

RED IN THE CORPUS OF CONTEMPORARY AMERICAN ENGLISH INTERPRETATION OF COLOR IN LAW AND POLITICS SECTION

Prihantoro

Universitas Diponegoro, Indonesia
prihantoro2001@yahoo.com

ABSTRACT

Color does not only correspond to light-spectrum-based visual identification of a concrete entity, but it may also symbolize abstract values conceived in a socio-cultural community. Some values are shared, but some others are distinctive. Even in one community, the meaning of a color might be multi-interpretable. For instance, the color of RED is understood as a 'stop' command when it comes on traffic light. However, in the compounds such as 'red specialist', 'red herring', or 'red army', we cannot take for granted that the <red> also means 'stop'. Both linguistic and meta-linguistic awareness are required to define what RED means. This paper seeks to describe the literal and non-literal meaning of RED in the Corpus of Contemporary American English (COCA), by focusing on laws and politic section. By using Keyword(s) in Context (KWIC) method, I have managed to retrieve 328 concordance lines where RED collocates with other token(s). I extracted the lines, including the extended context, for in depth interpretation. Both literal and non-literal meanings of the collocations are categorized into different classes, and the investigation indicates that in laws and politic science text, RED is a color that is often used to address negative polarities such as fallacy and ethnic discrimination.

Key words: Colors, red, negative polarity, corpus, meaning, keywords in context

How many colors do you know? The answer to this question might vary depending on two variables: physics and language. What we commonly understand is that there are 12 colors. However, a recent study in physics (Kinoshita et al, 2008) has showed color is so complex that 12 is merely a simplified figure. On the other hand, some studies in linguistics have shown that there are languages, where the users do not distinguish some colors that are distinctive in another language (let's say English). Consider Navajo language (Stea et al, 1972), where only one word is addressed to blue and green. Consider also Shona language that does not distinguish red and orange (De Bortoli & Maroto, 2001). Then how they distinguish the two colors? They technically do not. But when necessary, one of the strategies is using the color they know as an attributive to a

noun, such as; 'blue sky', 'blue like the sea water', 'green leaf', 'green like grasses' and etc.

Besides number of colors, serious attention has been given to the topic of color interpretation. Colors can be understood literally, or non-literally (metaphorically). The literal interpretation of RED¹ is as color itself (as described in physics studies as a particular composition of light spectrum that our eyes perceive). While to understand the literal meaning of RED requires mere visual and linguistic awareness, to understand the non-literal meaning of RED involves meta-linguistic awareness as well.

¹ In this paper, uppercased words like RED indicate a lemma, or the basic form that is present in the dictionary. When the words are surrounded by angle brackets <> they are the word forms, or the orthographic forms.

The focus of this paper is to retrieve expressions involving the lemma RED. It also seeks to describe and to categorize the meanings of red, both literally and metaphorically. By the end of this research, I will also determine the polarities of RED. The data for conducting this research is obtained from the social and politic section in the Corpus of Contemporary American English. The next section of this paper deals with related studies the importance of contexts in determining the meaning and polarity of a word. By the end of the next section, I will show the review of these studies and show how my paper can contribute more to the subject of discussion.

LITERATURE REVIEW

Context, Meaning, and Polarity

When human is exposed to speech or text, the information is processed gradually up to the point of comprehension (Fernandez & Cairns, 2011)². Comprehension means that human successfully decodes the information. To achieve this, linguistic, meta-linguistic awareness (or both of them) are required to do the decoding process completely. Compare example (1) and (2), and focus on the underlines:

(1) The supreme leader finally decided to resign

(2) Saddam can no longer fight the resistance. The supreme leader finally decided to resign

Article <the> in (1) indicates that the compound is definite. However, this information is not sufficient to decode to whom does <the> concern or who the supreme leader is. Compare to example (2) where the anaphoric referent is present. At this point, it is not difficult to understand the referent of <the>. However, you may recover the referent of <the> in (1) if the same knowledge is shared to other participants in the speech event. The knowledge is called meta-linguistic awareness.

In some cases, meta-linguistic awareness takes primary role in taking deduction. One of the examples is in the ambiguity resolution. In the

ambiguous sentence, 'I saw the man with the microscope', if you know what 'microscope' is, you will not deduce this sentence as a tool to see a person. Instead, you will deduce this sentence as there is a man with microscope that you saw. This may happen only when your understanding of microscope is the same as mine. So, what <black> color means to you? Our conception may be different. Black is often be interpreted as the color of death. However, in some areas in Indonesia, yellow is the color of the flag that you hoist when someone in your family passes away.

