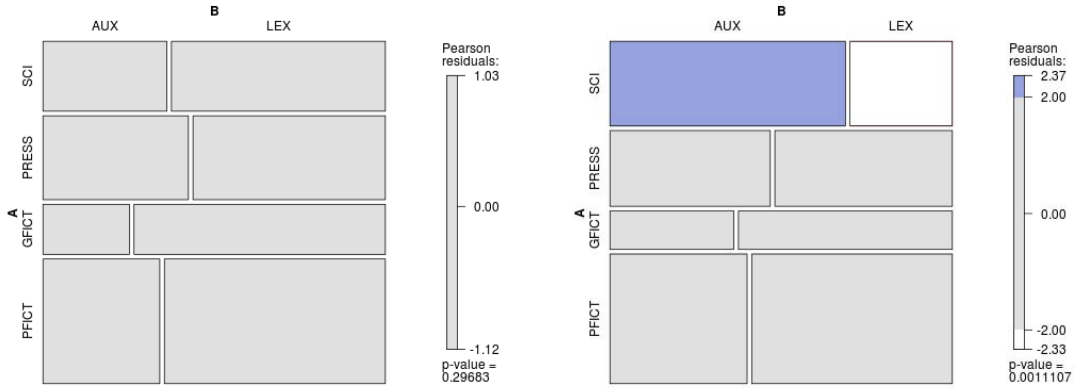
	modal AUX	lexical modals	Σ
Popular science, science (F, J)	22 36.7%	38 63.3%	60
Press (A, B, C)	31 43.1%	41 56.9%	72
General fiction and biographies, essays, belles lettres (G, K)	11 25.6%	32 74.4%	43
Popular fiction (L, M, N, P, R)	37 34.6%	70 65.4%	107
Σ	101	181	282

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

	modal AUX	lexical modals	Σ
Popular science, science (F, J)	46 69.7%	20 30.3%	66
Press (A, B, C)	28 47.5%	31 52.5%	59
General fiction and biographies, essays, belles lettres (G, K)	11 36.7%	19 63.3%	30
Popular fiction (L, M, N, P, R)	41 40.6%	60 59.4%	101
Σ	126	130	256

Table 7: Distribution of modals in four major groups of registers (1991, British and American English)

While the 1961-distribution does not deviate significantly from chance ($\chi^2 = 3.6909$, df = 3. p = 0.2968), the 1991-distribution is highly significant ($\chi^2 = 16.044$, df = 3, $p < .01$). As the mosaic plots in Diagram 6 show, this is, again, mostly due to the overrepresentation of modal auxiliaries (cf. the blue/dark grey top left corner), and the underrepresentation of lexical markers of epistemic modality (cf. the white top right corner), in scientific and popular scientific texts.



1961 1991

Diagram 6: Grammatical vs. lexical epistemic modal markers: 1961 vs. 1991

The results presented above suggest that it is mainly in scientific and popular scientific prose that the frequency of modal auxiliaries increased, relative to other indicators of epistemic modality. However, focusing on these registers alone would probably be premature. There is a clear cross-register trend towards the 'strengthening' of epistemically used modal auxiliaries at the expense of lexical strategies. This becomes apparent if we compare the proportions of modal auxiliaries among epistemic expressions in all register types. This comparison is provided by Diagram 7.

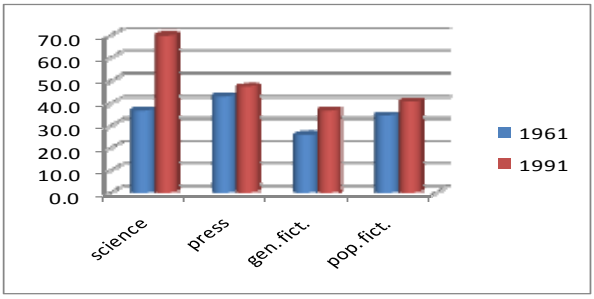


Diagram 7: Proportions of modal auxiliaries in four register types

The proportion of modal auxiliaries in epistemic function, in relation to lexical markers of epistemic modality, increased in all registers between 1961 and 1991. Table 8 shows the distribution of modal auxiliaries and lexical modals in the three register types 'Press', 'General fiction' and 'Popular fiction'. As Table 8 shows, the proportion of epistemically used modals rises from 35.5% to 42.1% between 1961 and 1991. The numbers are not significant at a five per cent level ($\chi^2 = 1.5715$, df = 1, p = .21), but they may very well be when more data are considered.

