Using Wmatrix to Explore Discourse of Economic Growth

Chunyu Hu[1]

[1] School of English for International Business, Guangdong University of Foreign Studies, China

Correspondence: Chunyu Hu, Associate Professor of School of English for International Business, Guangdong University of Foreign Studies, Guangzhou 510420, China. Tel: 86-20-3631-7183. E-mail: gwhcy@gdufs.edu.cn

Abstract

Growth is a concept of particular interest for economic discourse. This paper sets out to explore a small corpus of economic growth, which consists of articles from *The Economist*. The corpus software used in this study is a web-based tool Wmatrix, an automatic tagging software able to assign semantic field (domain) tags, and to permit the extraction of key words and key semantic domains by applying the keyness calculation to tag frequency lists. The results show that at 99% confidence (or $p < 0.01$), the cut-off of 6.63 (log likelihood value) indicates that there are 1051 positive keywords (including multiword expressions) and 80 key semantic domains generated from the corpus. It is evident that BRICs or the emerging economies/markets, in particular China and India, are a big concern regarding economic growth over the past years. A number of examples of possible ways forward in teaching methodology are presented.

Keywords: corpus, economic growth, discourse study, keyness analysis, Wmatrix Tools

1. Introduction

There has been a growing literature on the economic discourse over the past three decades. Some studies have dealt with the linguistic and discursive aspects of the widely read masterworks: Adam Smith's *The Wealth of Nations* (Bazerman, 1993; Henderson, 2006), Marshall's *Elements of Economics* (Del Lungo Camiciotti, 2005), and Keynes' *General Theory* (Bondi, 2010). Other studies have instead focused on the distinctive characteristics of key written economics genres used for the transmission and dissemination of knowledge: prediction in schematic structure in research articles (Dudley-Evans & Henderson, 1990), and economic metaphor in media articles or textbooks (Alejo, 2010; Boers, 2000; Charteris-Black, 2000; Henderson, 1982, 1994, 2000; Hu, 2014; Sun & Jiang, 2014; White, 2003).

While research on economic discourse can be based on researchers' intuitive knowledge (Tribe, 2015), there are a growing interest in corpus approaches to social sciences. Gaballo (2012), for example, explores the discourse of economics and the way it is used and organized in the business section, and in the finance and economics section of *The Economist* by bringing together methodologies of critical discourse analysis, systemic functional linguistics and corpus linguistics allowing the texts to be explored from different perspectives while providing multiple insights.

Economic growth is a big concern in world-wide economy, and not surprisingly, it is also a concept of particular interest for applied linguists. Unlike previous studies that are mainly concerned with *growth* metaphor (Crawford Camiciottoli, 2007; Henderson, 1994; White, 2003), this paper aims to explore discourse of economic growth by conducting keyness analysis with the aid of corpus tool Wmatrix.

2. Kyeness Analysis: From Keywords to Key Semantic Domains

Keyness analysis has a relatively recent history, which concerns the analysis of keywords, key part-of-speech, and key semantic domains. The term "keywords" is widely used to refer to words which are important in some way, either in individual texts or in a given culture. According to Stubbs (2010, p.23), there exist three different concepts of keywords, which derive from quite different academic traditions. The first sense of keywords derives from cultural studies (Williams, 1976), the second from comparative quantitative corpus analysis (Scott, 1996), and the third from work on lexico-grammar (Francis, 1993).

The first concept of keywords is labeled as "cultural keywords" (Williams, 1976; cf. also Wierzbicka, 1997). The meanings of words like *alienation*, *capitalism*, *family*, *fiction*, *hegemony*, *literature, media, tradition* etc. were

taken by Williams to represent the most distinctive features of contemporary western culture, by integrating synchronic and diachronic perspectives in a full appreciation of meaning. These words are 'key' because they are of social, cultural or political significance and they are the 'dictionary' of a culture and a social group. The first hints to these words can be traced back to Firth (1957), though he used the term "focal or pivotal words" instead of keywords.

The second concept of keywords has its root in John Sinclair's corpus-driven approach to language. In such a perspective, keywords are studied through their typical co-occurrence with other lexico-semantic units. For example, Stubbs (1996, p. 172) refers to 'cultural key words' as "words which capture important social and political facts about a community". The important feature for Stubbs is that the cultural and ideological implications of a keyword can be illuminated by an analysis of its collocation and semantic preference – the tendency of the word to co-occur with other words and with words belonging to a specific semantic category or field (see also Sinclair 2004). Stubbs (1996, p. 166) traces his efforts back to that of Firth (1935) on "focal and pivotal words" and to Williams' book on key words.

