VISION RESEARCH

An International Journal for Functional Aspects of Vision

Biochemistry & Cell Biology • Molecular Biology & Genetics
Anatomy, Physiology, Pathology & Pharmacology • Optics, Accommodation & Refractive Error
Circuitry & Pathways • Psychophysics • Perception • Attention & Cognition
Computational Vision • Eye Movements & Visuomotor Control

ISSN 0042-6989 | Volume 47 | Number 19 | September 2007

Available online at www.sciencedirect.com

ScienceDirect

Vision Research 47 (2007) 2499–2505

Vision Research

www.elsevier.com/locate/visres

Letter case and text legibility in normal and low vision

Aries Arditi *, Jianna Cho

Arlene R. Gordon Research Institute, Lighthouse International, 111 East 59th Street, New York, NY 10022, USA

Received 7 August 2006; received in revised form 14 June 2007

Abstract

It is thought by cognitive scientists and typographers alike, that lower-case text is more legible than upper-case. Yet lower-case letters are, on average, smaller in height and width than upper-case characters, which suggests an upper-case advantage. Using a single unaltered font and all upper-, all lower-, and mixed-case text, we assessed size thresholds for words and random strings, and reading speeds for text with normal and visually impaired participants. Lower-case thresholds were roughly 0.1 log unit higher than upper. Reading speeds were higher for upper- than for mixed-case text at sizes twice acuity size; at larger sizes, the upper-case advantage disappeared. Results suggest that upper-case is more legible than the other case styles, especially for visually-impaired readers, because smaller letter sizes can be used than with the other case styles, with no diminution of legibility.

Keywords: Low vision; Legibility; Typography; Visual acuity; Reading speed

1. Introduction

There is conventional wisdom, supported by some evidence and logic within the fields of typography and cognitive science, that asserts that text set in mixed upper- and lower-case is more legible than all upper-case (all capital letters). Typographers generally point to the fact that word shape is more distinctive with mixed- and lower-case than it is with all upper-case, a virtue that results from the fact that all upper-case characters are the same height and have no ascenders and descenders, whereas lower-case characters, which have both ascenders and descenders vary in both height and average position, arguably making words constructed with them more distinctive due to more variation in the height of word contours (see Fig. 1).

Miles Tinker, an authority on legibility and typography said "Lower-case letters have more 'character' in terms of variation in shape and the contrasting of ascenders and descenders with short letters. This leads to characteristic word forms that are much easier to read than words in all capitals" (Tinker, 1963; p. 34). Tinker found that while upper-case text was perceived at a greater distance, it had a 'retarding effect' on reading speed, especially for long intervals of reading, and was preferred by only 10% of readers, compared with 90% for lower-case text (Tinker, 1932; Tinker & Patterson, 1929).

The evidence from cognitive science comes from tachistoscopic experiments that suggest that letter identification follows word identification rather than preceding it. Cattell (1886) early on showed that with tachistiscopic presentation, words are recognized more accurately than letters, a phenomenon that in various guises and variations, has come to be known as the "word superiority effect." While there are alternate explanations of this and related effects, such as the greater ease with which letters are recognized within words than in isolation, it has been taken as evidence for a dominant role of word shape in word recognition, relative to letter recognition. Since lower-case words appear to have more distinct shapes than upper-case, there is the common belief that the word superiority effect is responsible for what is assumed to be the greater legibility of lower-case text.

There is a very sensible competing idea, however: that all upper-case text should be more legible since the letters are in general larger than in lower-case text. Enlarging nearly

* Corresponding author. Fax: +1 212 751 9667.
E-mail address: aarditi@lighthouse.org (A. Arditi).

0042-6989/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2007.06.010

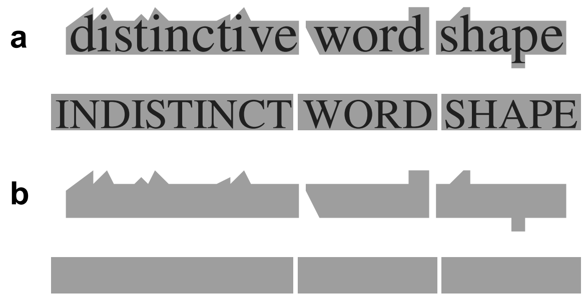

Fig. 1. Shapes of words in (a), outlined in (b), are more distinctive in mixed-case and lower-case than all upper-case text.

any small object makes it more visible, of course, and indeed nearly all optical vision aids rely on the enhanced visibility of magnified objects to achieve better visibility. Text set in visually small sizes in general, and low vision reading in particular, might be expected to benefit from the larger letter sizes of upper-case letters.

