

УДК 81.42

Search Engine Optimized Texts: Communicative, Pragmatic and Linguistic Aspects

Fyodor A. Nikolaev*

*St. Petersburg State University of Economics
21 Sadovaya Str., Petersburg, 191011, Russia*

Received 18.04.2016, received in revised form 27.08.2016, accepted 19.10.2016

This article describes some linguistic and communicative-pragmatic peculiarities of English SEO-optimized texts and defines the place of this type of texts in the genre structure of advertising discourse.

The article is based on the results of the comparative analysis of the two corpora. The target corpus contained SEO-optimized texts selected from the main pages of the websites of large retail companies; the referential corpus consisted of texts of ads from the web pages of periodicals. The results of the corpus data analysis allow to conclude that as a genre variation of the advertising discourse the SEO-text is characterized by several features unusual for historically formed "prototypical" advertising text: the extremely high number of nouns in relation to the total number of verbs, a low percentage of the lexical variability, a high number of adjectives in the superlative degree implementing evaluative function ("self-glorification"), which is not typical of modern dynamics of advertising texts.

Keywords: advertising discourse, SEO-text, corpus data analysis, technogenic addressee, discourse strategies, discourse tactics.

DOI: 10.17516/1997-1370-2016-9-12-2978-2985.

Research area: philology.

Introduction

Despite the fact that numerous works are devoted to the advertising discourse, all the genres of this type of discourse have not received detailed description in the linguistic literature. The advertising discourse is a multidimensional formation characterized by the significant genre diversity so its investigation and conceptualization are quite challenging.

The main difficulties of the description of the advertising genre are motivated, according to L.

A. Kochetova, by several factors: the dimension of the advertising discourse that uses different messaging channels, which leads to the semiotic complexity of the genre and its permanent transformations, modifications, etc.; the difficulty of genre identification of the advertising discourse, arising from the continuous development and differentiation of the genre system, which results in a diffusion of genres and innovations in the advertising discourse (as cited in: Karasik, 1992, p. 267).

© Siberian Federal University. All rights reserved

* Corresponding author E-mail address: Fyodor-nikolaev@yandex.ru

One of the relatively new and little studied genre of advertising media discourse is the so-called SEO-text. Search engine optimization is a complex of means to influence the position of the links on the website in the result of the search process. SEO-texts are placed mostly on the home page, in the section “About the company”, in the form of information articles in special sections of sites, in the form of news or catalogue texts, price lists.

The main feature of SEO- texts is that unlike the vast majority of advertising texts which are focused on the human perception, they are targeted at one more so-called technogenic addressee, namely, search engines (Google, Yandex, Bing, etc.). The task of SEO-copywriter is to create a text, which will not only be adequately perceived by a human addressee and will have a certain impact on them, but also will bring the website to the highest position of search results when the user enters a certain search query.

High website ranking in search results eventually leads to higher sales of a company, which owns a site. It is achieved by introducing the key words nominating the advertising objects in the texts located on a site (the words that fully or partially match the search request of the site user).

Theoretical framework and methods

The issue of discourse differentiation as well as the problem of discourse type features and functions is currently studied in linguistics. Advertising discourse and its genre system is a subject of interest for many researchers in the époque of globalization and “computerization”. The peculiarities of the way advertising discourse functions as well as its linguistic, stylistic, structural-compositional, typological and genre aspects have been thoroughly studied by L.A. Kochetova (Kochetova, 2013), E.N. Galichkina (Galichkina, 2013), V.I. Karasik

(Karasik, 1992), V.I. Kolokol'tseva (Kolokol'tseva, 2012), I. Forbes (Forbes, 1992), I.A. Guseinova (Guseinova, 2010).

We hypothesized that such type of texts as the SEO-optimized text (SEO – search engine optimization), present mostly on home page of a the web-site of a commercial organization operating in a highly competitive niche, has a number of significant differences from the “prototypical” advertising text.

We have created two corpora of texts with equal volume of words (approx. 50,000) and conducted an analysis of the lexical composition of each corpus. The first corpus consisted of SEO-optimized texts that not only have anthropogenic addressee, but also a technogenic addressee. The latter refers to the search engines, e.g. Google, Yandex, Bing. The texts were collected from the index pages of the web-sites of the USA and UK large trading companies operating in wide-spread industries: cars, real estate, food, furniture, accessories, etc. The referential corpus was created from the texts of ads placed on the web pages of periodicals covering the same industry, and with no additional anthropogenic addressee.

