

PECULIARITIES OF THE LEXIS USED IN FINANCIAL ENGLISH

Assistant PhD Candidate Daniela DĂLĂLĂU
“Petru Maior” University of Târgu Mureș

Abstract

The world of finance is a highly dynamic and competitive one. Thus, specialists working in this field have to face and cope with new and challenging communication situations all the time. Within this context, English has rapidly become the international language of finance. Initially, Financial English professionals focused solely on the functionality of the language and research into the characteristics of the core lexis used in Financial English was relatively limited. However, things have started to change recently. Therefore, the aim of this paper is to highlight some of the peculiarities of the lexis for specific purposes used in magazine articles covering financial topics (taken from *The Economist*) as compared to the general vocabulary used in articles covering common everyday issues (taken from *The Guardian*).

Keywords: Financial English, core financial lexis, corpus-driven analysis, keywords

1. Introduction

The ESP field has changed a lot over the years, switching from one area of interest to another. One of the most important changes is the one mentioned by Dudley-Evans & St John and it refers to a shift in importance as far as Business English is concerned “One major change has been the emergence of Business English as a major strand of ESP teaching. Early ESP work was dominated by English for Science and Technology.... However, in the 1990s ... the largest area of growth is Business English.” (Dudley-Evans & St John, 1998:31)

Things have continued to evolve and other more specific areas of the general world of business have developed rapidly. Taking into consideration that business is a rather broad term, which covered several sectors initially (such as providing goods and services, commerce, management, marketing, finance, accounting, etc.), Business English has proved to be insufficient for learners with more specific needs performing activities in more specific fields.

The financial sector is highly dynamic, competitive and marked by rapid changes. Thus, this has led to an increasing need for teaching and learning Financial English.

This internationalization of English in the world of finance is based on at least two major causes. Firstly, there has been an irreversible growth in financial institutions worldwide, and we refer here to retail banks, investment banks, insurance companies, etc., which have turned from national into international and set up subsidiaries everywhere in the world.

Secondly, the micro and PC revolution and computerization in general, as well as the extensive use of the Internet have provided the vehicles for the English language.

The developments mentioned above had major effects on the increasing need for using English in financial transactions on the financial markets and the stock exchanges worldwide. There have been several decisive aspects in the development of the world of finance, which are worth being mentioned here as they have determined the use of the English language as the “lingua franca” of finance at international level:

- the development and importance of Wall Street;

- the major role of City in Great Britain;
- the creation of the Single Market within the European Union;
- the adoption in Frankfurt, Luxembourg and Tokyo of English as the *lingua franca* for currency dealers and stock brokers;
- the increasing influence of international financial banking, aid agencies and organizations, the World Bank, the OECD (the Organization for Economic and Co-operation Development) is documented in their publications in English.

Taking into consideration all the things mentioned above, it is obvious that English is necessary for international financial transactions, among non-native speakers.

In the beginning, Financial English professionals focused on the functionality of the language and research into the characteristics of the core lexis used in Financial English was relatively limited. However, things have started to change recently. Thus, this paper aims to highlight some of the peculiarities of the lexis for specific purposes used in magazine articles belonging to Financial English.

2. Methodology

The present paper is based on a corpus-driven comparison and analysis of English press articles. In order to carry out the research presented in this paper, we used a computer software – *WordSmith Tools* version 6.0 and we compiled a corpus of approximately 62,000 words.

The corpus consists of two smaller corpora of similar sizes. Our aim was to create a specialised corpus, focusing on the language used in financial articles, and a general corpus with articles containing everyday general English. Sinclair (2001) explains the difference between a general corpus and a specialised corpus. The former provides the researcher with information about the “language as a whole”, whereas the latter provides information about the “characteristics of the genre”.

Firstly, we compiled a Financial English (FE) corpus, which is a selection of articles taken from *The Economist*. We focused on 30 articles chosen mainly from the *Finance and economics* section and we covered a period of six months, from January 2012 to June 2012.

Secondly, we compiled a General English (GE) corpus, consisting of 35 articles selected from the English newspaper, *The Guardian*, during the same period. The difference in the number of articles comes from the fact that we wanted to have approximately the same number of words in both corpora, and since some of the articles in *The Guardian* were shorter, we needed more articles.

