Title

Does enlarging font size facilitate English word and sentence reading in children as beginning readers?

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 45(45)

Authors
Liao, Weiyan
Hsiao, Janet

Publication Date
2023

Peer reviewed

Does enlarging font size facilitate English word and sentence reading in children as beginning readers?

Weiyan Liao (u3007255@connect.hku.hk)
Department of Psychology, University of Hong Kong
Pokfulam Road, Hong Kong

Janet H. Hsiao (jhsiao@hku.hk)
Department of Psychology, The State Key Laboratory of Brain and Cognitive Sciences, and Institute of Data Science, University of Hong Kong
Pokfulam Road, Hong Kong

Abstract

We examined how enlarging the font size from a regular size that young children were typically exposed to affected their reading performance and eye movement behavior. We showed that making the font size larger than a regular size impaired word pronunciation accuracy. This impairment was associated with non-verbal IQ but not changes in eye movement behavior, suggesting that it may be more related to the development of font-size dependent perceptual representations of words than information extraction strategies. The font size change also decreased children's gaze transition consistency among words during sentence reading, suggesting that it interfered with eye movement planning for implementing and developing visual routines for reading. These results suggested that young children may have initially developed font-size dependent perceptual representations for words and for eye movement planning during sentence reading before forming font-size independent reading behavior as observed in adults through increasing reading experiences with different font sizes.

Keywords: Word reading; sentence reading; font size; eye movements; EMHMM

Introduction

Reading is an essential skill in daily life. Children typically start to learn to read and gradually develop their reading skills and visual routines for reading during the first and second grades at primary school (Kiiveri & Maatta, 2012). Skills acquired during early learning can significantly affect learning at a later stage (Sohrabi, 2019). Thus, ways to facilitate children's reading development has been a central issue to researchers and educators. Some have proposed font size used during learning to read as an important factor for reading development (Tinker, 1968; Wilkins et al. 2009). It was suggested that the font size used in children's books should be between pt 14-18 for Grade 1 children (i.e., .28° to .36° of visual angle per letter under a viewing distance of 60 cm; the ideal reading distance is suggested to be between 38-63 cm, Message, 2014), and pt 12 for Grade 5 children (i.e., .24° of visual angle per letter; Tinker, 1959). Some have argued that enlarging the font size may facilitate children's reading. However, inconsistent results have been obtained in the literature, and this inconsistency may be related to children's age/learning stage. For instance, Wilkins et al. (2009) showed that Grade 3 (English-speaking) children as early learners spent less time to judge whether English sentences were semantically correct when the sentences were in a larger font size (26 pt) than in a smaller font size (22 pt). Similarly, Katzir et al. (2013) showed that Grade 3 children had higher comprehension scores when reading Hebrew texts in a larger font size (20 pt) than in a smaller font size (20% smaller than 20 pt). These results suggested that Grade 3 children had better performance with a larger font size. In contrast, Katzir et al. (2013) showed that older, Grade 5 children had lower comprehension scores when reading texts in a larger font size (13 pt) than in a smaller font size (i.e., 20% smaller than 13 pt), suggesting that enlarging font size may actually impair older children's reading performance. This phenomenon may be related to children's visual routine development for reading. Indeed, adult readers were shown to display more consistent eye fixation behavior during reading than children (e.g., Joseph et al., 2009; Yan et al., 2019), suggesting that children gradually develop an adult-like visual routine during learning to read (see also Hsiao et al., 2022). Once a visual routine is formed for a particular font size, as in the case of older children for the range of font sizes they have been exposed to (12-18 pt), deviation from this regular size may disturb eye movement planning, leading to less consistent eye fixation behavior and consequently impaired performance (Hsiao, 2011; Hsiao & Cheng, 2013).

In contrast, for young children as beginning readers, they may not have formed a consistent visual routine, and thus may have larger flexibility/lower consistency in eye movement behavior when the font size changes during reading. Consistent with this speculation, Scaltritti et al. (2019) showed that Italian children had a larger decrease in fixation duration as a result of font size increase during reading as compared with young adults. Similarly, Hughes and Wilkins (2000) showed that in reading a list of English words, children aged from 5 to 7 had slower reading speed when the font size decreased, whereas children aged from 8 to 11 did not show this effect. These results suggested that beginning readers' eye fixation behavior during reading may be influenced more by changes in font size.

