Vision Performance Institute

Technical Report

Effect of Character Spacing on Text Legibility

Yu-Chi Tai, Shun-nan Yang, and John Hayes, James Sheedy

Vision Performance Institute, Pacific University, Oregon, USA

Running title: Font size and optimal inter-letter spacing

Correspondence concerning this article should be addressed to Yu-Chi Tai, Vision Performance Institute, Pacific University in Oregon, Forest Grove, OR 97116. Email: ytai@pacificu.edu

Phone: 1.503.352.2289.

Abstract

This study investigated the effect of inter-letter spacing on letter and word recognition over a range of font sizes, font types, and different levels of text familiarity (word frequency). Inter-letter spacing affects the degree of lateral interference form neighboring letters. For commonly used fonts, character spacing is proportional to font size. Since lateral interference would be expected to operate over a fixed retinal area, we hypothesized that, relative to the default spacing for the common 10- to 12-point fonts, the optimal inter-letter spacing would be larger for smaller font sizes and smaller for larger font sizes. We also hypothesize that optimal spacing will be smaller for familiar words than for unfamiliar words because of top-down enhancement and that sans serif fonts will need wider spacing than serif fonts because of the extra contours.

Fifty-four subjects were recruited and tested in two tasks: text (letter and word) recognition and spacing preference. In the text recognition, subjects were asked to read aloud 840 words or the middle letter in each of 420 random-letter strings. Half of the words were high-frequency and the other half low-frequency. All words and letter strings were 5 letters in length, created in 72-point font rendered in Grayscale on a LCD screen, and viewed at various distances to simulate font sizes and spacings. After text recognition, subjects were asked to choose the preferred image among 12 layouts of a same word or letter string. These 12 layout differed by inter-spacing only.

Subjects responded faster and more accurately to bigger font, more familiar words, and with wider spacing. The effect of inter-letter spacing was stronger for smaller fonts and less familiar text (low-frequency words or letter strings) than for larger fonts. There was stronger effect of condensed spacing on sans serif fonts for regular and smaller fonts, but overall, the results were similar between font types.

Overall the results support the hypotheses and suggest the use of wider spacing for smaller font sizes and more condensed spacing for bigger fonts.

Introduction

In typography, inter-letter spacing (also called *tracking*) refers to the amount of white space between two letters. Unlike *kerning*, which is the special adjustment of spacing between two particular characters, inter-letter spacing generally applies to the overall spacing between all neighboring characters in a set and affects the information (character) density in a line or a block of text. The size of inter-letter spacing affects text legibility as it manipulates the amount of lateral interference among neighboring characters. Studies have found that the presence of neighboring characters interferes with the processing of individual characters (Yang, Tai, Hayes, & Sheedy, 2009) or words (Sheedy, Subbaram, Zimmerman, & Hayes, 2005; Tai, Sheedy, & Hayes, 2006). It has been identified that the size threshold for identifying single letters (i.e., a measure of legibility) is much smaller than the size threshold for identifying words, which is referred to as the "letter superiority effect" (Sheedy et al., 2005). By increasing the distance between characters, word legibility improved and gradually reached an asymptote close to single character legibility (Tai et al., 2006).

Current common practice is to have the same proportional spacing applied to a range of font sizes. Such proportional scaling maintains the spatial ratio for a character and its adjacent space; however, it may not be the best design to optimize reading performance. Since lateral interference apparently operates over a relatively fixed retinal distance regardless the size of the characters, smaller spacing between characters can cause larger lateral interference effects on word recognition. Empirical evidence indicated that the default spacing at common text of 10- to 12-point font is barely enough to avoid lateral interference for 10- and 12-point fonts, as shown by the existence of the letter superiority effect (Tai et al., 2007). Consequently, the proportional spacing will result in tighter spacing for smaller font sizes, which would be expected to cause measurable lateral interference. The interference may be further

exacerbated by the degradation of the small-font text image, due to fewer pixels displaying each character. At larger font sizes, the proportional spacing results in larger spacing and hence letter recognition should be free from the lateral interference of neighboring letters. For larger font sizes words and letters are expected to have the same response time and letter spacing likely can be reduced without compromising word legibility.