According to its functions, we used *Wordsmith* in three ways. Firstly, we used the *Wordlist* tool to generate a frequency list (see Annexe 2) for each corpus, which arranges all the words in the corpus in order of frequency and the number of times they appear. The *statistics* function can analyse features such as type/token ratio, word and sentence length, etc. Secondly, we used the *KeyWords* tool to find and identify words which are unusually frequent or infrequent in the FE corpus compared to the GE corpus (see Annexe 3). Thirdly, we used the *Concord* tool to find all the examples of a word in the corpus and show them in concordance lines. This tool was particularly helpful for seeing how certain words are used in

context, identifying common collocations and establishing if words are used with a general, sub-technical or technical meaning.

The results obtained with the three *WordSmith* tools are presented and discussed below both from a qualitative and a quantitative point of view.

3. Analysis and discussion

We first used the *WordList* tool to generate lists of the most frequent words found in the FE corpus as well as in the GE corpus (see Annexe 2) and to make a general statistical analysis aiming to compare features such as word and sentence length (see Annexe 1). The statistical analysis of the two corpora reveals that mean word length (in characters) is longer in the FE corpus than in the GE corpus. Although the FE corpus contains more sentences than the GE corpus, the mean (in words)/sentence is slightly lower in the FE corpus than in the GE corpus.

One interesting peculiarity, which the statistical analysis highlights, is that the FE corpus has a lower lexical diversity. This is obvious when comparing the number of tokens (words), types (distinct words) and the TTR (type/token ratio) in the two corpora.

Lexical diversity is defined as a measure of how many different words are used in a text and is often used by researchers as a synonym for *lexical richness*. However, authors like Read (2000) and Malvern et al. (2004) explain that the lexical diversity measure is only one part of the multidimensional feature of lexical richness. Thus, Read suggests other factors such as lexical sophistication, number of errors, and lexical density (Read 2000). The traditional way to calculate lexical diversity is the ratio of different words (types) to the total number of words (tokens), the so-called type-token ratio, or TTR.

One common aspect of the TTR measure is that text samples containing large numbers of tokens give lower values for TTR and vice versa. Put another way, this means that a longer text in general has a lower TTR value than a shorter text. Therefore, the use of the TTR is relevant for comparisons only when texts are of equal length.

Our comparison reveals that although there is a slight difference in the number of tokens between the two corpora (FE corpus – 30,419, GE corpus – 30,729), this is not enough to explain the difference in the number of types (FE corpus – 4,835; GE corpus – 5,718), and TTR (FE corpus – 15.89; GE corpus – 18.61) respectively.

As for the lists of the most frequent words found in each corpus, Annexe 2 presents the top 50 most frequent words in both the Financial and General corpora. It is worth mentioning that, after generating the lists using the *WordList* tool, we took into consideration only content words.

Comparing the two lists, we notice that there is little overlap between them. There are only four words that appear on both lists (*year(s)*; *say(s)/said*; *government(s)* and *see*). The verb *say* is the only content word in both lists that has mainly the same function, i.e. it is used to report what people state. Considering that these texts are newspaper articles, its usage is understandable. The verb *see* is used almost exclusively to make reference to financial statistics (e.g. *see chart*) in the FE corpus. The word *year* is frequently used, in the FE corpus, for reporting the extent to which financial aspects change (expressed as a percentage) over a period of one year (e.g. *Greek GDP shrinking by as much as 7% a year*). It is also found in collocations such as *bubble years*, referring to years in which an economic cycle is

characterized by rapid expansion followed by a contraction. The noun *government* appears in the FE corpus in a considerable number of collocations with a clear financial meaning (e.g. **government** bonds; **government**-bond yields; **government** debt; **government** borrowing; **government** accounting; **government** accounts).

The two word lists of both corpora were then used, with the help of the *KeyWords* tool, to generate the key words list. Key words are words that appear in a text more or less frequently. There are two types of key words. The ones that occur with greater than normal frequency are called positive key words, and those that occur with lower than normal frequency are known as negative key words. Comparing the Financial corpus to the General corpus produced a list of 92 key words with 74 positive key words and 18 negative keywords. Annexe 3 presents 55 positive key words and 6 negative key words that were identified.

