Thematic Organisation as a Criterion for Establishing English Text-types*

Mohsen GHADESSY
Baptist University, Hong Kong

ABSTRACT

In recent years there have been a number of attempts by applied linguists to establish what has been called English text-types. Biber and Finegan (1986), for example, provide a list of 9 text-types based on the frequency of some lexical items, phrases, and sentences. They use some advanced statistical methods, namely factor and cluster analyses to reach their conclusions. The purpose of this study is to show that Biber and Finegan's criteria are related mainly to what has been called the ideational (field) and, to a lesser extent, the interpersonal (tenor) meanings in a text. The consideration of a third kind of meanings in a text, i.e. textual meanings (mode) has been neglected by them. A detailed analysis of one feature belonging to the textual meanings, namely thematic development, in five different text-types given by Biber and Finegan, is then presented to show how different clustering of the same text-types may result. The argument of the paper is that Biber and Finegan deal with surface phenomena, i.e. vocabulary and a few grammatical structures as characteristics of English text-types. What is also needed, however, is a consideration of the underlying textual processes that contribute to the meaning(s) of a text. It is concluded that to-date the most complete and satisfying way to establish English text-types is provided by the concept of register and register analysis.

1. INTRODUCTION

With the ever-increasing popularity of discourse analysis and text-linguistics in recent years, a number of linguists have tried to establish what has been referred to generally as functional varieties and more precisely as registers (Halliday, 1994), genres (Swales, 1990) and text-types (Biber and Finegan, 1986) in the English language. All these linguists have been influenced

by a formal (structural), functional or a combination of formal and functional approaches to the analysis of discourse. Based on such a distinction, Schiffrin (1994) distinguishes between six different approaches to the analysis of discourse, namely speech act theory, interactional sociolinguistics, the ethnography of communication, pragmatics, conversational analysis, and variation analysis.

Crucial to the establishment of any functional variety is the definition of text and the criteria that have been used to delimit one functional variety from another. Some text-linguists (Swales, 1990; Bhatia, 1993; Biber, 1995) do not specifically define «text/a text» but their criteria for text analysis imply that they are following a formal/structural approach, namely, that a text is a unit larger than a sentence (clause) – in fact it is a combination of a number of sentences (clauses) or a number of elements of structure, each made of one or more sentences (clauses). In such cases the criteria for distinguishing between two texts are the presence and/or absence of elements of structure or types of sentences, clauses, words, and even morphemes such as -ed, -ing, -en in the two texts. Whether texts are analysed in terms of some elements of structure or a number of sentences (clauses) that can then be broken down into smaller units – a top-down analysis – or in terms of smaller units such as morphemes and words that can be put together to build the larger unit of text – a bottom-up analysis – we are still dealing with a formal/structural theory and approach to text analysis. As recent examples of this kind of approach we can cite Crystal (1991), Biber and Finegan (1986), and Biber (1988) for the more formal approach and Van Dijk (1986) and Swales (1990) for the more structural approach.

The rest of this paper will deal with units of analysis in register profiling (section ii), clause as the unit of analysis (section iii), thematic progression as a criterion (section iv), and a discussion of Biber and Finegan's clusters in the light of the findings for thematic organisation of five of their registers (section v). The conclusions of the study will then be presented in section vi.

2. LINGUISTIC FEATURES AS UNITS OF ANALYSIS

Crystal (op. cit.: 224) in his "stylistic profiling" advises that one should select "features" which have "the greatest variety identifying capability". For example the -eth verb ending is "a major characteristic of religious English". He proposes three evaluative criteria for each feature thus selected, i.e. (1) frequency of occurrence, (2) overall distinctiveness, and (3) level of precision for defining and identifying the feature. In relation to the above example, despite its low frequency of occurrence, -eth has a very high distinctiveness and it can be precisely described and identified. Each criterion is given "arbitrary values for calculating the stylistic distinctiveness of a variety or sample". A

crucial question that Crystal asks is "Where do the lists of features come from?" In his case they come from "a mixture of published descriptions and relevant analytical experience". In fact, the stylistic profiling of two functional varieties presented by him are based on a number of features first introduced in Crystal and Davy (1969).

