Vision Research 65 (2012) 21–24

Contents lists available at SciVerse ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Do serifs help in comprehension of printed text? An experiment with Cyrillic readers

Leyla Akhmadeeva [a], Ilnar Tukhvatullin [a], Boris Veytsman [b,*]

[a] Bashkir State Medical University, 3 Lenina Str. Ufa 450000, Russia
[b] School of Systems Biology and Computational Materials Science Center, College of Science, George Mason University, MS 6A2, Fairfax, VA 22030, USA

ARTICLE INFO

Article history:
Received 24 February 2012
Received in revised form 29 May 2012
Available online 13 June 2012

Keywords:
Reading
Text comprehension
Serif
Sans serif
Cyrillic

ABSTRACT

The role of serifs for the comprehension of printed text has been controversial in the literature. The analysis was often confounded by the fact that fonts used for comparison have many differences besides being serif or sans-serif. In this study we use fonts from the same foundry and meta family to assess the differences in reading comprehension for Cyrillic readers ($n = 238$). The results show no difference in the speed of reading and the comprehension between the serif and sans-serif texts. This conclusion is consistent with the ecological hypothesis recently formulated for font sizes.

1. Introduction

Several centuries of the Gutenberg era created a certain amount of lore about the role of different typographic devices for reading and comprehension. The rules about font shape, text size, spacing between the lines, paragraph indentations and many other features were carefully studied, transmitted from generation to generation and codified (see, e.g. the classical work by Bringhurst (2004)). Sometimes these rules were exalted to the status of moral imperatives, as reflected by the titles of books by Tschichold (1991) (*Essays on the Morality of Good Design*), and Spiekermann and Ginger (2002) (*Stop Stealing Sheep and Find Out How Type Works*). Perhaps one of the most often disputed typographic devices is the serif—a small detail at the end of the strokes forming a letter. Rejected by Bauhaus and Constructivist typographers in the beginning of 20th century (Bartram, 2004), they returned as an important part of the printing in the later period; it is very interesting to follow the evolution of one of the most influential typographers of the last century from "The New Typography" (Tschichold, 1998) to "The Form of the Book" (Tschichold, 1998). It should be said that the early discussions mixed the art (how to make the text aesthetically pleasing?) and the science (how to make the text easy to read and comprehend?) of book making. We cannot comment here on the art part of the equation, and will deal only with the science.

One of the most interesting scientific insights in the old arguments of typographers is the so called ecological hypothesis (Legge & Bigelow, 2011). Based on the analysis of many studies of the influence of type size on reading, the authors concluded that the print sizes actually used over the centuries in the book making are in the "comfort zone" for a normal vision reader, and the variations in the size are of low importance. It is tempting to extend this hypothesis on the other typographic devices, including serifs and assume that the latter really do not matter for reading.

There was a number of publications about the influence of serifs on legibility and readability of printed text (Arditi & Cho, 2000, 2005; Bernard et al., 2003). Interesting enough, the results were not conclusive: the authors found that the differences between serif and sans-serif fonts were rather small. This might be seen as an argument in favor of the generalization of ecological hypothesis. However, one problem with these studies was that the fonts used (the ubiquitous Times New Roman and Arial) have many differences besides one being serif, and another being sans-serif: they were designed with different goals in mind. A comparison of their performance is in fact a comparison of two complex entities with many different features. Is there a reason to think that the presence or absence of serifs is the main driver of difference in performances? It might be much more convincing to compare two fonts belonging to the same group, designed by the same artists with the same goals in mind, where the difference is just the serifs.

An exception is the work by Morris et al. (2002), where the authors used a set of specially designed fonts from the Lucida family which differed only by the presence and absence of serifs. The authors found that sans serif font was read 20% faster at very

* Corresponding author.
E-mail addresses: la@ufaneuro.org (L. Akhmadeeva), tukhvatullin.ilnar@yandex.ru (I. Tukhvatullin), bveytsma@gmu.edu (B. Veytsman).

0042-6989/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.visres.2012.05.013

low sizes, possibly close to the critical print size (Legge & Bigelow, 2011), but at higher sizes the difference was negligible. The comparison was done using rapid serial visual presentation method. However, the question of the influence of serifs on the reading speed in natural conditions and the text comprehension is still open.

Recently ParaType company released under a free license high quality fonts PT Serif and PT Sans (Farář, 2011). They have already been used for typesetting such high profile books as FAO Statistical Yearbook (FAO, 2012). Serif and sans serif fonts from this meta family are designed according to the same ideas. They have many common features (Fig. 1), and provide a good approximation to a controlled experiment on the importance of serifs. Note, however, that the strokes of the sans serif font have almost constant thickness, while the serif font provides a significant contrast between the thickness of vertical and horizontal strokes. On the other hand, it is customary in font design to have sans serif fonts with lower "typographic contrast" (Jaspert, Berry, & Johnson, 2009).