In addition to font size, typeface is another factor significantly affecting text legibility (e.g., Mansfield, Legge, & Bane, 1996; Sheet et al., 2005); however, it is not clear whether the effect of inter-letter spacing will be the same across font types, such as between serif and sans serif fonts. Some researchers believe that serifs enhance letter discrimination by making the special features more distinctive and decipherable and increasing the salience of the main strokes (e.g., McLean, 1980; Rubinstein, 1988); however, there are good reasons to argue that serifs are small and ornamental with little impact, or even being harmful for taking up space from the main strokes. While extra spacing is needed to accommodate the addition of serifs, the average spacing between two serif letters (from the nearest point of a glyph to another glyph) is usually no larger than sans serif letters, due to the process of kerning or the adjustment made to accommodate serifs from two glyphs within the same space (see Table 1). However, the distance between the main strokes could be potentially larger in serif fonts in order to accommodate the glyphs. The actual spacing between characters is determined by the font designers.

Table 1. Samples of text and spacing width - serif and sans serif fonts.

12-pt Sans serif fonts	
Corbel	The quick brown fox jumped over a lazy dog.
Calibri	The quick brown fox jumped over a lazy dog.
Candara	The quick brown fox jumped over a lazy dog.

Arial	The quick brown fox jumped over a lazy dog.
Helvetica	The quick brown fox jumped over a lazy dog.
Tahoma	The quick brown fox jumped over a lazy dog.
Futura	The quick brown fox jumped over a lazy dog.
Comic	The quick brown fox jumped over a lazy dog.
Frutiger	The quick brown fox jumped over a lazy dog.
Verdana	The quick brown fox jumped over a lazy dog.
12-pt Serif fonts	
Cantaur	The quick brown fox jumped over a lazy dog.
Garamond	The quick brown fox jumped over a lazy dog.
Baskerville	The quick brown fox jumped over a lazy dog.
Times New Roman	The quick brown fox jumped over a lazy dog.
Bodoni MT	The quick brown fox jumped over a lazy dog.
Cambria	The quick brown fox jumped over a lazy dog.
Constantia	The quick brown fox jumped over a lazy dog.
Georgia	**The quick brown fox jumped over a lazy dog.**
Boon Antiqua	The quick brown fox jumped over a lazy dog.
Palatino Linotype	The quick brown fox jumped over a lazy dog.

In this study we tested the effect of inter-letter spacing on letter and word recognition over a range of font sizes, font types, and different levels of lexical familiarity (word frequency). Our hypotheses are:

1. Relative to the default and proportional spacing for the common 10- to 12-point fonts, the optimal inter-letter spacing will be larger for smaller font sizes and smaller for larger font sizes.

2. Text recognition is affected by both bottom-up visual processing (e.g., character legibility) and top-down cognitive processing. Therefore the optimal spacing would be wider for unfamiliar words than for familiar words because feature identification would be more important for visual abstraction of unfamiliar words.

3. As there is more white space between the main strokes of two neighboring letters for sans serif fonts, we expected that the optimal spacing would be more condensed for serif fonts than for sans serif fonts.

The current study manipulated these factors to test their effects on optimal spacing with different types of stimuli. It is expected that the findings will help improve font design.

Methods

Subjects

Fifty subjects (34 females, age 18-60, average age: 28.17 yr old) were recruited from the Pacific University community for either monetary compensation or for course credit. They were all native English speakers, with normal or corrected-to-normal distance visual acuity of 20/20 or better for each eye, and without ocular pathology, oculomotor limitation, abnormality, or reading disabilities. Before participating in the experiment, each subject read and signed a consent form that was consistent with the Declaration of Helsinki and approved by the Pacific University Institutional Review Board for Human Subjects.

Material

Two types of stimuli were constructed to examine the possible separate effects of spacing on word legibility , including 840 words and 420 non-word letter strings, each containing five characters. The words were selected from the MRC Psycholinguistic Database (http://www.psy.uwa.edu.au/MRCDataBase/uwa_mrc.htm). Half of the words were high-frequency words (occurring 200 to 13339 times per 3,000,000 words); the other half were low-frequency words (less than 20 occurrences per 3,000,000 words). The non-word letter strings were composed of five letters with the middle letter randomly selected from a set of 11 letters (a, c, e, n, o, r, s, u, v, x, z) and the other four letters generated by a computer program under the constraint that all five letters were different and not converted to be a real English word.

All stimuli were displayed in 72-point lowercase font, rendered in Grayscale, half in Georgia (serif) and the other half in Helvetica (sans-serif). Both are proportional fonts (glyph width varies by the shape of individual characters) with similar width and height for glyphs of a same character. Seven inter-letter spacings were used in the study: -15, -10, -5, 0, +5, +10, or +15 pixels from the default spacing of 72-point font. The images were generated using a special program, ClearType Test (cttest.exe), developed by Microsoft. It allows precise control of the inter-character spacing up to 1/64 pixel. To create the effect of various font sizes without changing the pixel resolution of the testing stimuli, viewing distance was manipulated (50, 150, 300, 360, 450, and 600 cm) so that the resulted text angular size for each 72-point character would correspond to that of font sizes 72, 24, 12, 10, 8, and 6 point as viewed from 50 cm. Compared to the font sizes of 10- and 12-point commonly used for onscreening reading and writing, the 72- and 24-point fonts represent larger font sizes mostly used in Power Point presentation while 6- and 8-point fonts represent smaller font sizes read from in handheld devices or magazines.

