

# Lexical Analysis of Annual Reports

Blanka Klímová

*University of Hradec Králové*

## 1 Background

The article is based on a 5-year linguistic-stylistic research project on annual reports (ARs), particularly a model annual report from the *Bank of England*. The lexical analysis concentrates only on lexical words except auxiliary verbs *to be* and *to have* owing to their difficult differentiation.

## 2 Linguistic Analysis

The analysis includes the first 50 lexical words (appendix 1) found in the text of AR of the *Bank of England*. The keywords contain characteristic features of the language corpus and guide us into the main topics. Most of them are the words with the highest frequency in the text, however, this is not a rule (appendix 2). The keywords were formed on the basis of two corpora. One of them is the text of the *Bank of England Report and Accounts*; the bigger one is the text of Business Law topics, which is a part of the electronic programme—Legal Point 1.0a. The keywords were computerised by the Keyword programme of the WordSmith Tools.

### 2.1 Keywords

The analysis shows that in the set of 50 lexical keywords the majority is formed by nouns (35/ 70%), followed by adjectives (8/ 16%), abbreviations (6/ 12%) and numerals (1/ 2%). The interesting fact is that no verb appeared among the keywords.

On the basis of Roget's International Thesaurus, 1995, the keywords are divided into lexical-semantic fields (appendix 3) and they are described in their collocations with one sentence example. E.g.: **Permanence**—*this lexical field expresses stable/ unchangeable quality or state in a positive way: economic/ financial/ monetary/ price stability. The Bank's strategic aims in relation to financial stability are:...*

### 2.3 Adjectives

The semantic classification of adjectives includes the following dominant concept fields: **spatiality** (central, European, International), **characteristic of a sphere of activity and value** (monetary, financial, banking, operational) and **value** (gilt). The last adjective can also be assessed as a compound noun.

## 2.4 Abbreviations

In general lexis abbreviations form a small part. Therefore the presence of a considerable number of abbreviations is very important, and particularly, from the stylistic point of view as an indicator of a specific genre of AR. The tendency to abbreviate words is specific not only for conversation (slang) but also for professional speech and is motivated by language economy (Čermák and Filipec, 1995: 107). In the text of AR the abbreviations are predominantly technical and professional. They are as follows: *£M, HM, UK, EMU, plc/PLC, Ltd.*

## 2.5 Numerals

Numerals are represented in AR by a keyword numeral *million*. The reason is that a great number of figures in the financial statements are in millions as the Bank is one of the most important financial institutions in the world with a large annual turnover.

## 3 Stylistic Analysis

The stylistic analysis includes the following aspects: *stylistic classification of the keywords*, their *origin* and *metaphors*. Attention is also devoted to personal pronouns and to verbs, which do not occur among the first 50 lexical keywords, however, from the stylistic point of view they seem important.

### 3.1 Stylistic classification of keywords

On the basis of lexical analysis of the keywords it is possible to state that these keywords are the words which characterize the style of business English, and particularly, financial and banking genres. The proof is a comparison of these words with the *English-Czech Business Dictionary* (Collin, 1996), which does not contain the word *pages*. Interpretation of this word is necessary as it is used with reference to pages, especially in the *Directors' Report* or the *Report of the Auditors*.

The above-mentioned dictionary contains a lot of words from different areas of business English, such as *insurance, stock exchange* or *foreign trade*. Therefore a comparison of 50 keywords of the AR of the *Bank of England* with 50 keywords of four other ARs (*National Westminster Bank, Komerční Bank, IBM and Silicon Graphics*) was made to find out a typical lexis for texts of any AR. This lexis includes the following 22 keywords: *company's name* (e.g. IBM), *advice, bank/company, bank's/company's, currency, developments, director, finance, financial, international, management, million, month* (i.e. the month which ends the company's financial year), *note, operations, pages, staff, statements, structure, systems, treasury, year*.