One of the goals of our work was to use these fonts to compare the performance of serif and sans serif fonts for a Russian Cyrillic reader.

Another goal of the experiment was to inquire whether general laws like the ecological hypothesis, developed on the Latin typography material, are valid for the Cyrillic script. The Cyrillic typography has its own traditions, sometimes close to those of the Western typography, sometimes different. In particular, the use of serif and sans serif fonts in the Russian typography is very similar to the one in the Western tradition (see, for example, the discussion of the parallel processes in Russian and European typography in the first third of 20th century by Bartram (2004)). There are many books, including textbooks, printed every year in Russian and other Cyrillic languages. Thus an experiment with Russian readers is interesting both as a fundamental question, and as practical matter.

To conclude this section we discuss the rationale for the methods chosen. In many experiments on legibility and readability, a special device is used that demonstrates letters and words to the subjects with a controlled time of exposure (see the discussion of the protocol by Legge, Pelli, et al. (1985) and Legge, Rubin, et al. (1985)). These experiments gave a valuable insight into the speed of recognition of letter and word forms, which is useful for many practical applications, for instance traffic signals (see, e.g. Carlson & Holick, 2005; Garvey et al., 2004; Ullman et al., 2005). However, the reading of long texts by an educated person is a complex process that involves not just letters and words, but rather the comprehension of text as a whole (Legge et al., 2002). Therefore we chose to ask the subjects to read the texts with their "normal" speed and measure the results. A similar approach was used by dos Santos Lonsdale, Dyson, and Reynolds (2006), where the appropriateness of typography devices for quiz and exam texts was measured by simulating the exams.

a b c d e f g h i j k l m n o p q r s t
u v w x y z
a b c d e f g h i j k l m n o p q r s t
u v w x y z
а б в г д е ё ж з и й к л м н о п р
с т у ф х ц ч ш щ ъ ы ь э ю я
а б в г д е ё ж з и й к л м н о п р
с т у ф х ц ч ш щ ъ ы ь э ю я

Fig. 1. Lowercase Latin and Russian Cyrillic letters in PT serif and PT sans.

2. Materials and methods

Undergraduate students of Bashkir State Medical University (Ufa, Russia, 4th, 5th and 6th year of study, 188 females and 50 males) volunteered to participate in the experiment and gave their informed consent (see the disclosures). All were fluent Russian speakers and had normal or corrected vision. They were randomly separated into two groups and given the same text about the history of neurology in Russia (see Supplementary Materials). The first group ($n = 108$) got the text in PT Serif with the effective x-size 12 pt (this was a 12 pt font scaled 0.95 as recommended by Farář (2011), so the actual x-size was 11.4 pt), the second one ($n = 130$) got the text in PT Sans with the same x-size. The participants were asked to read the text at their most comfortable speed, and mark the point achieved after 1 min (as measured by the experimenter). Then they were given a questionnaire with 10 questions about the text (multiple choice, four options per question, see Supplementary Materials). The questionnaires were typeset with the same font and size as the texts, i.e. either PT Serif or PT Sans.

The number of words read per minute and the number of questions correctly answered were tabulated together with the vital statistics about the students (gender, age, year of study) as well as self-reported high school and university average grades (GPA).

3. Results and discussion

The average numbers of words read per minute and of correct answers with the standard deviations are shown in Table 1. As seen from this table, the difference between serif and sans serif variants is small. The results of the standard statistical tests (Venables & Ripley, 2010) are shown in Table 2. We see that the difference between serif and sans serif fonts is indeed not significant at 95% level. This is also illustrated by Figs. 2 and 3.

As seen from the figures, the variance in the speed of reading and the number of correct answers is large: it seems the reading skills of the students quite differ. This might be partially explained by the diverse social backgrounds of the students. However, the relatively large number of participants in the study allows one to measure small effects even in this diverse population. Indeed, the variation of the mean in the group with n participants scales as $1/n$, so with large n we can detect small difference between the mean values. A more refined analysis includes power computations (Cohen, 1988). They show (see the supplement) that in our tests we can reliably detect the effects larger than 0.366σ, where σ is the in-group standard deviation.

In our experiment the in-group standard deviation for the speed of reading was 49 words per minute, and the in-group standard deviation for the number of correct answers was 2.2 answers (see the Supplement). Thus we could reliably detect the difference in the speed of reading larger than 18 words per minute and the difference in the number of correct answers larger than 0.81. The measured difference was much smaller: 1.74 words per minute and 0.38 correct answers.

To further test the sensitivity of our methods we calculated the dependence of the results on other parameters: gender, age, GPA in school and university, and others (see Supplementary Materials). The results significantly depended on the university GPA and the

Table 1
Speed of reading and comprehension levels.

Font	Words per minute		Correct answers	
	Mean	Std. sev.	Mean	Std. dev.
Sans serif	207.9	49.7	5.35	2.20
Serif	206.2	48.5	5.07	2.20