Article

The Influence of Typography on Algorithms That Predict the Speed and Comfort of Reading

Arnold Wilkins [1]*, Katie Smith [2] and Olivier Penacchio [2,]

[1] Department of Psychology, University of Essex, Wivenhoe Park, Essex, Colchester CO4 3SQ, UK; arnold@essex.ac.uk (A.W.); kasmit@essex.ac.uk (K.S.)

[2] School of Psychology and Neuroscience, University of St. Andrews, St Mary's Quad, South Street St Andrews, Fife KY16 9JP, U.K. op5@st-andrews.ac.uk

* Correspondence: arnold@essex.ac.uk;

Abstract: The speed with which text can be read is determined in part by the spatial regularity and similarity of vertical letter strokes as assessed by the height of the first peak in the horizontal autocorrelation of the text. The height of this peak was determined for two passages in 20 fonts. The peak was unaffected by the size of the text or its content but was influenced by the font design. Sans serif fonts usually had a lower peak than serif fonts because the presence of serifs resulted in a more even spacing of letter strokes. There were small effects of justification and font- dependent effects of font expansion and compression.   The visual comfort of images can be estimated from the extent to which the Fourier amplitude spectrum conforms to 1/f. Students were asked to adjust iBooks to obtain their preferred settings of font and layout. The preference was predicted by the extent to which the Fourier amplitude spectrum approximated 1/f, which in turn was jointly affected by the design of the font, its weight and the ratio of x-height to line separation. The above algorithms can be usefully applied to any orthography to estimate likely speed and comfort of reading.

Keywords: font; spatial periodicity; discomfort; reading speed; autocorrelation; Fourier amplitude spectrum

1. Introduction

There have been many studies of the effects of typographic variables on reading [e.g. 1] including the effects of letter size (*x-height*), line spacing (*leading*) [2] and typeface (e.g. *serif* vs *sans serif* [3]. Text is a complex stimulus in which the effects of typographic variables interact together and with other variables such as familiarity [4] to determine the speed and comfort of reading. In this paper it is shown that despite this complexity, simple algorithms are able to predict both reading speed and choice of font. They do so by registering the extent to which the spatial periodicity of text interferes with vision. One algorithm measures the periodicity from letter strokes, whereas the other measures aspects of the periodicity from lines of text.

2. Horizontal spatial periodicity

Weaver [5] reported the case of a lady who experienced seizures when reading. She found that the seizures occurred only when she was reading material printed in Times or Palatino, and not when reading the same material printed in Arial or Verdana. Her observations were confirmed electrophysiologically.   Times and Palatino are fonts with serifs, whereas Arial and Verdana are sans serif fonts. Serifs have little effect on reading [3].   However, the fonts differ not only with respect to serifs but in the periodicity of the letter strokes, as illustrated in Figure 1.   In Arial and Verdana the space between the strokes within a letter is greater than that between letters, whereas in Times and Palatino the letter strokes are evenly spaced, giving the typeface a spatial regularity referred to by typographers as *rhythm*.
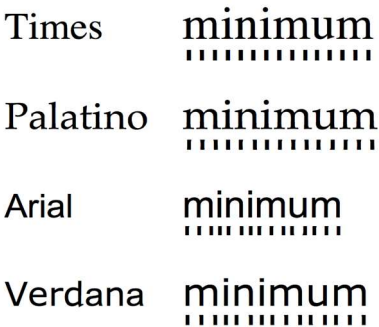
Times   minimum

Palatino   minimum

Arial   minimum

Verdana   minimum

Figure 1. The horizontal periodicity in the letter strokes for the word minimum printed in Times, Palatino, Arial and Verdana. The font size has been chosen to equate the x-height. The pattern beneath each word is that formed from the central body of the letters.

The rhythm of a typeface can be assessed using horizontal autocorrelation, the correlation of an image with a second version of itself, displaced horizontally by a small amount (lag). The mathematical technique can be understood by imagining text printed on the transparency of an overhead projector.   When two identical transparencies are superimposed and in register the light transmitted through them is at its maximum (the correlation is 1.0). When the top transparency is moved horizontally across the lower one the lag increases and the transmitted light decreases (the correlation decreases), reaching a minimum when a majority of letter strokes on the top transparency are in the spaces between letters on the lower transparency. As the top transparency is moved further, and the lag increases further, the transmitted light increases (and the correlation increases), reaching a maximum when the majority of letter strokes on one transparency lie on the neighbouring letter strokes on the other.

The height of the first peak in the horizontal autocorrelation is therefore a measure of the spatial periodicity of the typeface, dependent on both the shape and spacing of neighbouring letter strokes. Wilkins et al. (2007) showed that the height of the first peak in the horizontal autocorrelation of a word determined how striped in appearance the word was judged to be. More importantly, they showed that the height of the peak predicted the speed with which the word could be read aloud. Words with high peak were read about 10% more slowly than those with low. The peak also predicted the speed of silent visual search through a paragraph of text. Reducing the autocorrelation by compressing words near the middle and expanding them at the edges (leaving their length unchanged) increased reading speed, even though readers preferred to read an undistorted version of the text [6]. One of the reasons for the effects of spatial periodicity on reading speed concerned the ways in which the eyes move across text when reading.