In addition to the words and letter strings for the text recognition task, an additional 12 displays were made for the spacing preference task. Each display contained 12 images of the same word or letter string, which were presented simultaneously to the subject from the viewing distance of 50 cm. Half of the displays were presented in 72-point Georgia and the other half in 72-point Helvetica font. The 12 words or letter strings on each display included the following spacings in random order: -25, -20, -15, -10, -5, 0, +5, +10, +15, +20, +25, and +30 pixels from the default spacing. Each stimulus contained five letters, with all thin glyphs (e.g.,

Figure 1a and 1d), all wide glyphs (e.g., Figure 1b and 1e), or a combination of both thin and wide glyphs (e.g., Figure 1c and 1f).

jljlj	jljlj	silly	silly
jljlj	jljlj	silly	silly
jljlj	jljlj	silly	sily
jljlj	jljlj	sly	silly
jljlj	j	silly	sy
jljlj	jljlj	silly	silly

(a)   (d)

wmwmw	wmwmw	mummy	m u m m y
wmwmw	w m w m w	mummy	mummy
wmwmw	wmwmw	mummy	mummy
w m w m w	wmwmw	mummy	mummy
wmwmw	w m w m w	m u m m y	m u m m y
wmwmw	w m w m w	mummy	mummy

(b)   (e)

wlwlw	wlwlw	milky	milky
w l w l w	wlwlw	m i l k y	mky
wlwlw	wlwlw	mky	m i l k y
wlwlw	wlwlw	milky	mky
w l w l w	wlwlw	milky	milky
wlwlw	w l w l w	milky	milky

(c)   (f)

Figure 1. Examples of the testing stimuli in the spacing preference task. Figure (a), (b), and (c) represent tested letter strings in 72-point Georgia, with letters of thin glyphs (figure a), wide glyphs (figure b), and a combination of letters of thin and wide glyphs (figure c). Figure (d), (e),

and (f) represent images of words in 72-point Helvetica of thin, wide, and combination glyphs, respectively.

Apparatus

Experiment Builder software (SR Research, 2008) was used to script the experiment and control the progress of trials. A Pentium 1.8 GHz Dell PC and a 17" LCD monitor with a 60Hz vertical refresh rate were used to conduct the program and display the visual stimuli. A voice key (SV-1 voice key, Cedrus Corp., CA: San Pedro) was used to register the onset time of the subject's oral response. A digital distance vision chart (Smart Vision 20/20) was used to measure distance visual acuity prior to the experiment.

Procedure

The subject was briefed about the goals and the experimental procedures before informed written consent was obtained. Then they were screened to meet the criterion of 20/20 or better distance visual acuity for each eye, with habituated optical correction if needed.

Trials for word recognition and letter recognition were conducted separately, with 30 blocks for each task. There were 14 trials per block for letter recognition and 28 trials per block for word recognition. All images within a block were presented at a fixed viewing distance (50, 150, 300, 360, 450 or 600 cm), randomly selected for each subject. These viewing distances created text angular sizes corresponding to font sizes 72, 24, 12, 10, 8, and 6 point, respectively, as if viewed from 50 cm. Each viewing distance was repeated 5 times (5 blocks). The inter-character spacing and font type were randomly assigned to each trial, and words with high- and low-frequency were evenly dispersed within each block of word recognition task.

The participant stood in front of the LCD monitor at the assigned viewing distance, wearing a headset with the microphone of the voice key close to their mouth. For letter

recognition, each trial started with a preparation display of two crosses, one above and one below the area where the middle of the word or letter string was to be displayed. The experimenter prompted the subject to look at the fixation and be ready for the presentation of the testing stimulus. Then the experimenter pressed a button to initiate the testing display. After a brief delay (i.e. a stimulus onset synchrony, or SOA) that varied randomly from 1000 to 2500 milliseconds, a string of 5 letters was displayed with the middle letter centered at the screen and remained on until an oral response was received or when the 10-second waiting time was reached, whichever came first. The subject was asked to call out the middle letter of the string as quickly and as accurately as possible through a voice key, which registered the time of response. The experimenter determined the accuracy of the response and registered it into the program, which action triggered the onset of the two crosses for the next trial. In between trials there was a chance for break if needed. The same procedure was used for trials of word recognition with the revision that, instead of calling out the middle letter, the subject was asked to read aloud the whole word.