Comparing a sample of legal English with news reporting in the press, written discourse for both, the following feature categories are selected. Each category is then sub-divided for the actual number of features used. The categories and an example of a feature, in brackets, in each are as follows:

For legal writing: graphetics (unbroken format), graphology (words in capitals), grammar (very long sentences), semantics (lexical repetition). For newspaper writing: graphetics (range of type sizes), graphology (comma omission), grammar (short paragraphs), semantics (unusual word formation). The category grammar in each of the varieties has the largest number of features.

After assigning a value for each of the three criteria of frequency, distinctiveness, and precision, the total value for each feature is calculated and the results are presented in histographic profiles. The visual displays for the legal sample and the news report are given in Figure 1.

Biber and Finegan (1986) provide an initial typology of English text types based on a multi-feature/ multi-dimensional approach (MF/MD). An important distinction made by them is between "genre categories" and "text types". The former "... are used to characterise texts on the basis of external criteria", the latter are defined in terms of "linguistic characteristics of text themselves" (op. cit: 20). 41 linguistic features are selected for the purpose of analysis: "... we include all linguistic features that have been identified in previous research on functional markers of different styles, modes, or registers" (op. cit.: 26). Biber and Finegan (ibid) use a much larger sample of 545 texts coming from 16 spoken and written genres, namely spoken (face-to-face conversation, interviews, planned speeches, and so forth) and written (press, academic prose, professional letters, and so forth).

Initially they carry out a factor analysis of all the features in the samples. They state that "A factor analysis identifies linguistic features that co-occur with a high frequency in texts" which is "... an indication of a common communicative function shared by these features" (p. 27). In their analysis both the presence and absence of a feature(s) are significant and hence the factors have positive or negative loadings. Based on three observed significant factors, three dimensions are established for distinguishing the text types, namely "Interactive versus Edited Text", "Abstract versus Situated Content", and "Reported versus Immediate Style" (pp. 31-33). Genres are then plotted, on the basis of the factor scores, within the two poles of each dimension. For example, face-to-face conversation is at the pole of Interactive and press is at the pole of Edited Text with professional letters somewhere in the middle.
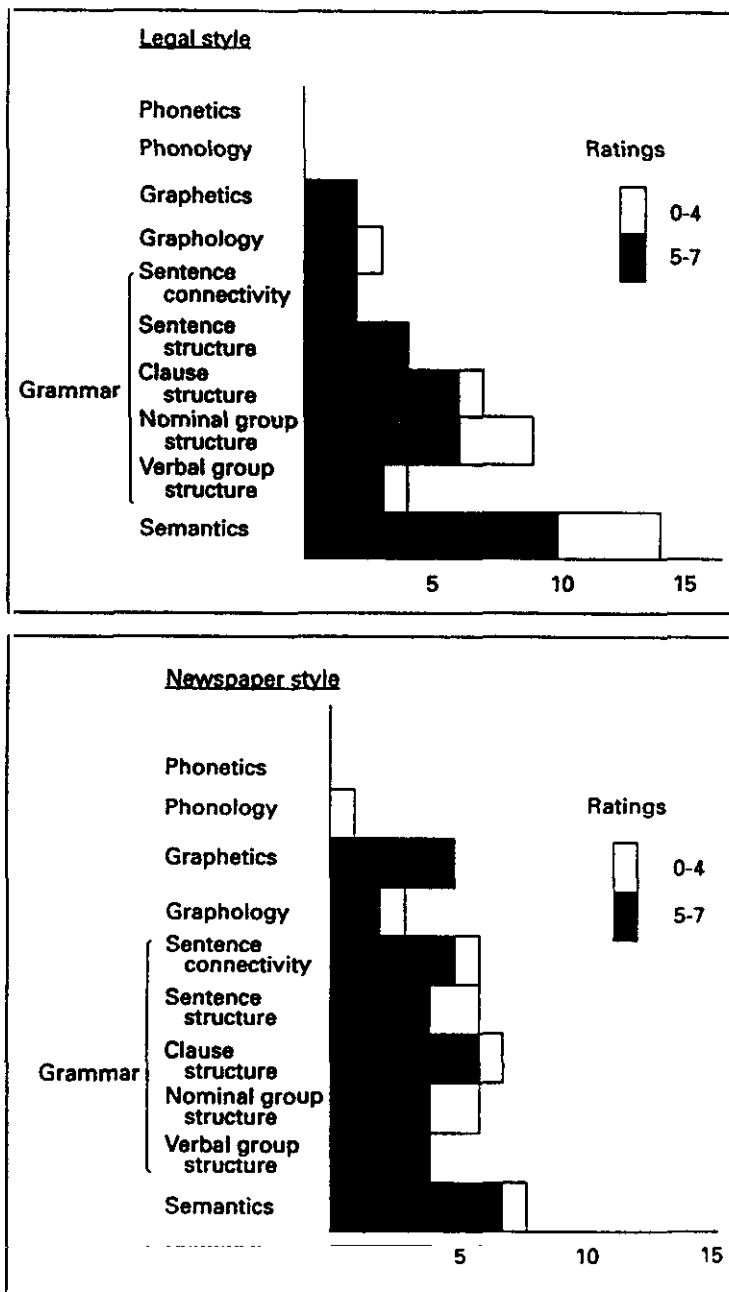
Mohsen Ghadessy