### 3.2 Origin of the keywords

The analysis includes 37 non-abbreviated keywords because the words of the same basis are considered to be the words of the same basis (appendix 4). The statistical data show that out of 37 words 32 are of Roman origin (86.5%). As a rule those are 3-syllable or longer words (22 words/ 59.5%), which are scientific or at least abstract in nature. They rank among

the higher style which is distinguished by its accuracy and refinement. The old-English words are 5/ 13.5%.

### 3.3 Metaphors

Metaphors do not appear in the text apart from one example—*lender of last resort*. The reason is that the text of an AR has to contain unambiguous information.

### 3.4 Verbs

It is true that verbs do not occur among the first 50 lexical keywords but among 175 other keywords they do. They are as follows 9 full verbs (*achieve, continue, include, increase, make, pay, publish, support, use*) and 2 modal verbs (*may, must*).

On the basis of semantics the verbs can be divided into **dynamic** (*achieve, continue, increase, make, pay, publish, support, use*) and **non-dynamic** (*include*). Interesting also is their distribution in the individual parts of the text of the AR of the *Bank of England* (appendix 5). The quantitative analysis of 11 keywords shows that the most frequent group of the verbs is the verb **include** found in the part *Notes to the Financial Statements* (26/ 26.5%) because in this part the condition of the company at the end of financial year is described: *Commitments on behalf of HM Treasury... are not included in these accounts as the Bank is concerned in such transactions only as agent*.

The second most frequent verb is the modal verb **may** (12/ 12.2%) which with its intrinsic modality describes what the bank may or may not do. With its intrinsic modality *may* is a bearer of higher stylistic value and therefore it is used in the language of administration (Dušková, 1988: 191): *In exceptional circumstances... the bank may act 'as lender of last resort' to financial institutions...*

The third group is represented by **continue** (16/ 16.3%) in the parts *Operational Reviews* (8/ 8.2%) and *Governor's Foreword* (6/ 6.1%) because these parts describe the constant effort of the company or its individual sections to achieve good economic results:... *the Printing Works continued to achieve significant reductions in unit costs throughout the period...*

The analysis shows that the key verbs in spite of their small representation divide the text of the AR into two parts: **dynamic** (50/ 51%) including *Governor's Foreword, Operational Independence for the Bank, The Bank's Purposes, The Bank's Strategy, The Court of Directors, Operational Reviews* and **stative** (26/ 26.3%) including *Directors' Report* and *Notes to the Financial Statements*.

## 4 Conclusion

The linguistic lexical analysis showed that in 28 cases (56%) there existed concordance between the keywords and the most frequent words in the text of the AR of the Bank of England, which is a proof that the text of the AR has its prior words leading into the key topics that within a framework of other varieties of business language would be obviously different because even the words which we would have expected from the financial point of view are among the most frequent words (*assets, cost/ s, profit, rate, securities, cash, services*) but they do not rank among the keywords.

A high frequency of nouns (35/ 70%) collocated with syntactic stative adjectives are typical of the text of the AR of the *Bank of England*. Moreover, the abstract nouns are in the majority, which is also a feature of professional style (Čechová, 1997).

From the quantitative point of view the most common semantic fields are: *act*, *communication*, *group* and *person*. From the qualitative point of view the meaning of individual fields gives evidence about the analytical, economic, informative and sociable functions of the text of the AR. These semantic fields are also basic concepts of any good company's management.

Semantic field *act* means to perform sociably beneficial work.

Semantic field *communication* is important for successful communication with the public, particularly with shareholders.

Semantic field *group* distinguishes subordination of one's own interests to public ones. The stress is put on teamwork.

Semantic field *person* marks the persons that are involved in the company's successful management.

The choice of the individual dominant fields of adjectives also testifies to the leading position of the Bank both in Europe and in the world, the activities of this institution and the overall positive picture of the language of the text of the AR. Furthermore, other concepts contribute to the positive character of the text of the AR. Even the concept *inflation* with its collocates indicated a positive picture of the text of the AR (*low inflation*, *reduce inflation*).