After completing the word and letter recognition tasks, the subject performed 12 trials of spacing preference. In each trial, 12 images of the same word or letter string were presented simultaneously on the LCD screen at viewing distance of 50 cm. The subjects were asked to choose the image that was easiest to read. The total testing time of the study was 70-90 minutes.

Data analysis

Subject response accuracy and latency for the word and letter recognition tasks and their preferences for the spacing preference task were recorded. Response accuracy was transformed by Arcsin ($P'i$ = arsin $\sqrt{P_i}$, when Pi is a percentage) to normalize the distribution before analysis and converted back after analysis before reporting. Response time was

transformed to log10 values for normal distribution and values outside the range of 99.4% were excluded from analysis. Both accuracy and RT were aggregated to calculate the mean for individual subject in each condition. Then Mixed Model Within-subjects ANOVAs were used to compare across inter-letter spacings, font types, and corresponding font sizes. The derived estimated means and their standard errors (with between subject variance removed) were used to plot figures.

Results

Accuracy

For letter recognition, the data were split by font type and font size before fed to Mixed Model Within-subject ANOVA to generate converged model. Table 2 shows the result of the testing effect. Accuracy is significantly affected by inter-letter spacing for both font types at all font sizes ($p=.001$ or smaller), except for letters viewed from 24-pt Georgia.

Similar results were obtained for word recognition. The data were split by font type and word frequency but the effects of font size and letter spacing on response accuracy were tested directly with Mixed Model Within-subject ANOVA. Table 3 shows significant main effects ($p<.0001$) of font size and spacing and their interaction ($p<.0001$), for both high- and low-frequency words with Georgia and Helvetica.

Table 2. Accuracy for letter recognition.

Type III Tests of Fixed Effects[a]

Font type	Font size	Source	df	or df	F	Sig.
Georgia	6	Intercept	1	50	8742.158	0.000
		spacing	6	300	28.683	0.000
	8	Intercept	1	50	13090.418	0.000
		spacing	6	300	8.177	0.000
	10	Intercept	1	50	38088.512	0.000
		spacing	6	300	4.043	0.001
	12	Intercept	1	50	49966.977	0.000
		spacing	6	300	6.623	0.000
	24	Intercept	1	50	50758.859	0.000
		spacing	6	300	1.747	0.110
	72	Intercept	1	50	45113.009	0.000
		spacing	6	300	6.388	0.000
Helvetica	6	Intercept	1	50	12352.148	0.000
		spacing	6	300	58.586	0.000
	8	Intercept	1	50	15388.355	0.000
		spacing	6	300	20.222	0.000
	10	Intercept	1	50	45586.474	0.000
		spacing	6	300	53.794	0.000
	12	Intercept	1	50	17907.026	0.000
		spacing	6	300	4.506	0.000
	24	Intercept	1	50	46371.583	0.000
		spacing	6	300	6.198	0.000
	72	Intercept	1	50	36665.938	0.000
		spacing	6	300	30.487	0.000

Table 3. Accuracy for word recognition.

Type III Tests of Fixed Effects [a]

Frequency	Fonttype	Source	Numerator df	Denominator df	F	Sig.
High frequency words	Georgia	Intercept	1	51.000	157481.378	0.000
		Font size	5	2091.000	10.378	0.000
		Spacing	6	2091.000	13.204	0.000
		Font size * Spacing	30	2091.000	13.049	0.000
	Helvetica	Intercept	1	51.000	182598.368	0.000
		Font size	5	2091.000	43.261	0.000
		Spacing	6	2091.000	38.999	0.000
		Font size * Spacing	30	2091.000	16.387	0.000
Low frequency words	Georgia	Intercept	1	51.000	17843.485	0.000
		Font size	5	2091.000	7.819	0.000
		Spacing	6	2091.000	10.297	0.000
		Font size * Spacing	30	2091.000	8.760	0.000
	Helvetica	Intercept	1	51.000	15710.377	0.000
		Font size	5	2091.000	7.654	0.000
		Spacing	6	2091.000	34.860	0.000
		Font size * Spacing	30	2091.000	7.398	0.000

Figure 2 shows the plot of the estimated means for response accuracy across all conditions. Response accuracy was increased with font size and reached the ceiling when the text was 12 point or larger. Word frequency affected accuracy as well. The more familiar high-frequency words were responded more accurately in general than low-frequency words and letter strings, especially with condensed spacing of -10 or more from the default. While the results were similar in both font types, Georgia trials were answered slightly better than Helvetica trials for smaller fonts with highly condensed spacings.

