USING BRITISH NATIONAL CORPUS IN LANGUAGE TEACHING AND LEARNING

I. Pekařová

Technická univerzita v Liberci,
Ekonomická fakulta, Katedra cizích jazyků,
Studentská 2, 461 17, Liberec, Česká republika
ivana.pekarova@tul.cz

Abstract

This contribution sums up the current experience with and insight into using the *British National Corpus* in classes of English at the Faculty of Economics at the Technical University of Liberec. Since the results so far have proved to be positive, the author formulates a plan how the used activities will be extended in the academic year 2009/2010.

1. Introduction

There were several reasons why the British National Corpus (BNC) was included into the repertoire of teaching aids which were available for classes of English as a foreign language at the Faculty of Economics at the Technical University of Liberec.

First, a lot of modern textbooks and dictionaries nowadays are based on corpora, or at least, they benefit from information collected in corpora. This makes them much more authentic, connecting them to the real language produced by concrete speakers, in contrast to the previously used teaching materials often focusing mainly on grammar and the written language. Some of the older course books (for example, English for Economists, 2002, or New International Business English, 2000 had a persistent tendency to teach English which did not seem to exist anywhere but in the language classroom. Second, in spite of the fact that modern textbooks are much more natural and they often work with concrete spoken utterances, the English language keeps developing in such a fast way that no textbook writer can really match in his or her work. Also, all the potential patterns of word combinations can never be captured in one teaching text, no matter how good and up-to-date it is. Third, textbooks like Business Objectives, 1993 still tend to focus more on the written language, which often results in a discrepancy between what students master successfully in class and how they thrive when involved in real communication acts outside of classrooms. And finally, international textbooks cannot include all individual needs of different speakers of English, and that is why there may not be enough good examples sufficiently illustrating the usage of lexical items introduced. Then teachers like us following old-fashioned textbooks and having a rare access to online sources or a reference library in the classroom are left almost empty-handed. They can rely only on their intuition, frequently producing hastily made examples, which can be anything but authentic. Or, they can utilize a limited range of language usage examples recorded in textbooks, whose limitations and date of origin usually affect inferences that can be drawn from them. That is why it seems that some tool for complementing our intuition would be extremely useful. I claim that we can use the corpus apart from various types of dictionaries for accomplishing this aim, and also for a more effective way towards creating reliable partial generalisations and working hypotheses about the language functioning.

Let us proceed now to a brief explanation about what the BNC is, where and how it can be obtained, and what features it offers to teachers and their students.

2. British National Corpus

BNC is a website comprising altogether 100 million English words in approximately 4,000 different text samples, collected between 1970s and 1993. It allows one to run a query against the collection of both written and spoken texts from a wide range of sources. The previously mentioned facts clearly indicate it being a monolingual, mixed and synchronic corpus. Any Internet user interested in studying the state of the present day English has a free access, for non-commercial purposes, to this representation of the current language available at the following address: http://www.natcorp.ox.ac.uk/

In the part focusing on the written samples, the extracts come from regional and national newspapers, periodicals, journals, specialist books and popular fiction, letters, memoranda, essays, academic discourse, and other kinds of texts. The spoken repertoire consists of transcriptions of informal conversations, and spoken language collected in different situations and contexts. Two hundred volunteers were used to collect the data; they had been sampled for age, sex, education, and accent.

Undoubtedly, there are other corpuses an educator can select from, and by no means do I claim that the BNC is the best and only one. No in-depth comparison was accomplished to gather enough evidence for such a statement. The bias for adhering to this corpus has been manifold and mainly based on identifying this corpus as the one that would meet our present needs in the best way. The choice was partially dictated by the fact that our newly adopted core textbook series The Business, 2008 were British, so the corpus was an unproblematic way to complement our main teaching materials without any exposure to risky areas, e. g. puzzling prepositional phrases recorded in the corpuses based on the American English usage, etc. Furthermore, this corpus is friendly enough even for a partially competent IT user not to add further overwhelming workload through a challenging interface, software and/or compatibility requirements, etc. Last but not least, in the context of our educational system which is not sufficiently saturated with money, it is very important that the simple search function of the corpus is not charged for.