	1961	1991
modal AUX	79 35.6%	80 42.1%
lexical modals	143 64.4%	110 57.9%
Σ	222	190

Table 8: Modal auxiliaries and lexical modals in 'Press' and '(General and Popular) Fiction'

4 Discussion of results

The data presented in Section 3 have shown that the use of epistemic modal auxiliaries increased considerably at the expense of lexical modal markers between 1961 and 1991 in both British and American English. This is interesting because it runs counter to the general trend for modals to gradually give way to semi-modals or quasi-modals (e.g. Krug 2000, Leech & Smith 2006, Mair 2006, Mair & Leech 2006, Collins 2009, Leech 2010). Our data suggest that this trend may apply only to modals with a deontic function. One way of interpreting the distributions in our sample is therefore to assume that modal auxiliaries underwent (and perhaps are still undergoing) a process of specialization, establishing themselves as grammaticalized markers of epistemic modality while gradually losing their

deontic function. Such a development of increasing specialization is generally not uncommon in grammaticalization processes. As Kytö (1991) has shown, English modals underwent such a development during the Early Modern English period, and this process might be continuing into the present day.

While the corpora investigated in the present study revealed a global increase in epistemically used auxiliaries (cf. Section 3.2), our register-specific investigations have shown that this increase is mainly due to extreme changes in the scientific and popular scientific registers, while the developments in the other registers were more moderate and in fact not significant at a five per cent level (cf. Section 3.4). This suggests that we are dealing with a register-specific effect. In this case the observed changes could reflect changes in 'discourse practice' rather than the grammatical system of English. We thus have to consider two alternative, though not mutually exclusive, explanations, i.e. the 'global reorganization hypothesis' and the 'stylistic hypothesis'.

The 'global reorganization hypothesis' is attractive because it is compatible with our general knowledge of processes of grammaticalization in the modal domain. As van der Auwera & Plungian (1998), among others, have shown, epistemic modality typically represents a later stage in the development of modal categories than deontic modality (cf. also Traugott 1989, Bybee, Perkins & Pagliuca 1994: 194-205). Considering the tendency for deontic modality to be more and more expressed with semi-modals and quasi-modals (as shown e.g. by Krug 2000, and the other studies cited above), we could assume that the 'division of labour' among modal expressions was reorganized between 1961 and 1991, and is perhaps still being reorganized. While deontic modality is increasingly encoded with 'less grammaticalized' types of expression, modal auxiliaries are leaving this domain, becoming increasingly specialized for epistemic modality. In other words, modal auxiliaries are 'pushed' further towards the epistemic end of the modal domain by semi-modals, and the entire cohort moves rightwards. This 'push chain' is displayed in a simplified form in (7).

(7)

LEXICAL MEANINGS	DEONTIC MODALITY	EPISTEMIC MODALITY
	semi-modals	modal auxiliaries
⇒	semi-modals	⇒ modal auxiliaries

The 'global reorganization hypothesis' is expected to be basically, though not entirely, independent of the factor 'register'. A specialization of modal auxiliaries to an epistemic function can be assumed to be part of a native speaker's competence, and it is unlikely to exhibit a strong genre bias (though there is of course always leeway for minor differences

between registers). This hypothesis accordingly stands and falls with the extent to which the genre bias pointed out in Section 3.4 will be (dis)confirmed in future studies. In these studies we will of course also have to investigate the expression of deontic modality in relation to epistemic modality within a given corpus or corpus fragment.

While a register bias is not easily reconciled with the first hypothesis outlined above, it is fully compatible with our second hypothesis. Various authors have observed a 'colloquialization' of certain written registers (e.g. Biber & Finegan 1992, 1997, Mair 2006). It is conceivable that specific registers – in particular, scientific prose – acquired more colloquial features, especially with respect to 'implicitness', which is associated with spoken and more speech-like discourse styles, while explicitness is typically associated with written discourse styles (see for instance Biber 1988). As pointed out in Section 1, we regard modal auxiliaries as less explicit indicators of epistemic modality, insofar as they tend to provide no information about the source of evidence, and are less precise with respect to epistemic strength.