The effect of spacing on response accuracy was more obvious in smaller fonts than in larger ones. For smaller fonts of 6 to 8 point, accuracy continued to improve as spacing increased to the default level for low-frequency words and for letter strings, where less help

from the top-down process was obtained; for high-frequency words, even condensed letter spacing like -10 point made no difference from the default condition. For common text size of 10- to 12-point fonts, accuracy was better and started to reach asymptote with spacing of -5 pixels, suggesting better tolerance of condensed spacing. For larger font sizes of 24 or 72 points, response accuracy was not affected, even with extremely condensed spacing of -15 pixels from the default.
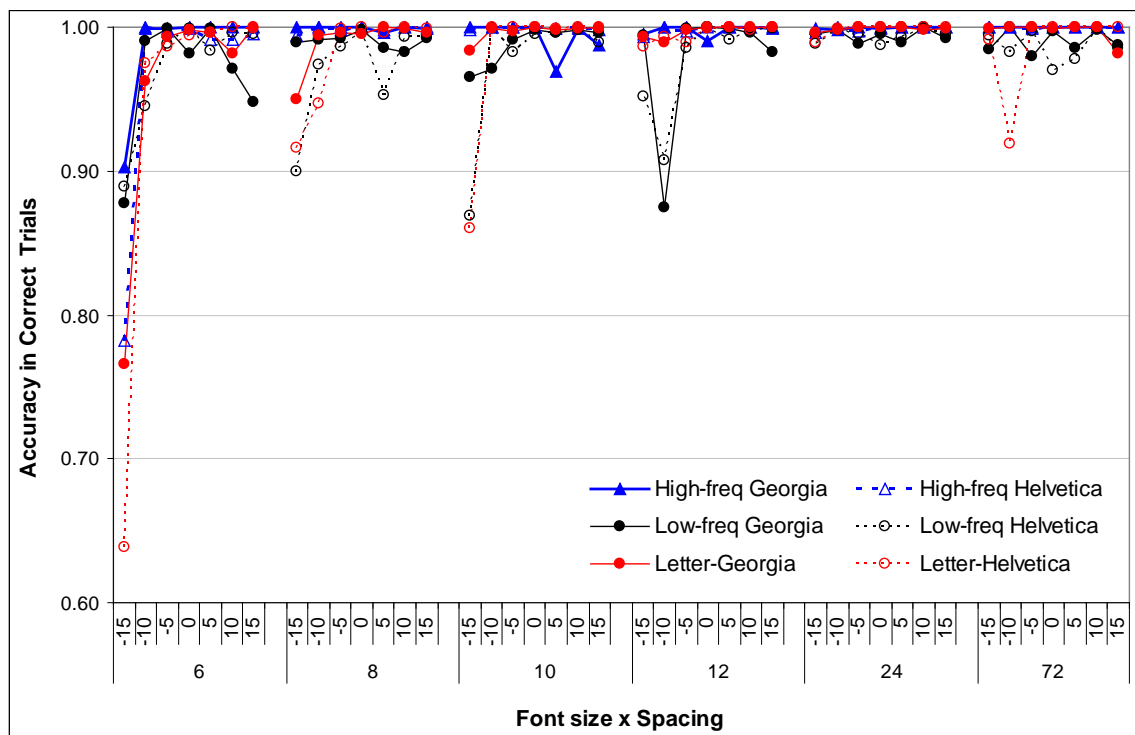

Figure 2. Accuracy for word and letter recognition

In conclusion, these findings suggest the following:

1. Accuracy of text recognition was affected by font size, word frequency, and inter-letter spacing.

2. Although similar, accuracy with serif font (Georgia) was slightly better than with sans serif font (Helvetica) when presented with smaller font size in highly condensed spacing.

3. Accuracy was strongly affected by the interaction of font size and spacing. Spacing with smaller fonts needs to be maintained at the default size or larger. For regular font size of 10 to 12 points, spacing can be more condensed, about -5 pixels from the default, and still maintain good accuracy. For bigger font sizes letter spacing can be decreased to -10 or more without sacrificing text legibility.

In short, the results suggest that the proper range of inter-letter spacing for accurate text recognition changes with font size: the bigger the font, the more condensed the spacing can be. Font spacing can be reduced to a slightly condensed size (e.g., -5 pixels) for larger fonts for more economical use of screen space.

