Perception, 2007, volume 36, pages 1788 – 1803

doi:10.1068/p5651

Stripes within words affect reading

Arnold J Wilkins, Jennifer Smith, Clare K Willison, Tom Beare, Alexandra Boyd, Gemma Hardy, Louise Mell, Charlotte Peach, Samantha Harper
Department of Psychology, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK;
e-mail: arnold@essex.ac.uk
Received 19 July 2006, in revised form 6 May 2007

Abstract. In a series of eight studies it is shown that the first peak in the horizontal autocorrelation of the image of a word (which captures the similarity in shape between the neighbouring strokes of letters) determines (i) the appearance of the words as striped; (ii) the speed with which the words are read, both aloud and silently; and (iii) the speed with which a paragraph of text can be searched. By subtly distorting the horizontal dimension of text, and thereby reducing the first peak in the horizontal autocorrelation, it is shown that the speed of word recognition can be increased. The increase in speed is greater in poor readers.

1 Introduction

Legros and Grant (1916) considered the similarity of shape between letters as providing an index of legibility. They measured the overlap when similar letter pairs were superimposed, and referred to the 'unique black area'. Typefaces with greater black area were considered to be less legible. These initial attempts to study word shape were never complemented by behavioural measures. Psychologists who have measured the effects of word shape on reading have considered the shape only in terms of the outline of a word. The extent to which the outline affects reading has been the subject of longstanding controversy (eg Beech and Mayall 2005; Cattell 1886; Haber et al 1983; Paap et al 1984). One of the enduring difficulties in this research has been to establish a simple way of classifying word or letter shape, other than by partial word fragmentation or gross manipulation by mixing case. As far as we are aware there has been no attempt to measure the significance for reading of the internal similarity of shape within a word.

Reading speed is affected by the semantic and syntactic content of the material, and these sources of variance mask any but the most substantial contributions from typography, at least in the short-term. Wilkins and coworkers (1996) required children to read randomly ordered common words and showed that the speed with which passages of text were read was highly reliable, and reflected visual aspects of the text such as its colour, a finding replicated in adults by Evans and Joseph (2002). Hughes and Wilkins (2000, 2002) showed that this reading task was sensitive to simple typographic parameters such as letter size.

In this paper we consider the shape of a word in terms of the component strokes of the letters within the word itself. We follow the initiative of Legros and Grant, but with respect to words as whole rather than their component letters. We measure the similarity of word shape in terms of the horizontal autocorrelation. To appreciate what the horizontal autocorrelation is, it may be helpful to imagine a word printed in opaque ink on an overhead transparency. Supposing two identical such transparencies are placed on an overhead projector. When the transparencies are in register, a maximum amount of light will be transmitted through the combined transparencies. If the top transparency is moved relative to the lower one, the overall level of light transmitted is affected—see figure 1a. If the top transparency is moved horizontally across

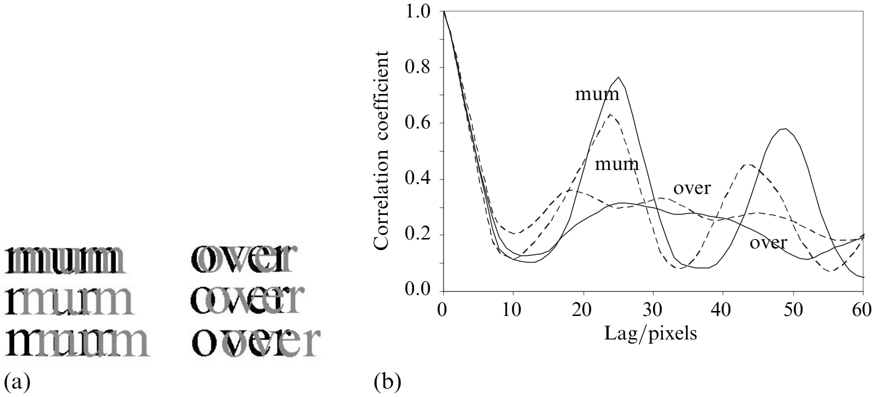

Figure 1. (a) Illustration of the basis of the horizontal autocorrelation of the image of a word. Two versions of a word are superimposed and the upper image (grey) shifted horizontally with respect to the lower. The horizontal autocorrelation is based upon the extent of overlap of the two images which varies as the images are shifted one with respect to the other. (b) Horizontal autocorrelation functions for the words 'mum' and 'over' set in Times New Roman 12 point (solid curves) and Arial 10 point (broken curves), digitised at 600 dots per inch. The x-axis shows the lag, ie shift of one image with respect to the other, on a scale in which 1 mm = 23.6 pixels.

the bottom transparency, the amount of light transmitted is initially reduced because the letter strokes in one version of the word block the spaces in the other version. As the displacement continues, however, and neighbouring letter strokes come into register, so the amount of light transmitted increases. As the top transparency is displaced still further, the amount of light transmitted once again decreases and then increases again. The light transmitted varies with horizontal position according to a function with peaks and troughs. This function is, in effect, the horizontal autocorrelation. In the following studies the horizontal autocorrelation is defined as the covariance of the density of corresponding pixels in the portion of the image common to the original and the shifted images, divided by the variance of the density of pixels in the original image. For an image M pixels high and N wide, the horizontal autocorrelation, r, is a function of the shift or lag, h:

$$r_h = \left[\sum_{i=1}^{M}\sum_{j=1}^{N-h}(d_{i,j+h} - \bar{d})(d_{i,j} - \bar{d})\right] \Bigg/ \left[\sum_{i=1}^{M}\sum_{j=1}^{N}(d_{i,j} - \bar{d})^2\right],$$

where each pixel has a density d. Note that the autocorrelation is unaffected by overall scale, but is affected by the size of the window relative to the size of the word, to the extent that this affects the variance of the pixel densities.