The third concept of keywords has much to do with stylometry in which keywords are statistically significant lexical items. Guiraud (1954, cited in Culpeper 2009, p. 32), for example, used 'mots-clés' (keywords, based on relevant frequency) to contrast 'mots-clés' (based on absolute frequency). It is, however, Mike Scott who has popularized the practice of keyword analysis in this line. Interestingly, Scott (1997) also relates his work to that of Williams in the 1970s in terms of its purpose. Unlike Williams' study that hardly paid attention to text and genre and left methodological tools for the analysis of meaning completely undiscussed, Scott's work is a text-focused one and adopts a standard procedure of keywords extraction. With the aid of KeyWords facility of his program WordSmith Tools designed by Scott, it is a relatively easy and rapid task to calculate the incidences of each and every single word in the target data as well as a comparative data set, to undertake statistical comparisons between incidences of the same words in order to establish significant differences, and finally to see the resulting keywords ranked according to degrees of significance of difference.

Keyword analysis and its extension (key cluster analysis) have been used in numerous studies and applied to a wide variety of research questions ranging from language education (Scott & Tribble, 2006) to sociolinguistics (Baker, 2010), economics (Bondi, 2010), and to literary stylistics (Mahlberg, 2013), to name only a few. While this line of research as practiced by Scott and many followers has exerted immense influence, it is not without its problems. To begin with, keyword analysis tends to deliver more keyword results than is possible for the researcher to analyze (Berber Sardinha, 1999). Second, relatively low frequency words may not be identified as key (Baker, 2004). Third, keywords "only focus on lexical differences, rather than semantic, grammatical, or functional differences" (Baker 2004, p.354).

Scott (2010, p.52) also believes that there are at least five restrictions and limitations regarding the status of keywords: 1) Keyness is not intrinsic to the word or cluster itself but is context-bound; 2) The size of context is a matter of choice; 3) Keywords act as pointers to specific textual aboutnesses and/or styles, and the reference corpus affects this in ways which are still not fully understood; 4) Keywords are statistically arrived at but are not fully and completely established. Definitive status for a whole set of KWs cannot be claimed; and 5) The procedure is far from being able to handle related forms which readers can easily distinguish.

One way to address the above problems is to conduct key part-of-speech and key semantic analyses which give rise to analytical categories that (1) are fewer than keywords, thus reducing the number of categories a researcher needs to take into account, and (2) group lower frequency words which might not appear as keywords individually and could thus be overlooked (Rayson, 2008).

3. Method

3.1 Data Used in This Study

The corpus data used in this study consist of about 41,800 words from 55 essays downloaded from *The Economist* by using WebCorp. The original WebCorp project was an experiment to see whether it was likely to develop a system to extract linguistic data from web text efficiently and present this to the linguist in as usable as fashion as it is presented in traditional corpora (Kehoe & Renouf, 2002). WebCorp attempts to treat the web itself as a corpus. Meanwhile, researchers can also use the web as a source of texts to build smaller corpora, as practiced in this study.

The system receives a word or phrase and other requirements from the user (*growth* in this study), passes these to a commercial search engine (Google, AltaVista, etc), and extracts the 'hit' pages from the search engine results. Below is the interface of WebCorp that was used to extract the essays including *growth* from the site *www.*

economist. com:

Figure 1. The interface of WebCorp

Each page is accessed and processed and the extracted concordances are presented to the user in a choice of formats. The Google Search API returned 64 hits (out of an estimated 54200). WebCorp accessed 64 web pages and generated 700 concordances, but only 55 pages were reliable and used as the source of corpus data.

3.3 Corpus Tool Used in This Study: Wmatrix

Wmatrix is a web-based tool for corpus analysis that combines techniques from corpus linguistics and natural language processing. It provides a web interface to the Constituent Likelihood Automatic Word-tagging System (CLAWS) and UCREL Semantic Analysis System (USAS) corpus annotation tools, and standard corpus-linguistic methodologies such as frequency lists, keyword lists, collocation and concordances.

