ENFOQUES TEÓRICOS Y PRÁCTICOS DE LAS LENGUAS APLICDAS A LAS CIENCIAS Y A LAS TECNOLOGÍAS

Ana Bocanegra Valle,
M. Carmen Lario de Oñate,
Paloma López Zurita
(Eds.)

Cádiz: Universidad de Cádiz, 1999

Depósito legal: S. 575. 1999
ISBN: 84-89609 – 32 – 2

Tesitex, S.L.
C/ Melchor Cano, 15
37007 Salamanca

2.16. MODAL FREQUENCY IN ENGLISH SCIENTIFIC TEXTS

Jordi Piqué-Angordans
J.-Vicent Andreu-Besó
Universidad de Valencia

Introduction

When dealing with scientific literature, authors often use modality to escape from an absolute involvement with what they believe may or may not be the answer to their questions. Their search to bridge gaps in existing research forces them to be cautious in the use of their language, and thus we find ourselves –whether translating or simply interpreting a scientific text– in the midst of fuzzy and imprecise language. This is why it is so important to have an accurate grasp of these ways of expression, both in terms of understanding the cold meaning of a given sentence but, most important of all, to try to capture the author's intention, which often goes beyond the written word, when he/she uses a given modal verb to express his/her message. Butler (1990: 139) has rightly observed that "modals present traps for the unwary or linguistically unsophisticated reader or writer of scientific texts," and hence, at times, a modal "is crucial for the interpretation of the text", and at other times, "the omission of the modal would make very little difference to the overall sense."

Modals have been the object of study for a good number of years, especially ever since Huddleston and colleagues (Huddleston, Hudson, Winter and Henrici, 1968; and Huddleston, 1971) analyzed them. In recent years, every now and then a new text appears in which modals are assessed and studied (Palmer, 1979 [1990] and 1986; Perkins, 1983; Coates, 1983, among others). However, we hope to shed some new light for both scholars and young writers alike, by approaching a study of modals using as a reference a large corpus of scientific abstracts. Besides, new advances in computer science and retrieval processes have made it easier to come up with larger corpora for the benefit of their analysis. Authors (e.g., Lock 1988; Salager-Meyer 1992) have already pointed out that the abstract section in a scientific paper is read more receptively than the RA itself. We might further add that full range research has been made available through repertoires of abstracts, often the first piece of bibliographical material researchers resort to as a very useful vehicle of information.

Modal Verbs and Modality

Authors agree on the fact that subjectivity is an essential criterion for modality, since it is concerned with attitudes and opinions, and thus not related to the verb alone but to the whole sentence. Modality, therefore, could be defined as "the grammaticalization of speakers' (subjective) attitudes and opinions" (Palmer, 1986: 16). Nevertheless, it is also important to consider, especially in scientific texts, the ways in which others may report the expression of their convictions. In addition, whether these convictions are subjective or reported, the study of modality cannot be confined to non-factuality; as Palmer (1986: 18) remarks, "there are good reasons for handling factual statements together with opinions and judgements."

Quirk et al. (1985: 221-230) list the following modal auxiliaries, together with their contracted and uncontracted negatives: *can/could, may/might, must, need, have (got) to, ought to* and *should, will/would, shall*. Besides *need*, Radford (1997: 516; author's own block letters) also includes *dare* "when followed by a **bare** (*to*-less) infinitive complement." Palmer (1979 [1990]) further brings up the idea of epistemic (EpM), deontic (DeM), and dynamic modality. However, for the purpose of this paper, following Palmer's *Mood and Modality* (1986), our analysis focused on the dual modality type, that is, EpM and DeM.

According to Palmer (1979 [1990]: 50), EpM is understood as "the modality of propositions, ... rather than of actions, states, events, etc." Modals are used in various and complex ways in scientific texts, especially making generalizations about what is possible and what is not, on the basis of

observation of what actually happens. This complexity of use is lessened when we consider EpM and the fact that it contains two basic degrees of possibility (*may*) and necessity (*must*). Thus it is the most clearly recognizable from the other two types of modality since it "is to be interpreted as showing the status of the speaker's understanding or knowledge" (Palmer, 1986: 51).