The issue of size complicates matters considerably in studies of legibiliity, however, since there does not seem to be any method for characterizing letter size that properly accounts for both upper- and lower-case letters. One may equate upper-case and lower-case character size by equating cap height (the height of a capital letter) to the x-height of lower-case letters. This generally results in findings of lower-case being more legibile than upper-case (Smith, Lott, & Cronnell, 1969). Since 12 of the 26 lower-case letters have ascenders and descenders that extend well above the x-height or below the baseline, respectively, this method of specifying letter size gives an unfair size advantage to lower-case letters. Characterizing letter size by overall font size (conventionally equal to cap height plus descent, at least for computer fonts), similarly, gives a size advantage to upper-case letters since 14 of the lower-case letters (those without ascenders or descenders) are smaller than virtually all the upper-case letters, in both breadth and height.

In this paper, we adopt the latter convention, of specifying letter size by font size, i.e. by the sum of cap height plus descent, which is usually specified in points. We do so because font size is usually specified in this fashion in typography and graphic design, without distinguishing between upper- and lower-case character size. This decision will allow us to make very practical conclusions that can be applied by any graphic designer. In the discussion, of course, we will consider the inherent size difference between letter cases, and weigh the potential advantage of enhanced word shape information in lower-case words against the letter size advantage of words set in upper-case.

2. Methods

We assessed relative legibility of different case conditions using three different criteria for legibility:

1. Size thresholds (visual acuity) for letter identification, measured with 5-letter, strings presented on a video monitor, using an up–down staircase (Levitt, 1971) with 0.05 log unit size steps. Size (or, inversely, distance) thresholds are probably the most common method for assessing text legibility (Tinker, 1963), and are widely used in applied settings such as highway signage, with lower size thresholds indicating higher legibility. We used two kinds of stimuli: random strings of all lower-, all-upper, and randomly selected case and 5-letter words, all upper- or all lower-case, randomly selected from the 2110 most frequent 5-letter words in English (Francis & Kucera, 1982).
2. Reading speeds using rapid serial visual presentation (RSVP). Higher legibility, by this criterion, allows faster reading. We measured reading speed using RSVP with small (two times acuity size) and large letters (roughly 10 times acuity size), using both all upper-case and conventional mixed-case text from an expanded MNREAD (Legge, Ross, Luebker, & LaMay, 1989) corpus. Reading speed is a less common measure of legibility but it is perhaps more representative of ordinary reading than is size threshold. And because RSVP can support extremely high rates of reading (Rubin & Turano, 1992), it has the potential to be more sensitive to subtle differences in legibility. RSVP reading was tested with individual sentences, whose speed was varied to determine the speed that supported a 50% correct (of words) reading rate.
3. Reading speeds using continuous reading of text passages taken from standardized tests (9th grade level). We included this condition to address possible differences between reading speeds with RSVP with those more commonly observed with continuous reading.

We also characterized participants' degree of vision loss with by assessing visual acuity with a transilluminated Lighthouse/ETDRS distance acuity chart. These measurements were also used in the computation of acuity reserve (see below).

2.1. Stimuli

2.1.1. Size thresholds: (Experiment 1)

In this experiment, random 5-letter strings were presented centered on a SONY Multiscan 520GS monitor, as black (3.6 cd/m^2) letters on a white (129 cd/m^2) background. Normally-sighted subjects viewed the screen optically folded through a front-surface mirror at an optical distance of 788.4 cm, so that letters were at least 100 pixels in height (from the top of an upper-case letter to bottom of the descent), or equivalently, for these fonts, 66.66 pixels in cap height. For these subjects, the letters were rendered in reverse on the screen to compensate for the mirror reversal. Subjects with low vision viewed the screen directly (i.e. with no mirror) at a viewing distance of 100 cm. Participants were seated comfortably in a chair, with head position fixed with a head and chin rest.

The random letter strings were constructed by sampling (with replacement) from the 26 letters of the English alphabet, and for the random case condition, then selecting the upper- or lower-case version of the letter randomly with probability 0.5.