Response time

Data of accurate trials were split by stimulus type (high-frequency words, low-frequency words, and letter strings) and font type (Helvetica and Georgia) before fed into Mixed Model Within-subject ANOVA. Table 4 shows that RT was significantly affected by the main effects of font size and inter-letter spacing ($p < .0001$) as well as their interactions ($p < .0001$) in all conditions.

Table 4. Response time for word and letter recognition.

Type III Tests of Fixed Effectsa

Accuracy	Stimulus type	Font type	Source	Numerator df	Denominator df	F	Sig.
Correct Trials	High-frequency words	Georgia	Intercept	1	50.979	150653.198	.000
			Font size	5	2090.743	147.610	.000
			Spacing	6	2090.743	96.721	.000
			Font szie *	30	2090.743	18.547	.000
		Helvetica	Intercept	1	51.000	151597.493	.000
			Font size	5	2091.000	115.470	.000
			Spacing	6	2091.000	91.521	.000
			Font szie *	30	2091.000	9.636	.000
	Low-frequency words	Georgia	Intercept	1	51.014	115340.693	.000
			Font size	5	2088.451	150.008	.000
			Spacing	6	2088.451	48.956	.000
			Font szie *	30	2088.451	20.913	.000
		Helvetica	Intercept	1	51.001	118377.666	.000
			Font size	5	2090.010	114.038	.000
			Spacing	6	2090.010	59.213	.000
			Font szie *	30	2090.010	15.097	.000
	Letter strings	Georgia	Intercept	1	50.002	136191.809	.000
			Font size	5	2044.121	481.920	.000
			Spacing	6	2044.119	134.919	.000
			Font szie *	30	2044.119	14.181	.000
		Helvetica	Intercept	1	50.000	135648.558	.000
			Font size	5	2048.037	406.713	.000
			Spacing	6	2048.037	193.388	.000
			Font szie *	30	2048.037	10.512	.000

Confined by the number of trials, we could not test the differences among stimulus type and between font types directly, nor could we locate the point of inflection (i.e., the point at which the curvature changes) statistically, as the statistical analysis cannot reach a converged model. However, several findings can be determined from Figure 3, the plot of the estimated mean RTs across conditions:

1. High-frequency words were responded faster than low-frequency words, which were recognized faster than identifying the middle letter in letter strings.

2. RT was similar between serif (Georgia) and sans serif (Helvetica) fonts, though occasionally it took longer to recognize a sans serif font at smaller font size.

3. In general, RT decreased as font size increased, especially for less familiar stimuli (low-frequency words and letter strings), for which less top-down information was available; however, the effect of font size was significantly attenuated by inter-letter spacing.

4. Similarly, RT decreased by the increase of inter-letter spacing for all font sizes, but the size of the effect was affected by font size.

    a. For smaller text like 6- or 8-point font, RT quickly decreased as spacing increased, especially for letter strings. The slope was steeper for condensed spacings and became flatter around the default spacing with smaller improvement as spacing continued to expand.

    b. For text with regular font sizes like 10 or 12 point, RT for high-frequency words was relatively stable when spacing was -10 pixels or larger, but continued to improve for low-frequency words and letter strings until reaching default spacing.

    c. For bigger font sizes (24 or 72 point), RT was less affected by inter-letter spacing, especially for familiar words.

5. An interesting observation is the RT pattern of low-frequency words. While it was more similar to high-frequency words in smaller font sizes, it became more comparable to the pattern for letter strings for larger font sizes.
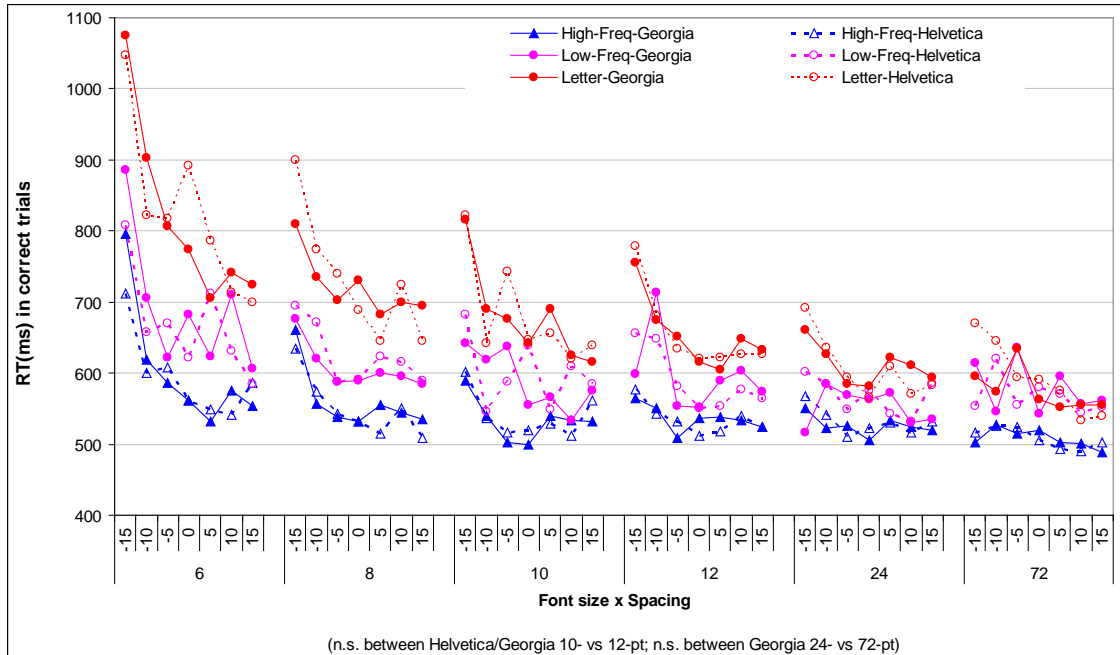
Figure 3. Response time for word and letter recognition