During a rapid eye movement (saccade), one eye generally leads and the other follows, resulting in a misalignment of the eyes that requires correction when the eyes come to rest [7]. Jainta, Jaschinski, & Wilkins [8] measured saccades and vergence movements when their participants read German sentences. When the words had a high first peak in the horizontal autocorrelation the eyes rested on each word for longer during the vergence eye movements that corrected the misalignment (vergence error). The realignment took longer with a spatially periodic word because the alignment was then more precise.

Little is known about the effect of typographic variables on the size of the horizontal autocorrelation, even in two studies in which the spatial periodicity has been shown to affect reading. The first of the present studies therefore measured the effect of typographic variables on the autocorrelation.

3. Study 1

Study 1 explored the way in which the first peak in the horizontal autocorrelation of text varies with font and layout. Twenty common fonts were selected, shown in the first column of Table 1.

84 **Table 1.** 20 fonts showing mean of the first peak in the horizontal autocorrelation (AC peak) and the
85 residuals after fitting a 1/f cone to the Fourier amplitude spectrum.

86

	Font	AC peak	Residuals x10^12
sans serif	Andale mono	0.196	3.92
	Arial	0.395	5.10
	Ayuthaya	0.283	4.62
	Booster Next Medium	0.310	3.96
	Calibri	0.410	3.38
	Gill Sans regular	0.403	4.74
	Helvetica	0.403	3.74
	Microsoft Sans Serif	0.408	3.76
	Open Sans	0.359	4.36
	Raleway medium	0.366	4.06
	Verdana	0.387	4.20
serif	Athelas regular	0.423	3.40
	Baskerville	0.434	4.40
	Bookman Old Style	0.427	4.91
	Cambria	0.432	5.28
	Century	0.413	5.30
	Charter Roman	0.426	3.70
	Garamond	0.398	5.16
	Lucida bright	0.433	4.05
	Times	0.445	5.03

87 *3.1. Comparison of 20 fonts*

88 3.1.1. Procedure

89 A passage from "Small House at Allington" by Anthony Trollop and a passage from
90 "Middlemarch" by George Eliot were generated using Microsoft Word for Mac 2011 version 14.6.8.
91 They were printed to the 15-inch Retina screen of an Apple Macbook Pro running OSX 10.11.6. The
92 passage was prepared in each of the 20 fonts listed in Table 1, nominal 10 point in size, with default
93 letter spacing, and with a ragged right margin and an interlinear spacing of 1.15 points. The page was
94 20.5 cm wide and 28.5 cm high and was saved with a resolution of 5.67 pixels per mm. The horizontal
95 autocorrelation of a central section of the page 512x512 pixels in size was obtained using Matlab with
96 iteration of the *corr2d* function.

Figure 2. The first peak in the horizontal autocorrelation for 20 fonts, measured with two different samples of text.

3.1.2. Results

Figure 2 shows a plot of the first peak in the horizontal autocorrelation for the 20 fonts.   As can be seen, the peak is very similar for the two passages of text, but differs for each font.   The absolute value of the difference in the autocorrelation for the two text samples averaged only 0.0092 (SD 0.0064), demonstrating that the method shows reliable differences attributable to the font design, and not the content of the text.   The mean value for each font is shown in the second column of Table 1.

As can be seen from Table 1, the font with lowest autocorrelation was a "typewriter" font in which each letter had the same width. Of the proportional fonts, the two with lowest autocorrelations were Open Sans and Raleway medium. Arial and Verdana had autocorrelations less than 0.4. Although these are all sans serif fonts, the remaining cluster of fonts with autocorrelations greater than 0.4 included Gill sans as well as many other fonts with serifs.   The difference in the first peak in the horizontal autocorrelation was significantly greater for the fonts with serifs than those without t(18)=3.45, p =.003.

3.1.3. Discussion

There were differences in the horizontal autocorrelation of a page of text that were due almost entirely to the font (typeface) and not the content of the text.   Evidently a single page of text provides for sufficient averaging to remove most of the variability from content.

3.2. Equating empty space

Although all the fonts had the same nominal point size, they differed in width and in the height of the central body of the letters (x-height). The x-height of Times nominal 14pt was similar to that of Google Open Sans nominal 12pt, for example. Although the line spacing was nominally 1.15pt, it differed from one font to the next, partly to maintain a similar ratio of x-height to line space.   Other things being equal, the larger the area of empty white page, the greater the autocorrelation. The variation between fonts therefore depended in part on this variation and not on the spatial periodicity between letters. Further, the smaller the size of the letters, the fewer the pixels used to create the letter form, and the greater the degradation in shape.

3.2.1. Procedure

127   In order to compare the autocorrelation of different fonts independently of these factors, the text
128   from Trollope was created in 28 point, and the size normalised to give an x-height of 27 pixels for all
129   fonts, using the imresize function in Matlab (bicubic interpolation), cropping the width of the
130   fragment at 1168 pixels.