Figure 1b shows the horizontal autocorrelation functions of two words, 'mum' and 'over' set in 12 point Times New Roman and 10 point Arial. [The words were originated in Microsoft Word, printed to an Adobe pdf file and imported with the use of anti-aliasing into Adobe Photoshop with 600 dots per inch (dpi) as an eight-bit 700 pixel wide by 80 pixel high grey-level image; the image was saved in tiff format and analysed using Matlab.] As can be seen from figure 1, the word 'mum' has several peaks indicating the spatial periodicity. The word 'over' has little such periodicity.

The first peak in the horizontal autocorrelation results from the similarity between the neighbouring strokes of letters, whether these strokes are components of the same letter or neighbouring letters. The second peak derives mainly from the similarities between neighbouring letters, in particular those letters that have more than one stroke.

The autocorrelation provides a simple way of measuring the extent to which the image of a word approximates a pattern of stripes. The word 'mum' for example consists of a repetitive pattern of vertical line elements. The word 'over' does not.

Patterns of stripes are known to induce perceptual distortions (Wilkins 1995). Individuals with migraine are particularly susceptible to these distortions (Huang et al 2003; Wilkins 1995). A person's susceptibility predicts the effects of coloured overlays in reducing distortions (Wilkins 2003) and improving reading speed (Evans et al 1994; Hollis and Allen 2006). The studies described here followed the first author's observation of a pupil who was unable to read text printed in font smaller than 14 point. She reported that the word 'mum' appeared to move, and she was unable to read it.

The first study shows that the horizontal autocorrelation predicts the appearance of words as striped. If people are asked to rate words on the basis of how striped they appear, then it is the first peak in the horizontal autocorrelation that best correlates with their judgments.

2 Experiment 1

2.1 *Participants*

One male and nine female students from the University of Essex aged 18–25 years (mean 21 years) took part. In this experiment, as in all others to be reported, participants wore glasses if they normally did so for reading. None was visually impaired.

2.2 *Materials*

We selected 40 words from the 100 most frequent words in the British National Corpus (Leech et al 2001), on the basis that they contained four or five letters.

Each word was originated on a Dell personal computer in lower case Times New Roman, 12 point (x-height 1.95 mm, measured from the printed copy with a magnified scale), with Microsoft Word under XP Professional 2002, from which was printed an Adobe pdf file, converted with 600 dpi resolution to a grey-level image with anti-aliasing under Adobe Photoshop, and saved as 800×137 pixel tiff tiles. We then analysed the tiff tiles using Matlab. (Anti-aliasing had only a small effect on the autocorrelation. When the autocorrelation was measured on binary images the results were closely similar.) With the exception of the word 'very' the autocorrelation function showed a clear initial peak with a lag of about 27 pixels (1.14 mm), and a second peak, usually lower in height, with a lag of about 50 pixels (2.12 mm). Subsequent peaks were not analysed.

2.3 *Procedure*

Each participant was asked to rate the 'stripiness' of each of the words on a scale from 0 (not at all stripy) to 10 (very stripy).

2.4 *Results*

The horizontal autocorrelation function was characterised for each word in terms of the value of the minimum at first dip or trough, the maximum at first peak, the minimum at second trough, and the maximum at second peak. The values of these parameters for each word were then correlated with the mean rating of 'stripiness'. Table 1 shows the Pearson product moment correlations between each of the four parameters and the mean ratings of 'stripiness'. The correlation between the first peak and the rating of 'stripiness' was significant ($p < 0.0001$), and significantly higher than the remaining correlations ($p < 0.05$).

Table 1. Pearson product moment correlations between the mean rating of the 'stripiness' of a word and the value of troughs and peaks in the horizontal autocorrelation.

	First trough	First peak	Second trough	Second peak
Correlation	0.201	0.688	0.359	0.286

2.5 *Discussion*

The first peak in the autocorrelation reflects the similarity between neighbouring strokes within a word (whether the strokes are from the same letter or neighbouring letters). It is this peak that is best correlated with the perception of the 'stripiness' of the word.

Striped images are known to result in perceptual distortion (Wilkins 1995), and individuals who report perceptual distortion of text usually read more slowly (Wilkins 2003). The next experiment shows that the first peak in the horizontal autocorrelation of a word predicts the speed with which the word is read.

3 Experiment 2

3.1 *Participants*

Seven male and twenty-five female students and staff from the University of Essex aged 17 – 48 years (mean 20 years) took part.

3.2 *Procedure*

We selected 22 monosyllabic words, four or five letters in length from the 100 most frequent words in the British National Corpus (Leech et al 2001). The words were originated on a Dell personal computer in Microsoft Word under Microsoft XP Professional 2002 and set in lower case Times New Roman 12 point and Arial 11.5 point so as to have similar *x*-height (2.0 mm) in a 900×82 binary image. The words were divided about the median with respect to the height of the first peak of the horizontal autocorrelation of the word. The first peak was here (and subsequently) defined as the maximum value in the range bounded by the first and second minima in the function obtained by averaging the autocorrelation across all words. Occasionally ($< 5\%$ of occasions) there were minor peaks in the individual autocorrelation functions that this measure avoided.

The words with high and low first peak were presented in separate paragraphs or columnar lists. The paragraphs consisted of eighteen lines single spaced (about 5 mm baseline to baseline), with the same words on every line in a different random order. The words in the lists were similarly constrained as to order, but were presented in a column spaced vertically 25 mm apart. A practice test was also prepared consisting of two-, three-, or four-letter monosyllabic words, selected without respect to their autocorrelation, none of which appeared in either of the test passages.

Participants were given 1 min in which to read aloud as many words as they could. Sixteen participants read the words presented in columns, and a different sixteen participants the words presented in paragraphs. Eight in each group read words set in Times New Roman and eight those set in Arial. The test began with the practice passage and was then followed by the high- or low-autocorrelation words in a random order, counterbalanced across subjects. The participants' scores comprised the total number of words correctly read in the appropriate sequence: words that were transposed, omitted, or incorrectly read were not counted.