3. Corpora as a learning and teaching aid for the teacher

Non-native teachers are learners of English as anybody else. The more advanced speakers of English they become, the more varied forms and tools for further language growth they need. Since a real spoken contact with native speakers is still in many institutions in the Czech Republic hardly as sufficient as most of us would benefit from, the collection of authentic texts is then a rather welcome substitution for the insufficient amount of other encounters. Similarly to all learners, even teachers can reinforce their knowledge and push it forward by this rather comfortable access to the source of information providing multiple data on lexical, semantic/pragmatic, syntactic, morphological, graphological and orthographical issues. As if the detailed lexical coverage was not a powerful tool as such, the corpus also lends itself to a shortcut into getting information on cultural issues, history, geography and various other areas. To become a truly competent language speaker, any learner of a foreign language needs to master these facts alongside communication skills. Educators and students possess a way for correcting and complementing their intuition about how certain language elements in English function. All of us, on different levels of language knowledge, formulate our own working hypotheses, and we can test them against the data recorded in the corpus, and generate new hypotheses or formulate further queries connected with the gained insight and evidence available.

4. Discussion of concrete results of one search

A student of English as the second language in year 2 (i.e. on A2 level according to <u>CEFR,</u> 2001) asked about the word "*racketeering*". Let us use this request for clarification as an illustration of how a corpus can support and reinforce learning. The word does not appear in any context in the core textbook (<u>The Business</u>) used in that class. An obvious source is the dictionary available in class, and then consulting various dictionaries accessible online via the computer in the particular classroom, which is a positive, but in our context rather exceptional, situation. The information presented in the hard copy was as follows:

racketeering noun [U] DISAPPROVING ; making money from a dishonest or illegal activity; They have been accused of racketeering.
racketeer noun [C] DISAPPROVING ; someone who makes money from a dishonest or illegal activity

(Cambridge Advanced Learner's Dictionary, p. 1169, or the accompanying CD-ROM, 2008)
It included some advice on the very basic derivation, it clearly suggested that the word had a negative connotation, and it introduced one colligation through the sample sentence where a verb was used together with a preposition with the searched for word. The definition was rather vague not explaining exactly if the person who had been accused of something bad was a pickpocket, or guilty of embezzlement, or of money-laundering.
The CD-ROM accompanying the dictionary also displays entries not in the full form with the definition, but in the form of a list and this is the result for the search in question:

black market n	scam n
contraband n	scalp v SELL 1
contraband adj, at contraband n	scalp v , at tout (v) SELL UNOFFICIALLY 3
gun-running n	smuggle v,smuggling n , at smuggle (v)
pass V GIVE 7	tout v SELL UNOFFICIALLY 3
racket n CRIME 3	traffic n TRADE 3
racket n CRIME 4	traffic v
racketeering n	trafficking n , at traffic (v)

run v TAKE 17, under the counter, sell sb a bill of goods, sell sb a pup, sell sb a bill of goods
This places the studied word in the wider context of other words belonging to the same semantic field, and it also enables contrasting the meaning to other similar ones. We gain a better opportunity to compose a mental map where the delicate shades of meaning should be placed.
In our common quest we continued with various types of online dictionaries. They offered the following summaries or definitions:

- *racketeering* - vydírání (Google dictionary)
- Main Entry: [2]*racketeer*, function: verb , date: 1928, intransitive verb, : to carry on a <u>racket</u> transitive verb : to extort money from (Merriam-Webster Online)
- Two meanings are distinguished – legal and business - in - yourdictionary.
  racketeering legal – n = a system of organized criminal extortion of money or favours from businesses through the use of violence, intimidating, or other illegal means; a pattern of illegal activity carried out by a criminal group or syndicate, including but not limited to fraud, murder, extortion, and bribery.
  Its business definition said - *racketeering* = carrying on an illegal activity, or racket, to make money. Racketeering is generally applied to structured groups such as child pornography rings or union members involved in labour racketeering.