131   3.2.2. Results

132   Figure 3 shows the relationship between the autocorrelation obtained using this method and
133   that originally obtained for the 10pt font. The correlation is 0.91.

134   3.2.3. Discussion

135   Evidently the first peak in the horizontal autocorrelation is largely independent of any image
136   changes associated with scale and spacing, at least over the range of scale examined. Google Open
137   Sans is surrounded by a circle and Times by a square in Figure 3. Note that Google Open Sans is
138   consistently a font with low autocorrelation, and Times one with high.

139
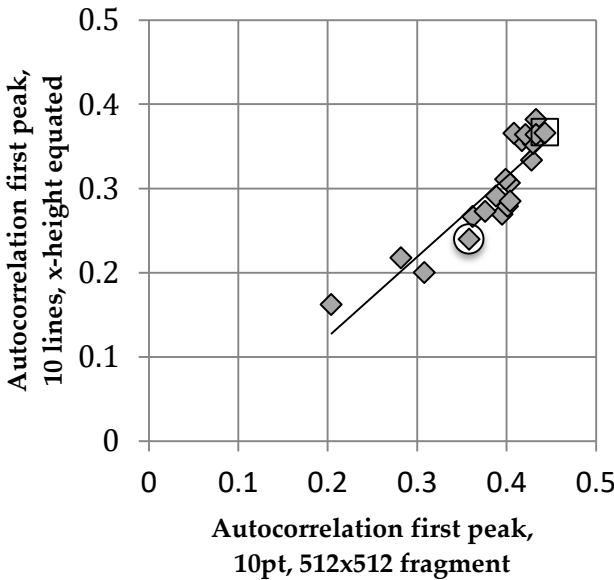
140   **Figure 3.** Correlation between text from Trollope 10pt with 1.15pt interline spacing and the same
141   passage normalised for x-height, for the 20 fonts listed in Table 1. The points for Open Sans and Times
142   are surrounded by a circle and a square respectively.

143   *3.3. Horizontal spacing*

144   3.3.1. Font compression and expansion

145   The spacing of letters may be expected to affect the lag at which the autocorrelation reaches its
146   peak, but the effect of expansion on the height of the first peak in the autocorrelation depends on the
147   intra- and inter-letter space and is not easily predicted. Wilkins et al. [6] showed that for Times the
148   autocorrelation is at its peak when the letter spacing is default: the peak decreases with both
149   expansion and compression. There was comparatively little change for Verdana.

150   3.3.1. Justification

151   The word processor used in the previous studies achieved right justification by varying the space
152   between but not within words, and the average spacing between words was increased when the text
153   was justified. The passage by Trollop was printed using Microsoft Word in one of four fonts (Times,
154   Century, Verdana and Open Sans). Each was printed in one of three nominal sizes (10pt, 12pt, 14pt)
155   with both right justification and ragged right margins, a total of 24 samples.

156   The autocorrelation of the 24 samples was obtained using the algorithm that analyzed 10 lines
157   without inter-line spacing. The effect of font size was inconsistent, but there was a small but
158   consistent effect of justification: it increased the first peak in the autocorrelation for all four fonts,
159   although only by an average of 3.3%.   The average spacing between words increased when the text
160   was justified, and the extra blank white page increased the autocorrelation.

4. Discussion of Study 1

162   The first peak in the horizontal autocorrelation of words has been shown to influence the speed
163   with which they can be read both in a list [6] and as connected meaningful sentences [8]. The peak is
164   high when vertical strokes of letters resemble each other and are evenly spaced, and words are then
165   read more slowly [6], [8]. Serif fonts generally have a higher peak, and that other aspects of text
166   spacing and layout have a comparatively small effect. Although the peak has been shown to influence
167   reading speed, familiarity with fonts design is, however, also likely to influence speed of reading [4],
168   and this may have a larger effect.

