Abstract

The study is to show how corpus-based methods map out the strategy for helping with the acquisition of idiomatic expressions and figures of speech. The aim of this approach is to provide students with scaffolds relying on which they will be able to develop their metaphoric competence and take fun out of learning at the same time.

Idioms represent an essential characteristic of effective communication. Harold E. Palmer determined phraseology as the key to fluency in a foreign language. On the other hand, mastering idioms presents severe difficulties related to remembering, semantization and performance.

Corpus-based approach provides a means of constructing a lexico-grammatical word profile: its general context, syntactic patterns, semantic prosody. All these aspects taken into account help learners come to grips with the grammatical company the idioms keep and their discourse. For example, to provide solid language skills and depth of cultural understanding many programs of university curricula require students to read classical books on a par with modern literature, which very often creates confusion as the students are unable to decipher which idiomatic expressions belong to the active or to the passive vocabulary of native speakers, which age groups use this or that idiom. Among Russian learners such idioms as it's raining cats and dogs, my cup of tea are very popular, though they might bring a smile or confusion to English speakers. The same story is with the French idioms un averti en vaut deux and mener quelqu’un en bateau.

In this regard, it is essential for a teacher to have an array of resources to maximize the learning process and to put across the importance of verifying the figurative usage of idiomatic expressions via language corpora, thus bringing diverse, reliable and authentic data in class. The use of corpora in class is known as data-driven learning, when learners act as researchers of the language rather than imitators.

As a basic premise, corpus-based approach to teaching should rest on several main principles: accessibility and clarity of tasks (guiding principle) as well as learners' motivation.

Taking into account all the benefits of data-driven learning we implemented corpus-based tasks in foreign language classes to see how they can help students to enhance their knowledge of the languages they study and their culture. The benefits of such tasks is their productivity. The outcome was the increased awareness of the aspects of the usage in sentences and connotation of phraseological units. Moreover it developed students' autonomy and cognitive skills.

Keywords: corpus, phraseological units, idioms, teaching a foreign language, corpus-based exercises.

1 INTRODUCTION

There has always been a strong interest to set expressions with fixed and constant structures which can be reproduced as ready-made units. In ancient times ready-made units were only regarded as a decorative device for stylistic purposes. With the appearance of the cognitive theory initiated by Lakoff and Johnson, phraseological units opened up as "mirrors" reflecting various spheres of our life connected with figurative and emotional perception of the world. The nominative nature of phraseological units is based on the emotions and attitude of the speaker towards different realia. The value of phraseological units lies in their creative individuality, which is compared by the Russian linguist Mokienko to a "national currency" in intercultural communication. Hence, foreign language learners being influenced by the psycholinguistic stereotypes find it hard to understand idiomatic expressions not to mention their acquisition. Let us look into the following examples: Ah well, horses for courses. Just because a plumber can mend your washing machine, it doesn't follow that he can mend your car as well. The individual parts of the first sentence do not contribute to the overall
meaning that it is important to choose suitable people for particular activities because everyone has different skills.

Students seem to understand every word in the sentences adorned with idioms, but still fail to grasp what the speaker wanted to say. Such ambiguity is typical not only of young learners, who tend to take words at their face value, but advanced learners as well. Still phraseological units fulfill such important functions as diversification of the language (e.g. to cheat - to pour banana oil, to take sb for a ride; to cause trouble - to put a cat among pigeons) and they make utterances more effective and succinct at the same time. A working awareness of figurative language, then, is a valuable tool for learners since it can facilitate comprehension, aid communicative interactions between native and non-native speakers extend vocabulary and help interpret and store new language items [1].

2 METHODOLOGY

As mastering formulaic language is notoriously difficult for learners, it has attracted researchers' attention to investigating the principals of selecting idiomatic expressions for students' study, seeking the ways and techniques of teaching that will build up students' knowledge of idioms. The books which we found the most useful in this respect are those based on teaching idioms through full-colour illustrations, accompanied by graded exercises, such as *Illustrated English Idioms* by Andrew Betsis and Lawrence Mamas. Traditionally effective methods of teaching idiomatic expressions are based on integrating reading, listening, writing, speaking activities together.