The first stage of annotation involves CLAWS, a part-of-speech tagger which assigns a part-of-speech (POS) tag or grammatical word classes to every word in running text with about 96-97 percent accuracy (Leech & Smith, 2000), e.g., 'NN1' for singular common noun and 'VM' for modal auxiliaries. The POS-tagged text is then fed into SEMTAG, which assigns semantic tags representing the general-sense field of words from a lexicon of single words and a list of multi-word combinations called idioms, e.g., 'as a rule'. Currently, the lexicon contains nearly 37,000 words and the idiom list contains over 16,000 multi-word units. An idiom list enables the corpus tool to identify any idiomatic expressions, usually non-decompositional sequences, and to assign a special set of tags to the words in that particular idiomatic phrase to denote a part-of-speech relation above the level of the word (Rayson, 2008). The semantic tagset, loosely based on the *LongmanLexicon of Contemporary English*, has a multitier structure with 21 major semantic fields and more than 232 subdivisions. For instance:

I MONEY & COMMERCE

I1 Money generally

I1.1 Money: Affluence

I1.2 Money: Debts

I1.3 Money: Price

I2 Business

I2.1 Business: Generally

I2.2 Business: Selling

I3 Work and employment

I3.1 Work and employment: Generally

I3.2 Work and employment: Professionalism

I4 Industry

Items not contained in the lexicon or idiom list are assigned a special tag, Z99. Antonymity of conceptual classifications is indicated by +/− markers on tags, e.g., N3.8+ (*Speed: Fast*) as opposed to N3.8- (*Speed: Slow*). Comparatives and superlatives receive double and triple +/− markers respectively, e.g., larger (N3.2++) and largest (N3.2+++).

Users can load their own corpus data, make frequency lists, generate key words, run concordances and calculate collocations via a web-browser interface. In addition, the corpus data is automatically tagged for part-of-speech and semantic fields. These two extra levels of annotation can then be used to generate tag frequency lists, tag concordances and collocations of tags and to apply the log-likelihood keyness statistic to the tag level. The log-likelihood statistic (LL) is employed by Wmatrix; only items with a LL value of equal to or more than 6.63 are considered to be statistically significant, since 6.63 is the cut-off for 99 percent confidence of significance.

Wmatrix enables researchers to annotate their data sets relatively easily and rapidly for both grammatical and semantic categories, and then to identify which categories are key. Recent studies include: Afida (2007), focusing on semantic domains in business English, and Archer et al. (2009), focusing on semantic domains in Shakespeare's plays.

3.3 Corpus Query and Analysis Procedures

Step 1: Upload the corpus of economic growth to the web-browser interface, and then the corpus data are automatically tagged for part-of-speech and semantic fields, generating a set of three frequency lists: a word frequency list, a part-of-speech (POS) and semantic field (domain) frequency list, as demonstrated in Figure 2.
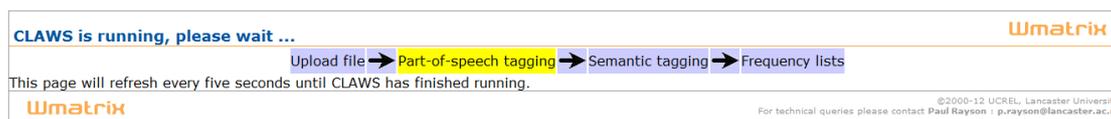
Figure 2. Upload the corpus data and automatically annotate the data via the web-browser interface

Step 2: In order to generate keyword list, key POS list, and key semantic domain list, choose the reference corpus for each component. In this study, BNC Sampler Written was chosen as the reference corpus, as shown in Figure 3.

Figure 3. Choosing reference corpus

The final step is to save all the relevant lists for further analysis. Since the key POS list is not the main concern of this study, only results from keyword list and key semantic domains list will be presented in the next section for further analysis.

4. Results and Analysis

4.1 Keywords

The results show that at 99% confidence (or $p < 0.01$), the cut-off of 6.63 would indicate that there are 1051

keywords (including multiword expressions such as *interest rates, current account, Mr Obam, long term, population growth, health care*). Table 1 presents some top keywords occurred in the corpus. The keywords ranking from 31 to 50 are *Chinese, inflation, 90%, 2011, slower, economist, IMF, global, 2012, year, more, spending, than, manufacturing, Ms, Singapore, urbanisation, property, trade*, and *consumption*.

Table 1. The top 30 keywords in the corpus of economic growth

Rank	Keywords	Freq.	%	RC. Freq.	RC. %	LL
1	growth	409	1.07	121	0.01	2087.15
2	economy	144	0.38	78	0.01	650.34
3	gdp	85	0.22	10	0	486.96
4	economic	128	0.34	134	0.01	476.05
5	china	91	0.24	52	0.01	405.62
6	economies	69	0.18	11	0	383.36
7	debt	83	0.22	56	0.01	354.22
8	india	55	0.14	29	0	250
9	firms	53	0.14	29	0	238.78
10	china's	33	0.09	0	0	213.56
11	productivity	51	0.13	40	0	208.44
12	countries	67	0.18	111	0.01	206.74
13	innovation	47	0.12	36	0	193.45
14	economists	31	0.08	4	0	176.06
15	rate	58	0.15	105	0.01	171.56
16	india's	27	0.07	1	0	166.19
17	is	631	1.66	8780	0.94	159.09
18	emerging	38	0.1	28	0	158.19
19	america	45	0.12	57	0.01	155.81
20	investment	57	0.15	127	0.01	151.31
21	fiscal	30	0.08	10	0	149.96
22	rates	46	0.12	82	0.01	137.09
23	tropical	29	0.08	15	0	132.42
24	capital	44	0.12	78	0.01	131.48
25	has	234	0.61	2380	0.26	129.51
26	indian	29	0.08	18	0	126.56
27	its	159	0.42	1316	0.14	126.03
28	world	98	0.26	566	0.06	123.86
29	brics	19	0.05	0	0	122.96
30	2010	26	0.07	12	0	121.83