2.1.2. RSVP reading (Experiment 2)

We used custom software to present each word of a sentence centered vertically and horizontally on the computer monitor, for a constant time interval. Text was black on white, as with the size thresholds. The participant read aloud each sentence as it was presented, prior to presentation of the next sentence.

2.1.3. Continuous reading (Experiment 3)

Four text passages of ninth grade-level reading difficulty, and approximately 400 words in length, were used. The subject read the text aloud continuously, while the experimenter timed the reading of the entire passage and recorded errors.

2.1.4. Font

We used TrueType Arial as the display font for the entire study. Arial was selected because it is found on most computers used for desktop publishing today, and because it has a large x-height, making it relatively less likely to produce legibility differences based on differences in relative size of upper and lower-case letters. Font point size for the reading speed measurements (Experiments 2 and 3) was set to an acuity reserve (Whittaker & Lovie-Kitchin, 1993) of 2, such that the x-height of the lower-case letters would subtend twice the visual angle of the letters of the visual acuity chart. Additional measurements (Fig. 6) were made with acuity reserve of 10 (10 times the size of the chart letters).

2.2. Participants

Normally-sighted participants were Lighthouse research staff (two participants), or recruited from the Lighthouse International Volunteer Service (two participants). Low vision participants were recruited through the Lighthouse Low Vision Service, and were identified for potential participation in the study by search of the Lighthouse Consumer Information System. Participant details are shown in Table 1, which also shows the specific experimental conditions each participant was tested in. All low vision subjects had clear ocular media and their visual acuity loss was due to macular dsyfunction. All participants, both normal and low vision, were naïve to the purposes of the experiment.

3. Experiment 1: Size thresholds

3.1. Procedure

Size thresholds were measured using a staircase method (Levitt, 1971) in which correct identification of at least 4 of 5 letters (in correct order) was required for a decrease in letter size on the subsequent trial, while no more than three letters correct elicited a size increase on the subsequent trial. For a 26-letter stimulus set size, this procedure converges on the 68.6 percent correct point on the psychometric function. Subjects were required to give 5 letter responses to all trials, and were encouraged to guess if they reported difficulty. On trials in which the size changed, the magnitude of the change was 0.05 log unit, half the size change from line to line on state-of-the-art visual acuity charts. Data prior to the 2nd reversal of each staircase were discarded, in order to concentrate the data used in the analysis close to the threshold.

For each subject, data were collected for random letter strings first, and words subsequently. Each case condition was run four times, in an order that did not favor any condition with respect to practice (see Table 2). Each run terminated after 15 staircase reversals. Since each condition was run four times, each condition's threshold was estimated from 52 staircase reversals (13 reversals from each run).

Table 2
Order of condition blocks within runs for random letters (nonense) and word identification conditions in Experiment 1

Run	Random letters	Words
1	All cap, all low, mixed	All cap, all low
2	Mixed, all low, all cap	All low, all cap
3	Mixed all low, all cap	All low, all cap
4	All cap, all low, mixed	All cap, all low

All responses were given verbally by the subject; the experimenter typed the responses into the computer, which then presented the next 5-letter string whose size was contingent on the subject's performance. Letters were presented continuously until subjects responded. Subjects were able to change their responses if they did so prior to the experimenter's finalizing the response to that line. This procedure results in a negligible lapse, or extraneous noise rate (Arditi, 2006).

3.2. Results and discussion

Log size thresholds (in arc min of visual angle) are shown as a function of letter case for three normally-sighted participants in Fig. 2 and the four participants with low vision in Fig. 3 The thresholds for each subject are geometric means of all the staircase levels visited (after the second reversal of each run); the number of measurements on which the thresholds were based ranged from 73 to 116. Standard errors (s.e.'s) about these means (which reflect accuracy of values in terms of proportion, rather than magnitude) were small; the maximum ratio of s.e. to threshold over all participants and all stimulus conditions was 0.041.