Figure 1.   Histographic profile of stylistic features: legal and newspaper style (Crystal, 1991

After the factor analysis, Biber and Finegan (ibid) carry out a cluster analysis "... to group texts on the basis of their use of linguistic features". Such an analysis would result in English "text types", which according to them are "... groupings of texts that are in fact similar in their linguistic form, regardless of external criteria" (p. 33). Thus 9 clusters with an "interpretive label" are established (pp. 38-39). Each cluster includes the genres and the percentages of texts which fall in the cluster. For example, cluster one labelled "Immediate Interaction" with a total of 44 texts includes the genres of face-to-face conversation (51%), telephone conversation (65%), interview (4%), and broadcast (5%). The composition of the clusters by genres is displayed in the Appendix.

3. CLAUSE AS THE UNIT OF ANALYSIS

According to Halliday (1982: 209) "Text is the process of meaning; and a text is the product of that process." The most important issue for a text linguist, then, is (1) given the product, how do we go about "breaking" it into its constitutive units, or (2) in "making", i.e. creating text, what meaning processes do we go through. These two questions are inseparable like the two sides of a sheet of paper. The Systemic Functional Grammar of Michael Halliday (1994) provides the tools for answering the above questions. Below, we will deal with how, according to SFG's guidelines, texts can be broken into their constitutive parts. This implies that we have to have a unit of analysis.

Unlike the linguists referred to above, the SFG's unit is the clause, which mediates between grammar and discourse and, like many other linguistic primes, has "psychological reality". Halliday (ibid) convincingly argues and shows that a clause and a text are similar in many different ways, namely both have structure, coherence, function, development, and character. He concludes that "It is no accident that it is possible to illustrate so many of the relations in a text by reference to relations in a clause." and "A clause is a kind of metaphor for a text - and a text for a clause". Support for the sentence (clause) as the unit of analysis also comes from Sinclair (1993) who states that "The sentence is regarded as the likeliest unit to carry the status of «text of the moment»".

Accepting a clause as the unit of analysis for register profiling has certain consequences, i.e. all the meaning-making processes in a clause, and subsequently in a text, have to be considered. Halliday's dictum that "text is the process of meaning" must guide our methodology for making or breaking texts. The meaning making processes comprise the operations of three metafunctions in any language, namely the ideational, the interpersonal, and the textual (Halliday, 1994).