Clearly most keywords in Table 1 fall into two semantic categories:

- Lexis relating to economic terms: *growth*, *economic*, *economy, economists*, *GDP*, *debt*, *productivity*, *investment*, *fiscal*, *rates*, and *capital*

- Lexis relating to places: C*hina*, *China's*, *America*, *India*, *Indian*, *tropical* and *world*

The concordancing lines of other keywords such as *emerging* and *BRICs* demonstrate that *emerging* is primarily used to modify economies and markets, and that *emerging economies/markets* correspond to BRICs including China, India, Brazil and Russia. In fact, *Russia, Russian, Brazil*, and *Brazil's* are all keywords in the corpus, except that their keyness is far lower than that of C*hina, China's, India*, and *Indian*, as shown in example 1-3.

(1) Europe and the U.K. actually contracted, while China (and several other **emerging** economies) grew notably less briskly. (The Economist 2012-12-17)

(2) Normally, prices slump during a Us recession but demand form **emerging** markets means that hasn't happened; the west has turned from being a price-setter into a price-taker. (The Economist 2013-01-24)

(3) IN THE decade after Jim O'Neill of Goldman Sachs coined the acronym "**BRICs**" in 2001, grouping together four big countries with the potential for sustained growth, the "B", Brazil, really put itself on the economic map. (The Economist 2013-05-07)

It is evident that BRICs or the emerging economies/markets, and China and India in particular, are big concern regarding economic growth over the past years.

4.2 Key Semantic Domains

For p < 0.01 with 1 d.f., the cut-off of 6.63 (log likelihood value) indicates that there are 80 USAS tags significantly overused and 68 tags significant underused. The underused semantic domains include *Pronouns*, *Anatomy and physiology*, *Kin, Clothes and personal belongs*, *Knowledgeable*, *Music and related activities*, *Religion and the supernatural*, *Warfare*, etc. Given the limited length, negative key semantic domains will not be discussed in the rest of this paper.

Table 2 presents the top 20 semantic domains. The occurrence of some semantic domains is not a surprising finding, because one would expect that a corpus of economic growth has much to do with *Business*, *Money*, *Measurement*, *Industry*, and the like. Domains like Geographical names can be easily identified by manually grouping the keywords, as presented in previous section. Nevertheless, the ability to extract key semantic domains demonstrates clearly the advantages of the comparison at the semantic level. Because one only needs to examine a much smaller number of key domains in the semantic comparison than the number of keywords, the overall trends are easier to identify.

Table 2. The top key semantic domains in the corpus of economic growth

Rank	USAS tag	Freq	%	RC. Freq	RC. %	LL	Semantic Domains
1	N3.2+	600	1.58	1259	0.14	1645.65	Size: Big
2	I2.1	521	1.37	1614	0.17	1128.48	Business: Generally
3	I1.1	479	1.26	1796	0.19	902.22	Money and pay
4	I1.3-	144	0.38	265	0.03	422.52	Cheap
5	Z2	868	2.28	9920	1.07	374.48	Geographical names
6	I1	265	0.7	1404	0.15	366.81	Money generally
7	N3.8	121	0.32	305	0.03	299.13	Measurement: Speed
8	N5	464	1.22	4958	0.53	232.24	Quantities
9	I1.2	177	0.46	1103	0.12	205.3	Money: Debts
10	I4	95	0.25	327	0.04	190.93	Industry
11	N3.8-	80	0.21	230	0.02	182.15	Speed: Slow
12	I2.2	235	0.62	2167	0.23	156	Business: Selling
13	T3-	229	0.6	2414	0.26	117.91	Time: New and young
14	N5+	306	0.8	3805	0.41	107.14	Quantities: many/much
15	M7	327	0.86	4238	0.46	102.22	Places
16	W1	119	0.31	1004	0.11	91.55	The universe
17	G1.1	275	0.72	3523	0.38	88.81	Government
18	A2.1+	304	0.8	4121	0.44	83.22	Change
19	I1.3+	38	0.1	132	0.01	75.86	Expensive
20	I1.1+	51	0.13	252	0.03	75.63	Money: Affluence