Since in English the same forms are used for both types of modality, EpM and DeM, there is a certain amount of difficulty distinguishing them. It must be held in mind, however, that while both seem to share the involvement of the speaker, EpM, on the one hand, is basically concerned with "language as information, with the expression of the degree or nature of the speaker's commitment to the truth of what he says" (Palmer, 1986: 121). DeM, on the other hand, "is concerned with language as action, mostly with the expression by the speaker of his attitude towards possible actions by himself and others" (ibid.). In addition, according to Simpson's (1990: 67) appreciation, DeM "is concerned with a speaker's attitude toward the desirability (or nondesirability) of certain actions or events." Based on his research, Simpson believes that while the epistemic system comprises a continuum of commitment ranging "from possibility to logical necessity" (ibid.; his italics), the deontic system, by contrast, "extends from permission through obligation to requirement."

Object of Analysis

In this study we took the term modal verb in its widest sense; thus, for our analysis and frequency count we included those auxiliaries understood by most authors as modals. According to Palmer (1979 [1990], and particularly 1989: 33), the following can be formally defined as modals and were included in our initial frequency analysis: *can, may, must*, *ought (to)*, *will, shall, be able to;* marginally, *need* and *dare* (including *might, could, would* and *should*); also those auxiliary verb properties peculiar to them, such as negative forms with *n't*, tags, etc. In a second phase of our analysis, we concentrated on six basic modals for the purpose of comparing our findings with those studied by Salager-Meyer (1992), namely *can*, *could*, *may*, *might*, *must*, and *should*. In addition, a collocational approach drew us closer to the understanding of both types of modality being studied.

The aim of this study, therefore, was to complement existing literature on modality across disciplines, such as the work by Huddleston (1971), Butler (1990), Simpson (1990), Salager-Meyer (1992), López Folgado (1994), among others. The study of sentences in which modal verbs are involved requires the use of tools which go beyond paper and pen. Thus the need for this experiment was further enhanced in the sense that it includes a thorough computational study of large data, versus what has been analyzed until now, that may be a more detailed analysis but based on smaller corpora. For example, Salager-Meyer (1992) bases her study on a varied corpus of 84 abstracts, 49 RAs, 21 reviews, and 14 case reports; Kourilová (1993) analyzes two different corpora of 50,000 words each, representing 5 RAs and two chapters from two different textbooks; in another study, Salager-Meyer (1993) uses 29 RAs, 21 reviews, and 14 case reports; Valero Garcés and Calle Martínez (1997), study 2 abstracts in Spanish vs. 2 in English, to name but a few.

Corpus Description and Materials

Our aim was an attempt to widen the scope of this type of analysis through the comparison of three different corpora: our first and largest corpus was made up of 13,131 medical abstracts (1-HS-ABS), representing a total of 2,452,723 words. For the purpose of comparing different chunks of texts, we also analyzed a second corpus of 20 health science RAs (2-HS-RAs), which amounted to a total of 77,497 words in length. And for further comparison, a third corpus was randomly selected from linguistic and literary texts (3-LIT/LING), for a total of 227,767 words.

Analysis of Results

Our initial frequency count of modality in the three corpora, both in health sciences (abstracts and RAs) and linguistic and literary essays, is shown in table 1. (The percentages shown are established in reference to the total number of modal verbs in each of the three corpora.)

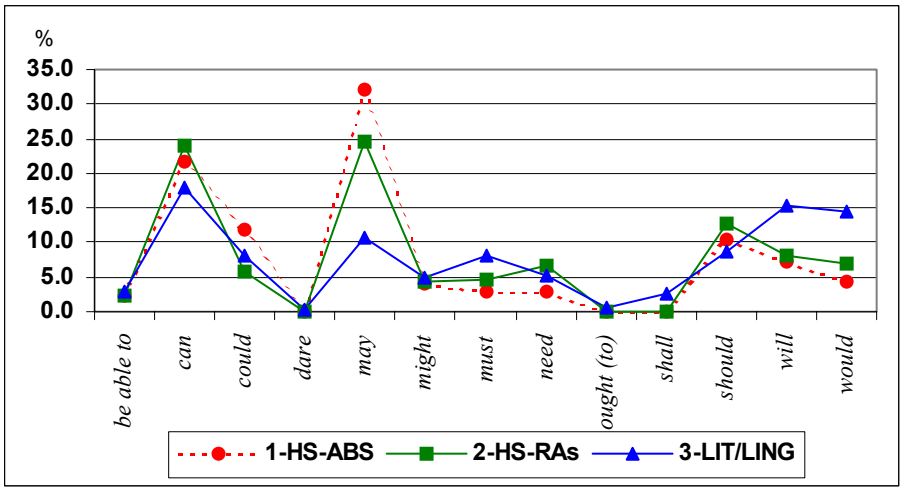
Table 1. Frequency of modals in the three corpora in percentages (relative figures).

