

## Statistical correspondence analysis: selection of modal verbs in four groups of texts

**Cándida Fernández Morales**

**Ángel Felices Lago**

Universidad de Granada

### Abstract

The aim of this paper is to observe the position of four defined groups of texts according to their modal verbs outlines. A statistical correspondence analysis is applied to establish the relationships between them. Our corpus of 822 modal verbs (498 non-remote modal verbs and 324 remote modal verbs) is taken from two different genres (advanced textbooks and specialized magazine) and two specific areas (business / marketing and finance). A total of 40 texts have been analysed (20 from textbooks and 20 from *The Economist*, a well-known semi-specialized journal). Such an analysis helps us to know the selection trend in each group, suggesting that a sub-genre classification must be considered within the textbooks due to their final aim and not in relation to their study areas. This is illustrated by giving a brief outline of how such a trend is influenced by the rhetorical aim of the sources the texts belong to. This paper is part of a broader research project on verb phrases in the groups of texts mentioned above (Fernández Morales, C. 2000).

**Key words:** statistical correspondence analysis data base modal verbs verb phrases statistical package for social sciences (SPSS)

### 1. Introduction

As far as we know, correspondence analysis is a statistical technique rarely used in linguistic analyses. However, we apply it in this paper for comparing the modal verb profiles of a selected group of texts in order to find out similarities or dissimilarities among them.

It is a well-known fact that comparative studies on modal verbs are important to show the writer's attitude or mood and, obviously, this can be applied to a group of texts. That is, the communicative aim of a group of texts exerts influence on modal verb selection in such a group. A comparative analysis of modal verbs as the one shown here not only helps us to differentiate groups of texts or find similarities among them but it also shows us how their tendencies of selection are related to the general aim of different text types. In other words, the groups of texts can be identified by their selection of modal verbs. The total number of modal verb phrases found in our texts (822) represents 19.6% of all analysed finite verb phrases. It may be considered a low percentage but high in comparison with previous studies in the previous century, such as Barber (1962) 16% and Salager-Meyer (1992: 97) who found only 9%. However, two of the three groups of texts analysed by Wingard (1981: 60) represented 10%, 26% and 21% respectively.

### 2. Corpus

Our corpus of modal verbs (822) has been selected from 40 texts (see appendix). Twenty have been taken from *The Economist* and the other twenty from advanced textbooks in

two subjects: business-marketing and finance. So we have four groups of texts to compare according to the source and area. One group is made up of 10 texts from the business section of *The Economist* and the other ten texts come from the finance section of *The Economist*. The remaining two groups consist of ten texts from advanced textbooks (finance and marketing). It is important to note that the research has been conducted with the introductory part of selected chapters from the textbooks, so that the extension of textbook texts is similar to *The Economist* articles. We have followed Quereda (1993) for the classification of modal verbs: central modals, (*can, may, shall, will, must, ought to, could, might, should* and *would*), marginal modals (*used to, dare\** and *need\**) and semi-auxiliary modals *be and have (is to, be going to, have to, had better\**, and *have got to\**). All these verbs have been introduced in a data base. However, the asterisk indicates the modal verbs which have not been found in our texts. The modal locutions: *be + adjective (be bound to, be up to, be able to etc.)* have not be included. A detailed specification of modal verbs and the method of analysis can be found in Fernández Morales (1999: 197-199)

### **3. Software tools and method of analysis**

Microsoft Access has been used to make our own data base. The modal verb phrases found in the four groups of texts and known as remote and non-remote have been collected taking into account their possible modifications with specification of the auxiliary modal verb and the text they appear in.

The Statistical Package for Social Sciences (SPSS) has been chosen to carry out correspondence analysis which shows the contingency table of the four text groups and modal verb types. The Excel Spreadsheet has been used to make the previous exploratory analysis, the graphic of profiles and the correspondence analysis using the SPSS output. Both statistical techniques, profile graphic and correspondence analysis, are two perspectives or different points of view to give the same information.