The positive key words list includes 39 of the top 50 most frequent words in the FE corpus (e.g. *bank(s)*, *euro*, *debt*, *capital*, *investors*, *equity*; *investment*; *growth*, *billion*, *GDP*, *markets*, *returns*, *countries*, *zone*, *bonds*, *funds*, etc.). Besides these, there are also nine other key words that seem to belong clearly to the world of finance and that can be part of the core financial lexis (*loans*, *deposits*, *ECB*, *liquidity*, *premium*, *IMF*, *prices*, *shares*, *creditors*, *currency*, *yield*). However, there are words such as *capital*, *growth*, *returns*, *losses*, *shares and rates* that can be classified as sub-technical words. Using the *Concord* tool may help us to decide if they are used with different meanings, as different parts of speech, or with different collocations in the two corpora. There are also five words (*private*, *zone*, *union*, *risk* and *crisis*) that do not seem to have a clear connection to the world of finance and *Concord* may shed some light on their usage in the FE corpus. The positive key words list contains three function words (*its*, *is*, *by*) that will be further investigated using the *Concord* tool.

The negative key words list contains mainly functional rather than content words, which can be divided as personal pronouns, possessive adjectives, conjunctions, contractions and auxiliaries, and they will be analysed using *Concord*.

Once we identified the most frequent words and the key words, the *Concord* tool was used to see how some of these words are used in context, especially whether they are used differently in the two corpora.

As mentioned above, there are five words that appear as positive key words on the list, but these are not clearly connected to finance (*private*, *zone*, *union*, *risk* and *crisis*).

The adjective *private* appears 76 times in the FE corpus. In all examples, it collocates with words related to finance such as:

- **private**-equity firms/ /performance/investments/executives/managers/ownership;
- **private** investors/borrowers/creditors/bondholders/debt/companies/capital/saving/pensions;
- **private**-sector finance/debt/bonds/tax/accounting rules/creditors/saving/losses;
- **private**-money creation;

In the GE corpus, *private* appears 10 times. In two examples, it collocates with *businesses* and *companies*:

- many institutions are looking more and more like **private** businesses, not public services;
- while **private** companies make substantial profits.

These are the only examples that may be somehow related to the financial world. In our opinion, they are rather related to the business world in general. In the rest of the examples, it collocates with words from general English such as:

- *private treaty/collection/assessment/individuals/train/detectives/road.*

As far as the noun *zone* is concerned, it collocates with types of currencies in all 49 examples in the FE corpus. The most used one is the *euro* in 48 cases, in examples such as:

- *The euro zone has also softened the terms of its loans to Greece.*

and it is used with the word *sterling* only once:

- *The Irish left the sterling zone.*

In the GE corpus, the word *zone* appears only once, and it collocates with *comfort*.

- *I had also existed within a comfort zone of my established skills...*

The noun *union* appears to be used quite differently in the FE and the GE corpora. There are only three examples of *union* in the GE corpus, which is too few compared to the 30 examples found in the FE corpus. Although all examples in both corpora share the same basic meaning of unity, the difference comes from the fact that *union* collocates with different words in both corpora. In the GE corpus, *union* is found in the following examples:

- *the lecturers' union;*
- *the Public and Commercial Services union;*
- *at the heart of the student union on the Woodhouse Lane campus.*

Of the 30 examples found in the FE corpus, eight refer to the *European Union* and this is not strictly related to the world of finance. However, the articles containing this collocation describe the financial situation within this area. In 15 examples, *union* collocates with banking, forming the collocation *banking union*.

Two examples were found for each of the following words:

- *currency union;*
- *fiscal union;*
- *monetary union.*

The word *risk* appears 48 times in the FE corpus, and in most of the examples, it is used in collocations with a clear financial meaning, such as *risk premium*, *equity risk premium*, *credit risk*, *risk-free interest rate*, *risk-averse investors*. However, the meaning of *risk* is more general in some examples (e.g. *it leaves the investor at risk from inflation*). In the GE corpus, *risk* appears only in five examples, all of them having a general meaning (e.g. *the issue most at risk; you risk wasting your time and money*).