5. Vertical spatial periodicity

170   Certain periodic patterns, particularly stripes, can be uncomfortable to look at. They can induce
171   perceptual illusions of movement, shape and colour. The patterns responsible for discomfort and
172   illusions have characteristics similar to those of patterns that induce electro-encephalographic
173   abnormalities, or seizures, in patients with photosensitive epilepsy, suggesting a neurological rather
174   than optical basis for the visual effects [9]. There are large individual differences in susceptibility to
175   the illusions and associated discomfort and these differences are related to a person's history of
176   headaches. There are several convergent lines of evidence that the visual cortex is hyperexcitable in
177   migraine headache. It is therefore parsimonious to explain the individual differences in susceptibility
178   to the visual effects of stripes in terms of individual differences in cortical excitability [10]. The
179   successive lines of text resemble a pattern with spatial frequency, contrast, and duty cycle sufficient
180   to induce discomfort and even seizures [11], [12]. This can best be understood using Fourier analysis.
181   Fourier analysis is a mathematical technique that decomposes images into components. The
182   components are typically (but not necessarily) sine-wave patterns with a variety of spatial
183   frequencies, orientations and phases.   Consider the small sample of the image in Figure 4. An
184   enlarged version of the sample is shown in the left insert and the profile of the luminance over space
185   is shown in the top graph. This waveform can be created by adding together the waveforms
186   numbered 1-5 immediately below. The peak-trough amplitude of these waves is directly proportional
187   to their wavelength. The spatial frequency, f, of the waves is the reciprocal of their wavelength, so
188   the amplitude is proportional to $1/f$.   This means that when the logarithm of the amplitude is plotted
189   against the logarithm of the spatial frequency a straight line with a slope of -1 results. It turns out that
190   most images from nature have a Fourier amplitude spectrum with a slope close to -1 [13].
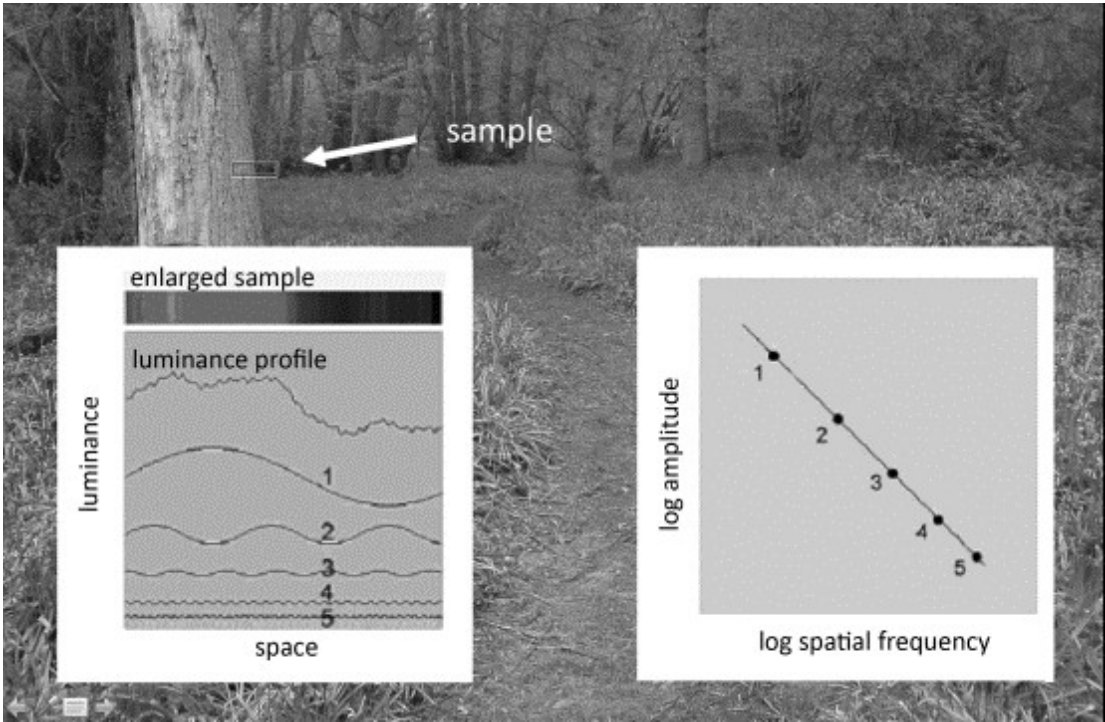
191

Figure 4. Illustration of the component waves in Fourier analysis. The variation in luminance over space (luminance profile) of the sample shown at the top and enlarged in the first row of the left hand inset can be thought as composed of the addition of the waves shown below, and numbered 1-5. The amplitude decreases with their spatial frequency as shown in the right-hand inset.

The visual system has adapted to process natural images. Field [14] argued that the bandwidths of channels tuned to spatial frequency are optimized for a 1/f amplitude spectrum. Their bandwidth remains constant when expressed on an octave scale, so that a similar amount of information is then carried by each channel. Atick & Redlich[15] have argued that the shape of the contrast sensitivity function enables images with a 1/f spectrum to be coded efficiently. The contrast sensitivity is low (the channel has low gain) for low spatial frequencies that have a high amplitude, and this conserves metabolic energy. In computational models of the visual system, striped patterns, which are rare in nature and do not conform to a 1/f structure, result in an excess of 'neural activity' and a non-sparse distribution of 'neural' firing [16]. Wilkins [18] reviewed neuroimaging studies and concluded that images that are uncomfortable to observe are generally associated with an elevated haemodynamic response, consistent with the computation models and suggesting that the discomfort is a homeostatic mechanism that avoids excessive cerebral metabolism.