The 'stylistic' hypothesis is compatible with the register variability that we observed, as such a change may affect certain genres to a greater extent than others. In its most extreme form, the hypothesis could even be restricted to the scientific register. Publication habits have changed dramatically in the past few decades, e.g. insofar as there seems to be an increasing tendency to publish work in progress and preliminary findings (as does the present paper, in fact). This tendency could favour more vague means of expressing epistemic judgements, which do not make explicit the epistemic source and the degree of probability attributed to the truth of a proposition. While the 'stylistic' hypothesis seems highly plausible to us, more detailed investigations have to be carried out in order to test it. In particular, a more fine-grained investigation of scientific texts would be needed, with a differentiation between science and popular science, perhaps between different branches of science, and with an eye to other details which can only be identified during the investigation itself.

5 Conclusions

Starting with the observation that there have been changes in the frequency of epistemically used modals in a corpus of English popular scientific prose (Kranich 2010), the present study has reported on the quantitative distribution of different types of modal expressions in parallel samples from the BROWN-family of corpora (BROWN, LOB, FROWN, FLOB). It has been shown that there is a marked increase in epistemic uses of modal auxiliaries, in comparison to corresponding lexical expressions. However, this development shows a considerable register

bias, being most pronounced in scientific and popular scientific texts. For the other registers, there seems to be a trend in the same direction, which is not, however, significant at a five per cent level for the data considered in this study. Further investigations will have to be carried out in order to determine the exact degree of register sensitivity of epistemic modal markers.

We have discussed two explanations for the changes observed in our samples (1961 vs. 1991). First, the frequency asymmetries may be symptoms of a more general reorganization of the English modal system. The second explanation regards the asymmetries as a matter of discourse styles. While we have conjectured that the 'stylistic hypothesis' is more likely, it has been pointed out that the two hypotheses under consideration are of course not mutually exclusive. Each of them may be responsible for specific aspects of the quantitative effects that we have observed. It could turn out, for instance, that only particular modal auxiliaries show a tendency towards specializing for the epistemic domain – like *might*, which is almost entirely restricted to an epistemic use already. Some modal auxiliaries that are overall decreasing such as *may* or *must* (cf. Mair & Leech 2006: 327) might be specializing currently in a parallel way. Alternatively, it is also conceivable that we are looking at a general yet weak trend of specialization of modals for the epistemic domain, and that this tendency multiplies in specific registers due to discourse effects such as the tendency to use modal hedges in publications reporting on preliminary findings. We, for our part, hope to be able to reduce the proportion of epistemically used modal auxiliaries, and hence of vaguely modalized statements, in follow-up work by zooming in on the questions that emerged in the course of our investigation.

References

Biber, Douglas (1988). *Variation across Speech and Writing*. Cambridge: Cambridge University Press.

Biber, Douglas and Edgar Finegan (1992). The linguistic evolution of five written and speech-based English genres from the 17th to the 20th centuries. In Rissanen, M., O. Ihalainen, T. Nevalainen and I. Taavitsainen (eds.), *History of Englishes. New Methods and Interpretations in Historical Linguistic*s, 688–704. Berlin: De Gruyter.

Biber, Douglas and Edgar Finegan (1997). Diachronic relations among speech-based and written registers in English. In Nevalainen, Terttu & Leena Kahlas-Tarkka (eds.), *To Explain the Present. Studies in the Changing English Language in Honour of Matti Rissanen*, 253–275. Helsinki: Société Néophilologique.

Bybee, Joan, Revere Perkins and William Pagliuca (1994). *The Evolution of Grammar. Tense, Aspect and Modality in the Languages of the World*. Chicago: The University of Chicago Press.

Collins, Peter (2009). *Modals and Quasi-Modals in English*. Amsterdam: Rodopi.

Declerck, Renaat (2009). Not-yet-factual at time t: a neglected modal concept. In Salkie, Raphael, Pierre Busuttil & Johan van der Auwera (eds.), *Modality in English. Theory and Description*, 31–54. Berlin: Mouton de Gruyter.