Repeated measures linear mixed effects modeling (Pinheiro & Bates, 2000) revealed that for both normally-sighted and low vision participants (whose size thresholds differed, $F_{1,6} = 58.82$, $p < 0.0003$), thresholds are lower for words than random letter strings ($F_{2,25} = 344.23$, $p < 0.0001$), evidencing a word superiority effect. In addition, thresholds for all caps conditions were lower than those for lower-case ($F_{2,25} = 12.19$, $p = 0.0002$) or mixed-case conditions. The mixed-case condition fell intermediate to the UPPER and lower-case conditions, for normally-sighted participants

Table 1
Characteristics and experimental conditions run, for the 9 participants of the study

Subject	Age	Diagnosis	Log MAR	Snellen equivalent	Experimental conditions run
CC	28	Normal vision	0.0	20/20	Size thresholds, reading speed (a.r. = 10)
KB	30	Normal vision	−0.2	20/13	Reading speed (a.r. = 2)
RA	20	Normal vision	0.0	20/20	Size thresholds, reading speed (a.r. = 2)
RL	34	Normal vision	−0.1	20/16	Reading speed (a.r. = 2)
LH	72	Normal vision	0.2	20/32	Size thresholds, reading speed (a.r. = 10)
IR	72	Diab. retinopathy	0.9	20/159	Size thresholds, reading speed (a.r. = 2)
LG	83	ARM	1.2	20/317	Size thresholds, reading speed (a.r. = 2)
MG	76	ARM	1.1	20/252	Size thresholds, reading speed (a.r. = 2)
SM	77	ARM	0.9	20/159	Size thresholds, reading speed (a.r. = 2)

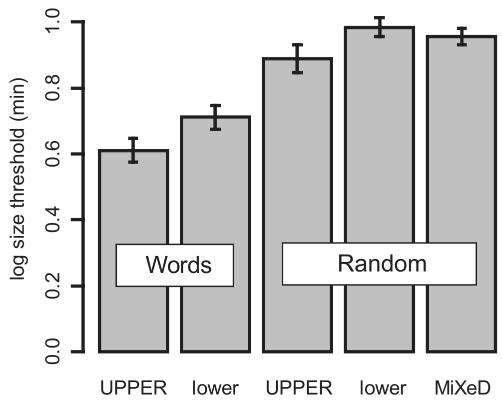

Fig. 2. Means of normally-sighted participants' ($n = 3$) log size thresholds (in arc min visual angle) for identification of words and random strings as a function of letter case condition. Error bars indicate ±1 s.e.m.

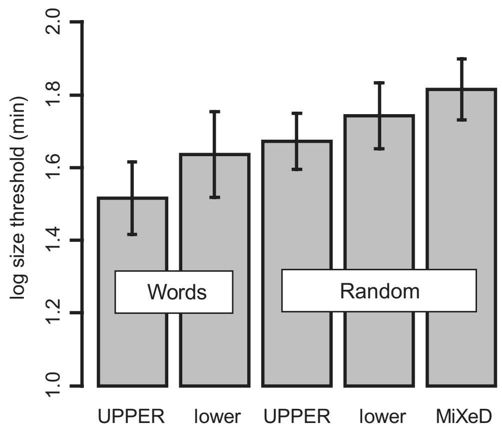

Fig. 3. Means of log size thresholds of participants with low vision ($n = 4$) as in Fig. 2.

but not those with low vision, whose thresholds were highest in the mixed-case condition.

One might expect that the mixed-case condition would fall intermediate to the upper- and lower-case conditions, as they did for the normally-sighted participants, simply because on average, the stimuli were of intermediate size. However, the task is also more difficult than either the upper- or lower-case conditions because the *a priori* probability of a correct guess is reduced from 1/26 to 1/52. In the present experiment, note also that the variability in the threshold estimates was considerably larger for the low vision participants than the normally-sighted participants, making our finding of the highest thresholds in the mixed-case condition for low vision participants less certain.

4. Experiment 2: Reading speed

In this experiment, we assessed the affect of letter case on reading speed using both the RSVP reading technique and continuous text presentation. Continuous presentation was of 'pages' composed of lines of text displayed on a monitor. This kind of reading is not unlike reading from a book or periodical, except that viewing distance and display luminance are better controlled.

Our rationale for using RSVP, in which words are presented one at a time in the center of the computer monitor, was that since it allows reading at higher speeds than with continuous verbal reading (Rubin & Turano, 1992, 1994), especially for normally-sighted readers (Rubin & Turano, 1992). RSVP, then, might plausibly be more sensitive to subtle differences in legibility. In order to further enhance this sensitivity, we used sentences from an expanded MNREAD corpus. These sentences are by design 56 characters long (including interior spaces) with roughly comparable comprehensibility (See Mansfield, Ahn, Legge, & Luebker, 1993 for details). Since the sentences are very short, readers can store most or all of each sentence in short-term memory, and report the sentence verbally without needing to maintain a high rate of verbal output, which might otherwise limit speeds.