Accordingly, we speculated that enlarging the font size from the regular sizes children are exposed to during reading may interfere with their visual routine development for reading, resulting in decreased eye movement consistency and worse reading performance, and this effect may especially affect beginning readers due to their less developed reading skills and visual routines for reading. Therefore, here we aimed to examine how enlarging the font

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society*.

size from a regular size during reading affected young children's reading performance and eye movement behavior, and whether the font size effect in eye movement behavior was associated with that in reading performance. We hypothesized that enlarging the font size from its regular size would interfere with young children’s visual routine for reading, leading to less consistent eye movements, and thus may affect their reading performance. In addition, we hypothesized that decreased eye movement consistency may be associated with impairment in performance.

To test our hypotheses, we conducted an eye tracking study of English word and sentence reading with primary grade 1 and 2 students who were beginning English readers. We compared children’s reading performance as measured in pronunciation accuracy and eye movement behavior between two font size conditions, a regular font size commonly used in Grade 1 and 2 textbooks, 18 pt (.29° of visual angle under the viewing distance of 60 cm), and a larger font size that was 31% larger than the regular font size (.38°) and beyond the range of the regular font sizes for young children. We measured participants’ eye movement behavior using summary statistics measures that were commonly used in the literature, including number of fixations and average fixation duration. In addition, to examine our hypothesis regarding eye movement consistency, we used Eye Movement analysis with Hidden Markov Models (EMHMM, Chuk et al., 2014), a machine-learning model-based approach for eye movement analysis, to quantify participants’ eye movement consistency during reading using entropy (Hsiao et al., 2020; Hsiao, Chan, et al., 2021; Hsiao et al., 2022. See the Methods for more details). In addition, we aimed to examine whether the font size effect in eye movement behavior was associated with that in reading performance, after controlling for the variance from reading-related cognitive abilities.

Methods

Participants

We recruited 55 Grade 1 and 2 students aged from 6 to 8 (Mean = 7.133, SD = .726) from local primary schools who had been learning English as a second language (female = 23). Children in Hong Kong typically begin to learn English when they start kindergarten at age 3 (Zhou et al., 2014). According to power analyses , a sample size of 34 was needed to acquire a medium effect size (f = .25) in a repeated-measure ANCOVA (on reading performance; β = .20; α = .05) with 1 group and 2 levels of measurement (regular vs. larger font size), and a sample size of 55 was needed to acquire a medium effect size (f^2 = .15) in a multiple linear regression analysis (on font size effect in reading performance; β = .20; α = .05) with 1 tested predictor (font size effect in eye movement behavior) and 10 control predictors (reading-related cognitive abilities).

Design

The design consisted of a within-participant variable font size (regular vs. enlarged). The dependent variables were participants’ reading performance as measured in percentage of pronunciation accuracy in word and sentence reading tasks, and eye movement behavior including summary statistics measures (including number of fixations and average fixation duration) and eye movement consistency as measured in entropy using EMHMM (see the EMHMM section below for more details). ANCOVA was used with participants’ IQ as assessed using the Raven’s Standard Progressive Matrices (RSPM; Raven, 1958) as a covariate.

We then used a multiple linear regression analysis to examine whether the font size effect in eye movement behavior, if any, was associated with the font size effect in reading performance after controlling for reading-related cognitive abilities. The font size effects in reading performance and eye movement behavior were measured using normalized difference in these variables, defined as (R – E)/(|R| + |E|), where R and E referred to the performance and eye movement behavior under regular and enlarged font size conditions respectively.

Materials & Procedure

Participants first performed English word and sentence reading tasks, followed by a list of cognitive ability tests were found to be related to reading performance (see below for details).

For the word and sentence reading tasks, the regular font size was determined using the recommended font size for children’s books. Accordingly, the font size of a children’s book would be between .28° to .36° of visual angle per letter (Tinker, 1959). In the current study, we measured the regular font size using the lower-case letter “x”, which subtended .29° x .29° of visual angle under a viewing distance of 60 cm. The enlarged font size as measured using “x” subtended .38° x .38° of visual angle.

English word reading task Participants read aloud, without any time limit, 40 English words (20 words for each font size condition) selected from Woodcock-Johnson III Tests of Cognitive Abilities – Test 10: Word Reading Fluency and the EDB English Language Curriculum wordlist (Woodcock et al., 2001). The words used in the two conditions were matched in word frequency, number of syllables, and word length, and were counterbalanced across participants. The words ranged from 3-6 letters (1 to 3 syllables) and were presented one at a time at the center of the screen in an ascending order of difficulty. In each trial, after the accuracy of the pronunciation was recorded by the experimenter, the stimulus disappeared and the next trial started. Pronunciation accuracy and eye movements were recorded.