	1-HS-ABS		2-HS-RAs		3-LIT/LING	
Modals	**f**	**%**	**f**	**%**	**f**	**%**
be able to	325	2.45	20	2.17	83	2.87
can	2,883	21.77	221	24.02	519	17.92
could	1,589	12.00	54	5.87	236	8.15
dare	0	0.00	0	0.00	10	0.35
may	4,263	32.19	225	24.46	313	10.81
might	524	3.96	40	4.35	139	4.80
must	367	2.77	43	4.67	238	8.22
need	373	2.82	60	6.52	150	5.18
ought (to)	5	0.04	0	0.00	16	0.55
shall	6	0.05	1	0.11	75	2.59
should	1,386	10.46	117	12.72	249	8.60
will	958	7.23	74	8.04	447	15.44
would	566	4.27	65	7.07	421	14.54
Totals	**13,245**		**920**		**2,896**	

A first reaction to these results should be drawn from the fact that corpora 2-HS-RAs and 3-LIT/LING yielded a similar use of modality, being much lower in corpus 1 (abstracts). In reference to the total number of words, the amount of modals, 920 in corpus 2-HS-RAs, and 2,896 in corpus 3-LIT/LING, represents that they were used 11.6 times per 1,000 words in corpus 2-HS-RAs, and 12.7 per 1,000 words in corpus 3-LIT/LING. In corpus 1-HS-ABS, however, the amount of 13,245 represents that only 5.4 modals per 1,000 words were used.

In reference to individual use of modals (Table 1), the differences appear more markedly. While *may*, *can*, *could* and *should*, in this order, appear to be most frequently used in abstracts (corpus 1), and *may*, *can* and *should* in RAs (corpus 2), authors of linguistic texts and literary critics (corpus 3) would appear to prefer the use of *can*, *will*, *would* and *may*, in this order. These results are graphically expressed in the figure 1.

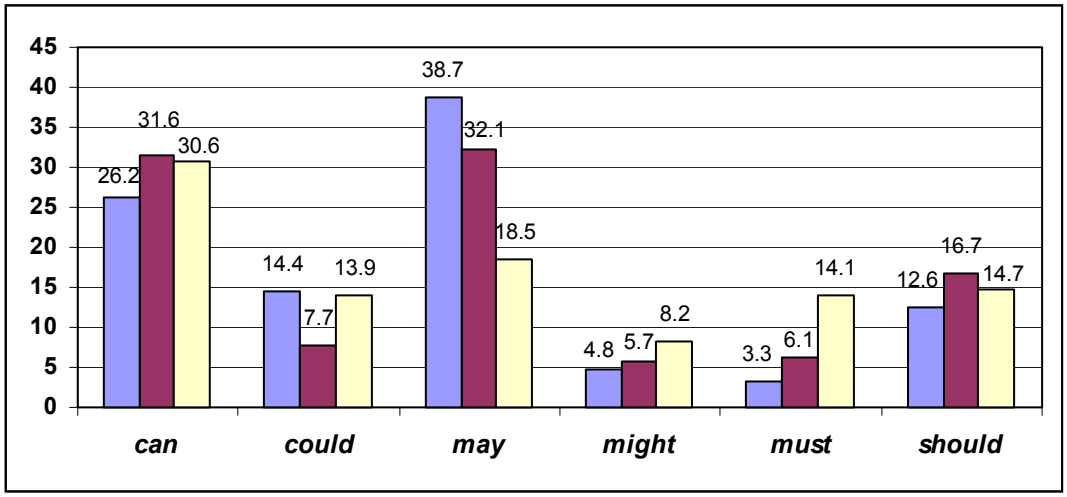
Figure 1. Frequency of use of modals in percentages (relative figures).