4.1. Procedure

4.1.1. RSVP

Word presentation rate, which was controlled by a desktop computer, was varied only between sentences, by an amount that was contingent on reading error rate. Because we had a limited number of 56-character MNREAD sentences (357), and wanted to obtain error rates for a range of presentation rates. Subjects were given practice on 60-character MNREAD sentences prior to testing. The experimenter determined informally during the practice phase the speed region in which the subject began to make errors, by increasing speed by 20% if no errors were made, and decreasing speed if errors were made. Once data collection began, the speed increments and decrements were reduced to 10%, and data collection proceeded in staircase fashion, such that if no errors were made, the speed was increased (by the experimenter); if no words in the sentence were correctly identified, the speed was reduced. Our goal was thus to obtain nonzero error rates for several presentation speeds, sampling a range of the sloping portion of the psychometric function. We obtained estimates of between 5 and 10 speeds for each of the two case conditions (upper and mixed-case), for each subject. Error rates (in characters per 56-character sentence) were then fit by fit by probit (Finney, 1971), to a cumulative Gaussian. Maximum reading speed was taken to be the speed in words per minute, at which 50% errors were made. Following the method of Carver (1976), speeds in words per minute were computed by assuming that each sentence was composed of 9.33 standard length words (each 6 letters in length) and dividing by the exposure time for the sentences.

It might be argued that the cognitive load of remembering the text strings might differentially affect younger and older subjects; however, we cannot test this since all of our younger subjects were normally-sighted and only one older subject had normal vision. But the comparison of interest here is letter case within normally-sighted and low vision groups, who obviously differ significantly in overall reading speed.

4.1.2. Continuous text

In the continuous reading speed measurements, simple text passages of ninth grade reading level and roughly 400 words in length, were presented on the screen with text wrapped and no hyphenation, one screenful at a time, with subsequent screens elicited by the experimenter, until the passage read was completed. The experimenter recorded errors, and the reading speed recorded was the number of words correctly read divided by the reading time in minutes.

Viewing distances and font sizes for both the RSVP and continuous reading measurements, were chosen to approximate an acuity reserve of about 2 for all participants, which is close to the maximum reserve our low vision observers had available to them due to their relatively poor reading acuity (Lovie-Kitchin & Whittaker, 2000). Subsequently, we assessed reading speed in two normally-sighted readers at much larger letter sizes (see Section 4.2 below).

4.2. Results and discussion

Fifty percent correct RSVP speed thresholds and error-corrected continuous text reading speeds for the three normally-sighted participants, are shown in Fig. 4. Data from the same conditions are shown for the four low-vision participants in Fig. 5. For both of measures, and both types of subjects, upper-case text produced faster reading speeds ($F_{1,20} = 5.530$, $p = 0.029$), again supporting the idea that all upper-case text is more legible than mixed-case text.

Note that while our low vision readers were reading with roughly the same acuity reserve as they had when reading under non-experimental conditions, the same was not true

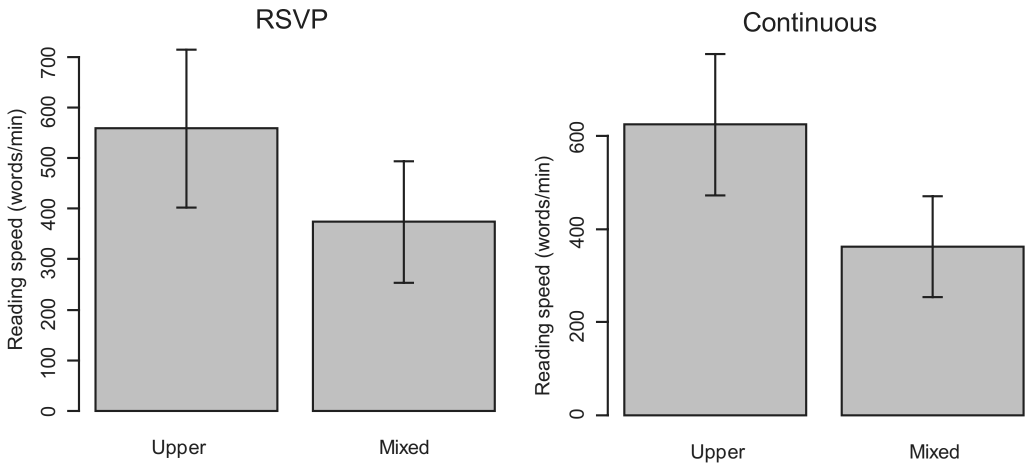

Fig. 4. Mean reading speeds for MNREAD sentences presented by RSVP (left) and continuous text (right) for the two case conditions for normally-sighted readers (acuity reserve = 2; $n = 3$).