English sentence reading task Participants read aloud, without any time limit, 40 English sentences (20 sentences for each condition) from Woodcock-Johnson III Tests of Cognitive Abilities – Test 9: Sentence Reading Fluency and the most widely used local English textbooks (Woodcock et al., 2001). The sentences ranged from 3-10 words and were presented in an ascending order of difficulty. The sentences used in the two conditions were matched in length and were counterbalanced across participants. The stimulus disappeared and the next trial started after the accuracy of the pronunciation was recorded by the experimenter. Pronunciation accuracy and eye movements were recorded.

Verbal & visuospatial one-back tasks These tasks measured the ability of working memory (Lau et al., 2010). In each task, there were 66 trials in total, with 22 trials in each block. In the verbal one-back task, participants were presented with a single number at the center of the screen for 1000 ms, followed by a blank screen for 2500 ms. Under a viewing distance of 40 cm, each number subtended a horizontal and vertical visual angle of 2.58° x 3.15°. Participants were asked to judge whether the presented number was the same as the last one presented. In the visuospatial one-back task, participants were presented with a blue square appearing at one of 4 possible locations on the screen for 1000 ms, followed by a blank screen for 2500ms. The four locations were top center, bottom center, and 5.58° to the left or right of the screen center. Under a viewing distance of 40 cm, each square subtended 10.00° x 10.00° of visual angle. Participants were asked to judge whether the stimulus was at the same location as the last one presented. Their accuracy and reaction time (RT) were measured.

Eriksen flanker task The task measured visual selective attention (Ridderinkhof et al., 1999). Participants were presented with five arrows arranged horizontally at the center of the screen, and were asked to judge the direction of the center arrow. Under a viewing distance of 40 cm, the flanking arrows subtended 1.86° x 1.86° of visual angle, and the target arrow subtended 1.29° x 1.29° of visual angle. In the congruent condition, the flanking arrow direction was congruent with the target arrow. In the incongruent condition, the flanking arrow direction was incongruent with the target arrow. Flanker effect in RT was measured as $(I – C)/(I + C)$, where I and C referred to the performance in incongruent trials and congruent trials respectively.

Tower of London task The task measured participants' planning ability (Phillips et al., 2001). There were 12 trials. Participants were presented with a goal board on the left and a moving board on the right of the screen. There were three sticks of increasing lengths from left to right that could hold one, two, and three balls respectively and three balls of different colors on each board. Participants needed to move the balls on the moving board to match the goal with as few steps as possible within two minutes. Average number of extra steps, planning time and executive time were recorded.

Raven's Standard Progressive Matrices The Raven's Standard Progressive Matrices (RSPM; Raven, 1958) Set A to C were used to measure participants' non-verbal IQ (Education Bureau, 1986). There were 12 trials in each set and the trials were presented in an ascending order of difficulty. In each trial, participants were presented with a visual geometric design with a missing piece and options to fill in the missing piece. Accuracy was recorded.

Rapid Automatized Naming (RAN) Following the Digit Rapid Naming subtest from the Hong Kong Test of Specific Learning Difficulties in Reading and Writing for Primary School Students (HKT-P(II); Ho et al., 2007), a total of 40 digits (randomly drawn from five digits 2, 4, 6, 7, and 9) were printed on a white A4 paper in a 5 x 8 matrix. Participants named the digits as fast as they could, and a stopwatch was used to record the total naming time. They repeated this procedure once. Average time was calculated.

Apparatus

All stimuli in the English word and sentence reading tasks were presented in black with a white background with a resolution of 1024 x 768 on a 17" LCD monitor. Eye movements were recorded with an Eyelink Portable Duo eye tracker (SR Research Ltd.). A chinrest was used to reduce head movement. Calibration and validation were performed before each block; recalibration took place whenever drift correction error was larger than 1° of visual angle.