3.3 *Results*

Table 2 shows the means (and standard deviations) for the number of words read per minute in the columns and paragraphs. Overall, words with high first peak in their autocorrelation were read 7% more slowly. An analysis of variance was undertaken with layout (column/paragraph) and typeface as between-subjects factors and autocorrelation first peak as a within-subjects factor. The analysis showed no significant between-subjects effects, but a large and significant effect of autocorrelation ($F_{1,28} = 176.4$, $p < 0.001$).

Words with a high first peak in the autocorrelation were rated as having a striped appearance (experiment 1) and were slower to read (experiment 2). These effects justified a more detailed investigation of the relationships between autocorrelation, typeface design, and word length.

Table 2. Means (and standard deviations) for the number of words read per minute in experiment 2, shown separately for words presented in columns and paragraphs, and for words with high autocorrelation at first peak and those with low.

Layout	Font	High first peak	Low first peak
Column	Times New Roman	126 (9)	134 (11)
	Arial	134 (8)	141 (8)
Paragraph	Times New Roman	128 (13)	145 (11)
	Arial	129 (17)	135 (19)

3.4 *Comparison of different typefaces*

1000 words were selected at random from the English Lexicon Project (Balota et al 2007) website at http://elexicon.wustl.edu. The words varied in length from three to seventeen letters and in frequency from 94 000 per 1 014 000 to < 1. The 1000 words were originated as before in Microsoft Word and set in each of the following fonts: Arial 10 point, Times New Roman 11 point, Palatino Linotype 11 point, and Lucida Sans Unicode 10 point. The use of different point sizes equated the x-height of each word (to within 7%), although the length varied with font by about 15%. From Microsoft Word an Adobe pdf file was created and digitised as a 900×103 pixel tiff file at 600 dpi, as before. The files were then analysed with Matlab.

The x-height of the words was similar, but the typefaces differed in the width of their characters. Figure 2 shows the autocorrelation functions for each typeface averaged over the 1000 randomly selected English words. The peaks and troughs of the functions occur at different lags owing to the different character widths for each typeface (although the typefaces had similar x-height, the letter width varied by about 5 pixels).

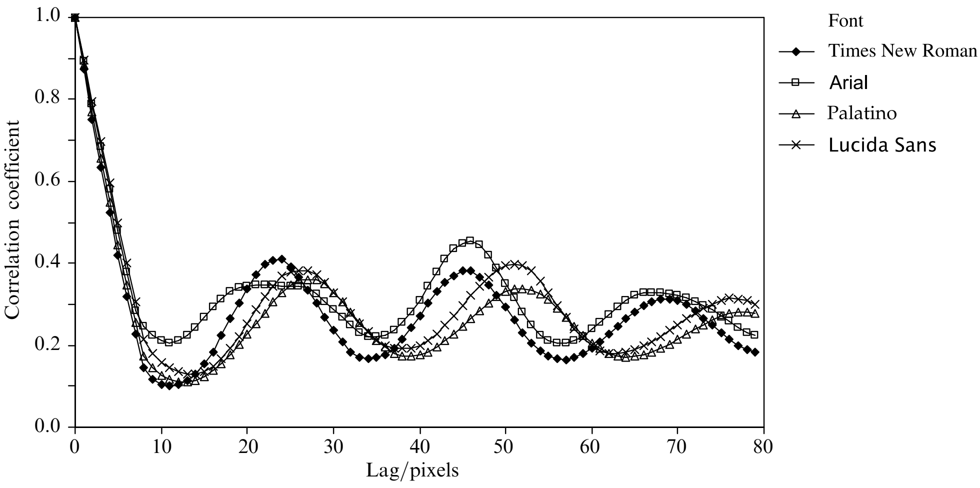

Figure 2. Horizontal autocorrelation function averaged for the same 1000 words set in various typefaces, shown in the legend. The words were sampled at random from those in the English Lexicon Project (see text).

The height of the peaks and troughs indicate how similar various spatial characteristics of the letter shapes are. Table 3 shows the intercorrelations between the height of the first trough, the first peak, the second trough, and the second peak, averaged for the 1000 words, the standard deviation being that for the four typefaces.

Note, first, that the height of the first trough and first peak are negatively correlated. A striped pattern would have an autocorrelation with a negative first trough

Table 3. Intercorrelations between the heights of the first trough, first peak, second trough, and second peak of the horizontal autocorrelation function of 1000 words. The means (and standard deviations) of the Pearson product moment correlation coefficients for four fonts, Times New Roman, Arial, Palatino, and Lucida Sans are shown

	1st peak	2nd trough	2nd peak
1st trough	−0.34 (0.16)	0.66 (0.03)	−0.08 (0.20)
1st peak		−0.41 (0.23)	0.27 (0.36)
2nd trough			−0.39 (0.20)

(where the dark stripes are matched with the neighbouring light stripes) and a positive first peak (where the light stripes are matched with the neighbouring light stripes). The negative correlation between the height of the first trough and that of the first peak shows that, to a lesser extent, this pattern is being shown by the words with striped appearance and high first peak. Note, second, that the heights of the first trough and second trough are strongly correlated ($r = 0.66$). These troughs are influenced partly by the spacing between letter strokes, which is itself partly a function of intercharacter separation. (When the separation between letters was increased by selecting a Times New Roman font 'expanded' by 1 point in Microsoft Word, the first trough decreased in size and showed a larger proportional change than the remaining peaks and troughs.)

In figure 3 each point represents a word of a given length, and its position is determined by the value of the first peak in the autocorrelation function for that word. The function was obtained for 1000 words set in Times New Roman 12 point, but similar scatterplots obtain with other typefaces and sizes. Notice that, for long words, the range of values decreases with the length of the word. When the word has twelve or more letters the constraint on the value of the first peak is evident. This is attributable to the constraint provided by the larger samples of letter strokes. Note that with shorter, more frequent, words there is little relationship between word length and the first peak of the horizontal autocorrelation function.