### **4. Results**

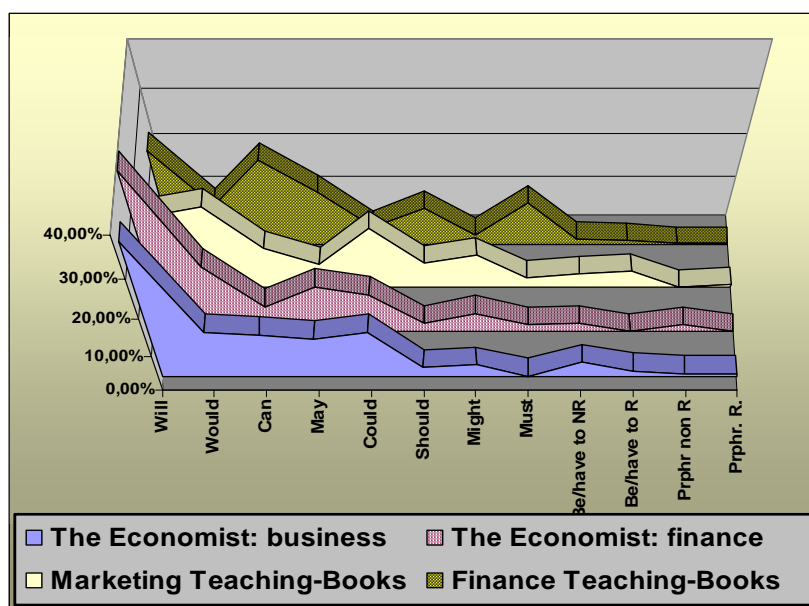
It has been found that *The Economist* has a higher frequency of modal verb phrases than the textbooks except for the modal remote passive verb phrase whose frequency is quite high in the marketing textbooks but hardly appears in *The Economist* texts. Comparing areas, Finance (both textbooks and *The Economist*) has the higher percentage of modal verb phrases. This might be due to the uncertain or insecure topic environment. Consequently, the author is inclined to use prediction and probability expressions. Below, results of profile graphic and correspondence analysis are shown: The same information about modal verb profiles is given but from different perspectives. From our point of view, the results of correspondence analysis are better because similarities and dissimilarities among the four groups of texts are evident with reference to modal verb frequencies.

## 4.1. Profile graphic

Figure 1 represents the modal verb profile for each one of the four groups of texts. Each profile is represented by a design. The profiles of modal verbs from *The Economist* business and finance texts are placed below. The marketing and finance textbooks are located above. There are also indications of modal verbs in the texts analysed (Notice NR for *non-remote* and R for *remote*) and percentages.

### Profiles

Figure 1.



### Commentaries

At first sight, looking at *The Economist* profiles we can observe that the two groups of texts from *The Economist* show a similar form (though not equal). This contrasts with the irregular aspect of profiles in marketing and finance textbooks. Look at both profiles, their forms are quite different. That is, there is a closer similarity between the profiles of *The Economist* texts and is not the case in the profiles of textbooks. Such similarities and dissimilarities or contrasts between the groups of texts are quite clear according to the frequencies of central modals (*will, would, can, may, could, should, might and must*) which show higher frequencies. Notice that remote modal verbs: *would, could* and *might* are more frequent in the marketing texts of textbooks than in remaining groups. This may be due to the number of case studies in marketing textbooks and consequently a clear tendency to remote verb phrases can be found. (Case study texts explain marketing strategies which a sample company has followed to improve profits, for instance, widening its market share, facing competitors or introducing a new product into the market. There is a history of the company or a reference to what happened to the company in the past and how it acted.) But, nevertheless, the graphic also shows that *might* appears in *The Economist* texts though percentages are lower in comparison with the marketing textbooks. In short, if in figure 1 we can see modal verb profiles of the

four groups of texts with their similarities and dissimilarities, we can also see the same information, but in a different way, using correspondence analysis (figure 2.).

