

## **Frequency of occurrence of English colour idioms in Bulgarian vs. native learners corpora vs. everyday native speakers communication**

Tsvetelina Vateva  
Medical University “Prof. Dr. Paraskev Stoyanov” – Varna,  
Department of Foreign Language Teaching, Communications and Sport  
Tsvetelina.Vateva@mu-varna.bg

### **Abstract.**

The research presented in this paper investigates the frequency of occurrence of idiomatic expressions containing the basic colour terms in English as demonstrated through the use of English by Bulgarian learners at the university level, native learners of English at the same level of proficiency, and native speakers of English.

The paper analyzes data obtained from the following 3 corpora: the Bulgarian Component of the International Corpus of Learner English (ICLE), the LOCNESS corpus, and the British National Corpus.

The analysis explores and draws conclusions regarding 3 main research questions:

1. Which are the colour terms that are most frequently used in idiomatic expressions?
2. Is there any significant difference in the frequency of use of colour idioms between the groups of Bulgarian vs. native learners at the same level of proficiency?
3. What is the frequency of use of the colour idioms used by both the Bulgarian and the native learners in the common everyday communication of native speakers?

**Keywords:** English colour idioms, Bulgarian learners, native learners, native speakers, frequency, corpora.

### **I. Introduction**

The initial aim of the corpora research presented in this paper was to establish and compare the frequency of occurrence of idiomatic expressions containing the basic colour terms in English as demonstrated through the use of English by Bulgarian learners at the university level, native learners of English at the same level of proficiency, and native speakers of English.

However, the results from the comparison between the occurrences of colour idioms in Bulgarian and native learners' use turned out to be limited by the contextual topics of the two learners' corpora explored, so the focus of the research had to be narrowed further to exploring only the frequency of occurrence of the colour idioms from the two learners' corpora in the common everyday use of English by native speakers.

### **II. Research Questions**

The research questions that were aimed for exploration via this paper were the following:

1. Which are the colour terms that are most frequently used in idiomatic expressions?
2. Is there any significant difference in the frequency of use of colour idioms between the groups of Bulgarian vs. native learners at the same level of proficiency?
3. What is the frequency of use of the colour idioms used by both the Bulgarian and the native learners in the common everyday communication of native speakers?

### III. Methods

The analyzed data was gathered from the LOCNESS corpus, the Bulgarian Learners' Corpus and the British National Corpus. Quantitative analysis was performed by entering the data from the two Learners' corpora in tables summarizing the number of examples containing colour idioms against the total number of examples containing the respective colour term (including both literal and idiomatic usages). Then, data regarding the frequency of occurrence of the idiomatic expressions from the two Learners' corpora in native speaker communication was collected from the British National Corpus. The following section presents the results from this research and a qualitative comparative analysis of the research findings.

Colour term	Number of examples containing the colour term	Number of examples containing idiomatic expressions with the colour term
white	33	4
black	34	4
red	20	0
green	11	0
blue	16	5
yellow	3	1
brown	9	0
purple	0	0
pink	2	0
orange	0	0
grey	9	8

### IV. Results and Data Analysis

The quantitative results from the research in the Learners' corpora are presented in the following tables:  
**Bulgarian Learners' Corpus:**

Evidently, the use of idiomatic expressions containing the basic colour terms in English is not very common among Bulgarian learners at the level of proficiency, at which the corpus was collected. Most examples contain the colour terms with their literal meaning denoting the respective colours. However, we should bear in mind that the corpus itself is limited to essays on particular topics, and that the contexts of those topics may not exactly coincide with all general contexts of use of idiomatic expressions containing colour terms.

Most interestingly within this corpus, 8 out of 9 instances of the term “grey” were used in idiomatic expressions – all with the meaning of “a depressing or miserable existence/life”. This fact again reflects the contextual restriction of the essay topics, but it may also be interpreted to reflect the tendency of Bulgarians to view their existence as miserable and constantly complain about it.

The most frequently used colour terms in idiomatic expressions by the Bulgarian learners were grey, blue, black, and white (in order from the colour with the most examples to the colours with the fewest examples).

#### LOCNESS Corpus:

Colour term	Number of examples containing the colour term	Number of examples containing idiomatic expressions with the colour term
white	58	6
black	60	5
red	3	0
green	7	3
blue	4	0
yellow	0	0
brown	3	0
purple	0	0
pink	0	0
orange	4	0
grey	3	2

Similarly to the Bulgarian Learners’ corpus, the native learners corpus results included primarily instances of literal usage of the basic colour terms in English. This should probably be attributed again to the fact that this corpus is also limited to essays on particular topics, and that the contexts of these topics may not exactly coincide with the contexts of use of idiomatic expressions containing colour terms.

The colours that were most frequently used in idiomatic expressions by the native learners were white, black, green, and grey (in order from the colour with the most examples to the colour with the fewest examples). The results from the research in the British National Corpus are summarized below in the form of a list of the idiomatic expressions from the 2 Learners' corpora which occur in the British National Corpus, each followed by the number of examples containing the idiomatic expressions.

**British National Corpus:**

white culture – 3  
white-collar – 226  
black market – 157  
Black Power – 29  
black school – 5  
green belt – 219  
greenhouse effect – 275  
grey area – 69  
grey life – 2  
grey home – 2  
grey daily routine – 1  
out of the blue – 203

The most frequently used idioms from this constricted subset in everyday native speaker communication were the following: *greenhouse effect*, *white-collar*, *green belt*, *out of the blue*, *black market* and *grey area*.

These results seem to be very straightforward in demonstrating that in everyday native speaker communication the most frequently used colour idioms are the least contextually restricted ones (“white-collar”, “out of the blue” and “grey area”) and the ones that pertain to widely popular topics of discussion like climate change (“greenhouse effect”, “green belt”) and illegal activities (“black market”).

## V. Conclusion

The straightforward answers to the research questions of this paper are the following:

The most frequently used colours in idiomatic expressions are: *grey*, *blue*, *black* and *white* (Bulgarian Learners' Corpus); *white*, *black*, *green* and *grey* (LOCNESS corpus), and as demonstrated by the sub-research in the British National Corpus – the most frequently used colour idioms are the least contextually restricted ones and the ones pertaining to widely popular topics of discussion.

There is no significant difference in the frequency of use of colour idioms between the groups of Bulgarian vs. native learners at the same level of proficiency.

The frequency of use of the colour idioms used by both the Bulgarian and the native learners in common everyday communication by native speakers is highest for the least contextually restricted idioms.

The implications of this research show that the wider context of use and the more general meaning of the expressions are reflected in their more frequent usage, whereas the narrower context of use and the more specific meaning of the expressions are reflected in their less frequent usage in everyday native speaker communication.

## References:

Bulgarian Component of the International Corpus of Learner English: <https://uclouvain.be/en/research-institutes/ilc/cecl/icle.html>  
LOCNESS Corpus: <https://uclouvain.be/en/research-institutes/ilc/cecl/locness.html>  
British National Corpus: <http://www.natcorp.ox.ac.uk/>