A main criticism of major attempts for register profiling, such as the one by Biber and Finegan, is that only the operations of the first two metafunctions are considered. The third metafunction does not have any role. This may be partially due to the ease with which linguistic features exemplifying the ideational and/or interpersonal metafunctions are selected. A reliable description of English Grammar, for instance that of Quirk *et al.* (1985), would provide the text linguist with all the features needed for this purpose. However, descriptions of processes for textual meanings are less well developed and hence more cumbersome to use, which lead to the observation by Schiffrin (1994:viii) that discourse analysis, namely investigation of the patterns beyond the sentence (clause), "... still remains a vast and somewhat vague sub-field of linguistics".

4.   THEMATIC ORGANISATION AS A CRITERION

One of the systems in the textual metafunction of language, as defined by Halliday (1994), is the Theme/Rheme options made for each individual clause in spoken and written language. Each clause conveys a message that has two parts, i.e. what comes first or the Theme, and what comes last or the Rheme. The Theme usually contains given information and the Rheme new information. It should, however, be pointed out that the system of Theme/Rheme is different from the system of Given/New in the Hallidayan approach. Halliday (1994:38) distinguishes between Theme/Rheme and Topic/Comment (Given/New) by saying that "...the label «topic» usually refers to only one particular kind of Theme ... and it tends to be used as a cover term for two concepts that are functionally distinct, one being that of Theme and the other being that of given". Halliday (op. cit: 61) continues that "The choice of clause Themes plays a fundamental part in the way discourse is organised; it is this in fact, which constitutes what is often known as the «method of development» of the text."

The "method of development" was introduced by Fries (1983), who argues that it deals with the lexico-semantic content of Themes (op. cit: 119). "Thematic content correlates with the method of development of a text (and the nature of the text)." More recently, Fries (1992a:1) has shown that "Thematic content correlates with different genres" and that "Thematic content correlates with different generic structures within a text". Other relevant research has also identified the thematic development of clauses in several other registers of English (Fries, 1992b, Ghadessy, 1995).

5.   RECONSIDERING BIBER AND FINEGAN'S CLUSTERS

The typology of English text-types by Biber and Finegan is based on " ... a cluster analysis, which groups together texts that are maximally similar in their linguistic characteristics irrespective of their genre classifications" (Biber and Finegan 1986: 20). In the production of any text choices are made constantly in different systems and at various levels. The units chosen by Biber and Finegan for their typology are almost all below the sentence/clause level, e.g. choices between tense and aspect, between pronouns, adverbs, and adjectives, between subordination, nominalisation and passivisation, between questions, infinitives and prepositional phrases. As such Biber and Finegan's typology does not say anything about the underlying processes involved in the **creation** of any text-type or register.

An analogy may be useful at this point. Suppose you were given the task of classifying 20 different residential properties. You may classify them by analysing the materials out of which the houses or apartments are made and subsequently classify them into a number of clusters based on the percentage of concrete, bricks, or wood in each construction. But there are other ways of classifying the residential properties, by the number of rooms, by the design of the structure, by the proximity to the downtown area, and so forth. In sum, Biber and Finegan's analysis gives us the answers to the question of what is in the text, namely the texts' ingredients − the materials in the building analogy − and not to the processes of meaning creation − the how, the why, the where, and so forth, in the building analogy. In Halliday's terminology, Biber and Finegan's analysis is not about "the process of meaning", it is about "the product of that process".

In order to show how a different clustering may result if criteria other than those given by Biber and Finegan (ibid.) are considered, five of the registers in their study were analysed to establish the similarities and differences between the grammatical and lexico-semantic properties of selected clause Themes. The registers included Academic Prose, Biography, Editorials, Official Documents, and Romances. Table 1 summarises Biber and Finegan's findings (pp. 38-39) on these registers. Numbers in bold show the highest percentage of the text-types in each genre.

Some conclusions based on the above data are as follows:

1.   Academic Prose (41%) and Official Documents (50%) have a lot in common and are good examples of the genre Formal Exposition.
2.   Biography (39%) is a good example of the genre Informal Informational Narrative.
3.   Editorials (78%) are very good examples of the genre Informal Exposition. Also Academic Prose (20%), Biography (22%) and Official Documents (30%) have some features in common in the same genre.