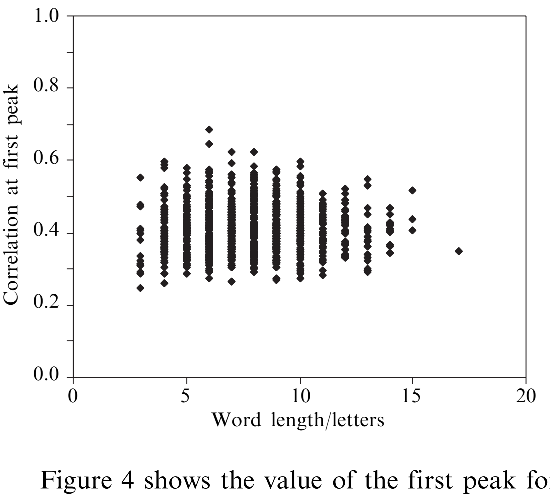

Figure 3. First peak in the horizontal autocorrelation of 1000 words set in Times New Roman 12 point shown as a function of word length.

Figure 4 shows the value of the first peak for 1000 words set in Times New Roman against the value of the first peak for the same words set in Palatino. The correlation is high (Pearson product moment correlation = 0.95), but note that the regression lies below the diagonal, indicating that, in general, Palatino has a lower value for the first peak.

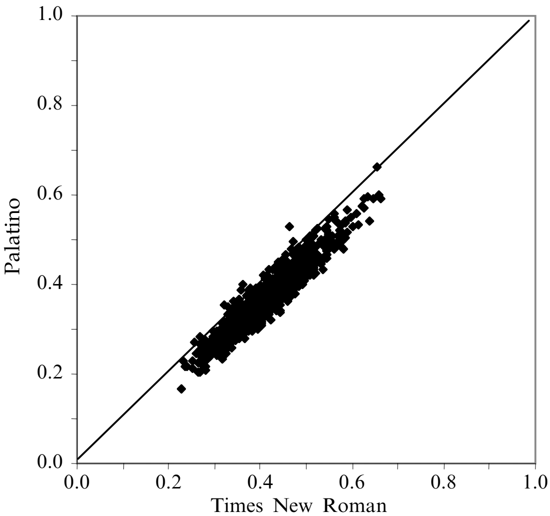

Figure 4. The values of the first peak in the horizontal autocorrelations of 1000 words set in Times New Roman and in Palatino.

The correlation between the value of the first peak in Times New Roman and in a sans-serif typeface, Arial, is shown in figure 5. The correlation is weaker, and note that the scatter indicates a narrower range of values for Arial. Arial also has a lower average first peak than Times New Roman. Despite the differences in mean values and ranges, the correlation is still moderate ($r = 0.68$), indicating that $0.68^2 = 46\%$ of the variance in the first peak is attributable not to the typeface but to the component words.

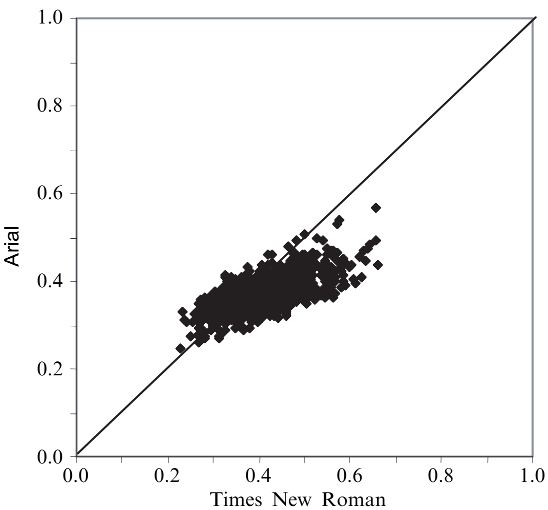

Figure 5. The values of the first peak in the horizontal autocorrelations of 1000 words set in Times New Roman and in Arial.

Figure 6 shows the correlation between the first peak for Times New Roman and that for Lucida Sans. Note that, although Lucida was designed to be clear to read, the values of the autocorrelation are similar to those for Times, notwithstanding the presence of serifs in the latter font.

The first peak in the autocorrelation is affected by the spacing of the letters. This is shown in figure 7 for the fonts Times New Roman (10 point) and Verdana (8 point), respectively, both with similar x-height. The same 1000 words were generated in Microsoft Word, as before, and the character spacing altered with the format–font–character spacing command. Note that the default spacing (0 point) gives the highest first peak in the autocorrelation for Times New Roman but not for Verdana.

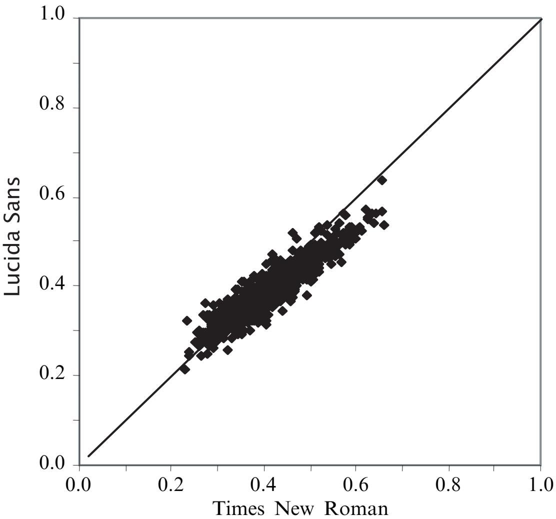

Figure 6. The values of the first peak in the horizontal autocorrelations of 1000 words set in Times New Roman and in Lucida Sans.

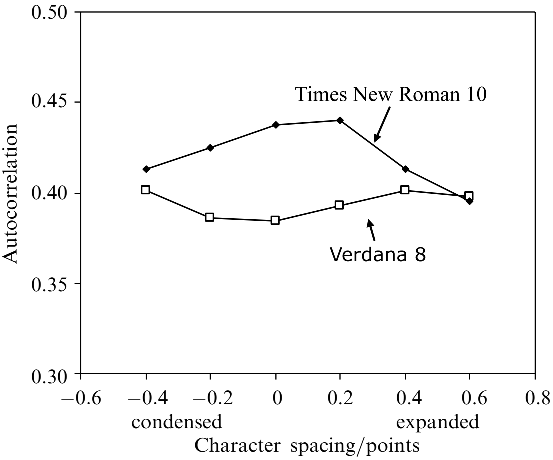

Figure 7. The average first peak in the autocorrelation functions for 1000 words set in Times New Roman 10 point and in Verdana 8 point, both with similar x-height. The average peak is shown as a function of intercharacter spacing, the default spacing being 0, and negative values signifying condensed spacing.