The noun *crisis* appears 31 times in the FE corpus and it collocates with words that have an obvious financial meaning (*banking crisis; debt crisis; euro crisis, financial crisis*). In the GE corpus, only one example was found and it has a general sense.

Using the *KeyWord* tool was useful for identifying a group of words that can be classified as sub-technical (*capital, growth, returns, losses, shares and rates*). Therefore, our assumption was that they were used differently in the two corpora, and in order to check it, we used the *Concordance* tool.

All 99 examples of *capital* found in the FE corpus used it in the financial sense. Only one example of *capital* was found in the GE corpus, and it referred to a capital city (e.g. *making the capital [London] the dourest place in the country*).

The word *shares* appears 25 times in the FE corpus, and it is a noun referring to shares in a company, in all the examples. Some of the most common collocations are:

- *preference shares*;
- *market shares*;
- *bonds and shares*;
- *to issue shares*;
- *to list shares*.

In the GE corpus, we found only one example of the word *shares*, but here it is used as a verb and it has a general meaning (*what he shares with his dad*).

The noun *rates* was found in 28 examples in the FE corpus. In all cases, it collocates with words clearly related to the world of finance (*interest rates*; *inflation rates*, *deposit rates*; *borrowing rates*, *lending rates*, *London inter-bank offered rates /LIBOR rates*). In the GE corpus, only two examples of *rates* were found, none of them related to finance (*divorce rates*; *lower rates of depression and other psychiatric disorders*).

As for the words *growth* and *returns*, the examples found in the GE corpus (two examples of *growth* and one example of *returns*) have a financial meaning as well. The word *losses* does not appear at all in the GE corpus, thus it is impossible to compare usage. However, it is worth mentioning that the meaning is clearly financial in all examples of these three words, in the FE corpus. Therefore, they may be part of the core lexis, especially collocations such as *credit growth*, *dividend growth rate*, *GDP growth*, *returns on assets*, *returns on equity*, *investment returns*, *private-sector losses*.

After analysing these words using the *Concord* tool, we can state that they may be part of the core financial lexis as sub-technical words.

The three function words (*its*, *is*, *by*), which appear on the positive key words list, were also investigated using the *Concord* tool. *Its* is used 150 times in the FE corpus and only 43 times in the GE corpus. In the FE corpus, *its* is used almost exclusively to refer to financial institutions, countries or continents. The main reason is that the financial articles cover topics related to the description and analysis of the financial situation or performance of different financial institutions or countries. However, in the GE corpus, *its* refers to a range of objects, places, and only occasionally to companies.

As far as *by* is concerned, it appears in 105 examples in the GE corpus and in 221 examples in the FE corpus. In most of the examples, in the GE corpus, it is used as a preposition, especially after passive structures for indicating who does something or what causes something. It is also used in some examples to refer to a particular moment or time (*by day*, *by now*, *by the time they reached Earth*). Beside these uses in the general corpus, in the FE corpus, *by* is extensively used for showing how much aspects related to the world of finance have changed or how much difference there is between things (*The MSCI global share index is up by almost 9%*; *cutting its benchmark interest rate by 50 basis points*).

As for *is*, it appears 524 times in the FE corpus compared to 309 times in the GE corpus. Although it is used in both corpora as an auxiliary verb, a verb and a linking verb, it is used more often as an auxiliary verb and as a linking verb in the FE corpus. As an auxiliary verb, it is used for forming the present continuous tense of verbs that describe actions that are in progress (*Portugal ... is injecting capital into its banks*; *Some of the additional money that*

the IMF is seeking to raise), but also for forming passive structures (*The government is expected to pump €10 billion; The FGD is funded by contributions from the banks*). As a linking verb, it is used for providing information and describing financial tools, institutions or situations (*Capital elsewhere is scarce; to convince investors that its debt is sustainable*).

Analysing the negative key words list, we notice that the “most” negative key words are personal pronouns (*I, he, we, she*), whose use is very limited in the FE corpus. *I* only appears four times in the FE corpus compared to 371 times in the GE corpus. In one case it appears in the name of a bank (Citi *f/i*), and in three examples, it is used as personal pronoun between inverted commas, when quoting the exact words of certain people.