Nowadays corpus-based research is gaining momentum. Language corpora are actively used by linguists, interpreters and text-book writers. Obviously corpus-oriented approach to teaching a foreign language consolidates students' knowledge and develops their cognitive skills. However, the majority of students of language departments know very little about corpora. In 2016 U. Proshina, an EFL instructor, conducted a survey among first and second-year students of Moscow State University in order to learn if they used corpora while studying English. Only one in twenty-three students knew what corpus is [2].

As far as phraseology is concerned and its study through corpora we consider the works by P. Ŏurčo [3], U. Römer [4], H.W.Halawachy [5] and P.V. Sysoyev [6] to be very important.

3 RESULTS

3.1 Corpus linguistics and teaching a language

What is a language corpus? The term is defined differently by different scholars, though, in the main it is a large amount of machine-readable data of actual language that includes the collections of literary and non-literary text samples to reflect on both the synchronic and diachronic aspects of a language [7]. The most accessible one is the Internet, which can help learners search new words (the ones dictionaries do not contain) and example sentences where idioms are used. Moreover The Google Ngram Viewer (an online search engine) measures the frequency of the language unit in Google's text corpora.

For example, traditional textbooks on stylistsics and lexicology of the English language recommend using the word handsome if we want to tell a man that he is attractive and pleasant. But nowadays we can often come across the phrase a beautiful man and a handsome woman. Referring to Google Ngram Viewer and looking for the frequency of the use of phrases a beautiful man, a beautiful woman, a handsome man and a handsome woman, students noticed that a beautiful man occurs 100 times in colloquial speech, literature, popular magazines, newspapers, and academic texts, a beautiful woman - 1311, a handsome man - 554 and a handsome woman - 98 times. Looking at the graph students presumed that the use of the word combination a beautiful man is likely to continue rising. Then they dig deeper. In class, under the teacher's guidance the students were asked to visit the COCA (Corpus of Contemporary American English) and study the concordance. And here is the data they collected: a beautiful man is not less masculine than a handsome man; a handsome man is a traditionally good-looking guy who fits all the generic qualities of a pleasant-looking man. A beautiful man has something different and unique. A beautiful man is more of a work of art than a handsome man. Sometimes, though, You're beautiful is said to a man to recognize that he has done something for you that you appreciate, but that has nothing to do with his looks. It's like saying Thank you.
Such linguistic tasks arise students' curiosity and motivate them to use more complex corpora - national corpora. In our English classes we used the following corpora: BNC (British National Corpus), WebCorp, The Bank of English (Collins Corpus), ANC (American National Corpus), COCA (Corpus of Contemporary American English).

In retrospect we must admit that implementation of corpus-based exercises in a class can be hampered by some practical difficulties. Firstly, not all language classes are conducted in computer rooms or research centers, where there is a computer for each student. Secondly, corpus-based lessons make it difficult to incorporate other activities. Not to be discouraged by a new approach to teaching a language and not to intimidate students by new technology, we agreed on a monthly lesson devoted to the study of idioms and idiomatic expressions based not entirely on corpora but etymological references, various listening and speaking activities as well.

### 3.2 Corpus-based tasks for teaching phraseological units

Let us consider some basic corpus-based exercises connected with the analysis of the data provided by corpora.

1. **Using language corpora to measure the frequency of phraseological units.**

   **Task:** Study the frequency of the idiomatic expression *it's raining cats and dogs* in British English and American English through Google Ngram Viewer (Table 1, 2).

   **Table 1. British Corpus**

   ![Graph](image1)

   **Table 2. American corpus**

   ![Graph](image2)
According to the data provided, students infer that American English speakers use *it's raining cats and dogs* more often while the British English corpus shows its popularity has declined since its peak in the 1940s.

As an extension students can compare different language corpora to increase their language competence and cultural competence as well.

Task: In Russian there's an elephant in a china shop, while in English there's a bull in a china shop. In the French language there can be both an elephant and a bull. What will Ngram Viewer reveal? Which language has a higher frequency usage of the idiom in question?

Using language corpora to observe the distribution of the same idiom in different discursive functions (spoken, fiction, academic, magazine, newspaper) to reinforce the idea that idiomatic expressions' realm is not obligatory informal or semi-formal.