With the exception of the two points with greatest expansion in Verdana, the differences between all neighbouring points in figure 7 were very highly significant across words.

A comparison of the first peak in the autocorrelation functions of the same 1000 words was undertaken for the following fonts: Times New Roman, Arial, Sassoon Primary, Tahoma, and Verdana (set in 12 point with default spacing and with the same overall image size, 680×118 pixels). Verdana had the lowest first peak.

The height of the first peak in the autocorrelation (but not its horizontal position) should be largely independent of scale. This was confirmed by a comparison of the 1000 words set in Times New Roman and Verdana fonts, 8, 9, 10, 11, and 12 point, digitised at 600 dpi in 700×130 pixel images. The average for Times New Roman was 0.452 with a standard deviation across font size of 0.014. The average for Verdana was 0.356, standard deviation 0.017.

3.5 *Discussion*

The height of the first peak in the horizontal autocorrelation function was associated with the rated 'stripiness' of a word (experiment 1) and with the time it took to read the word (experiment 2). The height of the peak varied with typeface and with words independently of type size, but there were similarities between words across typefaces which accounted for about 50% of the variance in the first peak of the horizontal autocorrelation.

Arditi and Cho (2005) quote McLean as asserting that sans-serif type is intrinsically less legible than seriffed type because some of the letters are more like each other than letters that have serifs. The data in figures 4 and 5 show that, in the case of the serif typeface, Times New Roman, and the sans-serif typefaces, Arial and Lucida Sans, the assertion is borne out. In Times New Roman the letters are more 'like each other', at least in respect of their higher horizontal autocorrelation, and in experiment 2 reading was, if anything, faster with the sans-serif font, Arial. The present data are consistent with those of Arditi and Cho (2005) who showed that when serifs are added or removed from a typeface there is little change in reading speed.

Thus far, the effect of the autocorrelation on reading speed has been demonstrated for only a small sample of words that included letters with ascenders or descenders. We therefore repeated experiment 2 using monosyllabic lower-case words without ascenders and descenders.

4 Experiment 3

4.1 *Participants*

One male and seven female students and associates of the University of Essex aged 18 – 32 years (mean 21 years) took part.

4.2 *Materials*

39 four- or five-letter monosyllabic words without ascenders or descenders were selected from the 200 most frequent words in the English language (Leech et al 2001). The words were set in Times New Roman 12 point (x-height 2.0 mm), ranked with respect to the value of the first peak in the horizontal autocorrelation function, and divided into two groups about the median.

Paragraphs were prepared consisting of 15 lines, each with the same 14 words per line in a different random order, one having words with high first peak and one words with low. A practice test was given as in experiment 2. The two experimental conditions, high and low autocorrelation, were presented in a random order, counterbalanced across participants. Again participants read the passages aloud as rapidly and as accurately as they could. The reading speed was determined as before.

4.3 *Results*

On average, 124 (SD = 18) words per minute were read from the passage comprising words with high autocorrelation at first peak, and 139 (SD = 20) per minute from the passage of words with low first peak, a difference of 12%. A paired-sample t-test showed that the difference in reading speed was stastically significant ($t_7 = 5.5$, $p = 0.001$ one-tailed) with 11% variance explained.

4.4 *Discussion*

The results of the above experiments suggest that words with a high first peak in the horizontal autocorrelation function are read aloud more slowly than those with low, and that these differences are not attributable to the presence or absence of ascenders and descenders. The results of the second experiment indicate that the differences between words may possibly be more pronounced when they are presented in a paragraph than in a widely separated list. In a paragraph, the vertical components of the stripes from words above and below the words being read may be important. If so, then the spacing of lines in a paragraph should influence the size of any effect of the striped properties of the words.

The next experiment was designed to see whether the effects are more marked if the paragraph is closely spaced rather than widely spaced.

5 Experiment 4

5.1 *Participants*

The participants were twelve male and twelve female students and staff at the University of Essex aged 20 – 59 years (mean 27 years).

5.2 *Materials*

50 high-frequency three- or four-letter words, without ascenders or descenders, were selected from the British National Corpus (Version 2, http://www.natcorp.ox.ac.uk). The words were set in Times New Roman 12 point and Geneva 10 point, x-height 2.0 mm, originated as before in Microsoft Word and printed to Adobe pdf files which were then converted with Adobe Photoshop to tiff images with a resolution of 300 dpi. The images were analysed with Matlab. The horizontal autocorrelation of each word was obtained and the words were divided about the median value of the first peak in the horizontal autocorrelation function into two groups: 19 with high first peak (main man mean men more move must name new now own room run same some time turn want war) and 19 with low (act are car care case cost its next once rate see seen set six sure too use view west). Microsoft Word was then used to create 32 passages of randomly ordered words, half comprising words with high first peak and half with low, half in Times New Roman and half in Geneva. Half the passages were presented with typical spacing (5.0 mm baseline-to-baseline) and half with wide spacing (9.7 mm baseline-to-baseline). All the passages were ranged left with ragged right margins.

A $2 \times 2 \times 2$ factorial design was used, giving eight conditions (2 of line spacing, 2 of font, and 2 of autocorrelation peak). Four different passages were available for each of the eight conditions, one selected at random for each participant so as to counterbalance the differences between passages across participants. This gave eight passages of text for each participant, one for each condition of line spacing, font, and autocorrelation peak.

5.3 *Procedure*

Participants were given the eight passages of text in a random order and asked to read the words in the passage aloud as quickly as possible. They were instructed not to use their finger as a guide across the page although they could use it on the left-hand margin if they wished. Participants were given 1 min in which to read as many words as they could, and their reading speed scored as previously.