Mohsen Ghadessy

TABLE 1

Percentage of each register included in Biber and Finegan's 9 genres.

Registers Genres	Academic Pro	Biography	Editorials	Official Docu.	Romance
Immediate Interaction	0	0	0	0	0
Formal Exposition	41	12	2	**50**	0
Informal-inter actional Text	0	1	0	0	0
Present Reportage	4	4	15	0	0
Informal Infor mational Narr.	5	**39**	4	0	0
Informal Exposition	20	22	**78**	30	0
Interactional Narrative	0	0	0	0	0
Formal Expo- sition + Narra.	**30**	18	0	20	0
Imaginative Narrative	0	4	0	0	**97**

4. Academic Prose (30%) is a good example of the genre Formal Exposition and Narration. Also, there is some similarity between Biography (18%) and Official Documents (20%) in the same genre.

5. The best example of the genre Imaginative Narrative (97%) is Romance.

Samples of the following five registers, collected independently of Biber and Finegan's materials, were then analysed for thematic organisation. Statistics on number of clauses were as follows:

TABLE 2

Total number of analysed clauses for each register

Texts	Academic Prose	Biography	Editorials	Official documents	Romance
No of Clauses	395	171	241	206	273

Halliday's guidelines (1994) for Theme selection were followed with some modifications based on Fries (1992a,b) and Ghadessy (1995). These include two major categories of properties given below, namely grammatical and lexico-semantic, which approximate what Berry (1996) has referred to as "ThemeF: a grammatical means of prioritising the meanings of discourse ThemeMs" and "ThemeM: a priority set of meanings reflecting the writer's main concern for a particular clause", respectively. Numbers in bold show the highest % of Theme type present in each register.

TABLE 3

Grammatical properties of the selected Themes
(% out of total clauses for register)

	Academic Pro.	Biography	Editorials	Official D.	Romance
Simple Theme	59.5	**74.5**	56.5	62.2	61.2
Multiple Theme	40.5	25.5	**43.5**	37.8	38.8
Textual Theme	35.2	24	**41.5**	30.1	34.8
Interper. Theme	5.3	1.8	3.3	**7.7**	5.1
Unmarked Idea. Theme	80	63.7	**81**	70.4	70
Marked Idea. Theme	20	**36.3**	19	29.6	30
Clause as Theme	3	4.7	**6.6**	2.9	4.8
Ellipted Theme	2	**6.5**	3.3	3.9	6.3
Predicated Theme	0	0	0	.5	0
Thematic equative	0	0	0	0	0
Grammatical metaphor	9	5.8	**17.5**	16.5	1.8

Examples of the above Theme types are as follows. The whole Theme in each case is underlined; bold is used to show the part under focus. All examples come from the data.

Simple Theme: *Free trade* is not merely a fashionable concept.
Multiple Theme: *Then he* tried .... , *Potentially these corpora* enable ....
Textual Theme: **But** the time for prevarication is over, ...
Interpersonal Theme: **Potentially** *these corpora* enable ...
Unmarked Ideational Theme: *The existence of English-language corpora* has ....
Marked Ideational Theme: *During the late 1930s and 1940s* Kennedy ....
Clause as Theme: *If the Government is serious about tackling air pollution,* it ....
Ellipted Theme: *Marco* gave a little leap of delight, *then* – ran off into ....
Predicated Theme: **It** is our hope that as ....
Thematic Equative: *What the Government wants is* .... (made up example)
Grammatical Metaphor: **Investigation** *of such questions* did not ....