The most significant difference (LL value 1645.65) in the semantic comparison is for the tag *N3.2+* representing the semantic field *size* (big). Under this umbrella term are words like *big, large, vast, expanding, growing, expanding, profound, substantial*, etc, as exemplified below:

(4)　Wage deals struck by **big** firms with big unions are imposed on others by bargaining councils. (2013-06-01)

(5)　As a result, **large** parts of the developing world will narrow the income gap between themselves and richer nations. (The Economist 1997-06-12)

(6)　But the world will need wisdom and stamina to reap the potentially **vast** benefits. (The Economist 1997-06-12)

(7)　China is a **vastly** more prosperous and expansive country than it was 20 years ago. (The Economist 2012-05-26)

(8)　Asia's third-**largest** economy **expanded** by 5% in the year to March…(The Economist 2013-06-29)

The second most significant difference (LL value 1128.48) indicates the overuse of the semantic domain of *Business: Generally* (I2.1) in the corpus of economic discourse. Upon examining the concordance for this tag (part of which is shown in Figure 1), it can be seen that under this umbrella term are words like *economy, economies, economist, business, company, firm, productivity, infrastructure, financial services*, etc.

	521 occurrences.		Extend context
think anew about the world	economy	. In the past , differences	1 More | Full
s.. better than many modern	economists	. He noted that sea-based tr	2 More | Full
England's relatively high	productivity	to the advantages of natural	3 More | Full
isease and low agricultural	productivity	in sub-Saharan Africa kept p	4 More | Full
ew more slowly than coastal	economies	(being entirely landlocked	5 More | Full
hosting and cloud-computing	firms	, such as Google and Amazon	467 More | Full
The level of comfort Indian	companies	get from Singapore is unmatc	468 More | Full
ration Centre . When Indias	economy	thrived, in 2003-08, so di	469 More | Full
Indias balance of trade in	business	and financial services has s	470 More | Full
ce of trade in business and	financial services	has slipped into modest defi	471 More | Full

(9)　The problem doesn't exist in America and other big **economies** because the large number of **companies** in such places—the US has about 6m employers to Canada's 1m—means no one **firm** will stand out when investment statistics are given by sector or region. (The Economist 2013-06-13)

(10) THE "**productivity** paradox" has been solved. Robert Solow, a Nobel laureate in economics, once famously observed that "you can see the computer age everywhere but in the **productivity** statistics". (The Economist 2003-09-11)

(11) Film buffs now view Swiss dream-sequences as cheesy, but India's big offshore hubs are more in fashion than ever. They present a mirror image of India's red tape, weak **infrastructure** and graft. (The Economist 2013-08-10)

The third most significantly used domain in the corpus of economic growth is *Money and pay* (I1.1). Under this umbrella term are words like *income, investment, saving, GDP, advances, capital, remittances, spoils*, *afford*, as shown in the sampling concordancing lines:

479 occurrences.			**Extend context**
loping world will narrow the	income	gap between themselves and ri	1 **More** | **Full**
graphyfor instance , through	investment	in transport and communicatio	2 **More** | **Full**
vernments with high rates of	saving	and low spending relative to	7 **More** | **Full**
and low spending relative to	GDP	. Each rise of ten percentage	8 **More** | **Full**
s found to raise the overall	GDP	growth rate by one percentage	11 **More** | **Full**
ates fall first , as medical	advances	reduce infant mortality and e	12 **More** | **Full**
free access for exporters to	capital	and intermediate goods , atte	13 **More** | **Full**
y heavily on flows of worker	remittances	: Bolivia , Burkina Faso , El	17 **More** | **Full**
t how to divide the economic	spoils	than how to enlarge them . Bu	35 **More** | **Full**
le exchange rate and can not	afford	to let its rural masses settl	38 **More** | **Full**

Needles to say, the semantic domains are closely related. The 7th most significant semantic category is labeled as *Measurement-Speed* (N3.8). There are 121 occurrences regarding this domain (N3.8), of which 80 instances fall into the 11th most significant category *Speed: Slow* (N3.8-), and 23 instances fall into the 21th most significant category *Measurement-Fast* (N3.8++). That is, the words and phrased used to describe *slow* is five time more frequent than those to describe *fast*, indicating that the recent concern of economic growth is its slowdown of the speed. Below are some examples containing words that symbolize the category *Speed: Slow*.