Preference

Table 5 lists the tally of individual spacings chosen as the preferred image under various

conditions. Chi-Square was used to test the goodness of fit between the observed and expected

distributions. In this case, the preference for each spacing level is expected to be the same

(1/12*52=4.33). The Chi-Sqaure result shows stronger preference for default spacings and

unfavorable for the condensed spacing of -10 pixels or more for both words and letter strings, in

Georgia or Helvetica, with thin or wide glyphs or their combinations. Figure 4 illustrates the

distribution of spacing preference across conditions. These results agree with the accuracy and

RT in text recognition, suggesting that images with default or close-to-default spacings (0 ~ +10)

were preferred, especially for word stimuli. For letter strings, preference for wider spacings was

higher than for word stimuli. However, these results were derived when viewed from 50 cm

distance of 72-point fonts. The results for different text angular sizes may vary.

Table 5. Chi-Sqaure analysis for spacing preference task

Characteristics *Choice Crosstabulation

Count

Stimulus	Font type	Glyph Characteristics		Choice -25	-20	-15	-10	-5	0	5	10	15	20	25	30	Total	Chi-square (N=12-1)	p-value
Letter string	Georgia	Characteristics combination																
				0*	0*	0*	0*	3	8*	7	12*	7	8*	2	5	52	42.154***	0.000
			fat	0*	0*	0*	1*	4	10*	14*	11*	5	3	2	2	52	57.8***	0.000
			thin	0*	0*	1*	0*	1*	8*	10*	11*	4	4	7	5	51	41.470***	0.000
		Total		0	0	1	1	8	26	31	34	16	15	11	12	155		
	Helvetica	Characteristics combination																
				0*	0*	0*	0*	0*	8*	15	9*	10*	6	2	2	52	66.615***	0.000
			fat	0*	0*	0*	0*	3	7	8*	13*	4	4	7	6	52	42.154***	0.000
			thin	0*	0*	0*	0*	0*	0*	10*	9*	13*	8	3	9*	52	64.308***	0.000
		Total		0	0	0	0	3	15	33	31	27	18	12	17	156		
Word	Georgia	Characteristics combination																
				0*	0*	0*	0*	8*	13*	23*	3	1*	1*	1*	2	52	127.538***	0.000
			fat	1*	0*	0*	4	13*	20**	9*	2	0*	0*	0*	3	52	104.923***	0.000
			thin	1*	0*	0*	0*	3	15*	11*	11*	5	4	2	0*	52	68.462***	0.000
		Total		2	0	0	4	24	48	43	16	6	5	3	5	156		
	Helvetica	Characteristics combination																
				0*	0*	0*	1*	5	18*	11*	9*	3	2	2	0*	51	82.882***	0.000
			fat	0*	0*	0*	3	11*	17*	8*	10*	1*	1*	0*	1*	52	83.231***	0.000
			thin	0*	0*	0*	0*	3	12*	21*	10*	1*	2	2	1*	52	110.462***	0.000
		Total		0	0	0	4	19	47	40	29	5	5	4	2	155		

* p<.05; ** p<.001;   *** p<.0001

Discussion and Conclusion

The current study examined the hypotheses that, relative to the default inter-letter spacing, the optimal spacing would be: (1) larger for smaller font sizes and smaller for larger font sizes, (2) smaller for familiar text (e.g., high-frequency words) than unfamiliar text (e.g., low-frequency words or random letter strings), and (3) smaller for serif fonts and wider for sans serif fonts.