If participants in a speech event do not share the same concept, then the recognition of the referent is most likely to fail. This often happens in translation error.

(3) I am feeling blue today
Hari ini saya sedang sedih (sad)
*Hari ini saya sedang biru**(literally blue)

In English, the color BLUE has negative polarity (sad). Preserving the color in Indonesian translation makes the sentence semantically not compositional (although it is grammatically correct). A color survey in USA³ has indicated that people interpret colors to positive, negative polarity or neutral. Aerim et al (2010) built a machine readable dictionary to assign polarities to digital Korean text for sentiment analysis. They believed that there are some lemmas that fall to 'flexible' categories. Their argument is that the polarity may shift to [+] or [-] following the companion words. As an example, the polarity of BIG in 'mobile phone's big screen' is [+], while in 'big problem' the polarity is [-].

A recent study from (Yassine & Jeesun, 2014) revealed that even for a lemma that is inherently [+] or [-], the companion word may change the polarity as in 'very perfect' where the value is [+] and 'too perfect' where the value is [-]. This follows Firth (1957) tradition, as also shared by corpus linguists. Compare 'a white wizard' and 'a white sheep'. In 'a white wizard', you are most likely to deduce 'white' as good (as opposed to evil). However, this reasoning for this is because of it is the noun it specifies <wizard>. In 'white sheep', you will

² This process is reversed when producing output

³ See Richard, et al (2005)

resort to its literal meaning that there is a sheep and the color is white (the inherent color of a sheep is white). If that so, what about 'black sheep'? It is uncommon for a sheep to be black. In this case, you are most likely to resort to individual difference in a common group, such as a member of family who behaves differently from others. But consider this sentence, 'The biological experiment successfully modified sheep's DNA to give birth to black sheep'. In this case, you will most likely resort to a sheep that is actually black.

In literary texts, we may find many referents that are expressed metaphorically. Literary text is often contrasted to academic text. On one side, academic language is known to be straightforward. On the other side, its readers are segmented (not all people can understand the word choice). Therefore, only people with shared meta-linguistic awareness may understand the meanings. Unlike the previous works, my paper seeks to describe both the linguistic composition and the referents of expressions involving RED in COCA, which is claimed to be the most balanced and representative corpus of American English (Davies, 2008).

METHODOLOGY

Research Corpus and Automatic Retrieval

Research Data

The data in this paper is obtained from the Corpus of Contemporary American English (Davies, 2008). A Corpus itself is a collection of texts that is organized in a way that enables users to retrieve information based on queries (Mc Enery & Hardie, 2012). This might resemble the performance of search engines when you browse internet. However, syntax queries for a corpus are linguistically formulated and works beyond character similarity match.

The difference between using a search engine and a corpus interface might be described as follow. When queries are sent to internet, for instance STRIKE, it retrieves all documents where the character match is near or exactly 100%. It does not distinguish STRIKE as a verb or

noun, a named entity, an acronym etc. But when you use corpus interface, you can also retrieve linguistic information annotated in the text. For instance, you can browse different word forms of lemma STRIKE such as, strike, strikes, stroke, striker etc. Some more advanced queries on annotated corpus allow you to retrieve semantic information such as features, polarities, and word relation (synonym, antonym, hypernym, etc).

Data Collection

COCA is composed of texts from different types and this corpus covers a wide range of sections written (or transcribed) in American English. Registered users may access the website (<http://corpus.byu.edu/coca/>). I retrieved lemma RED in the corpus and selected one word form of the highest frequency hits. As for this research, I focused the retrieval on social and politic section of newspaper texts. Because there are different text types (spoken data transcript, fiction, academic, newspaper, magazine etc), the retrieval was controlled to focus to Academic section: law and politic texts.

The result of the retrieval was displayed on concordance lines. Each concordance lines were analyzed to decode the meaning of RED. Each word form of RED that conceives associative meanings was assigned to different category as opposed to literal meaning. The lists of possible non-literal meanings that are similar were categorized to specific classification. I finally assigned polarity based on this categorization to conclude this research.

FINDINGS AND DISCUSSION

Literal and Metaphorical Meanings of RED

The retrieval for lemma RED has managed to discover seven results. In figure 1, result number seven (R.E.D) is not shown, as it is not the word form of lemma RED. This seems to be an acronym and, COCA misretrieved it. Among the six results, <red> is the highest frequency word. Therefore, the research was focused on this word form.