#### 4.2. Correspondence analysis

In a correspondence analysis each one of the four text groups is placed in a position according to their modal verb profiles. The rhombuses represent modal verbs. The circles represent each one of the four groups of texts being placed in a position within the graphic according to the modal verb profile contributions to factors 1 and 2 (the axes of the plane) respectively (see figure 2).

#### Commentaries

The correspondence analysis allows us to observe similarities and dissimilarities among the texts analysed according to their positions on the graphic. The greater the distance among the circles the greater the contrast. In figure 2 we can observe that the position of the two sections of *The Economist* are close to each other because their modal profiles are similar, whereas marketing and finance textbooks texts are widely separated. Their circles are opposite. And, similarly, both areas of *The Economist* texts are far away from the two areas of the textbooks. At first sight, this can be explained by saying that *The Economist* texts have a common goal, and follow the editorial guidelines. On the contrary, textbooks share the same purpose: to educate readers, but they are also different because finance texts are more theoretical than marketing texts, mostly case studies.

#### Correspondence Analysis

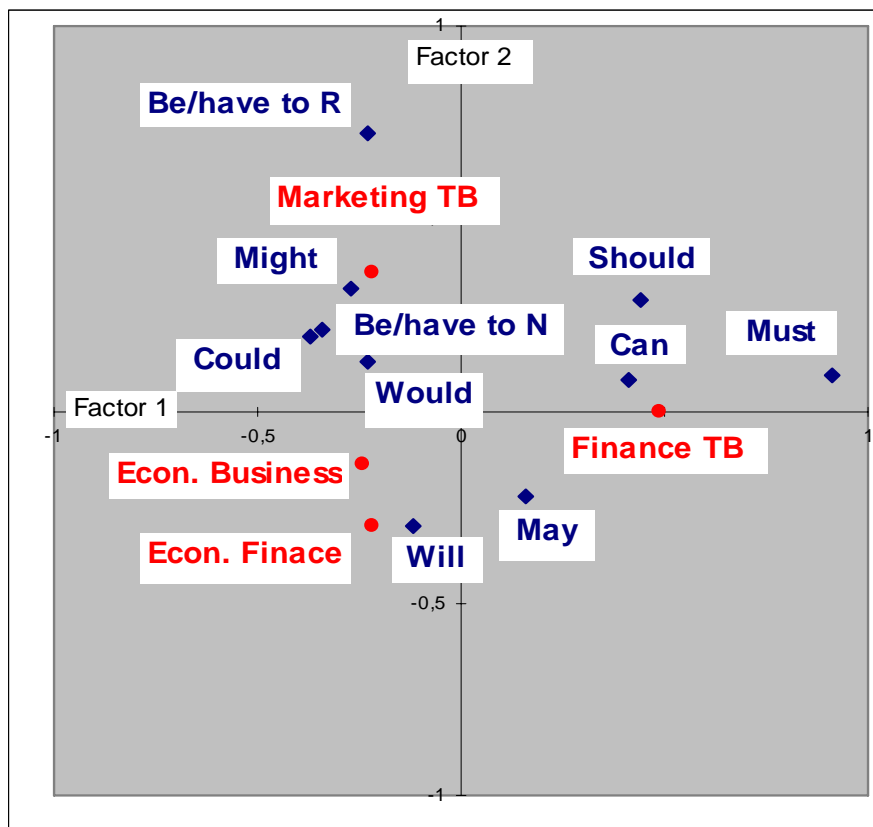


Figure 2.

### 4.3. Interpretation of graphs.

The position of rhombuses on the right, where *should*, *must*, *can* and *may* are placed, shows advice, suggestion, logical prediction and deduction, and the texts of finance textbooks have a clear tendency to use them. Notice their neutral position on axis *y*. They oppose the remote probability above, on the left, where *could*, *would* and *might* are placed and the marketing texts positioned. Below, on the left, appears *will*, which means a likely future prediction and where *The Economist* texts are positioned. So we have, on the right, suggestion, advice and logical prediction, and above, on the left, predictions about past and remote probability; below, neutral forms of predictions.