Another layer of information provided here says: See also bribe, bribery, extort, extortion, and fraud. Mentioned In : Wextortion , Wnepotism, Wprosecute , Racketeer Influenced and Corrupt Organizations Act , $Racketeer Influenced and Corrupt Organization Act $Taft-Hartley Act

Let us compare this information to what we get from the BNC. In the simple search function, it generates a random selection of 50 solutions. The search can be repeated, bringing some other examples and listing again some of those already disclosed. In this case there were 60 examples found altogether; 5 of them, chosen at random just for the illustration purposes, appear below as copied from the BNC screen.

CCC 1690 One of the government's liaisons with racketeering is partially documented.
CCC 1733 THE VIOLENCE, CRIME AND RACKETEERING ARE ALLOWED TO CONTINUE.
CCC 1766 Racketeering and syndicate crime have been invaluable to the government in that
CCC 1773 A government which condones and orchestrates crime, racketeering and violence as a means of alienating and isolating terrorist or political opponents from their host populations will, in the long run, pay the price of alienating itself from the general public.
CCC 1787 The killings, maiming, destruction of property, crime and racketeering continued through the autumn and winter.

The first piece of information is about the frequency of use and about the source from where the example of usage was recorded. The studied word, which the teacher's intuition might associate mainly with legalese, as was my case, and thus assume it to be present mainly in courtroom transcripts, was very often used in newspapers and journals (Guardian, The Economist, The Daily Telegraph, The Best of Sunday Times Travel), in specialised books, like one on British intelligence services, but also, rather surprisingly, in texts on music, films, and also in documentary programmes, for example about the topic money. This piece of information on the word's distribution is rather significant. It reveals that examples of usage which teachers base purely on their intuition may be not only with low authenticity, but teachers may be mistaken in their judgement about contextual associations, register, etc. Then, they could offer sentences unlikely to be heard in real communicative interchanges.

The structure of the corpus eliminates the danger of stereotypical associating English with typically white Anglo-Saxon protestants only, which may still be the case in some teaching texts. In Business Objectives, for example, one meets only one Chinese businessman, several photographs of Japanese businessman and one member of the cabin crew of a plane, and learns about an exotic place suitable as a venue for a conference. In contrast, the BNC is peopled by "inhabitants" from different places round the world, like the Philippines, Burma, Belfast, USA, and many others. The recorded examples come from various sources and thus we introduce different social, cultural and political issues, from stories about football players to those on rappers. Such an approach provokes discussion and empowers education towards citizenship. It invites students to apply critical thinking while they try to infer some hypotheses and conclusions. Simultaneously, it guarantees that we will eventually cover a topic interesting for every one of our students, which is a powerful motivational aspect. Also, when we come across a subject any of our students are much more knowledgeable about, they can teach others and share their extra knowledge, which is a needed contribution towards teaching future citizens and employees to work independently, creatively, but also to learn with and from others. It is clear that we cannot provide all the information we would want, but our aim is to enable students to formulate and to possess some partial generalisations about English that are memorable and relevant. This may materialize by introducing them to really typical examples

of the language usage. Evaluating the corpus search results gives them many more examples than just one dictionary search or an odd sentence offered by the book or the teacher. They practise working things out for themselves when searching in a wider context, not just the scope limited by the textbook's focus.