The focus of the analysis of quantitative data is on the lexical variation in each corpus compared to the historically formed prototype of the advertising text.

Discussion

The direction of the genre system developments of modern English advertising discourse lies in the transformation of genres, due to the development of the means of mass media that facilitate the emergence of genres, which have fundamentally heterogeneous semiotic structure enabling broadcasting visual, audio and graphic information. The transformation of genres takes place in computer-mediated discourse, and the specific difference of it from other genres of advertising discourse is determined by a number

of important factors: 1) hypertextual component; 2) the feedback; 3) depth deployability of the text as a result of transition by hypertext links; 4) features of verbal form (Karasik, 1992, p. 291).

Advertising genres of computer-mediated discourse allocated on inductive basis differ from print and TV advertising in metalinearity (term by Galichkina E.N.) (Galichkina, 2013, p. 291), i.e. the presence of hypertext, which allows deploying textual structure making it possible to switch to other texts. In computer-mediated communication the recipient responds to the advertising message through its activation and transition to the target site so the Internet communication is not only interactive and implementing various integration levels of the recipient but also individualized in terms of production and perception of information. The addressee appears to be not a passive recipient of information, but rather actively involved in its selection by activating hyperlinks.

Thus, the description of the advertising discourse in computer-mediated communication is based on the model of interactive mass communication. The main features of this model are: transformation of the subject-object relations of traditional mass communication in the subject-subject relations through the feedback which in the interpretation of some researchers is seen as dialogic one (Kolokol'tseva, 2012, p. 142).

The main function of the studied texts is persuasive one and it has a dual effect, which manifests itself in chronological order. Primarily, the SEO-text must satisfy the recipient with completeness of the information provided therein, as well as be adequately perceived from the point of view of the textuality criteria by both addressees – technogenic and anthropogenic. Such texts should be of interest to its recipient, they should intrigue and encourage them to stay on this site and not to leave for the competitors'

sites. After the first goal of SEO-text is achieved, its second feature of interactivity starts its work being manifested in hyperlinks. Hyperlinks allow the recipient to move to another page of the site, making them face a new set of textual and visual information, bringing the addressee to the target action – making a purchase – and thus implementing the marketing strategy of “customer journey”, i.e. a way, which a customer goes through from the page with SEO-text to the checkout page.

SEO-texts differ from a “prototypical” modern advertising text by a range of parameters. Under the prototype we understand in this case a representative of its category (sample, standard), which fully corresponds to its features. As noted by J. Forbes, “the closer the representation of the genre to its prototypical image, the easier it is for the text to realize its certain genre. The farther it is from the central prototypical image, the less well-defined will be the implementation of the genre” (Forbes, 1992, p. 378). One of the understandings of the prototype characterizes it as prior in time to the rest of the samples and the underlying other variations. Text approach to the genre allows you to organize the discourse, to identify “centers of genre organization”, which meet the main goals of discourse and are characterized by communicative stability and prototypicality (Guseinova, 2010, p. 28). The presence of these genres allows us to outline a discourse as a separate independent communicative entity. Other genres arise in later periods of discourse functioning, are borrowed from other spheres of communication, their presence contributes to the expansion of the discourse limits as well as its transformation into other types. The SEO-optimized text, despite the fact that it has many features of “prototypical” advertising text (convolution, signality, hierarchiness, evaluation, instructivity (Karasik, 1992, pp. 41-42)), cannot be attributed to the prototype of the advertising discourse because of the “technogenic” addressee.