In addition, our findings in regard to *can*, *may*, and *should*, three of the most commonly used modals in scientific writing,[1] are particularly significant when compared to other researchers' findings. As we indicated above, we compared our results to those drawn by Salager-Meyer (1992). Taking into account the six modal verbs mentioned, she gives the following percentages for the first three most frequently used (p. 98): 37.7% (*may*), 23.2% (*can*), and 21.7% (*should*). Using Salager-Meyer's criteria of analysis of the six modals, our frequency count (see figure 2 below) in the corpus of abstracts (1-HS-ABS) yielded the following percentages for the same three modals: 38.7% (*may*), 26.2% (*can*), and 12.6% (*should*).

In regard to research papers (corpus 2-HS-RAs), however, our findings show an almost even percentage of use of the first two modals –i.e. 32.1% (*may*) and 31.6% (*can*) (see Table 1)–, followed at a distance by *should* with 16.7%. In Salager-Meyer's (1992: 97) corpus of research papers, in turn, the use of *may* almost doubles that of *can* and *should*.

Figure 2. Percentage of use of six modals (relative figures).



As expected, the findings from our third corpus (3-LIT/LING) show a different picture, with *can* (17.9%) being the most widely used modal verb, followed by *will* (15.4%) and *would* (14.5%), as shown in table 1. Comparing these results with Salager-Meyer's findings, *may* and *should*, with 18.5% and 14.7%, respectively (figure 2), while *can* would be preferred by linguists and literary critics with 30.6%.

It is also noteworthy the difference shown in the use of *may* and *must*: while in corpora 1 and 2 *may* represents over 30% of use, in corpus 3 it represents just over 18%. A reversed trend, however, is shown with the modal *must*.

In an attempt to distinguish the type of modality (EpM or DeM), we studied collocationally the frequency of appearance of each modal verb listed (a few samples are given in figures 3 and 4 below). Our results basically agreed with previous studies in terms of an absolute predominance of EpM in health science texts. Kourilová (1993), for instance, emphasizes this predominance, while Varttala (1999: 184) points out that modal auxiliaries are "the word-class most often associated with epistemic meaning." DeM, in turn, appeared with a certain degree of frequency in corpus 3-LIT/LING, which also agrees with Simpson (1990) in his analysis of literary-critical discourse where he finds, along with EpM, numerous examples of DeM.

[1] Adams Smith (1984) finds *may* and *should* as the most frequent, while *can* is rare. In fact, the possibility of *can* being used in epistemic modality expressions is not shared by all researchers (e.g. Coates, 1983). This may partially explain why Varttala (1999: 184) considers *could*, *may* and *might*, but not *can*, as "the most central epistemic modal auxiliaries expressing tentative possibility." Huddleston (1971), however, finds the modal *can*, whether in epistemic or deontic mode, to be common in physics papers.

In addition to the high frequency of modality, it is important to stress the fact that in corpora 1-HS-ABS (health science abstracts) and 2-HS-RAs (health science RAs) EpM predominates. These epistemic expressions basically point at the possibility derived from the author's knowledge and understanding of the subject dealt with, based on his/her research (see figure 3).

Figure 3. Epistemic modality (EpM) collocates (from corpora 1-HS-ABS & 2-HS-RAs).

As with the first hypothesis, it	**could**	be argued that this prediction ...
the prognostic contribution of CD4+ cell count	**could**	not be fully evaluated because only 30% ...
Recent research shows that patient behavior	**may**	be misleading and result in underestimating ...
The currently available generic measures	**may**	not provide the necessary breadth ...
Self-monitoring and peer support	**might**	provide cues necessary to correct ...
Measuring fluid intake as a continuous variable	**might**	also reduce measurement error.
A similar complementation approach	**should**	also be useful for revealing the rules of RN ...
These results suggest that the effect of DDAVP	**should**	be reassessed in the treatment of ...

On the other hand, corpus 3 (linguistic and literary essays) introduces a good number of DeM samples throughout. Figure 4 shows a few examples:

Figure 4. Deontic modality (DeM) collocates (from corpus 3).