EMHMM

In the word reading task, following Liao and Hsiao (2021), we used the EMHMM version where ROIs were discovered in a data-driven fashion (e.g., Chuk et al., 2017; Chuk, Crookes et al., 2017; Chan et al., 2020; An & Hsiao, 2021; Hsiao et al., 2021; Liao, Li et al., 2022; Hsiao et al., 2022; Zheng & Hsiao, 2022). More specifically, we summarized a participant's eye movements in a font size condition using an HMM including person-specific regions of interest (ROIs) and transition probabilities among the ROIs estimated from the participant's data. The Variational Bayesian Expectation Maximization algorithm (Bishop, 2006) was used to determine the optimal number of ROIs for each HMM from a preset range 1 to 6 since the longest words had 6 letters. Each model with a different preset number of ROIs was trained for 300 times, and the model with the highest data log-likelihood was used. The models reflected individual differences in both ROI shape/location choice and gaze transitions among the ROIs.

In the sentence reading task, following previous eye movement studies on reading where ROIs were typically predefined to be on individual words with the assumption that words are the basic functional units of sentence processing (e.g., Perfetti, 1985), we predefined an ROI on each word and summarized a participant's gaze transition probabilities among words using an HMM for each sentence (Chuk et al., 2020; Cho et al., 2022a; 2022b; 2022c; Wang et al., 2023). The model thus reflected individual differences in gaze transitions among words but not ROI shape/location choices.

We quantified participants' eye movement consistency using entropy of the HMMs (e.g., Hsiao, et al., 2020; Hsiao, Chan, et al., 2021; Hsiao et al., 2022). Entropy is a measure of predictability; higher entropy indicates higher randomness/lower consistency (Cover & Thomas, 2006). In addition to overall entropy, to examine the temporal dynamics of the font size effect, we included marginal entropy of the first fixation, conditional entropy of the second fixation given the first, and conditional entropy of the third fixation given the second (Hsiao, Liao et al., 2022).

Results

English word reading task

Repeated-measures ANCOVA revealed a font size effect in word reading accuracy, $F(1, 53) = 4.178$, $p = .047$, $\eta_p^2 = .089$: participants had higher accuracy when reading words in the regular than in the enlarged font size (Figure 1a).

In summary statistics of eye movement measures, there was a font size effect in number of fixations, $F(1, 53) = 7.220$, $p = .010$, $\eta_p^2 = .144$: participants had more fixations when reading words in the regular than the enlarged font size (Figure 1b). No effect was observed in average fixation duration, $F(1, 53) = .652$, $p = .424$.

An example HMM that summarizes a participant's HMM under the regular font size condition is shown in Figure 2. It shows that when reading a word, the participant typically started from looking at the area between the word beginning and the word center (ROI 1, 82%), and then transitioned to the word center (ROI 2, 94%). Then, the participant may stay looking around the word center (ROI 2, 52%) or look towards the word beginning (ROI 3, 45%). When we examined the entropies of participants' HMMs, there was no font size effect in overall entropy, $F(1, 53) = .133$, $p = .717$, marginal entropy of the first fixation, $F(1, 53) = .010$, $p = .921$, conditional entropy of the second fixation given the first, $F(1, 53) = .026$, $p = .873$, or conditional entropy of the third fixation given the second, $F(1, 53) = .279$, $p = .600$.

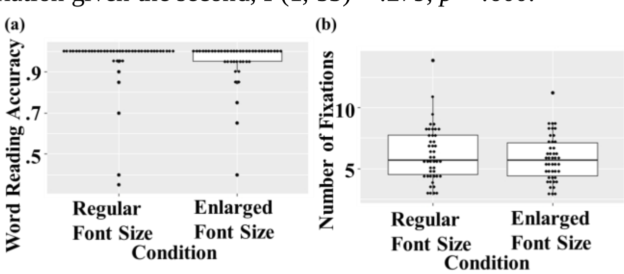
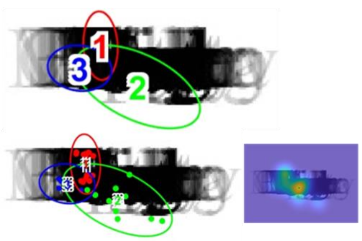


Figure 1: (a) Word reading accuracy. (b) Number of fixations.



	1	2	3
Prior	.82	.09	.09
1	.03	.94	.03
2	.03	.52	.45
3	.76	.12	.12

Figure 2: Example HMM summarizing a participant's eye movements under the regular font size condition. The background image was created by overlaying all word images. Ellipses show ROIs as 2-D Gaussian emissions. Table on the right shows transition probabilities among the ROIs. Priors show the probabilities that a fixation sequence starts from the ellipses. The smaller images show the assignment of actual fixations to different ROIs and the corresponding heatmap.