With the aim of identifying the linguistic characteristics of SEO-optimized texts we have created two text corpora. In the first corpus (containing 51,937 words) we included the SEO-optimized texts from the home pages of the sites belonging to various commercial entities in the UK and the USA. The main criterion for the choice of the business area of these organizations was the relevance and prevalence of the goods and services offered by these companies, as the increased demand creates competition among market players, stimulating firms to invest in advertising and brand promotion in the Internet. In the corpus of the SEO-optimized texts we included the texts from the index pages of websites of companies from such areas as: clothing, cars, real estate, skateboards, sports goods, textiles, books, food, furniture, accessories, shoes, music, tickets, etc. The created referential corpus we contains the texts of ads collected from the online editions of various newspapers and magazines. The main difference between the referential advertising texts corpus and the SEO-texts corpus is the location of their source. Both the corpora, however, contain texts describing the same advertised objects, e.g.:

Find your ideal home to buy using our unique search facilities of properties available from 18,243 estate agents (SEO-text with a hyperlink)

Purchasing Real Estate? A Chicago Title Land Trust is the first step in protecting the privacy of your most valuable and visible asset, your real estate (an ad from an online magazine).

The “prototype” of a text ad is a historically formed construct which implies a diachronic approach to the study of advertising texts. Because of the ever growing pace of technology development, entailing significant changes in the textual information representation and general trends to the creolization of transmitted messages, this prototype will undergo some changes that

can be observed in the analysis of the referential corpus data.

The coefficient of the lexical variation, demonstrating the ratio of the number of unique words (number of word types) to the total number of words in the corpus (number of word tokens), amounts in the SEO-texts corpus to only 12.14 with the normal ratios for written text of 50-70 (see :<http://www.slinfo.com/type-token-ratio/>).

Type-Token Ratio = (number of types/ number of tokens) * 100

$$6307/51937 * 100 = 12.14$$

It appears that such a low figure can be expected from such type of texts which is characterized by extremely high frequency (density) of keywords – nouns which primarily name the advertised objects.

The coefficient of lexical variability of the referential corpus is 17.24, which exceeds the ratio of the corpus of SEO-texts by 5.1, but is much lower than the index of Type-Token Ratio, peculiar to conventional printed texts.

$$8624/49997 * 100 = 17.24$$

The above-mentioned finding explains the significant quantitative predominance of nouns over verbs in the corpus, which is not typical of the “prototypical” text ad that has no anthropogenic addressee and is not designed to strengthen the efficiency of text in search engines. (cf.: Table and Table 2), and also indicates the impact of the online location on “unidirectional” text ad, distancing it from the historically formed prototype. Based on these data it is possible to speak about a global transformation of the advertising texts of the online discourse in lexical-grammatical and structural-compositional aspects.

A relatively high proportion of adjectives in relation to the total number of words is peculiar for SEO-texts, as well as for advertising texts in general. As noted by E.M. Wolf, adjectives are one of the main layers of the vocabulary, intended to express evaluation (as cited in: Forbes, 1992, p. 6).

Table 1

Total amount of words	Nouns	Verbs with full meaning			Adjectives		
		4344 (8.4%)			5124 (9.9%)		
		The finite form; Present Simple; infinitive forms	Participle I, Gerund forms	Participle II	Positive degree	Comparative degree	Superlative degree
51,937	13,177	2,611	930	803	4,615	211	298
100%	25.3%	5.1%	1.8%	1.5%	8.9%	0.4%	0.6%

Table 2

Total amount of words	Nouns	Verbs with full meaning			Adjectives		
		5850 (11.7%)			6299 (12.6%)		
		The finite form; Present Simple; infinitive forms	Participle I, Gerund forms	Participle II	Positive degree	Comparative degree	Superlative degree
49,997	10,500	3,248	1,347	1255	5,723	417	159
100%	21%	6.5%	2.7%	2.5%	11.5%	0.8%	0.3%

The evaluation strategy is often characterized as the main strategy of advertising discourse; its purpose is to convince the addressee in the advantages of the offered goods or services.

The number of adjectives in English advertising texts remains relatively constant for almost the whole history of English advertisement. The proportion of adjectives (in percentage terms) relative to total number of words amounts at the end of the 18th century to 6.7%, in the mid-19th century to 7.0% and in the late 20th century – to 6%. Only in the beginning of the 21st century a significant reduction in the number of adjectives in English advertisement (to 4.2%) is observed (Kochetova, 2013, p. 218). In the analyzed corpus we have found 5,124 adjectives, which is 9.9% of all words in the corpus. Thus, the number of evaluative adjectives in the SEO-optimized texts significantly exceeds the average number for the advertising discourse. In the referential corpus, at the same time, this number is 12.6%, which proves the influence of the “online” format on the various sub-genres of the advertising discourse.