Task: Study the distribution of the idiomatic expression *cook the books* in different functional styles in COCA (Corpus of Contemporary American English) (Table 3) and identify in which registers it is used more often and least of all.

<table>
<thead>
<tr>
<th>Table 3. Concordance of the idiom “cook the books”</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Task:</strong> Study the distribution of the idiomatic expression <em>cook the books</em> in different functional styles in COCA (Corpus of Contemporary American English) (Table 3) and identify in which registers it is used more often and least of all.</td>
</tr>
</tbody>
</table>

It is worth observing that due to inattention some students keyed in *cook books* (neglecting the definite article) and as a result got different data. For teachers it is actually a blessing in disguise. It gives ground to discuss the main criterion of an idiom - its fixedness. Idioms have to be used correctly or they have no meaning. Paul Hancock illustrates it humorously in his book *Is That What You Mean, Too?* [8]. He takes 50 common errors in the use of idioms and shows what meaning idioms have if their fixedness is even slightly disturbed.

Also idiomatic expressions are prone to transformations, especially proverbs. A proverb is a brief popular saying that gives advice about how people should live or that expresses some commonplace truth or a belief that is generally thought to be true; it's a form of folk poetic creativity. The proverbs that deny the norms and rules of behavior asserted by the proverb are called antiproverbs. In English, for the designation of anti-proverbs, there is the term perverb, which is a combination of perverse and proverb, which literally means a spoiled proverb. For example, the perverb *Do not byte off more than you can view* derives from *Do not bite off more than you can chew*. The meaning of the transformed one is that it is necessary to download the information that we really need, which will be used, but not just lie in stock. Although fluent speakers will not find it difficult to comprehend such transforms, a learner who is used to “pure”
idioms can conclude that creative and unpredictable cases can occur, especially in newspaper and magazine headlines.

Task: Find the examples of deformations of the proverb *an apple a day keeps a doctor away* in COCA and in BNC (Table 4). Key in a day keeps and note down concordance you get. Think of any other proverbs and find their perverbs.

**Table 4. Variations of the proverb “an apple a day keeps a doctor away” (BNC)**

<table>
<thead>
<tr>
<th>ID</th>
<th>Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>is a day keeps the blues away in the world. That's the motto of car salesmen who</td>
</tr>
<tr>
<td>2</td>
<td>is a day keeps the blues away. It's the motto of car salesmen who</td>
</tr>
<tr>
<td>3</td>
<td>is a day keeps the blues away in the world. That's the motto of car salesmen who</td>
</tr>
<tr>
<td>4</td>
<td>is a day keeps the doctor away. It's the motto of car salesmen who</td>
</tr>
</tbody>
</table>

4 Another valuable task is to ask student to guess the meaning of an idiom.

Task: Study the concordance, find the idiomatic expression and guess what it means (Table 5).

**Table 5. Concordance of the idiom “off the table”**

<table>
<thead>
<tr>
<th>SEARCH</th>
<th>FREQUENCY</th>
<th>CONTEXT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All the above mentioned activities lead to the realization of the necessity to master the language for describing graphs, trends and bar charts. Undoubtedly, it develops argumentative skills and critical thinking skills in general. Another advantage found in such approach to teaching a foreign language is that it allows students of different abilities and interest to feel comfortable in class doing conventional tasks based on corpora (multiple choice, true/false, error correction, gap filling), while others can give vent to their research abilities. Data-driven learning turns the traditional deductive approach to an inductive one, when students’ autonomy is increasing, students are taught how to find their own answers.

4 **CONCLUSIONS**

Foreign language teachers are now trying to find balance between traditional methods and innovative ones. Innovations in teaching very often involve technology, though it should not be applied simply because it exists or seems to be cutting edge. It should be used to solve a pedagogical problem.
The aim of this paper is to raise awareness of the opportunities language corpora provide. Many teachers and learners tend to underestimate the corpus resources that exist and how they can be used to facilitate language teaching and learning. We wanted to show that such language units as phraseological expressions can be taught and remembered better through explanation and analysis.

REFERENCES