Hyland, Ken (1994). Hedging in academic writing and EAP textbooks. *English for Specific Purposes* 13 (3): 239–256.

Hyland, Ken (1996). Writing without conviction? Hedging in science research articles. *Applied Linguistics* 17: 433-454.

Klein, Wolfgang (1994). *Time in Language*. London: Routledge.

Kranich, Svenja (2010). Translating epistemic modality. *There might perhaps be some problems*. Paper presented at *MIMS* (*Multilingual Individuals and Multilingual Societies*), Hamburg, 06-08 October 2010.

Kranich, Svenja (2011). To hedge or not to hedge: the use of epistemic modal expressions in popular science in English texts, English–German translations, and German original texts. *TEXT & TALK* 31 [1], 77–99.

Kratzer, Angelike (1991). Modality. In von Stechow, A. & D. Wunderlich (eds.). *Semantics. An International Ahdnbook of Contemporary Research*, 639-650. Berlin: Mouton de Gruyter.

Krug, Manfred. (2000). *Emerging Modals: A Corpus-based Study of Grammaticalization.* Berlin: De Gruyter.

Kytö, Merja. (1991). *Variation and Diachrony, with Early American English in Focus: Studies on CAN/MAY and SHALL/WILL* Bern, etc.: Peter Lang.

Larreya, Paul (2009). Towards a typology of modality in language. In Salkie, Raphael, Busuttil, Pierre & Johan van der Auwera (eds.), *Modality in English. Theory and Description*, 9–29. Berlin: Mouton de Gruyter.

Leech, Geoffrey (2010). Where have all the modals gone? On the declining frequency of modal auxiliaries in British and American English. Plenary talk given at the *Modality in English IV*, Madrid, 9–11 September 2010.

Leech, Geoffrey & Nicholas Smith (2006). Recent grammatical change in written English 1961-1992: Some preliminary findings of a comparison of American with British

English. In Antoinette Renouf and Andrew Kehoe (eds.), *The Changing Face of Corpus Linguistics*, 186–204. Amsterdam: Rodopi.

Mair, Christian (2006). *Twentieth-Century English: History, Variation and Standardization*. Cambridge: Cambridge University Press.

Mair, Christian and Leech, Geoffrey (2006). Current change in English syntax. In Bas Aarts and April MacMahon (eds.), *The Handbook of English Linguistics*, 318–342. Oxford: Blackwell.

Nuyts, Jan (2005). The modal confusion: on terminology and the concepts behind it. In Klinge, Alex & Henrik Høeg Müller (eds.), *Modality. Studies in Form and Function*, 5–38. London: Equinox.

Palmer, Frank (2001). *Mood and Modality*. Cambridge: Cambridge University Press.

Traugott, Elizabeth Closs (1989). On the rise of epistemic meanings in English: an example of subjectification in semantic change. *Language* 65: 31-55.

van der Auwera, Johan & Vladimir Plungian (1998). Modality's semantic map, *Linguistic Typology* 2.1, 79-124.

van der Auwera, Johan, Ewa Schalley & Jan Nuyts (2005). Epistemic possibility in a Slavonic parallel corpus: A pilot study. In Petr Karlik & Björn Hansen (eds.), *Modality in Slavonic languages: new perspectives*, 201–217. München: Sagner.

von Fintel, Kai (2006). Modality and language. In Donald M. Borchert (ed.), *Encyclopedia of Philosophy – Second Edition*. Detroit: MacMillan Reference USA.

Verstraete, Jean-Christophe (2007). *Rethinking the Coordinate-Subordinate Dichotomy. Interpersonal Grammar and the Analysis of Adverbial Clauses in English*. Berlin: Mouton de Gruyter.

White, Peter R.R. (2003). Beyond modality and hedging: A dialogic view of the language of intersubjective stance. *TEXT* 23: 259–284.

White, Peter R.R. & Motoki Sano (2006). Dialogistic positions and anticipated audiences – a framework for stylistic comparisons. In Karin Aijmer & Anne-Marie Simon-Vandenbergen (eds.), *Pragmatic Markers in Contrast*, 189–214. Amsterdam: Elsevier.