TABLE 4

Lexico-semantic properties of the selected Themes
(% out of total clauses for register)

	Academic Pro.	Biography	Editorials	Official D.	Romance
Speaker/hearer	**5.3**	0	0	0	2.1
Major Text Par.	10.9	43.9	18.6	10.7	**47.5**
Object, Portion of S.	9.9	4.1	**17.8**	1.9	12
Abstract Concept	54.5	16.9	37.3	**56.3**	17.8
Process	4	2.3	3.7	**7.3**	3.5
Time	2.3	**21**	10.8	4.8	4.6
Location	4.6	1.1	1.2	**4.8**	1.7
Manner	4	3.5	2.9	**9.7**	5
Cause	1.3	2.3	1.6	0	**3.2**
Condition	3.3	4.6	**5.8**	4.3	3.5

Examples of the above Theme types are as follows:

Speaker/hearer: *Next* **we** discussed some issues .... *You*'ve got to go back to Milan ....
Major Text Participant: *The youngest of the nine children, Edward* was ....
Object, Portion of Scene: *Paris* has an important international role to play ....
Abstract Concept: *Diplomacy* is generally more effective than publicity ....
Process: *To provide students with an education* .... which is ....
Time: *In 1946* Kennedy was elected as a ....
Location: *In the Senate* Kennedy established a reputation for ....
Manner: *Ironically,* **by weakening Carter,** Kennedy helped elect ....

Cause: ... *and (2)* **because the classes must be ....** , decisions are required ....
Condition: *If the Government is serious about tackling air pollution,* it is ....

Based on the above two tables, a cluster analysis was then carried out to determine the similarities and differences between the five registers. Tables 5 and 6 present the results.

TABLE 5

Clusters of five registers based on the grammatical features of the Themes

Cluster 1:	Academic Prose, Editorials
Cluster 2:	Official Documents, Romance
Cluster 3:	Biography

TABLE 6

Clusters of five registers based on the lexico-semantic features of the Themes

Cluster 1:	Biography, Romance
Cluster 2:	Editorials, Official Documents
Cluster 3:	Academic Prose

6.   DISCUSSION

A cluster analysis based on the present research provides different answers for the five above registers. Thus, in terms of the grammatical properties of Themes, Academic Prose and Editorials make one cluster, i.e. they are similar in terms of the 11 grammatical features of the selected Themes. This finding is nowhere shown in Biber and Finegan's results, namely Table 1. Looking at the columns for Academic Prose and Editorials in Table 1, 41% of the AP texts are included in the genre Formal Exposition whereas 78% of the Editorials is in Informal Exposition which, incidentally, includes 20% of the AP texts also. Official Documents and Romance make the second cluster in terms of grammatical features of the Themes in Table 5. Again Biber and Finegan's findings are totally different if we look at the columns for Official Documents and Romance in Table 1. Official Documents are the best example of the genre Formal Exposition. On the other hand, Romances are the best example of the genre Imaginative Narrative. Biography is on its own; although the present findings indicate that it is nearer the second cluster than the first.

In terms of the lexico-semantic properties of the selected Themes, Biography and Romance make one cluster, a finding that is at odds with Biber and Finegan's finding if we look at the columns for Biography and Romance in Table 1. Biography is the best example of the genre Informal Informational Narrative while Romance is included in the genre Imaginative Narrative. Editorials and Official Documents make the second cluster for lexico-semantic features in Table 6. We can say that there is some similarity between this finding and Biber and Finegan's results if we look at the respective columns in Table 1. Although 78% of the Editorials are included in the genre Informal Exposition, 30% of the Official Documents are also included in this genre. However, Official Documents are the best example of the genre Formal Exposition. In cluster three, Academic Prose stands on its own in terms of the lexico-semantic features of selected Themes.

7.   CONCLUSIONS

The notion of Theme as "prioritising" certain kinds of meaning (Berry, 1996) is crucial in any investigation of English registers, genres, and text types. We have seen, for example in Table 4, that in fiction animate Themes as major text participants are prioritised, that in academic prose and official documents abstract Themes play a dominant role, and that time adverbials as Themes are frequent in biographies. Texts are not simply collections of words, phrases, and sentences, i.e. products. They are processes of "creating and sharing of meaning between two participants" (Sinclair, 1993). In his most recent book, Biber (1995) rightly states that "Additional linguistic features could also be included in future analyses", namely "information structure, cohesion, coherence, and rhetorical organisation". Thematic organisation can also be included in this list as a feature. However, the crucial question is whether or not the inclusion of such features would alter the cluster configurations already established by Biber and Finegan (1986) and Biber (1995).