Generally, the keywords summarize everything that the Bank represents, builds on or aims at: *stability*, *development*, *analysis*... At the same time the key position of these words is to put into the reader's mind a certain *image* of the given institution.

It is possible to say that the collocation of the keywords divides the text of the AR into two meaningful units: **the idea—promotional unit** including *Contents*, *Governor's Foreword (Chairman's Statement)*, *Operational Reviews*, *The Board of Directors*; and **the analytical—economic unit** including *Directors' Report*, *Report of the Auditors*, *Financial Statements*, *Notes to the Financial Statements*.

The statistical comparative analysis of keywords of different ARs showed that there was a relatively small number of them but with a frequent occurrence, which is evidence of effort for stereotypical and economical expression.

The stylistic analysis confirmed that within varieties/ genres of business language one can speak about the language of the AR because the keywords and their collocates showed that in comparison with the business legal corpus there existed 44% words which were not commonly used in the language of business.

The lexis of the text of the AR of the Bank of England thus corresponds to its function. The objectivity and reliability were achieved by unambiguous designation, often by multiwords. The language of the AR is distinguished by several style markers, such as abstractivity, fixed terms, state-of-art terms (e.g. *monetary policy*, *monetary stability*, *financial statements*...), and numerals. Economical effort is expressed by abbreviations that are also symbols of text repetition. It is interesting to state that the issue of keywords analysis is still sometimes neglected (compare Těšitelová 1987 or Kořenský 1999).

The keywords and their collocates are important not only for characteristics of the language of the AR but also for teachers and textbook/ materials writers of this field. Several keywords, which are irreplaceable in the AR, are not found in the teaching materials (Tullis and Trappe 2000 or Cotton et al. 2000); e.g. *note*, *gilt*, *HM Treasury*.

Finally, the analysis of the keywords made it possible to create a dictionary of ARs based on the keywords of five ARs (*Bank of England*, *National Westminster Bank*, *Komerční Bank*, *IBM and Silicon Graphic*).

Appendix 1

50 lexical keywords of the AR of the Bank of England

Order	Word	Frequency (BANKOF)	% Ratio	Frequency	%Ration (Ref. corpus)
1	BANK	282	1.2	21	0.03
2	BANK'S	133	0.57	0	0.00
3	MONETARY	103	0.44	1	0.00
4	FINANCIAL	134	0.57	27	0.03
5	BANKING	78	0.33	1	0.00
6	£M	67	0.29	0	0.00
7	STAFF	71	0.30	2	0.00
8	MILLION	64	0.27	0	0.00
9	FEBRUARY	60	0.26	1	0.00
10	ENGLAND	55	0.23	0	0.00
11	YEAR	107	0.46	46	0.06
12	CENTRAL	42	0.18	0	0.00
13	MARKET	73	0.31	26	0.03
14	TREASURY	39	0.17	0	0.00
15	ANALYSIS	44	0.19	3	0.00
16	OPERATIONS	42	0.18	3	0.00
17	HM	35	0.15	0	0.00
18	EUROPEAN	34	0.14	0	0.00
19	GOVERNOR	33	0.14	0	0.00
20	STABILITY	33	0.14	0	0.00
21	DIVISION	48	0.20	10	0.01
22	MANAGEMENT	50	0.21	14	0.02
23	BANKS	38	0.16	4	0.00
24	GILT	29	0.12	0	0.00
25	UK	28	0.12	0	0.00
26	SUPERVISION	33	0.14	3	0.00
27	DEPUTY	29	0.12	1	0.00
28	HEAD	26	0.11	0	0.00
29	PAGES	25	0.11	0	0.00
30	NOTE	39	0.17	11	0.01
31	EMU	23	0.10	0	0.00
32	SIR	23	0.10	0	0.00
33	POLICY	60	0.26	38	0.05
34	INFLATION	22	0.09	0	0.00
35	OPERATIONAL	22	0.09	0	0.00
36	RESOURCES	22	0.09	0	0.00
37	INTERNATIONAL	26	0.11	2	0.00
38	ADVICE	23	0.10	1	0.00
39	STATEMENTS	50	0.21	29	0.04
40	DEVELOPMENTS	20	0.09	0	0.00
41	SYSTEM	35	0.15	11	0.01
42	CURRENCY	22	0.09	1	0.00
43	LTD	19	0.08	0	0.00