5.4 *Results*

Table 4 shows the reading speed for the eight experimental conditions. The reading speed for words with low first peak in the horizontal autocorrelation was on average 9% greater than for words with high. A three-way repeated-measures analysis of variance revealed a significantly higher reading speed on passages comprising words with low first peak ($F_{1,23} = 54.4$, $p < 0.001$) with 37% variance explained. Overall, reading speed was significantly higher for Geneva than for Times New Roman ($F_{1,23} = 10.8$, $p = 0.003$) with 2.4% variance explained. There was no effect of line spacing and there were no significant interaction terms.

Table 4. Means (and standard deviations) of the reading speed in words per minute for the eight experimental conditions of experiment 4: word spacing, typeface, and first peak in horizontal autocorrelation function.

Spacing	Font	High first peak	Low first peak
Wide	Times New Roman	140 (19)	151 (20)
	Geneva	144 (18)	156 (22)
Typical	Times New Roman	139 (20)	155 (23)
	Geneva	142 (16)	155 (25)

5.5 *Discussion*

The effect on reading speed of the first peak in the horizontal autocorrelation was confirmed with words without ascenders and descenders. There was no effect of line spacing, suggesting that any proximity of vertical contours from one line to the next was not critical for the effect. The values of the autocorrelation first peak for words printed in Geneva were on average lower than those for words printed in Times New Roman. In general, words printed in Geneva had lower first peaks in the horizontal autocorrelation and were read more quickly than words printed in Times, even though the x-height and line spacing were the same.

In all the above experiments the participants were required to read the words aloud. Subsequent experiments were designed to see whether articulation contributed to the effects on reading speed.

6 Experiment 5

6.1 *Materials*

In this experiment only words printed in Times New Roman were used. Sixteen new passages were created from the same vocabulary of 32 words used in experiment 4. The practice passage was created from 19 new words. Otherwise, the procedure resembled that of the previous experiment.

6.2 *Procedure*

The participants were those who had taken part in the previous experiment. They were asked to read the passage silently, searching for a pair of target words selected by the experimenter so that they occurred only once in the passage. They were asked to tell the experimenter as soon as they had found the words and the time taken to do so was recorded. If participants reached the end of the text without finding the words they were asked to indicate this to the experimenter. Participants who failed to find the words after 1 min were stopped and asked to indicate the position they had reached. Participants were instructed not to use their finger as a guide and they were also instructed not to skim-read.

6.3 *Results*

The notional visual search speed was calculated in terms of the number of words in the passage prior to the target word pair divided by the time taken to find the word pair. On average 278 (SD = 64) words were scanned per minute in the low first peak condition compared with 226 (SD = 41) words per minute in the high first peak condition, a difference of 23%. The difference between the conditions was significant ($t_{23} = 7.2$, $p < 0.001$) with 69% variance explained.

6.4 *Discussion*

The size of the first peak in the horizontal autocorrelation affected the speed of silent visual search, demonstrating that the effect on reading is not dependent on articulation. The effect size was at least as large in this experiment as in the previous experiments. In the next experiment a similar search task was used but the instructions differed. The subjects were encouraged to find the specified word pair as quickly as they could using any search strategy they wished.

7 Experiment 6

7.1 *Participants*

Four male and four female students and staff at the University of Essex aged 19–46 years (mean 29 years) participated. They had not taken part in previous experiments.

7.2 *Procedure*

The participants were given the same practice passage and four passages of randomly ordered words as used in the previous experiment. The participants were instructed to find a specified pair of words as quickly as they could using any search strategy they wished.

7.3 *Results*

The average search speed was 465 (SD = 191) words per minute for the low first peak condition and 358 (SD = 146) words per minute for the high, a difference of 30% that was significant ($t_7 = 4.26$, $p = 0.002$), explaining 10% of the variance.

7.4 *Discussion*

The results of experiments 1–6 show that the first peak in the horizontal autocorrelation function of the image of a word is a measure of word shape that predicts the appearance of the word as striped and, more importantly, affects the speed with which words are read, both aloud and during silent search, regardless of the search strategy.

Having obtained evidence that the similarity of neighbouring letter shapes within words affects the speed with which the words are read, we made a preliminary attempt to reduce the first peak by distorting the word image continuously along its horizontal dimension.

7.5 *Distorting the word image to reduce the first peak*

When a word is read, the eye had a tendency to fixate a point slightly to the left of the centre of the word (Brysbaert and Nazir 2005; O'Regan 1990). For this reason we decided to compress the horizontal dimension of the word at this point and to expand the word elsewhere so as to leave the length of the word unaffected. The function used to transform the horizontal position of pixels is shown in figure 8 and had the following equation:

$$y = 12.742x^6 - 40.568x^5 + 47.427x^4 - 24.264x^3 + 4.8715x^2 + 0.7909x$$

where x is the original horizontal position of a pixel and y its transformed position.

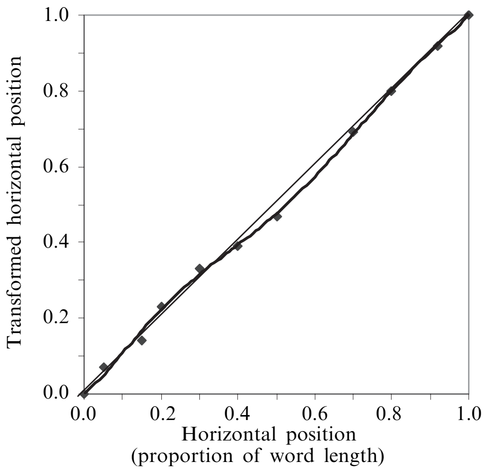

Figure 8. The distortion of the horizontal position of pixels in words used in experiments 7 and 8.

An example of the application of this transformation is shown in figure 9.

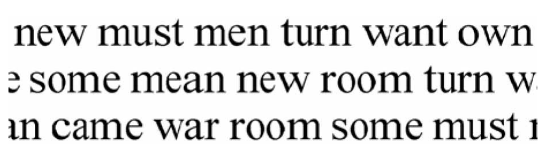

Figure 9. An enlarged fragment of one of the distorted passages used in experiments 7 and 8.