The other important question relates to the unit of analysis. "Additional linguistic features", as mentioned by Biber (op. cit.), will have to have sentence/clause as the unit of analysis and not linguistic features below them. It IS possible to identify "additional dimensions", as Biber suggests. However, the inclusion of these, as shown with the thematic analysis, may upset the conclusions based on the other dimensions.

The conclusion to be drawn from all of this is that Biber and Finegan (1986) and Biber (1995) have shown us a reliable and down-to-earth way of establishing English registers by their pioneering work using the multi-feature/multi-dimensional approach. Their criteria are necessary but not sufficient. Also the unit of analysis in their research has to be extended to accom-

modate the additional linguistic features that belong to the domain of discourse. In this way, we can provide a more valid profile of English registers based not only on the ingredients of the products, i.e. texts, but also the processes of meaning that underlie them.

Language Centre, Baptist University,
224 Waterloo Road, Kowloon, Hong Kong.
fax: 852-23397590
Phone: 852- 23397046
E-mail: mohsen @net1.hkbu.edu.hk

NOTES

* This paper was presented at the Eighth Euro-International Systemic Workshop at The Nottingham Trent University in July 1996.

REFERENCES

Berry, M. (1996). What is Theme? *Meaning and Form: Systemic Functional Interpretations.* M. Berry, C. Butler, R. Fawcett, G. Huang (eds.) Norwood, New Jersey: Ablex.1-64.

Bhatia, V.K. (1993). *Analysing Genre: Language Use in Professional Settings.* London: Longman.

Biber, D. (1988). *Variation Across Speech and Writing.* Cambridge: Cambridge University Press.

Biber, D. (1995). *Dimensions of Register Variation: A Cross-linguistic Comparison.* Cambridge: Cambridge University Press.

Biber, D. and Finegan, E. (1986). An initial typology of English text types. *Corpus Linguistics II: New Studies in the Analysis and Exploitation of Computer Corpora.* Jan Aarts and Willem Meijs (eds.). Amsterdam: Rodopi B.V.

Crystal, D. (1991). Stylistic profiling. *English Corpus Linguistics.* Karin Aijmer and Bengt Altenberg (eds.). London: Longman.

Crystal, D. and Davy, D. (1969). *Investigating English Style.* London: Longman.

Fries, P. H. (1983). On the status of Theme in English: arguments from discourse. Papers in Textlinguistics, Vol. 45, Janos S. Petofi and Emel Sozer (eds.), *Micro and Macro Connexity of Texts.* Hamburg: H. Buske.

Fries, P. H. (1992a). Theme, methods of development, and texts. Unpublished manuscript.

Fries, P. H. (1992b). Information flow in written advertising. *Proceedings of Georgetown University Roundtable on Language and Linguistics.* Washington: Georgetown University.

Ghadessy, M. (1995). Thematic development and its relationship to registers and genres. *Thematic Development in English Texts.* M. Ghadessy (ed.). London: Pinter Publishers.

Halliday, M. A. K. (1982). How is a text like a clause? S. Allen (ed.). *Text Processing.* Stockholm: Almqvist and Wiksel.

Halliday, M. A. K. (1994). *An Introduction to Functional Grammar.* (Second edition). London: Edward Arnold.

Quirk, R. *et al.* (1985). *A Comprehensive Grammar of the English Language.* London: Longman.

Schiffrin, D. (1994). *Approaches to Discourse.* London: Blackwell.

Sinclair, J. (1993). Written discourse structure. *Techniques of Description: Spoken and Written Discourse.* Sinclair *et al.* (eds.), London: Routledge.

Swales, J. (1990). *Genre Analysis.* Cambridge: Cambridge University Press.