It therefore possible to argue that images are processed inefficiently by the brain if they are unnatural and do not have a 1/f amplitude spectrum. According to this hypothesis images are uncomfortable to look when they are processed inefficiently and require excessive metabolism. Observers have been asked to rate the discomfort from images of modern art and of filtered visual noise or shapes[18], [19]. For all categories of image, the discomfort was minimal for those images with a 1/f Fourier amplitude spectrum. An algorithm with no free parameters can account for more than 25% of the variance in observers' ratings of discomfort [20]: a cone with a slope of 1/f was fitted to the two-dimensional Fourier amplitude spectrum, and the residuals were weighted by a contrast sensitivity function gleaned from the literature. Note that patterns of stripes have a Fourier amplitude spectrum that is not well fit by a cone, consisting of a few component spatial frequencies. They are therefore patterns that produce some of the largest residual scores, and they are uncomfortable to look at.

The algorithm has not hitherto been applied to text, so in the next study we assessed its effectiveness in this context.

6. Study 2.

Study 2 applied the algorithm to text and showed that it predicted the choice people make when observing text in electronic books. These books, presented on an electronic screen, permit readers to manipulate the font and size and to choose from a limited range of colours.

6.1. Predicting readers' choice of text characteristics

6.1.1. Participants

15 participants, 7 female and 8 male, aged 18-63 took part.

6.1.2. Procedure

Each participant was shown publications from iBooks. The books were presented in double-page format on an Apple MacBook Pro with a 15-inch Retina screen. The participants were asked to "make the text as easy to read as possible" using the iBook tools. With the tools, one of eight fonts could be selected (Original, Athelaf, Charter, Georgia, Iowan, Palatino, Servarek and Times New Roman) in a variety of sizes. The text could be black on a white background or on a sepia background, or the text could be white on a black background. When the participants had finished their settings the image of the screen was saved. Each participant adjusted two texts, one from "The Other Woman" by Laura Wilson and one from "The Angel" by Katerina Diamond. The texts were originally presented in their default settings: Original and Times New Roman with an x-height of 2.3mm and 2.2mm respectively, and line spacing of 6.8mm and 4.9mm. The work was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Informed consent was obtained.

6.1.3. Results

The final settings of texts had x-heights that ranged from 1.9mm to 2.9mm, with line spacing that ranged from 4.4mm to 8.4mm. There was no consistent change in the x-height or line spacing relative to the original text: 52% and 63% of settings showed an increase. The ratio of x-height to line spacing ranged from 31% to 46% and also did not change consistently, decreasing in 56% of cases. Although there was no consistent change in typographic parameters, the adjusted text nevertheless showed a consistent change in the residuals calculated by the algorithm: 29/30 settings (14/15 participants) showed a decrease as compared to the original, indicating the parameters of the text were altered in combination so as to be closer to those of natural images. Seven of the 15 participants changed the background colour from white to sepia and one used a black background for one of the two texts. The residuals were lower for the sepia background because of the reduction in contrast. Nevertheless, for all the eight participants who chose a white background the residuals were lower for the adjusted text than for the original in every case, averaged for the two samples.

6.1.4. Discussion

In all cases but one participants adjusted the iBook text in a direction predicted by the final version of the algorithm from [20]. Although the x-height and line separation did not change consistently, the text was adjusted in a manner to which the algorithm was sensitive. The adjustments were such as to make the text more like images from nature, a change consistent with a reduction in discomfort. Given that the algorithm was successful in predicting choice of typographic variables, it was important to ascertain more systematically what variables affected the output of the algorithm.

6.2. Fonts and the algorithm output

6.2.1. Procedure

The passages from Trollope and Eliot were set in the 20 different fonts given in Table 1, 10pt in size, with a line spacing of 1.15 pt. They were analysed using the final algorithm of [20].

266   6.2.2. Results

267   The residuals obtained for the two text samples are expressed as a scatterplot in Figure 5. The
268   correlation explains 87% of the variance, and indicates that the residuals are affected primarily by the
269   font, with little contribution from the textual content. The absolute value of the difference between
270   the text samples averaged 5.0% (SD 2.7%). Table 1 shows the average of the residuals for the two
271   samples. Among the conventional fonts, Open Sans had one of the lowest residuals.

272
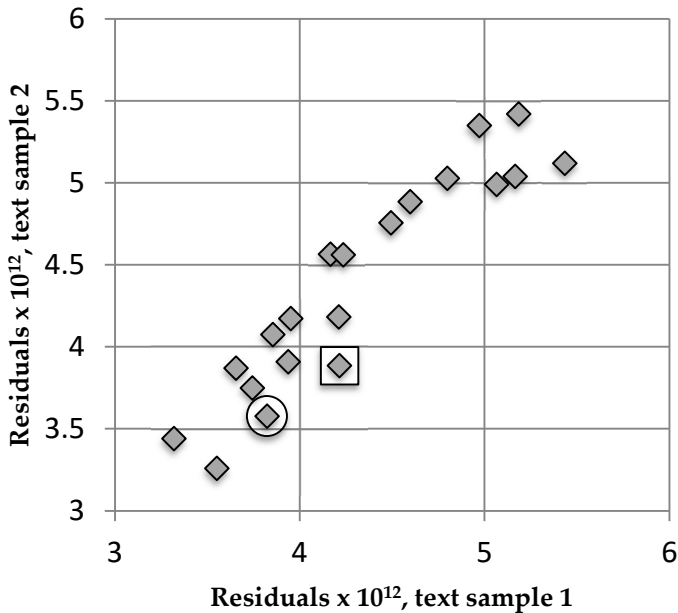
273   **Figure 5.** Residuals for the two samples of texts from Eliot and Trollope for each of the 20 fonts listed
274   in Table 1. The lower the residuals, the more similar the text to images from nature. The points for
275   Open Sans and Times are surrounded by a circle and a square respectively.