He appears in 22 examples in the FE corpus, and it is used mostly to quote the opinions of financial experts. On the contrary, it is used in 178 examples, in the GE corpus. The use of *she* is even more limited in the FE corpus, as it appears only 4 times compared to 93 times, in the GE corpus. It is obvious that there is a difference in gender, which might suggest that the world of finance is hardly accessible to women. The only women mentioned, in the FE corpus, are Mrs. Christine Lagarde and Mrs. Angela Merkel.

As for *we*, it appears 18 times in the FE corpus, compared to 159 times in the GE corpus. In almost all cases, it is used between inverted commas and it refers to a financial institution, company or government. On the contrary, in the GE corpus, it is used to refer to a group of people (family, friends, students, etc.), but also to people in general.

4. Conclusions

In spite of the fact that the corpus used for analysis was not large enough to be able to establish “core language” with any degree of certainty, the comparison made between the FE corpus and the GE corpus highlights some noticeable differences between the two corpora.

As demonstrated in the paper, a number of words and collocations stand out as clearly belonging to the world of finance. Using the Keyword and Concord tools, some of these were classified as technical and sub-technical. Although words presented in the lists may form part of the core financial lexis, these lists are far from exhaustive. It needs to be mentioned that there are words (e.g. *inflation, tax, dividend, monetary, income, revenues, bondholders*, etc.) that appear twenty times or fewer in the financial frequency list, and because of this, they were not taken in consideration when the programme generated the key words list. Thus, they do not appear on the key words list. However, this does not mean that they are not part of the core lexis of Financial English, either as technical or sub-technical words.

Annexe 1. Statistical analysis of the two corpora

	Financial English	General English
Tokens	30,419	30,729
Types (distinct words)	4,835	5,718
Type/token ratio (TTR)	15.89	18.61
Mean word length (in characters)	4.82	4.64
Sentences	1,560	1,531
Mean (in words)/sentence	19.5	20.07
1-letter words	948	1,208

2-letter words	5,008	5,095
3-letter words	5,372	6,156
4-letter words	5,402	5,457
5-letter words	3,659	3,574
6-letter words	2,889	2,633
7-letter words	2,789	2,412
8-letter words	1,904	1,694
9-letter words	1,380	1,221
10-letter words	879	759
11-letter words	386	402
12-letter words	220	220
13-letter words	111	142
14-letter words	33	33
15-letter words	17	13
16-letter words	4	4
17-letter words	3	4
18-letter words	0	0
19-letter words	0	1
20-letter words	0	1

Annexe 2. Top 50 most frequent words in the Financial and General word lists

Financial word list						General word list					
Word	Freq	%	Word	Freq	%	Word	Freq	%	Word	Freq	%
banks	212	0.68	says	47	0.15	said	116	0.37	day	28	0.09
euro	127	0.41	GDP	46	0.15	time	99	0.32	home	28	0.09
bank	110	0.35	market	45	0.15	people	62	0.2	long	28	0.09
debt	103	0.33	fund	44	0.14	says	59	0.19	see	28	0.09
capital	99	0.32	public	41	0.13	years	51	0.16	go	27	0.09
year	92	0.30	returns	40	0.13	children	50	0.16	hope	27	0.09
European	89	0.29	bonds	38	0.12	new	50	0.16	make	27	0.09
money	84	0.27	central	37	0.12	life	46	0.15	country	26	0.08
private	76	0.25	rate	36	0.12	year	46	0.15	want	26	0.08
investors	74	0.24	high	35	0.11	work	43	0.14	education	25	0.08
banking	73	0.24	interest	35	0.11	like	42	0.14	mother	25	0.08
countries	70	0.23	term	35	0.11	get	39	0.13	dad	24	0.08
investment	70	0.23	funds	33	0.11	last	39	0.13	good	24	0.08
equity	68	0.22	governments	33	0.11	never	39	0.13	great	24	0.08
firms	67	0.22	chart	32	0.10	way	39	0.13	part	24	0.08
government	67	0.22	deficit	32	0.10	know	34	0.11	right	24	0.08
growth	62	0.20	crisis	31	0.10	family	33	0.11	take	24	0.08
billion	56	0.18	assets	30	0.10	government	33	0.11	little	23	0.07
markets	56	0.18	buy	30	0.10	think	33	0.11	parents	23	0.07
economy	49	0.16	costs	30	0.10	made	32	0.1	running	23	0.07
financial	49	0.16	union	30	0.10	support	30	0.1	social	23	0.07
zone	49	0.16	credit	29	0.09	women	30	0.1	student	23	0.07