In the multiple linear regression predicting the font size effect in word reading accuracy, we added all reading-related cognitive abilities in step 1, and the model was not significant, with $F(11, 44) = 1.499$, $p = .204$, $R^2 = .440$. In step 2, we added the font size effect in number of fixations, and the results showed that it did not explain additional variance, $\Delta R^2 = .047$, $p = .191$, and the font size effect in number of fixations was not a significant predictor, $\beta = -.261$, $p = .191$, with $F(12, 43) = 1.582$, $p = .176$, $R^2 = .487$.

In an explorative analysis, we examined whether the font size effect in word reading accuracy was correlated with any cognitive abilities. The results showed that the font size effect in word reading accuracy was correlated with RSPM accuracy, $r(54) = .375$, $p = .014$, suggesting that young children with higher non-verbal IQ had more decrease in word reading accuracy due to enlarged font size (Figure 3). In contrast, the font size effect in number of fixations was not correlated with any cognitive abilities.

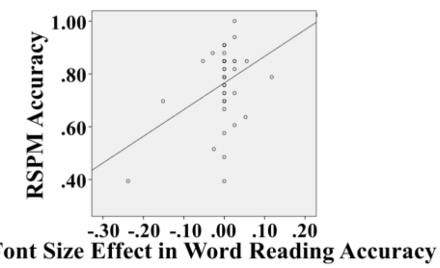


Figure 3: Correlation between font size effect in word reading accuracy and RSPM accuracy.

English sentence reading task

Repeated-measures ANCOVA revealed no font size effect in sentence reading accuracy, $F(1, 53) = .039$, $p = .845$.

In summary statistics of eye movement measures, there was no font size effect in number of fixations, $F(1, 53) = .213$, $p = .647$, or average fixation duration, $F(1, 53) = 1.304$, $p = .260$.

An example HMM that summarizes a participant's gaze transitions among words when reading a sentence is shown in Figure 4. In entropies, there was a font size effect in marginal entropy of the first fixation, $F(1, 53) = 12.036$, $p = .001$, $\eta_p^2 = .236$, conditional entropy of the second fixation given the first, $F(1, 53) = 19.282$, $p < .001$, $\eta_p^2 = .331$, and conditional entropy of the third fixation given the second, $F(1, 53) = 18.909$, $p < .001$, $\eta_p^2 = .327$: participants had lower entropies when reading sentences in the regular than enlarged font size (Figure 5). No font size effect was observed in overall entropy, $F(1, 53) = 1.974$, $p = .168$.

	1	2	3	4
Prior	.19	.57	.16	.08
1	.24	.34	.22	.20
2	.11	.62	.16	.11
3	.20	.31	.29	.20
4	.24	.27	.24	.24

Figure 4: Example HMM summarizing a participant's gaze transitions among words when reading a sentence. Ellipses show ROIs as 2-D Gaussian emissions. Table on the right shows transition probabilities among the ROIs. Priors show the probabilities that a fixation sequence starts from the ellipses. The smaller images show the assignment of actual fixations to different ROIs and the corresponding heatmap.

In an explorative analysis, we found that the font size effect in marginal entropy of the first fixation was correlated with flanker effect in RT, $r(54) = -.371$, $p = .037$, and verbal one-back accuracy, $r(54) = .350$, $p = .027$, suggesting that young children who had better visual selective attention and verbal working memory had more decrease in consistency of the first fixation due to enlarged font size (Figure 5a & b). The font size effect in conditional entropy of the second fixation given the first was not correlated with any cognitive abilities. The font size effect in conditional entropy of the third fixation given the second was correlated with Tower of London executive time, $r(54) = .376$, $p = .022$, suggesting that young children with worse executive planning ability had more decrease in the consistency of the third fixation given the second due to enlarged font size (Figure 5c).