8 Experiment 7

8.1 *Participants*

Nine male and twenty-one female students and staff at the University of Essex, aged 18 – 57 years (mean 29 years) took part.

8.2 *Materials*

19 three- or four-letter monosyllabic words without ascenders and descenders were selected at random from the Maclure Bar Reading Test grade 4 (Haag-Streit, UK), specified for readers aged 7 – 8 years. The words were originated in Times New Roman 13 point, with Microsoft Word, and printed to an Adobe pdf file and converted to a tiff file with Adobe Photoshop with a resolution such that the x-height of each letter was 47 pixels and the average width of each letter about 50 pixels. A Matlab program was used to reposition each pixel horizontally according to the equation shown above (approximating position to the nearest pixel). The mean of the first peak of the horizontal autocorrelation function for the original (untransformed) words was 0.54 (SD $= 0.11$) and for the transformed words 0.45 (SD $= 0.08$), a significant difference ($t_{18} = 7.377$, $p = 0.03$, one tailed) which accounted for 17% of the variance.

Matlab was used to generate images of randomly ordered sequences of the words as tiff files that were then printed with a 600 pdi laser printer with a magnification such that the x-height was 1.95 mm. Four passages of untransformed and four passages of transformed words were generated, each with 15 lines of 15 randomly ordered words, measuring overall 70 mm $\times$ 144 mm.

The words were of high frequency, with no ascenders and descenders. The following was a typical sequence: mean more can now name time man came run some mean new move must same.

8.3 *Procedure*

The order in which the passages were administered was counterbalanced so that each participant received four passages of each version, and the sequence of words used was balanced across participants.

Because the passages were not obviously different, it was possible to conduct the study with a double-masked design, neither the tester nor the participant being aware of the identity of the various conditions. Participants were asked to read aloud as many words as possible from each of the passages within 1 min. As before, their scores comprised the total number of words correctly read in the appropriate sequence.

8.4 *Results*

For the untransformed condition, an average of 144 (SD $= 19$) words were read per minute. Slightly more words were read in the transformed condition (145, SD $= 19$) although the difference was not significant for the group as a whole ($t_{29} = 1.00$, $p = 0.163$, one-tailed). However, four participants in the group who identified themselves as 'dyslexic', showed an average 9% increase in reading speed in the transformed condition ($t_3 = 2.75$, $p = 0.04$, one-tailed). For this reason the experiment was repeated with a group of individuals who were less proficient readers, children aged 11 – 15 years.

9 Experiment 8

9.1 *Participants*

Thirty-four boys and twenty-two girls aged 11 – 15 years (mean: 13 years 5 months) from two secondary schools took part. The children were selected by their teachers as having general literacy difficulties.

9.2 *Materials*

Four pairs of passages with identical sequences of words were prepared as in the previous experiment.

9.3 *Procedure*
The procedure was similar to that in the previous experiment. When the participants had finished reading all four passages, they were asked to rank the passages in order of their clarity.

9.4 *Results*
The children read an average of 72 (SD = 22) words per minute of the undistorted text and 75 (SD = 22) of the distorted text, an increase that averaged 4.0% and was significant across participants ($t_{55} = 2.96$, $p = 0.005$, two-tailed). The participants made fewer errors on the distorted text than on the undistorted text in terms of missed words, missed rows, and mispronounced words. The number of errors per passage averaged 2.0 for the undistorted text and 1.7 for the distorted text ($t_{55} = 4.21$, $p = 0.0001$, two-tailed). The error analysis demonstrates there was no speed – accuracy trade-off. Despite the superior performance with the distorted text, the rankings showed that participants thought that the undistorted text was clearer to read. The mean rank of the undistorted text was 2.29 and that of the distorted text 2.71, a significant difference ($z = 2.41$, $p = 0.016$; Wilcoxon signed ranks test).

10 General discussion
The internal shape of a word affects the speed with which it is read. One aspect of the shape is the striped appearance due to the similarity between neighbouring letter strokes, as measured by the first peak in the horizontal autocorrelation. The striped aspect differs between typefaces and between words. Words with high first peak in the autocorrelation are read more slowly both aloud and silently, the more so if they are generated in a font, such as Times New Roman, with a generally high autocorrelation. This applies to words presented singly and in a paragraph, regardless of the line spacing. Reducing the similarity between letter strokes by distorting words or varying the letter spacing can increase speed in poor readers, even if the number of characters per unit area (type economy) is unchanged.

It has been shown that distortions seen in a pattern of stripes resemble those seen in text partly because of the lines of text (Wilkins 1995). The present results would indicate that some of the distortions may also be due to stripes from the letter strokes. The stripes responsible for distortions have a spatial frequency in the range 1 – 10 cycles deg^{-1}, and the repetitive lines of text and the repetitive letter strokes both have spatial frequencies within this range. The distortions are maximal for gratings with spatial frequencies close to 3 cycles deg^{-1} and at these spatial frequencies cortical activation is increased, abnormally so in those suffering from migraine, who are particularly susceptible to distortions and visual stress (Huang et al 2003). Visual stress has been attributed to a patchy hyperexcitability of the visual cortex combined with a failure of intracortical inhibition when the pyramidal neurons in orientation columns are strongly stimulated and compromise shared inhibitory mechanisms (Wilkins 1995, 2003). Individuals with visual stress, who see many distortions, are more likely than others to improve their reading speed with coloured overlays (Evans et al 1994; Hollis and Allen 2006). The coloured filters are thought to reduce the effect of cortical hyperexcitability by redistributing the excitation so as to avoid particularly hyperexcitable regions, a proposal consistent with (i) preliminary imaging studies (Wilkins et al 2004) and (ii) the benefits from coloured filters seen in a variety of neurological disorders that involve cortical hyperexcitability (Wilkins 2003). The present findings support this interpretation: they show that the extent to which words approximate patterns of stripes predicts the speed with which they are read, particularly by those whose reading is relatively dysfluent. The findings therefore tighten the links between susceptibility to distortions, visual stress, reading speed, and reading difficulty.