risk	48	0.15	economies	29	0.09	say	29	0.09	told	23	0.07
see	48	0.15	fiscal	29	0.09	school	29	0.09	book	22	0.07
years	48	0.15	loans	28	0.09	things	29	0.09	services	22	0.07

Annexe 3. Key words in the Financial word list

Key word	Freq	%	RC. Freq	RC. %	Keyness	Key word	Freq	%	RC. Freq	RC. %	Keyness
Positive key words											
banks	212	0.68	0		294.79	fiscal	29	0.09	0		40.24
euro	127	0.41	0		176.42	loans	28	0.09	0		38.85
debt	103	0.33	0		143.04	financial	49	0.16	6	0.02	38.4
capital	99	0.32	1		127.66	central	37	0.12	2		38.34
bank	110	0.35	4	0.01	123.65	governments	33	0.11	1		38.15
investors	74	0.24	0		102.73	deficit	32	0.1	1		36.83
banking	73	0.24	0		101.34	losses	26	0.08	0		36.08
European	89	0.29	3		101.29	crisis	31	0.1	1		35.5
equity	68	0.22	0		94.4	deposits	25	0.08	0		34.69
investment	70	0.23	1		88.05	assets	30	0.1	1		34.18
firms	67	0.22	1		83.97	ECB	24	0.08	0		33.3
growth	62	0.2	2		71.03	liquidity	23	0.07	0		31.91
billion	56	0.18	1		69.04	fund	44	0.14	7	0.02	29.96
GDP	46	0.15	0		63.84	costs	30	0.1	2		29.43
markets	56	0.18	2		63.1	premium	21	0.07	0		29.14
its	150	0.48	43	0.14	63.08	market	45	0.15	8	0.03	28.55
countries	70	0.23	7	0.02	59.95	IMF	20	0.06	0		27.75
zone	49	0.16	1		59.59	prices	20	0.06	0		27.75
private	76	0.25	10	0.03	57.52	shares	25	0.08	1		27.6
is	524	1.69	309	1	57.06	rates	28	0.09	2		26.92
money	84	0.27	16	0.05	50.83	creditors	19	0.06	0		26.36
economy	49	0.16	3		49.22	currency	19	0.06	0		26.36
returns	40	0.13	1		47.49	yield	19	0.06	0		26.36
funds	33	0.11	0		45.79	union	30	0.1	3		25.68
bonds	38	0.12	1		44.82	Negative key words					
chart	32	0.1	0		44.4	I	4	0.01	371	1.2	-477.47
by	221	0.71	105	0.34	42.51	he	22	0.07	178	0.57	-138.92
rate	36	0.12	1		42.15	we	18	0.06	159	0.51	-129.19
risk	48	0.15	5	0.02	40.42	his	13	0.04	128	0.41	-108.85
economies	29	0.09	0		40.24	she	4	0.01	93	0.3	-101.19
credit	29	0.09	0		40.24	said	16	0.05	116	0.37	-85.57

Bibliography

- Dudley-Evans, T. & St John, M. (1998). *Developments in ESP: A multidisciplinary approach*. Cambridge: Cambridge University Press.
- Hutchinson, Tom & Waters, Alan. (1987). *English for Specific Purposes: A learning – centered approach*. Cambridge: Cambridge University Press.
- Kennedy, Cris and Bolitho, Rod. (1984). *English for Specific Purposes*. London: Macmillan Publishers.
- Malvern, David, Brian Richards, Ngoni Chipere & Pilar Duran. (2004). *Lexical diversity and language development: quantification and assessment*. New York: Palgrave Macmillan.
- Read, John. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Sinclair, John McH. (2001). Preface. In M. Ghadessy, A. Henry and R. L. Roseberry (eds.). *Small corpus studies and ELT*, vii–xv. Amsterdam: John Benjamins.