Order	Word	Frequency (BANKOF)	% Ratio	Frequency	%Ration (Ref. corpus)
44	GOVERNORS	19	0.08	0	0.00
45	PLC	19	0.08	0	0.00
46	MARKETS	29	0.12	7	0.00
47	ISSUE	29	0.12	7	0.00
48	FINANCE	24	0.10	3	0.00
49	DIRECTOR	44	0.19	25	0.03
50	STRUCTURE	18	0.08	0	0.00

## Appendix 2

### The most frequent lexical words of the AR of the Bank of England

Order	Word	Frequency	Order	Word	Frequency
1	BANK	282	26	DEPARTMENT	39
2	FINANCIAL	134	27	NOTE	39
3	BANK'S	133	28	TREASURY	39
4	YEAR	107	29	BANKS	38
5	MONETARY	103	30	COST	38
6	OTHER	89	31	SERVICES	38
7	BANKING	78	32	ACCOUNT	35
8	MARKET	73	33	COMMITTEE	35
9	STAFF	71	34	HM	35
10	DIRECTORS	71	35	NEW	35
11	£M	67	36	PROFIT	35
12	MILLION	64	37	RATE	35
13	FEBRUARY	60	38	SYSTEM	35
14	POLICY	60	39	EUROPEAN	34
15	WORK	57	40	GOVERNMENT	34
16	ENGLAND	55	41	MEMBERS	34
17	COURT	51	42	SECURITIES	34
18	MANAGEMENT	50	43	YEARS	34
19	STATEMENTS	50	44	GOVERNOR	33
20	DIVISION	48	45	STABILITY	33
21	EXECUTIVE	47	46	SUPERVISION	33
22	ANALYSIS	44	47	COSTS	30
23	DIRECTOR	44	48	SUPPORT	30
24	OPERATIONS	42	49	ACT	29
25	ASSETS	41	50	CASH	29

### Appendix 3

act – analysis, division, finance, issue, operations, supervision  
 artifact – market/ s, structure, system, currency  
 cognition – *policy*  
 communication – advice, note, pages, statements  
 group – Bank/ Bank’s, management, staff, Treasury  
 location – *England*  
 person – Deputy, Director, Governor/ s, head, Sir  
 possession – money, resources  
 process – developments, inflation  
 state – stability  
 time – February, year

### Appendix 4

#### Origin of 37 non-abbreviated keywords

ord.	English-Norman words	French-Latin words	Italian-Latin words	Old-English words
1	director	advice	Management	England
2	-	analysis	-	gilt
3	-	Bank/ s, Bank’s, banking	-	head
4	-	central	-	staff
5	-	currency	-	year
6	-	Deputy	-	
7	-	developments	-	-
8	-	division	-	-
9	-	European	-	-
10	-	February	-	-
11	-	Finance, financial	-	-
12	-	Governor/ s	-	-
13	-	inflation	-	-
14	-	international	-	-
15	-	issue	-	-
16	-	Market/ s	-	-
17	-	Monetary	-	-
18	-	Million	-	-
19	-	Note	-	-
20	-	operational, operations	-	-
21	-	Pages	-	-
22	-	Policy	-	-
23	-	Resources	-	-
24	-	Sir	-	-
25	-	stability	-	-
26	-	statements	-	-
27	-	structure	-	-

28	-	supervision	-	-
29	-	system	-	-
30	-	Treasury	-	-
total	1 word	30 words	1 word	5 words