276   6.2.3. Discussion

277   The font had a clear effect on the size of the residuals of the images of a page of text. The nominal
278   size of the font was held constant, although there was variation in x-height, which may have
279   contributed to the differences in residuals.

280   *6.3. Font weight*

281   There are large differences between fonts in the residuals obtained when a 1/f cone was fitted to
282   the two-dimensional Fourier amplitude spectrum of the page using the algorithm [20]. If the page
283   resembles a pattern of horizontal lines, giving large residual scores as a result, then the contrast of
284   that "grating" will be influenced by the *weight* of the font, that is, the average thickness of the letter
285   strokes.   In order to estimate the weight of the 20 fonts, the sentence "The quick brown fox jumped
286   over the lazy dog" was printed in each of the fonts. The central section of the line from the baseline
287   to the top of those letters without ascenders comprised a rectangular section of the page. The weight
288   was estimated as the number of black pixels divided by the total pixels in the rectangle. The weight
289   for sentences printed in 11, 12 and 14 point were calculated and the correlations between the estimates
290   obtained was high (>0.967).   The estimates for 11 point varied from 15% to 24% as shown in Figure
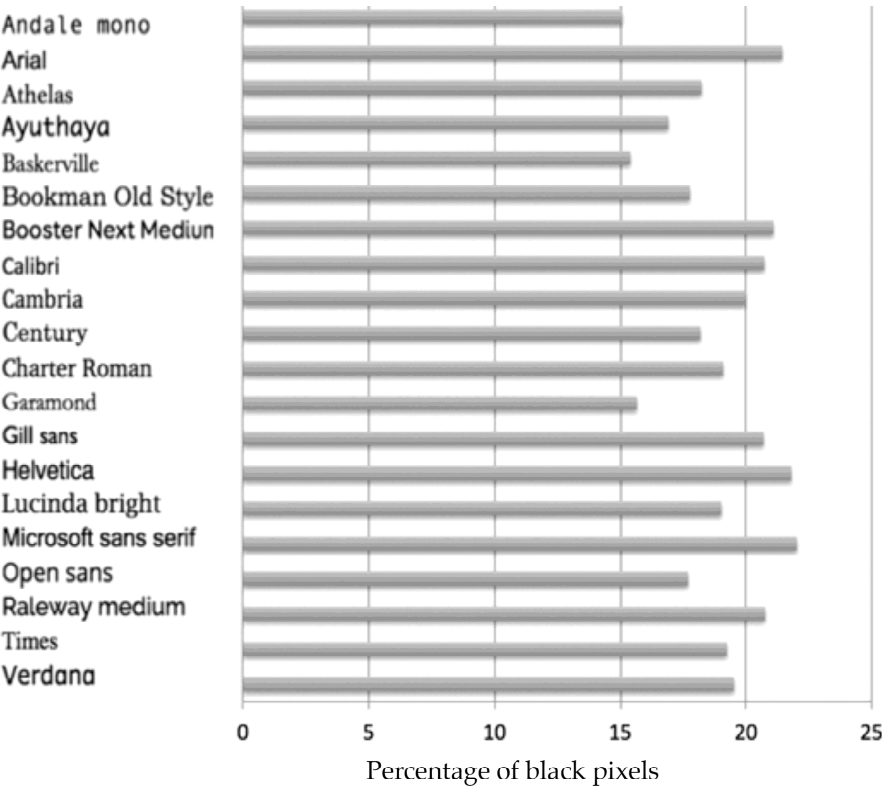291   6.

Figure 6. Weights of 20 fonts as estimated from the percentage of black pixels in a line of text.

There was a moderate but significant correlation ($r=0.624$, $p=.003$) between the algorithm output and the weight of the font as defined above, shown in Figure 7. The effect of font weight on the algorithm can therefore be understood in terms of the contrast of the striped lines or "grating" formed by the lines of text.

The next study investigated the effects of the parameters of text layout.

6.3.1. Procedure

Seventy two samples from Trollope were set in Century, Times, Open Sans and Verdana, 10, 12, and 14 point in size, with 1, 1.15 and 1.5 point inter-line spacing, with and without justification. They were analyzed using the algorithm from [20]. The samples of text with expanded letter spacing were also analysed.