	<input type="checkbox"/>	CONTEXT	FREQ
1	<input type="checkbox"/>	RED	90129
2	<input type="checkbox"/>	REDS	2740
3	<input type="checkbox"/>	REDDER	293
4	<input type="checkbox"/>	REDDEST	61
5	<input type="checkbox"/>	REDDING	38
6	<input type="checkbox"/>	REDDED	4

Figure 1. <red> as the high frequency word form

COCA is composed of five grand section/ text types: spoken, fiction, magazine, newspaper and academic. As I have commented previously, this research is focused on academic section (7160). See figure 2.

SECTION	ALL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
FREQ	90831	8926	32056	24533	18156	7160
PER MIL	195.61	93.39	354.49	256.73	197.96	78.62
SEE ALL SUB-SECTIONS AT ONCE						

Figure 2. Academic Section

History	Education	Geog/SocSc	Law/PolSci	Humanities	Phil/Rel	Sci/Tech	Medicine	Misc
935	196	724	328	1286	305	2322	463	530
76.36	20.76	44.75	38.14	107.83	45.25	164.97	69.10	124.51

Figure 3. Law and Politics

CLICK FOR MORE CONTEXT		<input type="checkbox"/>	[?]	SAVE LIST	CHOOSE LIST	-----	CREATE NEW LIST	[?]
1	1996	ACAD	PerspPolSci	A	B	C	are causing real problems for bureaucrats who insist on categorizing all Americans as colors -- red , yellow, black, or whi	
2	1996	ACAD	WorldAffairs	A	B	C	Union, which the Kremlin master had no doubt was in the works. The Red army was not, however, to march in a prevent	
3	1996	ACAD	WorldAffairs	A	B	C	the mouths of Stalin's closest cronies, of a general plan to unleash the Red army westward, first against its immediate w	
4	1996	ACAD	WorldAffairs	A	B	C	the famous Marshal Zhukov, one of the most successful Soviet commandants of the victorious Red Army in World War II	
5	1995	ACAD	PerspPolSci	A	B	C	, private foundations, and patrons; the political fallout from World War I the Red Scare, and the Cold War; the need to di	
6	1995	ACAD	PerspPolSci	A	B	C	as always harsh and brutal. As his Zarathustra put the matter, a " red judge " would precede the new thousand-year kin	
7	1995	ACAD	WorldAffairs	A	B	C	put together. Its record stands in startling contrast to the achievements of the International Red Cross and the Western I	
8	1995	ACAD	WorldAffairs	A	B	C	Warsaw. The Poles' goal was to liberate their own capital before the approaching Red Army could defeat the German troi	
9	1995	ACAD	IntAffairs	A	B	C	result, they grow to expect additional allocations, and bailouts of firms in the red become institutionalized as regular and	
10	1995	ACAD	IntAffairs	A	B	C	tourist hotels providing employment for redundant labor, government bureaus that produce little more than red tape and	
11	1995	ACAD	ABAJournal	A	B	C	: Philip Morris delivered the 25 boxes of documents on hard-to-photocopy paper, dyed dark red , that additionally seeme	
12	1995	ACAD	ABAJournal	A	B	C	Francisco-based consultant who counsels lawyers looking for alternative careers. # " Money is the red herring, " she say	
13	1995	ACAD	CanadaLaw	A	B	C	cost-sharing basis by competitors. Joint activity by actual or potential competitors raises possible antitrust red flags. Thr	

Figure 4. Example of Concordance Lines

Frequency hits in academic section shows 7160 hits. However, the academic section is divided into several sub sections. My focus in this paper is on law and politic, so I went deeper to the law and politic section as shown by figure 3.

As I have mentioned previously, the result of KWIC based query is presented in the form of concordance lines. Figure 4 presents us with some of the concordance lines (total 328 lines). Each concordance line conceives year of publication, section, source, and target keyword <red> with left and right context. The sample for this research was 25%, which means it focused on 82 lines. The sampling was taken randomly. The context for each concordance is left and right context. However, this context can

be expanded up to the textual level.

Significant improvements are seen on both the source and the text information. The source displays date, publication information, title, author and the source itself. It also expanded the coverage from just left-right context to the level of body text.