(12) When giants **slow down**: The most dramatic, and disruptive, period of emerging-market growth the world has ever seen is coming to its close (The Economist 2013-7-27)

(13) Inflation is relatively modest, but wages **lag** far behind. (The Economist 2013-8-10)

(14) In the first quarter of 2013 **sluggish** sales to the United States, by far Mexico's largest export market, helped reduce growth to a modest 0.8% compared with the same period in 2012. (The Economist 2013-05-25)

(15) Since 1970, however, the observed rate of technological advancement has **slowed** sharply, despite the temporary bump from the internet (The Economist 2012-9-8)

(16) Nevertheless, there are signs that spending, though continuing to rise, is growing more **slowly**. (The Economist 2013-5-11)

Interestingly, in contrast to the slowdown of emerging markets, America's economy was reported to grow faster:

(17) Evidence is mounting that America's GDP grew **faster** in the second quarter than the initial estimate of 1.7%, and has accelerated since. Healthy retail sales, rising production orders and low jobless claims all suggest that growth could be around 2.5%. (The Economist 2013-08-17)

Closely related domains also include the 38th most significant category *Healthy* (B2+) and the 44th most significant category *Health and disease* (B2).

(18) After a slump in the 1980s and 1990s, middle-income economies rebounded in the 2000s and have maintained a **healthy** growth rate more recently. (The Economist 2013-04-13)

(19) Over the longer run a **healthy** rate of nominal GDP growth would absolutely make reductions in public debt easier. (The Economist 2013-04-24)

(20) As a result, the **healthy**-nominal-GDP route to deleveraging would probably entail a bout of inflation above recent levels. (The Economist 2013-04-24)

(21) AS MOST of Africa begins to prosper, the continent's biggest economy is **faltering**. Figures released on May 28th showed that GDP in South Africa rose at an annualised rate of just 0.9% in the first quarter. (The Economist 2013-06-01)

(22) Slow growth and a sliding currency are alarming symptoms of a deeper **malaise** (The Economist 2013-06-01)

The metaphor that ECONOMY IS AN ORGANISM (or more specifically ECONOMY IS A PERSON) can be easily traced when examining the significant domains.

5. Discussion

The analysis of discourse of economic growth from *The Economist* reveals important aspects concerning current status quo of world economy. First, it is highly evident that BRICs or the emerging markets/countries (especially China and India) have played an active and leading role regarding economic growth over the past years. Another concern is the slowdown of the emerging markets which once created a miracle over the past decades.

The starting point to investigate a specialized corpus is normally to conduct keyword analysis. Influential as it is, the keyword analysis has a number of problems. By extending keyness analysis from word level to grammatical and semantic levels, some problems can be resolved. The apparent merit of Wmatrix lies in the mere fact that it allows macroscopic analysis (the study of the characteristics of whole texts) to inform the microscopic level (focusing on the use of a particular linguistic feature) and thereby suggesting those linguistic features which should be investigated further (Rayson, 2008). Furthermore, Wmatrix enables researchers to identify some of the significant semantic domains which are virtually unlikely or at least difficult at the word level. Third, an important semantic class can group together these low frequency words that would otherwise have been missed. For instance, words like *lag* and *sluggish* are quite infrequent and are thus not keywords, but they are still grouped into the 11th most significant category is *Speed: Slow* (N3.8-). The final advantage of Wmatrix briefly mentioned in Section 4 is its capacity to help researchers to quickly access economic metaphors or the absence of some kind of metaphor like *war* metaphor in this study.

Nevertheless, Wmatrix is not without its problems. Since the semantic tagger is automatic, there are some mistakes in the process. Error rates quoted for the semantic tagger quoted are 91% (Rayson, Archer, Piao & McEnery, 2004). So when carrying out the manual analysis of the results the researcher should also take account of possible tagging errors. Some errors, however, seem to has its roots in the initial semantic classification. As shown in Table 2, the 4th most significant category is Cheap (I1.3-). When the concordance output is under scrutiny, it is found that of total 144 occurrences there are 134 instances of the word *economic*. Nevertheless, it seems unreasonable to regard *economic* in the sampling concordances as the same semantic field with the word *cheap* or *duty-free*.

144 occurrences			Extend context
resulted in vast differences in	economic	performance ; in the future , po	1 **More** | **Full**
of law. Since Smith 's time ,	economic	development has continued to ref	2 **More** | **Full**
ange rates , moderate tariffs ,	duty-free	access for exporters to capital	7 **More** | **Full**
poverty rates tumbled . Gaping	economic	imbalances fuelled an era of fin	115 **More** | **Full**
ations, enormous quantities of	cheap	new labour became accessible . A	116 **More** | **Full**

Still, there are words or phrases that seemingly belong to one semantic domain but are not grouped into a certain category.