6.3.2. Results

There was no relationship between the residuals obtained and the number of words per page ($r=0.03$). The residuals were unaffected by the right margin, ragged or justified. The x-height also had little effect. There was a weak tendency for texts with large inter-line spacing to have lower residuals, but by far the strongest relationship between the output of the algorithm and the typographic variables concerned the ratio of x-height to line spacing. This relationship accounted for 79% of the variance, see Figure 8. Pages in which the text was widely spaced relative to the height of the letters were evidently more similar to images from the natural world, and potentially more comfortable.
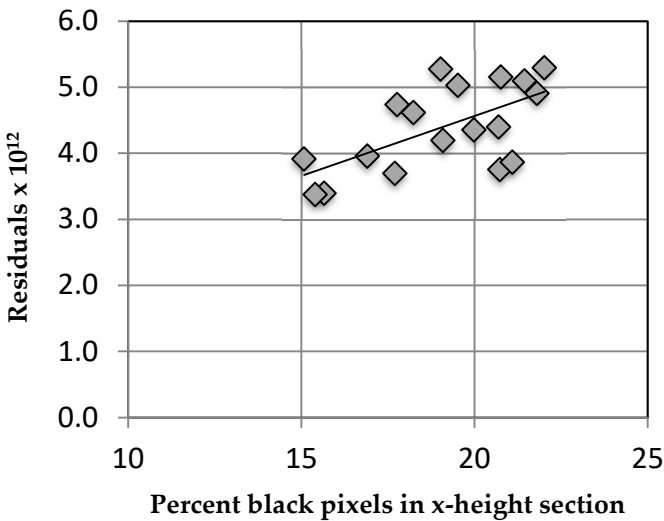
Figure 7. The correlation between font weight and the output of the algorithm (residuals) [20].

The ratio of x-height to line spacing can be likened to the duty cycle of the "grating" formed by horizontal lines of text. The lower the ratio, the greater the departure from a 50% duty cycle and the lower the Fourier power of the "grating" formed by the lines of text. In current typographic practice the ratio of x-height to line spacing is typically 35%-45%. The present findings suggest that this might usefully be decreased by increasing line spacing.
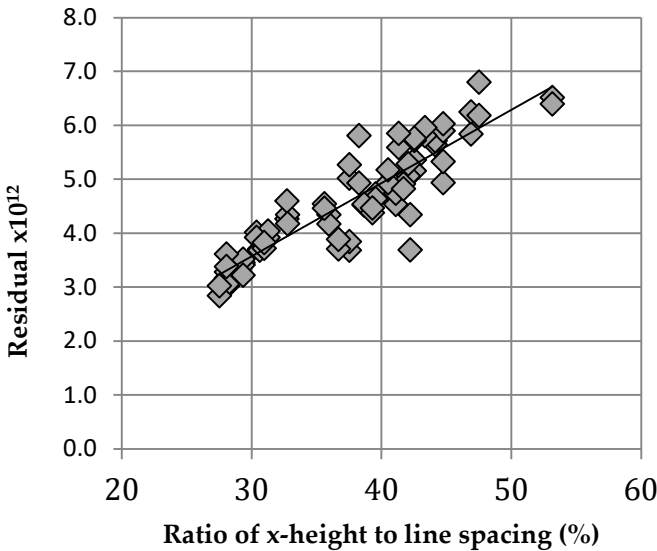
Figure 8. Relationship between the output of the algorithm (residuals) [20] and the ratio of x-height to line spacing for 72 texts, with various fonts, font size and spacing, with and without justification. The lower the residuals, the more similar the text to images from nature..

For both texts, from Trollope and Eliot, and for the two fonts, Times and Open Sans, the effect of expanding the font was to decrease the residuals progressively, and to do so by a similar amount. One possible reason was that the expansion reduced the average letter density and thus the contrast of the "grating" formed by the successive lines of text; an effect similar to that of the weight of the font.

The differences in residuals between fonts in Table 1 are considerable. The font with one of the lowest residuals, Open Sans, had a ratio of x-height to line spacing of 43%, which was large, relative

330 to other fonts. Note that iBooks currently permit the selection of font but *not* the ratio of x-height to
331 line spacing, except in so far as line spacing co-varies with font.   The differences in the iBook
332 selections were not therefore attributable to line spacing but to other aspects of the typefaces that the
333 algorithm was also successful at capturing.

7. General discussion

335 Two mathematical properties of text have been studied: (1) the spatial periodicity of the vertical
336 strokes of letters, which affects reading speed, and (2) the degree of departure from the Fourier
337 amplitude spectrum typical of natural images, which determines visual comfort of images. Overall,
338 the correlation between these two measures for the 20 fonts in Table 1 is only 0.082, suggesting that
339 the measures are independently useful. Open Sans is one of the few fonts that performs well on both
340 measures.
341 There is convergent evidence that in migraine, and perhaps in headache more generally, the
342 visual cortex is hyperexcitable [21], and that as a result, individuals can be particularly susceptible to
343 discomfort from patterns of stripes [9], including stripes in text [11]. In this paper we have shown
344 how the stripes in text are affected by the selection of font and by text layout. With the advent of
345 electronic text it is possible without additional cost to choose fonts with little *rhythm* and to increase
346 the spacing between lines of text. Currently the selection of fonts and spacing is insufficient to
347 accommodate individual preference.
348 The height of the first peak in the horizontal autocorrelation varies considerably for different
349 fonts, but is high for Sassoon Primary, a font used in schools. Sassoon Primary is read more slowly
350 than Verdana by school children [22], as predicted by the higher autocorrelation of Sassoon Primary,
351 notwithstanding the children's greater familiarity with this font. The question arises as to whether
352 Open Sans might provide a better font for children than those currently in use. There are three aspects
353 of the font that would then need to be changed, however. The q would need a tail to distinguish it
354 from p. Similarly, the b and d would need to be distinguished, perhaps by removing the base of the
355 stem of the b or adding a small tail to the stem of the d. Finally the upper case i and lower case L
356 would need to be distinguished, for example by adding a small tail at the base of the lower case L.
357 These changes could be very slight and would in any event have little or no effect on the
358 autocorrelation or residuals. The alterations would be simple to make and would be admissible
359 within the terms of the open license.