## Appendix 5

### Overview of the key verbs in the individual parts of the AR of the Bank of England

parts of AR	1		2		3		4		5		6		7		8		Total	
	abs	%	abs	%	abs	%	abs	%	abs	%	abs	%	abs	%	abs	%	abs	%
<i>continue</i>	6	6.1	-	-	-	-	-	-	-	-	8	8.2	1	1.0	1	1.0	16	16.3
<i>support</i>	-	-	-	-	-	-	1	1.0	-	-	-	-	-	-	-	-	1	1.0
<i>include</i>	-	-	-	-	-	-	-	-	-	-	7	7.1	3	3.1	16	16.3	26	26.5
<i>achieve</i>	5	5.0	-	-	2	2.0	-	-	-	-	2	2.1	-	-	-	-	9	9.2
<i>publish</i>	3	3.1	1	1.0	-	-	-	-	-	-	6	6.1	-	-	-	-	10	10.2
<i>increase</i>	-	-	-	-	-	-	3	3.1	-	-	-	-	-	-	-	-	3	3.1
<i>pay</i>	-	-	-	-	-	-	-	-	-	-	1	1.0	-	-	1	1.0	2	2.0
<i>use</i>	-	-	-	-	-	-	-	-	-	-	1	1.0	-	-	-	-	1	1.0
<i>make</i>	3	3.1	-	-	-	-	-	-	-	-	3	3.1	1	1.0	1	1.0	8	8.2
<i>must</i>	3	3.1	-	-	-	-	-	-	-	-	-	-	-	-	1	1.0	4	4.1
<i>may</i>	-	-	-	-	-	-	2	2.1	1	1.0	1	1.0	2	2.0	12	12.2	18	18.4
<b>Total</b>	<b>20</b>	<b>20.4</b>	<b>1</b>	<b>1.0</b>	<b>4</b>	<b>4.1</b>	<b>4</b>	<b>4.1</b>	<b>1</b>	<b>1.0</b>	<b>29</b>	<b>29.6</b>	<b>7</b>	<b>7.1</b>	<b>32</b>	<b>32.7</b>	<b>98</b>	<b>100</b>

#### Explanation to the headings of the table:

abs – absolute number

- 1 *Governor's Foreword*
- 2 *Operational Independence for the Bank*
- 3 *The Bank's Core Purposes*
- 4 *The Bank's Strategy*
- 5 *The Court of Directors*
- 6 *Operational Reviews*
- 7 *Directors' Report*
- 8 *Notes to the Financial Statements*



## Note

In the overview the parts Contents, The Executive, Report of the Auditors, Financial Statements and Addresses and Telephone Numbers are not included as no verbs appear in them.

## Works Cited

- Annual Report of Komerční Bank for 1996. (1996), Prague.
- Bank of England Report and Accounts 1997. (1997), London.
- Collin, P. H. (1996), *English–Czech Business Dictionary*, Prague: P&P Centrum.
- Cotton, D. et al. (2002), *Market Leader*, England: Pearsons Education Limited.
- Čechová, M. et al. (1997), *Stylistika současné češtiny* [Stylistics of Contemporary Czech], Prague: ISV.
- Čermák, F. and J. Filipec (1995), *Česká lexikologie* [Czech Lexicology], Prague: Academia.
- Dušková, L. et al. (1988), *Mluvnice současné angličtiny na pozadí češtiny* [A Grammar of Contemporary English at the Background of Czech], Prague: Academia.
- IBM Annual Report 1995. (1995), USA.
- Kořenský, J., F. Cvrček and F. Novák (1999), *Juristická a lingvistická analýza právních textů* [Juristic and Linguistic Analysis of Legal Texts], Prague: Academia.
- National Westminster Life Assurance Limited Report and Accounts 1994. (1994), Poole.
- Roget's International Thesaurus*. (1995), UK: Harper Collins Publishers.
- Scott, M. (1997), *WordSmith Tools*, Version 2.0, UK.
- Silicon Graphics Annual Report 1996. (1996), USA.
- Těšitelová, M. (1987), *Kvantitativní lingvistika* [Quantitative Linguistics], Prague: SPN.
- Tullis, G. and T. Trappe (2000), *Insights into Business*, London: Longman.