No one with self-consciousness	**can**	ever do anything drastic in life ...
This second trend does not go, and	**cannot**	go, further back than Ibsen's works ...
But critical (discourse) analysis	**cannot**	remain indifferent to questions of truth ...
It is conventional and one	**may**	doubt that it was written with full conviction.
This may not be your cup of tea, and you	**may**	consider that I am making mountains our of ...
The point is made that the revolutionaries	**must**	make a determined effort to do without ...
... any irrelevant or unnecessary facts	**must**	be ignored in order to grasp the essential ...
... strength has caught up with ours we	**shall**	have to fight a third world war, which ...

Conclusion

Basically, EpM predominates in health science texts in general, through which the author expresses a possibility based on his/her subjective conjectures. The rhetorical importance of the density of epistemic modality resides precisely in the degree of author involvement; while at the same time, the more speculative a text is considered to be, the more EpM predominates. This type of modality, along with various hedging devices, opens the pragmatic spectrum for the author of distinguishing necessity from possibility, and thus it becomes both obvious and relevant to see it more abundant in specialist texts than in popularizations and, as we have seen, in texts from a different type of speciality, such as texts from literary criticism and linguistics.

DeM, however, particularly present in corpus 3, would refer to an expression of necessity, derived from the subjective belief that an obligation must be imposed on self or on others. In this sense, the frequency of DeM expressions in the literary and linguistics texts we have analyzed has been more clearly recognizable as a specific device, along with EpM, in their written expression. The degree of difficulty in determining the exact meaning of deontinc or epistemic modality depends upon several variables, namely, the specific topic or degree of speculation, the author's commitment to a particular opinion on the subject studied and its expression, and finally the reader's understanding of such circumstances. Furthermore, the linguistic analysis may or may not correspond to the true intention of the author. Thus we believe that the data we have presented above deserves further study in terms of clearly ascertaining the rhetorical function of different modal expressions, especially in regard to an in-depth analysis of these two types of modality across disciplines.

Bibliography

Adams Smith, D. E. (1984). “Medical discourse: aspects of author’s comments.” *The ESP Journal* 3: 25-36.

Butler, C. S. (1990). “Qualifications in science: Modal meanings in scientific texts.” In Nash (Ed.), pp 137-170.

Coates, J. (1983). *The Semantics of the Modal Auxiliaries*. London and Canberra: Croom Helm.

Huddleston, R. D. (1971). *The Sentence in Written English: A Syntactic Study Based on an Analysis of Scientific Texts.* Cambridge: CUP.

Huddleston, R. D., R. A. Hudson, E. O. Winter & A. Henrici (1969). *Sentence and Clause in Scientific English.* London: University College of London.

Kourilová, M. (1993). "Epistemic modality in written scientific discourse." *Unesco ALSED-LSP Newsletter* 15: 4-18.

López Folgado, V. (1994). “Modalization in English scientific texts.” In S. Barrueco et al. (Eds.). *Lenguas para Fines Específicos (III). Investigación y Enseñanza*. Alcalá de Henares: Universidad de Alcalá, Servicio de Publicaciones, pp. 177-187.

Nash, W. (Ed.) (1990). *The Writing Scholar. Studies in Academic Discourse.* Newbury Park: Sage Publications.

Palmer, F. R. (1979 [1990]). *Modality and the English Modals*, 2nd ed. London and New York: Longman.

Palmer, F. R. (1986). *Mood and Modality*. Cambridge: Cambridge University Press.

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

Quirk, R., S. Greenbaum, G. Leech & J. Svartvik (1985). *A Comprehensive Grammar of the English Language*. Harlow, Essex: Longman.

Radford, A. (1997). *Syntactic Theory and the Structure of English. A Minimalist Approach*. Cambridge: Cambridge University Press.

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: 93-113.

Salager-Meyer, F. (1993). "Imprecision and vagueness (hedging) in today's medical discourse: Courtesy, coyness, or necessity?" *the ESPecialist* 14: 1-13.

Simpson, P. (1990). “Modality in literary-critical discourse.” In Nash (Ed.), pp. 63-94.

Valero Garcés, C., & M. C. Calle Martínez (1997). "Contrastive rhetoric in ESP: A cross-linguistic analysis of finite verb profiles in English and Spanish medical abstracts." *Unesco ALSED-LSP Newsletter* 20: 22-36.

Varttala, T. (1999). “Remarks on the communicative functions of hedging in popular and specialist research articles on medicine.” *English for Specific Purposes* 18: 177-200.