### 4.4. Reasons for such a selection of modal verbs.

It can be said that in knowledge modality (*will* and *may*). *will* is the most widely used in *The Economist* texts. As *The Economist* texts generally offer up-to-date news and try to attract readers' attention, the authors of the articles make predictions quite probable to happen in a near future, so the readers can test whether the predictions happen or not. So *will* is quite frequent. Prediction statements are a way to make readers become interested in the story as they can test in a short time whether predictions are true or not. In *The Economist* the functions of prediction and probability are quite frequent as they refer to a near future.

The tentative (more hypothetical) prediction which includes remote modal verbs (*would*, *could* and *might*) is more frequent in marketing textbooks, mainly in the case-study texts. The tendency of remote modal verbs in these texts is due to past predictions about a past or future in past events. As it has already been said, the topics are "real" facts about companies that carried out marketing techniques to improve their profits on some occasion in the past.

The influence modality, *should* and *must*, is likely to appear in finance textbook texts as they are quite frequent in instructions and the writer tries to influence the readers to do something. *Should* as an advice and *must* meaning *it's necessary* or *you're obliged*.

## 5. Conclusions

On the one hand, correspondence analysis has shown the position of the four text groups referring to their modal verb profiles better than the frequency profile graphic. When comparing texts, a correspondence analysis is a good statistical method to observe the tendencies of text groups as well as their linguistic similarities or dissimilarities. And such a method can be applied to other linguistic features and text types. On the other hand, correspondence analysis has shown the existence of two sub-genres the textbooks selected: theoretic and practice instruction (a wider study on verb phrase selection confirms that). The general purpose for which the texts were designed exerts a powerful influence on writers' selection of modal verbs. Finally, the results of the selection of modal verbs can be applied to ESP teaching if the instructor designs the appropriate activities.

## Appendix

### Finance textbooks

- Copeland, T., T. Koller & J. Murrin (1990). *Valuation: Measuring and managing the value of companies*. New York: John Wiley & Sons
- Keown, J. A., D. F. Scott, J. D. Martin & W. J. Petty (1996). *Basic financial management*. New York: Prentice Hall.
- Livingston, M. (1990). *Money & capital markets: financial instruments and its uses*. Englewood Cliffs: Prentice Hall.

### Marketing textbooks

- Johnson, G. & K. Scholes (1984). *Exploring corporate strategy: text and cases*. Nueva York: Prentice Hall.
- Swayne, E. L. & P. M. Ginter (1989). *Cases in strategic marketing*. New York: Prentice Hall

### *The Economist*

- The Economist* (1992). London 323, 7762; 325, 7783.
- The Economist* (1993). London 326, 7795; 7796.
- The Economist* (1996). London. 341, 7984; 7986; 7987; 7988; 7990; 7992.

## References

- Barber, C. L. (1962). Some measurable characteristics of modern scientific prose. In: Swales (1985) *Episodes in ESP*. 1-16. Oxford: Pergamon Press.
- Fernández Morales, C. (2000). *Un ejemplo de análisis de registros: la frase verbal en la revista The Economist*. PhD thesis. Servicio de Publicaciones Universidad de Granada.
- Quereda Rodríguez-Navarro, L (1997). *A morphosyntactic study of the English verb phrase*. Granada: Servicio de Publicaciones Universidad de Granada
- Salager-Meyer, F. (1992). A text-type and move analysis study of verb tense and modality distribution in medical English abstracts. In: *English for Specific Purposes*, 11, 2, 93-113. New York: Pergamon Press.
- Wingard (1981). Some verb forms and functions in six medical texts. In: *English for academic and technical purposes: studies in honour of Louis Trimble*. 53-64. Selinker, Tarone and Hanzely (eds.) Rowley: Newbury House Publisher.