Color

There are some examples when <red> refers to its literal meaning. Consider some entities like <red sea>, <red brick>, <red meat>. The nouns in the examples are attributed with <red> as they are literally and visually red. Red sea is the sea that might change color to red because it is densely populated with red algae. The same reasoning might apply to red brick and red meat (raw meat). Now consider <red cross> and <red crescent> in example (4):

(4) The existence of an apparatus capable of mounting the required responses is a function of the growth of the world relief system, in large measure represented by the UN system, the **Red Cross** and **Red Crescent Societies** and major non-governmental organizations such as Oxfam.

What is interesting about the examples is the red itself might be interpreted both literally and non-literally. It is true that the symbols of these institutions are colored red⁴. Then why we do not resort to a cross sign that is red and a crescent (in the sky) that is red? The presence of the companion words like <world relief system> and <required responses> has triggered our meta-linguistic awareness to suggest that these two institutions are responsible for rescue missions, and prevent us to resort to false assumption.

Ethnic Discrimination

Black and white are two colors that are fully contrast. These colors are often used to address two different polarities, like good and evil,

positive and negative, rich and poor etc. Ethnicity was once a sensitive issue In America, as well as in South Africa. Even in South Africa, ethnic segregation was authorized by Apartheid system (Downing, 2004). Now I will give you an example of how color in USA discriminates⁵ ethnicities. See example (5):

(5) Growing numbers of interracial pairings, along with the new terminology, are causing real problems for bureaucrats who insist on categorizing all Americans as colors — **red**, yellow, black, or white

That color is the representation of Ethnicity in USA is also acknowledged by (Yanow, 2003). In her book, *Constructing Race and Ethnicity in USA*, she further described the referent of those colors: red for American (Indian), Yellow for (Asian), black for African, and white for Caucasian. However, she also admitted that in the recent days the segregation gap has become small and smaller. Few years later after her publication, USA has its first African-American President, Barack Obama.

Socialist-Communist

In COCA, some expressions involving <red> are related to socialism and communism. Although the world war has ended long time ago, this stigma does not end and still last up to now. Consider example (6):

(6) The project was resuscitated in the 1970s by its staunchest advocates —the so-called " **red specialists** " who were trained in the Soviet Union as engineers and scientists and are China's greatest proponents of " grandiose" engineering and energy projects

The association of <red> to communism and socialism is much related to the two polarities (in the past) between USSR (Union of Soviet Socialist Republic) and USA. It is also associated to Republic of China where the governance is run by communist party. And in fact, the color

⁴ This part is quite tricky. For example, <red face> is often addressed to anger. At this point, the meaning seems non-literal. However, the face of someone who is angry may also be red; therefore, literal deduction is not entirely wrong.

⁵ The word <discriminate> does not always relate to social discrimination, but it also means 'to differentiate' as 'the respondents successfully discriminated [f] and [v] after several attempts'

of the flag of republic of China is also red. An elite soviet army was even named <red army> as shown by example (7):

(7) Georgii Zhukov, later the famous Marshal Zhukov, one of the most successful Soviet commandants of the victorious **Red** Army in World War II, was the commander he selected for his anticipated thrust to the west

In (6) and (7), <red> is marked by quotation, and upper case respectively. In (7), the initial font is uppercased as it is a proper noun. In (6), it is given a quotation as the writer wants us to 1) focus and 2) resort to metaphorical meaning that red is not only color, but it also signifies socialist-communist ideology.

Fallacy

Red can also be associated to fallacy, as in the compound <red herring>. Herring is a kind of fish; and there is an actual herring, in which the color is red. A question may rise why herring is used, or why fish is involved in such negative referent. You surely remember the expression such as 'there is something fishy (suspicious)'. But now, let us focus on <red herring>:

(8) I have come to regard economic espionage as something of a **red** herring in this debate — something that, although wanted by neither the government nor business, provides a distraction from the more compelling practical questions at hand.

The compound <red herring> in example (8) indicates a kind of fallacy, where you propose an idea that mislead or not relevant to the subject of talk. We can understand from the word choice <debate>, that there are at least two parties involved in the speech event. The term <distraction> that is used by the speaker clearly suggested that the topic (economic espionage) that the opposing party is trying to propose is not relevant to his proposition.

No-cross-Rules

Red means stop. This is an almost universal color meaning and used in some idiomatic

expressions in law and politic science text in COCA. In (9), Ehud Barack clearly said <would not cross>. The extended context implied that Israel would cease (stop) attacking Gaza.

(9) In July 2000, prior to his departure for the Camp David summit, Israeli Prime Minister Ehud Barak stipulated to his Cabinet the four " **red** lines " he would not cross during negotiations with the Palestinians.