Strict prescriptions in respect of self-glorification and excessive self-praise (which

may be perceived as unreliable, difficult to verify, or flattering) are treated as communicative norms of contemporary English advertising discourse. This kind of prescriptions reduces the frequency of adjectives in the superlative degree. This trend is accompanied by a general decline in the number of adjectives of positive evaluation, caused by the increasing significance of the informative tonality and the aspiration for the advertising message objectivity (Kochetova, 2013, p. 222).

The number of adjectives of positive evaluation in the superlative degree in the analyzed corpus equals 0.4% of all the words of the corpus and 5.6% of the total number of adjectives.

The number of adjectives in the superlative degree in the referential corpus equals 0.3% of all words of the corpus and only 2.5% of the total number of adjectives. These figures are significantly lower than the data obtained after the analysis of corpus with SEO-texts.

Thus, the implementation of the evaluation strategy in SEO-optimized texts does not correspond to modern trends. These texts do not tend to increase the reliability. They show

redundancy of the adjectives of general evaluation that implement the tactics of “self-glorification”. The evaluative redundancy of this type of texts occurs due to the basic pragmatic orientation on instant capture of consumers’ attention; considerations of ethical and aesthetic aspect remain beyond the copywriter’s attention

Another important factor that differentiates the studied type of texts from the conditional prototype is an extremely high proportion of nouns in relation to the total number of words in the text (25.3%). The abundance of nominative structures in the text indicates normally the considerable social distance between the addresser and the addressee of the advertising message, which is not peculiar to modern advertising texts. However, in the case of SEO-optimized texts it happens rather due to the presence of the additional technogenic addressee and high density of key words (nouns).

The selection of the most frequently used nouns in the corpus is quite untypical of modern

English advertising texts of any communication channel. See Table 3.

The utilitarian value “quality” is transmitted in the researched type of texts explicitly (by the lexeme “quality”) which is not characteristic of English advertising discourse at the present stage of its functioning. In modern advertising texts the quality of goods and services is mainly described by such adjectives as: *excellent, perfect, ideal, approved, first-class, superb*, etc. Quality appears as an objective property of goods, which is estimated by institutional subjects and is embodied in such lexemes and word combinations as *winning, award, award-winning, the best car of the year* and others. To idea of “quality” is also realized in the advertising texts by mentioning of awards, titles, first place in the ratings, the leading position in the polls, awarded first places, the recognition of the best product of the year in a particular category or class and so on (Kochetova, 2013, pp. 329-330).

Table 3

Frequency, 1 corpus	Frequency, 2 corpus	Word
118	71	quality
115	27	offer
107	45	range
98	9	selection
89	15	order
82	26	collection
74	28	price
64	6	prices
47	18	customers
47	2	items
45	11	customer
41	31	offers
40	32	market
37	30	choice
37	31	needs
36	0	arrangements

The most frequent attributes of “quality” identified in the target corpus are neither typical of the “prototypical” modern advertising text. The lexeme *quality* in many cases is used in attributive phrases with adjectives of general evaluation “high” (24%), “highest” (10), “good” (7), “premium” (6), “top” (5), e.g.:

Our mission is to supply good quality handmade dresses at an affordable price.

We pride ourselves on having a huge stock of high quality guitars and amplifiers from some most famous brands.

Our source specializes in premium quality fur in latest designs and styles.

We design ourselves and source directly from the manufacturer to bring our customers the best collection of top quality furs at wholesale prices.

It is important to note, however, that adjectives of general evaluation in advertising texts, created to be perceived by the human recipient, are typically used in combination with nouns which name the advertising objects, but not with the lexeme *quality*. The location of evaluative adjectives in the predicate position in the rheme part of a phrase is also their current functional peculiarity (Kochetova, 2013, p. 216). The analysis of the corpus showed that this communicative tendency is not peculiar of SEO-optimized texts, as evaluative adjectives function in them mainly in the attributive position, e.g.:

Penny Skateboards make premium plastic skateboards, letting you chase the sunshine and good time.

We make and model the best suits you need.