What were the results of the analysis of individual examples recorded in the BNC? First, semantic preferences have been clearly suggested by the frequency of encountering the searched-for entry in the company of several words. It is rather interesting that *racketeering* very strongly collocates with *crime*, and seemingly only in the order given below. It almost seems to have achieved the status of an irreversible binomial; while in the three-word combinations, the order of individual items tends to be interchangeable. In the category of other criminal acts often collocating with the word, there can be listed the following lexemes*: drug trafficking, smuggling, cocaine, money laundering, black market, contraband, killings, theft, maiming, extortion, bribes, tax/bank/securities fraud, use of firearms, hostage-taking and torture.* When *crime and racketeering* collocate with adjectives, the most frequently used ones were: *cynical, neighbourhood, fiscal, minor, scandalous, moral, sheer, and labour*. The topmost occurrence of collocations with the adjective *federal* proved to be somewhat unexpected as did the oxymoron-like expression *polite racketeering*. Apart from the adjectives, a wide repertoire of other words, with a much smaller frequency, can be found in the attributive position: *examples of, investigations into, a good deal of, charges of, potential of, liaison with, victim of, short of, 12 counts of , trial on, jail sentence for*.
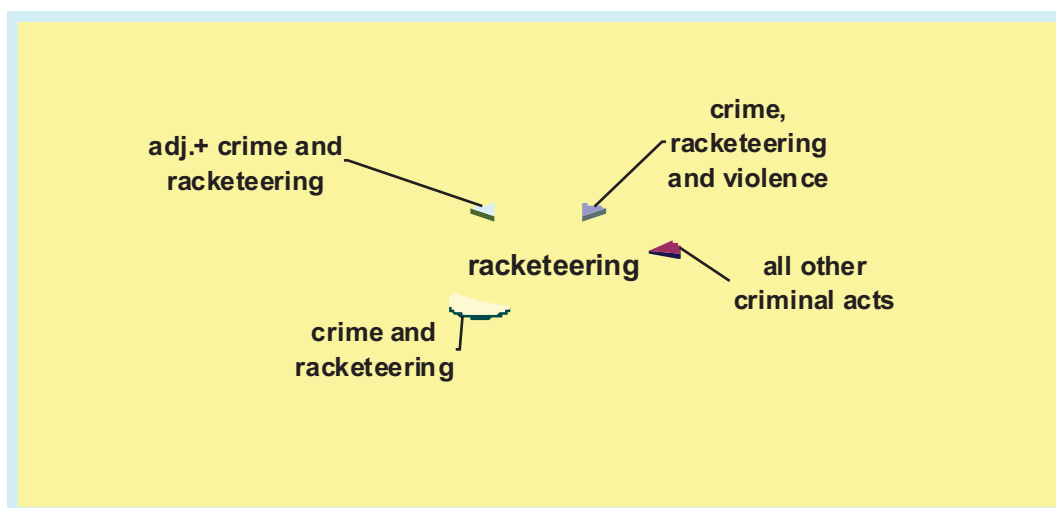
There is a noticeable frequency of collocations with the following rich scale of verb phrases:
be charged with, be sued for, convicted of/convicted on at least one count of racketeering , found guilty on racketeering charges, sentenced to imprisonment on racketeering charges, acquitted of; the less frequently used ones *: promote, condone, orchestrate, be involved in.*
The word *racketeering* converts easily and is used widely as an adjective itself, collocating most frequently with words like *charges, activities, allegations, suit, landlord, case*, etc.
It was clearly seen by both students and the teacher that from the linguistic point of view the results obtained when consulting the corpus were sophisticated and complex, sometimes even too much, and the teacher had to intervene.

5. Conclusion
We started our search about the meaning of the word *racketeering* and we learned so much in the process we could not have hoped for. Apart from this true well of outcomes directly

connected with studying English, one learns a lot about the social, political and economic context too. In our search, the legal system of the United Kingdom was introduced by showing various important players in that world, like *prosecutors, co-defendants, witnesses*, different stages of the prosecution – *sentence, charges, plead guilty, life imprisonment without parole*; then we learnt about the existence of some important documents like the *Racketeering Influenced and Corrupt Organisations Act*; we encountered stories of books or films and many other interesting types of information, totally unexpected at the beginning of the query. Some of this information was linguistically too difficult for A2 students, but it was appreciated fully by more advanced groups of learners.

Working with the corpus we repeatedly saw parts of society and geographical distribution of places where *racketeering* was more likely to be expected. The examples form texts published a long time ago and references to people like Michael X prove that while the language keeps developing, the human nature more or less remains unchanged. *Racketeering* is truly not a new issue in the human history, irrespective of what we may have thought.