(10) If so, let us work on it and add to it Canadian football. Is it defining the acceptable methods of protection of culture, something like what was done in the WTO on subsidies, **red** light, green light?

(11) The ABA played a key role, for instance, in efforts by a number of bars to block the Federal Trade Commission's attempt to regulate lawyers as " creditors " under the so-called **red** flags rule developed by the FTC to help detect identity theft

What Ehud Barak wanted to signify here (9) is four details in a rule that he will not cross during the negotiation. In (10), it makes use the 'traffic light' association which I believe is quite universal where red means 'no' or 'stop' and green means 'yes' or 'proceed'. In (11), 'red flags' are used to mark an unfortunate event, in this case, theft.

Discomfort

I also discover some expressions that <red> is often associated to a discomfort. In (12), <red ink> refers to an undesirable situation. The subsequent clause 'limited prospects for survival' confirms this.

(12) We were drenched in **red** ink and faced with limited prospects for survival.

The situation that <red> signifies in (12) has caused a discomfort. It must be understood negatively, as the impact is limited prospect for survival. The word red is sometimes associated with something bad, or discomfort such as in red zone (dangerous). There is an expression in Indonesian where you get bad grade, which is

nilai merah red mark, where in the past, teachers wrote bad grade with red ink.

Emotional Stress

The term 'red face' is a common expression shared across cultures to express that one is under a circumstance of certain emotional stress. One that might make it universal is the face is literally red; which is caused by the increase of stress-related hormones on face skin pigment. The stress may involve anxiety, embarrassment, anger, romantic situation or etc.

(13) A bit of upfront planning can avoid a lot of **red** faces later when a cloud provider's doors are locked and the remote servers storing your data start popping up for sale on eBay

In (13), the red faces are to be avoided as people get stressed when they are dealing with computer problems, in this case, providers and servers.

Complexities

The expression 'red tape' is used to show multi-layered regulations that people have to go through before doing something that is considered unnecessary. One however, must go through this, otherwise will not achieve his/her goal. This is usually related to bureaucracy.

(14) They are also increasingly willing to travel, live and work abroad, in part out of frustration with the corruption and **red** tape that continue to stymie entrepreneurial ambitions inside Russia

In (13), 'red tape' is the subject defined in the relative clause construction as something that stymie or stop you to do something, in this case, entrepreneurship in Russia.

So far we have discussed the multi-facets meanings of <red> in politics and laws section in the academic COCA. We know understand that meanings might change with regard to the

collocates. There are some expressions where <red> literally means a color. However, when they appear in uncommon combination, such as 'red tape', 'red herring', 'red army' the meaning might change.

Besides the collocation, what is more important is readers' knowledge. What makes people can understand this, is the shared background with the text or the text writer. The term 'red tape' can possibly be understood as a tape that is red in color if the reader does not have the background knowledge required. The term 'red herring' can literally be associated to a herring (fish) that is red.

People however are aware of sentence context and investigate this context first when they are unsure of the meaning. When the meaning they project seems to be odd, they will not resort to literal meaning (though not to the correct metaphorical meaning yet). As for this, they can consult existing resources to confirm the meaning.

Figure 5 shows the multi-facets meanings of <red> in COCA. Here, we can understand that meaning is always negotiated not only by linguistic factors, but also by speaker-readers relation in terms of their knowledge.

CONCLUSION

Lemma RED might be used to attribute both abstract and concrete entity. The sample analysis has shown that the polarity of <red> to these entities may be flexible. This follows the finding from Aerim, et al (2010). The color <red> might literally denote a color as in <red sea>.

However, the investigation has shown that the dominating discoveries about the polarity of <red> in COCA are negative. The color <red> is often associated to racial discrimination and fallacy

The number of shift to positive polarity is present but in insignificant quantity, such as 'rescue' in <red cross> and <red crescent>.