Another dominant utilitarian value of the advertising discourse is “price”. Being dominant for this type of discourse, the concept “price”, however, is rarely implemented in modern English advertising texts through the use of the key lexeme *price*. As noted by L.A. Kochetova,

“at the end of the 18th century explicit nomination of price was found in 51.8% of advertising texts, in the early 20th century however only in 22.5%. Since the first half of the 20th century the lexeme *price* has not been normally used” (Kochetova, 2013, p. 325). In modern advertisement when specifying the price it is common practice to use the digital designation or paraphrasing with, for example, the adjective *affordable*. This tendency is not typical of SEO-texts. The utilitarian concept “price” is actualized in the majority of the corpus contexts by the lexeme *price*. This lexeme is used 138 times throughout the corpus, while in the referential corpus we have found only 34 uses. The absence of the direct explication of price and quality can serve as an indicator of high status of the manufacturer and product, which is probably functionally irrelevant for SEO-texts.

Conclusion

Based on the statistical data obtained we can conclude that the SEO- texts have significant differences from the historically formed prototype of the advertising text as well as certain similarities.

Due to their focus not only on human but also technogenic addressee SEO-texts are characterized by a low degree of lexical variation, a high proportion of nouns and adjectives in relation to the total number of words. The means of implementation of the evaluative strategy in SEO-optimized texts do not correspond to the dynamics of modern advertising discourse. These texts do not exhibit the tendency of the increasing reliability. The revealed evaluation redundancy of this type of texts can be explained by its basic pragmatic orientation on capturing instant consumers’ attention. The considerations of ethical and aesthetic character remain beyond the attention of the copywriter.

References

Galitchkina, E.N. (2013). *Compyuternaya communicatsiya: lingvisticheskiy status, znakovye sredstva, zhanrovoye prostranstvo* [Computer communication: linguistic status, symbolic means, genre space]. Volgograd, Peremena.

Guseynova, I.A. (2010). *Communicativno-pragmaticheskiye osnovaniya zhanrovoy sistemy v marketingovom discursе: avtoreferat* [Communicative-pragmatic foundations of the genre system in the marketing discourse: extended abstract of Doctor Sci. (Philology) thesis]. Moscow.

Forbes, I. (1992). Review of Linguistic Categorization: Prototypes in Linguistic Theory by John Taylor, *In Journal of Pragmatics*, 18 (4).

Karasik, V.I. (1992). *Yazyk sotsialnogo statusa otsenki* [The language of the social status of assessment]. Moscow, Institut Yazykoznaniiya RAN.

Kolokol'tseva, V.I. (2012). *Internet-communicatsiya kak novaya forma rechy* [Functional Internet communication as a new speech form]. Moscow, Flinta, Nauka.

Kochetova, L.A. (2013). *Angliyskiy reklamnyy diskurs v dinamicheskom aspekte* [English advertising discourse in dynamic aspect]. Volgograd, Volgogradskoye nautchnoye izdatelstvo.

SEO-оптимизированные тексты: коммуникативные, прагматические и языковые аспекты

Ф.А. Николаев

Санкт-Петербургский государственный
экономический университет
Россия, 191023, Санкт-Петербург, ул. Садовая, 21

В статье описываются некоторые языковые и коммуникативно-прагматические особенности англоязычных SEO-оптимизированных текстов; определяется место данного типа текстов в жанровой структуре рекламного дискурса.

Статья основана на результатах сравнительного анализа двух корпусов текстов. Целевой содержал SEO-оптимизированные тексты, отобранные с главных страниц веб-сайтов крупных торговых компаний, референциальный корпус состоял из текстов рекламных объявлений на веб-страницах периодических изданий. Результаты анализа корпусных данных позволяют заключить, что, будучи жанровой разновидностью рекламного интернет-дискурса, SEO-оптимизированный текст характеризуется рядом особенностей, не свойственных исторически сложившемуся «прототипическому» рекламному тексту: чрезвычайно высоким количеством существительных по отношению к общему количеству глаголов; низким уровнем лексической вариативности; высоким количеством прилагательных в превосходной степени, реализующих оценочную тактику «самовосхваления», не свойственную современной динамике рекламных текстов.

Ключевые слова: рекламный дискурс, SEO-текст, корпусный анализ, техногенный адресат, дискурсивные стратегии, дискурсивные тактики.

Научная специальность: 10.00.00 – филологические науки.