(23) Having grown by 2.3% a year between 1995 and 2002, it grew by 4% annually in the following eight years. But Brazil then **ran out of puff**. It grew by a disappointing 2.7% in 2011, and a dismal 0.9% in 2012. (The Economist 2013-05-07)

As illustrated by the context, the phrase **run out of puff** may as well be grouped into the *Speed: Slow*.

6. Conclusion

This study has contributed to the ESP (or more accurately English for Business Purposes, EBP) field in several ways. Firstly, it introduces EBP teachers to a very useful and powerful corpus tool Wmatrix for extracting keywords and identifying key semantic domains. Secondly, the analysis based on the small corpus reveals an important trend of the world economy. Finally, it bridges the new gap between economists and linguists that calls upon more profound dialogue between two seemingly remote disciplines as advocated by McCloskey (1985) three decades ago.

It has to be noted that this study is by nature descriptive and demonstrative due to its limited data. Further research can, building on the methodology used in this paper, explore a much bigger corpus of economic growth, or compare two corpora of economic growth from articles published from different journals or magazines.

Acknowledgements

References

Afida, M. A. (2007). Semantic fields of problem in business English: Malaysian and British journalistic business texts. *Corpora, 2*(2), 211-239. http://dx.doi.org/10.3366/cor.2007.2.2.211

Alejo, R. (2010). Where does the money go? An analysis of the container metaphor in economics: The market and the economy. *Journal of Pragmatics, 42*, 1137-1150. http://dx.doi.org/10.1016/j.pragma.2009.08.013

Archer, D., Culpeper, J., & Rayson, P. (2009). Love — 'a familiar of a devil'? An exploration of key domains in Shakespeare's comedies and tragedies. In D. Archer (Ed.), *What's in a wordlist? Investigating word frequency and keyword extraction*. London: Ashgate.

Baker, P. (2004). Querying keywords. *Journal of English Linguistics, 32*(4), 346-359. http://dx.doi.org/10.1177/0075424204269894

Baker, P. (2010). *Sociolinguistics and corpus linguistics*. Amsterdam/Philadelphia: John Benjamins.

Bazerman, C. (1993). Money talks: The rhetorical project of the Wealth of Nations. In W. Henderson, T. Dudley-Evans, and R. Backhouse (Eds), *Economics and language* (pp. 173–199). London: Routledge.

Berber, S. T. (1999). Using KeyWords in text analysis: Practical aspects. DIRECT Working Papers 42, São Paulo and Liverpool. Retrieved from http://www2.lael.pucsp.br/direct/DirectPapers42.pdf

Boers, F. (2010). Enhancing metaphoric awareness in specialized reading. *English for Specific Purposes, 19*(2), 137-147. http://dx.doi.org/10.1016/S0889-4906(98)00017-9

Bondi, M. (1999). *English across genres: Language variation in the discourse of economics*. Modena: Il Fiorino.

Bondi, M. (2010). Perspectives on keywords and keyness: An introduction. In M. Bondi & M. Scott (Eds.). *Keyness in texts* (pp. 1-18). Amsterdam: John Benjamins. http://dx.doi.org/10.1075/scl.41.01bon

Crawford Camiciottoli, B. (2007). *The language of business studies lectures*. John Benjamins B.V. http://dx.doi.org/10.1075/pbns.157

Culpeper, J. (2009). Keyness: Words, parts-of-speech and semantic categories in the character-talk of Shakespeare's *Romeo and Juliet*. *International Journal of Corpus Linguistics, 14*(1), 29-59. http://dx.doi.org/10.1075/ijcl.14.1.03cul

Del Lungo Camiciotti, G. (2005). Editorial strategies and metadiscourse in Marshall's *Elements of Economics and Industry*. In J. Bamford & M. Bondi (Eds.), *Dialogue within discourse communities: Metadiscursive perspectives on academic genres* (pp. 67-85). Tubingen: Max Neimeyer Verlag. http://dx.doi.org/10.1515/9783110933222.67

Firth, J. R. (1935). The technique of semantics. *Transactions of the Philological Society*, 36-72.

Firth, J. R. (1957). A synopsis of linguistic theory, 1930–1955. *Studies in Linguistic Analysis,* 1-32. Special Vol., Philological Society. Oxford: Blackwell.

Francis, G. (1993). A corpus-driven approach to grammar: Principles, methods and examples. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds), *Text and technology* (pp. 137-56). Amsterdam: John Benjamins.

Gaballo, V. (2012). The language of business, economics and finance: A corpus-driven, analytical discourse approach. Macerata: EUM.