Van Dijk, T. (1986). News schemata. *Studying Writing: Linguistic Approaches.* C. Cooper and S. Greenbaum (eds.). Beverley Hills, CA: Sage Publications. 155-186.

SOURCES OF DATA

1.   For Academic Prose: Biber, D. and E. Finegan. (1991). On the exploitation of computerised corpora in variation studies. *English Corpus Linguistics.* Karin Aijmer and Bengt Altenberg (eds.), London: Longman.

2.   For Biography: O'Brien, Steven G. (1991). *American Political Leaders.* Oxford: ABC-CLIO Inc.

3.   For Editorials: *South China Morning Post.* (8-12 May 1995). Hong Kong.

4.   For Official Documents: *Review Document for the Bachelor of Arts Honours Degree in Arts and Social Sciences: The Bass Programme.* Vol. 1 (1995), pp. 1-10, Baptist University.

5.   For Romances: Kemp, J. (1993). Marco of Milano, a romance. *South East Asia Writes Back.* London: Skoob Pacifica Anthology No. 1.

APPENDIX

Composition of the clusters by genres ('*' marks cases where more than 50% of a genre occurs in a single cluster. The numbers in parentheses give the percentage of texts from each genre occurring in that cluster).

ONE	44 texts	Immediate Interaction
29	Face-to-Face Conversations	(51 %) *
13	Telephone Conversations	(65 %) *
1	Interview	(4 %)
1	Broadcast	(5 %)

TWO	66 texts Formal Exposition
33	Academic Prose (41 %)
15	Official Documents (50 %) *
9	Biography/Essay (12 %)
4	Professional Letters (40 %)
2	Popular Lore (5 %)
2	Hobbies (5 %)
1	Editorials (3 %)

THREE	49 texts Informational-Interactional
20	Interviews (8 %) *
9	Face-to-Face Conversations (16 %)
4	Telephone Conversations (20 %)
4	Spontaneous Speech (44 %)
3	Prepared Speech (33 %)
3	Professional Letters (30 %)
3	Popular Lore (7 %)
1	Broadcast (5%)
1	Biography/Essay (1%)
1	Hobbies (3%)

FOUR	66 texts Present Reportage
18	Press (41 %)
16	Broadcasts (84 %) *
10	Hobbies (26 %)
9	Popular Lore (20 %)
4	Editorials (15 %)
3	Biography/Essays (4 %)
3	Academic Prose (4 %)
1	Prepared Speech (11 %)

FIVE	51 texts Informal Informational Narrative
30	Biography/Essays (39%)
7	Popular Lore (16 %)
6	Press (14 %)
4	Academic Prose (5%)
2	General Fiction (7 %)
1	Editorials (4%)
1	Prepared Speech (11 %)

SIX	102 texts Informal Exposition
22	Hobbies (58 %) *
21	Editorials (78 %) *
17	Biography/Essays (22 %)
16	Academic Prose (20 %)
9	Official Documents (30 %)
9	Popular Lore (20 %)
5	Press (11 %)
3	Professional Letters (30 %)

SEVEN	30 texts Interactional Narrative
19	Face-to-Face Conversations (33 %)
4	Spontaneous Speech (44 %)
3	Telephone Conversations (15 %)
2	Interviews (9 %)
2	General Fiction (7 %)

EIGHT	72 texts	Formal	Exposition with Narra
24	Academic Prose (30 %)		
14	Biography/Essays (18 %)		
12	Press (27 %)		
12	Popular Lore (27 %)		
6	Official Documents (20 %)		
3	Hobbies (8 %)		
1	Prepared Speech (11 %)		

NINE	65 texts	Imaginative	Narrative
28	Romance Fiction (97 %) *		
24	General Fiction (83 %) *		
3	Biography/Essays (4 %)		
3	Press (7 %)		
3	Prepared Speech (33 %)		
2	Popular Lore (5 %)		
1	Spontaneous Speech (11 %)		
1	Broadcast (5 %)		