Here is the summary of our findings:

1. Response accuracy for word and letter recognition was significantly affected by font size, familiarity of the stimuli, and inter-letter spacing.

a. Response accuracy reached a ceiling for all text conditions presented in bigger font sizes (e.g., 24- or 72-point fonts). This suggests that when the character is big and clear, the spacing can be further condensed than default for more efficient use of display space, at least within the range of spacings tested in our text recognition task.

b. For text presented in smaller sizes, wider spacing enhanced recognition accuracy and speed, especially for unfamiliar text.

c. While Georgia and Helvetica are both legible font types, accuracy was slightly higher for Georgia when presented in small font sizes with condensed spacing, probably benefited from the distinctive serifs and the extra white space around the serifs. Sans serif fonts, on the other hand, with strokes of similar width, are harder to make out with smaller font sizes of limited pixels.

2. Response time for text recognition was also significantly affected by font size, familiarity of the stimuli, and inter-letter spacing.

a. RT decreased with larger font sizes and text familiarity.

b. RT for serif (Georgia) and sans serif (Helvetica) fonts were similar, but occasionally faster RT was observed for serif fonts with smaller fonts.

c. RT was particularly high for small fonts with condensed spacing, and decreased quickly as spacing increased. Small fonts with condensed spacing are particularly vulnerable to lateral inhibition.

d. Low-frequency words and random letter strings were more affected by close spacing than high-frequency words, suggesting the reliance of clear visual features for unfamiliar text.

e. For regular sized text, RT reached asymptote for word stimuli when spacing was close to the default range, but continued to decrease for letter strings as spacing increased.

f. RT for high-frequency words in bigger font sizes was relatively unaffected by spacing, but was improved with wider spacing for low-frequency words and letter strings.

3. When vieiwng 72-point font from 50 cm distance, most subjects preferred images with default or close-to-default spacing (0 ~ +10), especially for word stimuli; preference for wider spacing increased when presented with letter strings or for characters with thinner glyphs.

These results show that inter-letter spacing can affect how quickly and accurately text is perceived. Legibility and, presumably, reading performance could be enhanced if character spacing was varied according to the format (e.g., font size and font type) and the content (e.g., text or information familiarity) of the displayed text. For smaller font sizes, greater inter-letter spacing improves legibility (recognition accuracy) and decreases processing time (RT). Therefore, an additional whitespace around the character glyph will be preferred to reduce lateral inhibition effects. For common reading font sizes (10 and 12), the default spacing works well; even slightly condensed spacing (-5 pixels or so) will not harm the text processing. For larger fonts, however, inter-letter spacing is not critical to text recognition, especially for familiar words; on the contrary, excessive white space could result in the impression of isolated letter strings and break the continuity of the word shape. In these conditions, inter-letter spacing could be tighter than the current default value in order to use the limited screen resolution more efficiently.

One thing needed to be pointed out is that the current findings were derived by increasing viewing distance to create smaller angular text size. Although this accurately produced the appropriate text size and same proportional spacing across font size, it over-represented the pixel resolution of the smaller font sizes and could have favored reaction times and target identification for the smaller fonts compared to the larger fonts.

References

Mansfield, J. S., Legge, G. E., & Bane, M. C. (1996). Psychophysics of reading Xv: Font effects in normal and low vision. Investigative Ophthalmology and Visual Science, 37(8), 1492–1501.

McLean, R. (1980). The thames and hudson manual of typography (Vol. 1). London, UK: Thames and Hudson Ltd.

Rubinstein, R. (1988). Digital typography: An introduction to type and composition for computer system design. Boston, MA: Addison Wesley.

Sheedy, J.E., Subbaram, M. V., Zimmerman, A.B., & Hayes, J.R. (2005). Text legibility and the letter superiority effect, *Human Factors,* **47** (4), pp. 797–815.

Tai Y, Sheedy JE, Hayes JR. Can threshold legibility predict suprathreshold performance? Optom Vis Sci 2007; e-abstract 070058

Tai, Y.-C., Sheedy, J., & Hayes, J. (2006). Effect of letter spacing on legibility, eye movements, and reading speed [Abstract]. *Journal of Vision*, *6*(6):994, 994a, http://journalofvision.org/6/6/994/, doi:10.1167/6.6.994.

Yang, S.N., Tai, Y.C., Laukkanen, H., Sheedy, J. (2009). Effects of ocular transverse chromatic aberration on near foveal letter recognition. Vision Research, 49(23), pp.2881-90.