8. Conclusion

361 The general principles and methods described in this paper can be applied to all languages and
362 all orthographies. For example, spatial periodicity has been shown to affect the reading of Chinese
363 logograms [23]. The principles can be used to guide the design of any typeface and any written
364 material. They suggest, surprisingly, that font design has the potential to affect reading fluency and
365 comfort appreciably.

References

373 1.   Tinker MA. *Lagibility of Print.* Iowa State University Press, Ames Iowa; 1963.

374 2.   Williamson, H. *Methods of Book Design.* 2013. 121-123. p. Available from:
375   http://ir.obihiro.ac.jp/dspace/handle/10322/3933

376 3.   Arditi A, Cho J. Serifs and font legibility. *Vision Res.* 2005;

377   4.   Beier S, Larson K. How does typeface familiarity affect reading performance and reader preference? *Inf*
378   *Des J.* 2013;

379   5.   Weaver DF. Font specific reading-induced seizures. *Clin Neurol Neurosurg.* 2014;125:210–1.

380   6.   Wilkins AJ, Smith J, Willison CK, Beare T, Boyd A, Hardy G, et al. Stripes within words affect reading.
381   *Perception.* 2007;36(12).

382   7.   Liversedge SP, White SJ, Findlay JM, Rayner K. Binocular coordination of eye movements during
383   reading. *Vision Res.* 2006;46(15):2363–74.

384   8.   Jainta S, Jaschinski W, Wilkins AJ. Periodic letter strokes within a word affect fixation disparity during
385   reading. *J Vis.* 2010;10(13):2.

386   9.   Wilkins A, Nimmo-smith I, Tait A, Mcmanus C, Sala SD, Tilley A, et al. A neurological basis for visual
387   discomfort. Brain. 1984;107(4).

388   10.   Braithwaite JJ, Broglia E, Bagshaw AP, Wilkins AJ. Evidence for elevated cortical hyperexcitability and its
389   association with out-of-body experiences in the non-clinical population: New findings from a pattern-
390   glare task. *Cortex.* 2013;49(3).

391   11.   Wilkins AJ, Nimmo-Smith MI. The clarity and comfort of printed text. *Ergonomics.* 1987;

392   12.   Wilkins AJ, Nimmo-Smith I. On the reduction of eye-strain when reading. *Ophthalmic Physiol Opt.*
393   1984;4(1).

394   13.   Tolhurst DJ, Tadmor Y, Chao T. Amplitude spectra of natural images. *Ophthalmic Physiol Opt.* 1992;

395   14.   Field DJ. What Is the Goal of Sensory Coding? *Neural Comput* [Internet]. 1994;6(4):559–601. Available
396   from: http://www.mitpressjournals.org/doi/10.1162/neco.1994.6.4.559

397   15.   Atick JJ, Redlich AN. What Does the Retina Know about Natural Scenes? *Neural Comput.* 1992;4(2):196–
398   210.

399   16.   Hibbard PB, O'Hare L. Uncomfortable images produce non-sparse responses in a model of primary
400   visual cortex. *R Soc Open Sci.* 2015;2.

401   17.   Wilkins AJ. A physiological basis for visual discomfort: Application in lighting design. *Light Res Technol.*
402   2016;48(1).

403   18.   Juricevic I, Land L, Wilkins A, Webster MA. Visual discomfort and natural image statistics. *Perception.*
404   2010;39(7).

405   19.   Fernandez D, Wilkins AJ. Uncomfortable images in art and nature. *Perception.* 2008;37(7).

406   20.   Penacchio O, Wilkins AJ. Visual discomfort and the spatial distribution of Fourier energy. *Vision Res.*
407   2015;108.

408   21.   Aurora SK, Wilkinson F. The brain is hyperexcitable in migraine. *Cephalalgia.* 2007.

409   22.   Wilkins A, Cleave R, Grayson N, Wilson L. Typography for children may be inappropriately designed. *J*
410   *Res Read.* 2009;32(4).

411   23.   Cheong Y., So RH, Wilkins AJ. How vertical stripes affect recognition of Chinese characters. In: *Vision*
412   *Sciences Society.* 2016.