Henderson, W. (1982). Metaphor in economics. *Economics, 18*, 147-153.

Henderson, W. (1994). Metaphor and economics. In R. E. Backhouse (Ed.), *New Directions in Economic Methodology* (pp. 343-367). London: Routledge. http://dx.doi.org/10.4324/9780203204085.ch16

Henderson, W. (2000). Metaphor, economics and ESP: Some comments. *English for Specific Purposes, 19*(2),

161-173. http://dx.doi.org/10.1016/S0889-4906(99)00035-6

Henderson, W. (2006). *Evaluating Adam Smith: Creating the Wealth of Nations*. London: Routledge.

Henderson, W., & Dudley-Evans, T. (1990). Introduction: The analysis of economics discourse. In T. Dudley-Evans & W. Henderson (Eds.), *The language of economics: The analysis of economics discourse. ELT Document 134*, (pp. 2-15). London: MEP/British Council. http://dx.doi.org/10.1016/b978-0-408-02680-2.50013-5

Henderson, W., & Hewings, A. (1990). A language of model building? In T. Dudley-Evans & W. Henderson (Eds), *The language of economics: The analysis of economics discourse. ELT Document 134*, (pp. 43-54). London: MEP/British Council.

Hu, C. Y. (2014). A Corpus-based analysis of bubble metaphor in economics. *Journal of PLA University of Foreign Language, 1*, 18-31.

Kehoe, A., & Renouf, A. (2002). WebCorp: Applying the Web to linguistics and linguistics to the Web, in: Proceedings of WWW 2002, Honolulu, Hawaii. Electronic publication. Retrieved from http://www2002.org/CDROM/poster/67

Leech, G., & Smith, N. (2000). *Manual to accompany The British National Corpus (Version 2) with Improved Word-class Tagging*. Retrieved from http://ucrel.lancs.ac.uk/bnc2/bnc2postag_manual.

Mahlberg, M. (2013). *Corpus stylistics and Dickens's fiction*. New York: Routledge.

McCloskey, D. N. (1985). *The rhetoric of economics*. Madison: The University of Wisconsin Press.

Rayson, P. (2008). From key words to key semantic domains. *International Journal of Corpus Linguistics, 13*(4), 519-549. http://dx.doi.org/10.1075/ijcl.13.4.06ray

Rayson, P., Archer, D., Piao, S. L., & McEnery, T. (2004). The UCREL semantic analysis system. In *Proceedings of the workshop on Beyond Named Entity Recognition Semantic labeling for NLP tasks* in association with *4th International Conference on Language Resources and Evaluation* (LREC 2004), 25th May 2004, Lisbon, Portugal (pp. 7-12). Paris: European Language Resources Association.

Scott, M. (1997). PC analysis of key words — and key key words. *System, 25*(2), 233-245. http://dx.doi.org/10.1016/S0346-251X(97)00011-0

Scott, M. (2010). Problems in investigating keyness, or clearing the undergrowth and marking out trails. In M. Bondi & M. Scott (Eds.), *Keyness in texts* (pp. 43-57). Amsterdam: John Benjamins. http://dx.doi.org/10.1075/scl.41.04sco

Scott, M., & Tribble, C. (2006). *Textual patterns: Key words and corpus analysis in language education*. Amsterdam: John Benjamins. http://dx.doi.org/10.1075/scl.22

Sinclair, J. (2004). *Trust the text: Language, corpus and discourse*. London: Routledge.

Stubbs, M. (1996). *Text and corpus linguistics*. Oxford: Blackwell.

Stubbs, M. (2010). Three concepts of keywords. In M. Bondi & M. Scott (Eds.). *Keyness in texts* (pp. 21-42). Amsterdam: John Benjamins. http://dx.doi.org/10.1075/scl.41.03stu

Sun, Y., & Jiang, J.L. (2014). Metaphor use in Chinese and US corporate mission statements. *English for Specific Purposes, 33*(1), 4-14. http://dx.doi.org/10.1016/j.esp.2013.07.005

Tribe, K. (2015). *The Economy of the word: Language, history and economics*. Oxford: OUP. http://dx.doi.org/10.1093/acprof:oso/9780190211615.003.0002

White, M. (2003). Metaphor and economics: The case of growth. *English for Specific Purposes, 22*(2), 131-151. http://dx.doi.org/10.1016/S0889-4906(02)00006-6

Wierzbicka, A. (1997). *Understanding cultures through their key words*. Oxford: Oxford University Press.

Williams, R. (1976). *Keywords: A vocabulary of culture and society*. London: Fontana.

Copyrights