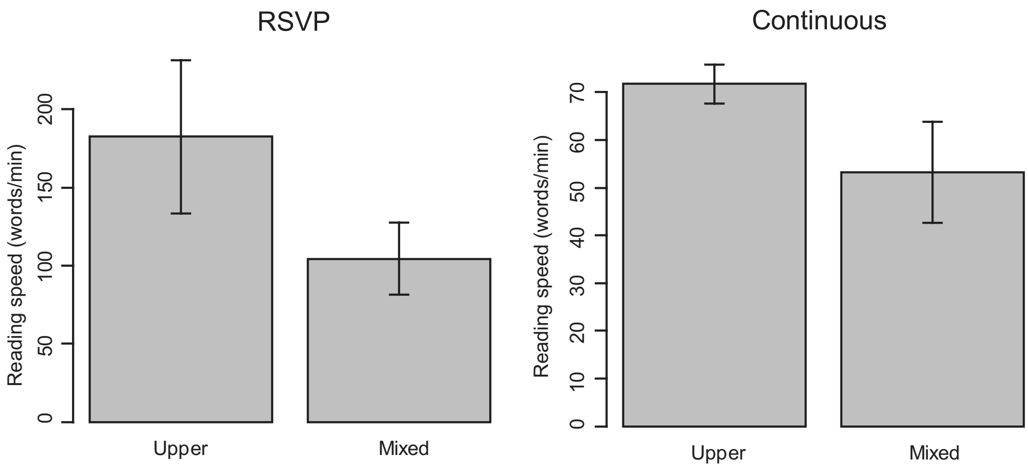

Fig. 5. Average reading speeds for MNREAD sentences presented by RSVP (left) and continuous text (right) for the two case conditions for low vision readers (acuity reserve = 2; $n = 4$).

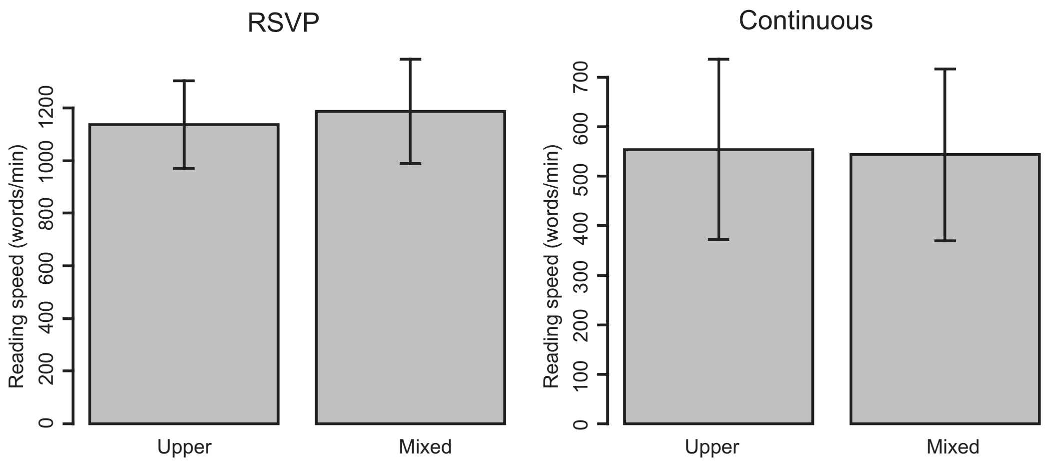

Fig. 6. Average reading speeds for MNREAD sentences presented by RSVP (left) and continuous text (right) for the two case conditions for two normally-sighted readers (acuity reserve = 10; $n = 2$).