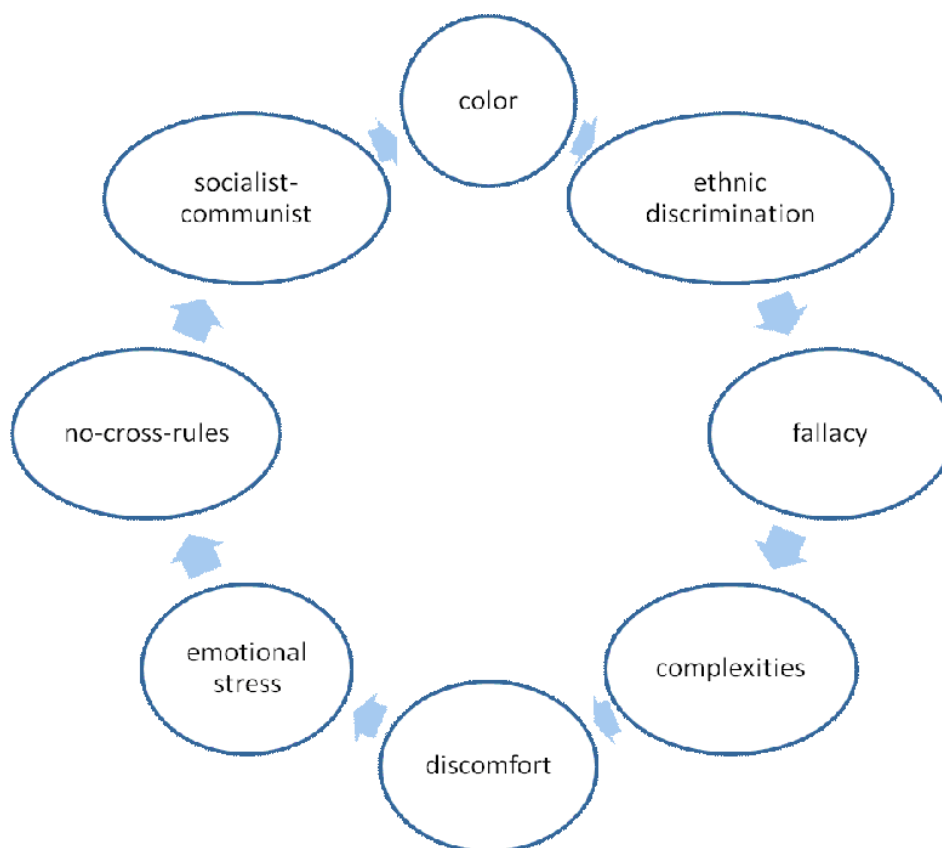


Figure 5. The Meaning of <red> in Politics and Laws Section in Academic COCA

The determination of polarity may vary from person to person. For instance, socialism-communism is not always considered negative, and democracy is not always considered positive. However, the most challenging part of this that there are some compounds that may fall to both literal and metaphorical meaning as they are visually red. To sum up, this paper has shown that the process of understanding the meaning of lemma RED in Social and Politic Science section in COCA strongly requires both linguistic and meta-linguistic awareness. This paper also suggests that other colors are also investigated for further research to widen the coverage of this research topic.

REFERENCES

Aerim, A., Laporte, E., & Jeesun, N. (2010). *Semantic Polarity of Adjectival Predicates in Online Reviews*. Seoul International Conference in Linguistics (pp. 20-27). Seoul: Korea Linguistic Society.

Davies, M. (2008). *American Corpus*. Retrieved August 9, 2014, from The corpus of contemporary American English (COCA): <http://www.americancorpus.org>

De Bortoli, M., & Maroto, J. (2001). Translating colours in web site localisation. *Proceedings of the European Languages and the Implementation of Communication and Information Technologies (Elicit) conference* (pp. 41-77). Paisley: University of Paisley.

Downing, D. (2004). *Apartheid in South Africa*. German: Heinne Library.

Fernandez, E.-M., & Cairns, H.-S. (2011). *Fundamentals in Psycholinguistics*. UK: Blackwell.

Firth, J. (1957). *Papers in linguistics 1934–1951*. London: Oxford University Press.

- Kinoshita, S., Yoshioka, S., & Miyazaki, J. (2008). Physics of structural colors. *Reports on Progress in Physics* 71(7), 076401.
- Mc Enery, T., & Hardie, A. (2012). *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.
- Richard, J. C., Hull, J., & Proctor, S. (2005). *Interchange 3rd Edition*. Cambridge: Cambridge University Press.
- Stea, D., Carson, D.-H., & Wisner, B. (1972). Navajo Color Categories And Color Discrimination-Experiment In Relation Between Language And Perception. *Anthropology UCLA*, 4(2), 27-38. *Anthropology UCLA*, 4(2), 27-38.
- Yanow, D. (2003). *Constructing Race and Ethnicity in USA*. New York: ME Sharpre Inc.
- Yassine, F., & Jeusun, N. (2014). Study on the Sentiment Polarity Types of Collocations for too and very. *Journal of Philology* (1), 23-32.