The present results are, however, open to other interpretations. During reading, the vergence angle of the two eyes is corrected after each saccade (Radach and Kennedy 2004), a process that may be dependent on the similarity of the image in each eye. Images with high horizontal autocorrelation can be inappropriately matched to one another, comprising vergence control. It remains to be seen whether the speed with which vergence is adjusted after a saccade is dependent upon the autocorrelation of the fixated word. Such an effect, if obtained, would help to explain the effects on reading speed observed in this paper.

Although in experiment 8 the distortion of a commonly used typeface resulted in a small increase in reading speed, we do not wish to give the impression that distortion of this kind is a suitable manipulation for improving type design—far from it. The design of a typeface is a complex issue involving many interacting factors, only a few of which have been brought to light in the current studies. The studies have shown that reading speed is affected by design features that are far more involved than the size of the typeface, the presence or absence of serifs, the boldness of the letter strokes per se, etc.

Many fonts were designed within the constraints set by metal typesetting. Digital typography gives designers greater freedom. The width of the body of a letter can vary, and the variation be determined by context. It is possible to diversify the shape and proximity of neighbouring letter strokes in particular contexts, the better to prevent repetitive stripes. The present findings suggest that the alteration could be subtle and almost unnoticeable, and yet improve reading speed, at least in dysfluent readers. Letters such as m, n, u, and h are common in words with high autocorrelation. Attention should be paid to the design of these letters, especially in words in which they appear jointly or in the context of letters such as i, j, and t.

Times New Roman is among the most commonly used fonts, but it has very high autocorrelation peaks. Verdana provides an alternative font with some of the lowest peaks, even when font size is reduced so as to give the same type economy. In future work we plan to compare a large number of fonts, including those we redesign, to see the extent to which their autocorrelation will predict reading speed.

Acknowledgments. We thank Kalvis M Jansons for the suggestion of using autocorrelation and Sean Jackson for help with software in preliminary studies. We thank the participants—staff and students at the University of Essex, Colchester; Colne Community School, Brightlingsea; and Deben High School, Felixstowe.

References

Arditi A, Cho J, 2005 "Serifs and font legibility" *Vision Research* **45** 2926 – 2933

Balota D A, Yap M J, Cortese M J, Hutchison K A, Kessler B, Loftis B, Neely J H, Nelson D L, Simpson G B, Treiman R, 2007 "The English Lexicon Project" *Behavior Research Methods* **39** 445 – 459

Beech J R, Mayall K A, 2005 "The word shape hypothesis re-examined: evidence for an external feature advantage in visual word recognition" *Journal of Research in Reading* **28** 302 – 319

Brysbaert M, Nazir T, 2005 "Visual constraints in written word recognition: evidence from the optimal viewing – position effect" *Journal of Research in Reading* **28** 216 – 228

Cattell J M, 1886 "The time taken up by cerebral operations" *Mind* **11** 277 – 282, 524 – 538

Evans B J W, Cook A, Richards I L, Drasdo N, 1994 "Effect of pattern glare and coloured overlays on a simulated reading task in dyslexics and normal readers" *Optometry and Vision Science* **71** 619 – 628

Evans B J W, Joseph F, 2002 "The effect of coloured filters on the rate of reading in an adult student population" *Ophthalmic and Physiological Optics* **22** 535 – 545

Haber L R, Haber R N, Furlin K R, 1983 "Word length and word shape and sources of information in reading" *Reading Research Quarterly* **18** 165 – 189

Hollis J, Allen P M, 2006 "Screening for Meares – Irlen sensitivity in adults: can assessment methods predict changes in reading speed?" *Ophthalmic and Physiological Optics* **26** 566 – 571

Huang J, Cooper T G, Satana D, Kaufman D I, Cao Y, 2003 "Visual distortion provoked by a stimulus in migraine associated with hyperneuronal activity" *Headache* **43** 664 – 671
Hughes L E, Wilkins A J, 2000 "Typography in children's reading schemes may be suboptimal: evidence from measures of reading rate" *Journal of Research in Reading* **23** 314 – 324
Hughes L E, Wilkins A J, 2002 "Reading at a distance: implications for the design of text in children's big books" *British Journal of Educational Psychology* **722** 213 – 226
Leech G, Rayson P, Wilson A, 2001 *Word Frequencies in Written and Spoken English Based on the British National Corpus* (Harlow, UK: Pearson Education)
Legros L A, Grant J C, 1916 *Typographical Printing Surfaces; The Technology and Mechanism of Their Production* (London: Longman, Green and Co)
O'Regan J K, 1990 "Eye movements and reading", in *Eye Movements and Their Role in Visual and Cognitive Processes* Ed. E Kowler (Elsevier, Amsterdam)
Paap K R, Newsome S L, Noel R W, 1984 "Word shapes in poor shape for the race to the lexicon" *Journal of Experimental Psychology: Human Perception and Performance* **10** 413 – 428
Radach R, Kennedy A, 2004 "Theoretical perspectives on eye movements in reading: past controversies, current deficits, and an agenda for future research" *European Journal of Cognitive Psychology* **16** 3 – 26
Wilkins A J, 1995 *Visual Stress* (Oxford: Oxford University Press)
Wilkins A J, 2003 *Reading through Colour* (Chichester, Sussex: John Wiley and Sons)
Wilkins A J, Huang J, Cao Y, 2004 "Visual stress theory and its application to reading and reading tests" *Journal of Research in Reading* **27** 152 – 162
Wilkins A J, Jeanes R J, Pumfrey P D, Laskier M, 1996 "Rate of Reading Test: its reliability, and its validity in the assessment of the effects of coloured overlays" *Ophthalmology and Physiological Optics* **16** 491 – 497

ISSN 0301-0066 (print) ISSN 1468-4233 (electronic)

PERCEPTION

VOLUME 36 2007

www.perceptionweb.com