for the normally-sighted readers, who typically read at many multiples of their threshold letter size. As an afterthought, we decided to assess reading speed at a visual size corresponding to more typical reading conditions for normally-sighted readers. We chose an acuity reserve of 10 because Whittaker & Lovie-Kitchin (1993) analysis suggested that this resulted in the highest reading speeds for normal readers. The data, shown in Fig. 6, suggest that the upper-case advantage disappears, when text is large enough. This is also consistent with Legge, Pelli, Rubin, & Schleske (1985) finding that there is a relative plateau in the function relating reading speed to print size in the mid- to large-print size range; that is, there is a substantial range of print sizes above the critical print size, at which reading speed changes very gradually. Since size itself has little effect on reading speed in this region, small differences in size associated with letter case will also have little or no measurable effect. Note, however, that we used a font (Arial) with a relatively large x-height, which would tend to minimize differences based on relative size of upper- and lower-case letters.

5. Conclusion

Our finding that size thresholds for upper-case text were lower than those for lower-case text in Experiment 1 are not surprising, and corroborate the findings of Tinker (1963) that at great viewing distance (as simulated by small visual size), upper-case text is more legible, even in a font with a relatively large x-height, which might be expected to minimize upper- and lower-case differences. Other fonts, which typically have smaller x-heights, might be expected to show upper-case text to have even greater relative legibility. Contrary to Tinker's findings, and the conventional wisdom, is the result that upper-case text is more legible in terms of reading speed, for readers with reduced acuity due to visual impairment, and in normally-sighted readers when text is visually small. This result may have practical significance as well; it suggests that, apart from economic considerations of how much space a given sample of text occupies, letter size determines legibility for low vision readers and for those viewing visually small text; and when point size is fixed, upper-case text is simply more legible, albeit less aesthetically appealing, than lower-case.

Acknowledgments

We thank Gordon Legge, Steve Mansfield, Beth O'Brien and Lee Zimmerman for the expanded MNREAD corpus used in this study. The work was partially supported by NIH Grants EY12465, AG14586, and EY015192, and grants from the Hoffritz and Pearle Vision Foundations.

References

Arditi, A. (2006). Lapses in the verbal reporting task. *Vision Research, 46*(8-9), 1327–1330.

Carver, R. P. (1976). Word length, prose difficulty, and reading rate. *Journal of Reading Behavior, 8*, 193–203.

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

Finney, D. J. (1971). *Probit analysis* (3rd ed.). Cambridge: Cambridge University Press.

Francis, W. N., & Kucera, H. (1982). *Frequency analysis of English usage lexicon and grammar*. Boston: Houghton-Mifflin.

Legge, G. E., Pelli, D. G., Rubin, G. S., & Schleske, M. M. (1985). Psychophysics of reading I. Normal vision. *Vision Research, 25*, 239–252.

Legge, G. E., Ross, J. A., Luebker, A., & LaMay, J. M. (1989). Psychophysics of reading VIII. The Minnesota low-vision reading test. *Optometry and Vision Science, 66*(12), 843–853.

Levitt, H. (1971). Transformed up–down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49*, 466–477.

Lovie-Kitchin, J., & Whittaker, S. G. (2000). Prescribing near magnification for low vision patients. *Clinical & Experimental Optometry, 82*, 214–224.

Mansfield, J. S., Ahn, S. J., Legge, G. E., & Luebker, A. (1993). A new reading-acuity chart for normal and low vision. *OSA Technical Digest, 3*, 232–235.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-Plus*. New York: Springer Verlag.

Rubin, G. S., & Turano, K. A. (1992). Reading without saccadic eye movements. *Vision Research, 32*(5), 895–902.

Rubin, G. S., & Turano, K. (1994). Low vision reading with sequential word presentation. *Vision Research, 34*(13), 1723–1733.

Smith, F., Lott, D., & Cronnell, B. (1969). The effect of type size and case alteration on word identification. *American Journal of psychology, 82*, 248–253.

Tinker, M. (1963). *Legibility of print*. Ames, lowa: Lowa State University Press.

Tinker, M. A. (1932). The influence of form of type on the perception of words. *Journal of Applied Psychology, 16*(4), 167–174.

Tinker, M. A., & Patterson, D. G. (1929). Influence of type form on speed of reading. *Journal of Applied Psychology, 13*(6), 205–219.

Whittaker, S. G., & Lovie-Kitchin, J. (1993). Visual requirements for reading. *Optometry and Vision Science, 70*(1), 54–65.