By introducing students to the basics of working with this self-access learning aid, we can empower them for a more independent and more responsible approach to language learning. Working with the data, they can try to find out more about the language and its culture.

How to proceed from here? Due to the positive outcomes of the BNC searches done by the teacher with the students, a list of searches connected with the studied lexemes has been designed; students will be expected to cover the task independently, this time without the teacher in the role of an intermediary, and later on to present the results to others both orally and in the written form. The lexical items assigned to be studied are of two types. In one type of searches students are expected to focus on new words they do not know, while in the other type there are words they have encountered and which may belong to their active vocabulary. Even in the second type of task, the search will bring illustrations of what knowing the word really means. It is expected that every single case will illuminate additional layers of meanings of the supposedly known items they have not come across yet. It will be done both in classes on the lower and on the higher levels to get results from various cohorts of students. Only actual results will reveal if the outcomes will be on the same quality and if different types of students will face different problems in the process. This collection of data will be significant in identifying further the potential quality of student learning taking place on different levels of knowledge and communicative competence.

References

[1] ČAPKOVÁ, H. et. al.: English for Economists. Prague: Ekopress, 2002.
[2] JONES, L., ALEXANDER, R.: New International Business English. Cambridge: CUP, 2000.
[3] HOLLETT, V.: Business Objectives. Oxford: Oxford University Press, 1993.
[4] The Common European Framework of Reference (CEFR). Strasbourg: Council of Europe, 2001.
[5] RICHARDSON, K. et al.: The Business. Oxford: Macmillan, 2008.
[6] Cambridge Advanced Learner's Dictionary. Cambridge: CUP, 2008.
[7] <URL:http:// http://www.natcorp.ox.ac.uk/>
[8] <URL:http:// http://www.google.com/dictionary>
[9] <URL:http:// http://www.merriam-webster.com/>
[10] <URL:http:// http://www.yourdictionary.com/>

<div style="border">

USING BRITISH NATIONAL CORPUS IN LANGUAGE TEACHING
AND LEARNING

</div>

LEHREN UND LERNEN MIT DEM BRITISH NATIONAL CORPUS

Dieser Beitrag fasst die bisherigen Erfahrungen mit der Awendung des *British National Corpus* (BNC) in Englischenkursen an der Ökonomischen Fakultät der Technischen Universität Liberec zusammen. Da die bisherigen Ergebnisse sehr positiv zu schätzen sind, erwähnt die Autorin ihr Vorhaben, die auf dem BNC beruhenden Aktivitäten sowie die Erweiterung ihrer Nutztung im akademischen Jahr 2009/2010 fortzusetzen.

WYKORZYSTANIE BRYTYJSKIEGO KORPUSU NARODOWEGO DO NAUCZANIA I UCZENIA SIĘ

Niniejszy artykuł stanowi podsumowanie dotychczasowych doświadczeń przy wykorzystaniu *Brytyjskiego Korpusu Narodowego* na zajęciach języka angielskiego dla studentów na Wydziale Ekonomii Uniwersytetu Technicznego w Libercu. Jakoże obecne efekty były pozytywne, autorka w zakończeniu przedstawia sposób, w jaki planuje kontynuację wdrażania działań opartych na BNC oraz rozbudowywanie skali ich wykorzystania w roku akademickim 2009/2010.

VYUŽITÍ BRITSKÉHO NÁRODNÍHO KORPUSU PRO VÝUKU A UČENÍ SE

Tento příspěvek shrnuje dosavadní zkušenosti s použitím *Britského národního korpusu* v hodinách anglického jazyka u studentů na Ekonomické fakultě TU v Liberci. Protože současné výsledky byly pozitivní, autorka v závěru nastiňuje, jakým způsobem plánuje pokračovat v zavádění aktivit založených na BNC a rozšiřovat škálu jejich využití v akademickém roce 2009/2010.