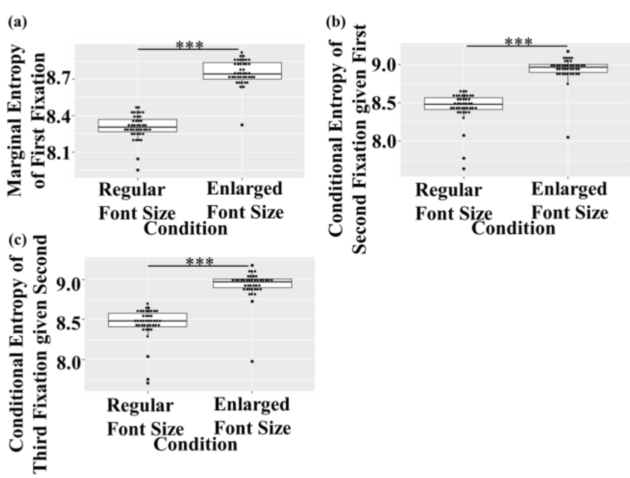


Figure 5: (a) Marginal entropy of the first fixation. (b) Conditional entropy of the second fixation given the first. (c) Conditional entropy of the third fixation given the second (*$p < .05$; **$p < .01$; ***$p < .001$).

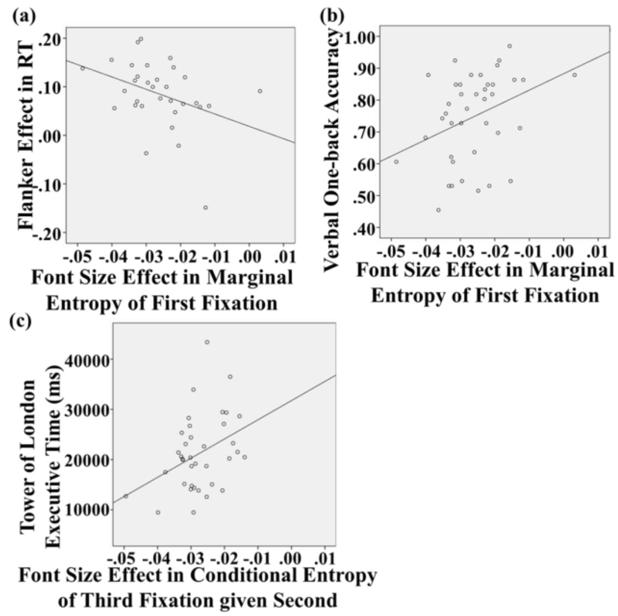


Figure 6: (a) Correlation between font size effect in marginal entropy of the first fixation and flanker effect in RT. (b) Correlation between font size effect in marginal entropy of the first fixation and verbal one-back accuracy. (c) Correlation between font size effect in conditional entropy of the third fixation given the second and Tower of London executive time.

Discussion

Here we examined whether enlarging the font size from a regular size impaired young children's reading performance and decreased their eye movement consistency due to the interference from the font size change to their eye movement planning behavior/visual routine during reading, and whether the font size effect in eye movement behavior was associated with that in reading performance. In word reading, consistent with our hypothesis, young children had decreased pronunciation accuracy when reading words with the enlarged than the regular font size. They also had a decreased number of fixations when reading in the enlarged font. However, these two font size effects were not associated. In contrast to our hypothesis, no font size effect was found in eye movement consistency. In sentence reading, in contrast to word reading, no font size effect was observed in pronunciation accuracy. However, young children had higher entropies/less consistent gaze transition behavior among words when reading sentences in the enlarged than regular font size.

Our results suggested that enlarging the font size from the regular font size did not facilitate young children's reading performance. In contrast, it may impair children's word reading performance. Consistent with our finding, Hughes and Wilkins (2000) showed that children aged 5 to 7 had slower reading speed when the font size decreased from a regular font size for their ages with. Note that some previous studies examining font size effects on children's reading performance did not report whether the font sizes used were commonly used for the participants' ages or not. For example, in Wilkins et al. (2009), both the smaller and larger font sizes used (20 pt vs. 26 pt) were larger than the suggested font sizes for Grade 1 children (14-18 pt). Note also that they focused the examinations on reading comprehension, whereas in our study we examined pronunciation. Future work may examine effects of font size change from a regular size on reading comprehension.

The impaired performance in word pronunciation accuracy due to enlarged font size suggested that in isolated word reading, deviation from a regular font size affected word identification in young children/beginning readers. However, in contrast to our hypothesis, enlarging the font size did not decrease eye movement consistency. Instead, it only decreased the number of fixations, although this effect was not associated with the impairment in performance. Thus, the impairment may be related to a mismatch of the perceived word to children's internal representations of the word that may be font size dependent, rather than interference to the learned visual routine for information extraction. In addition, we found that the font size effect in word reading accuracy was correlated with children's non-verbal IQ as measured in RSPM: the higher the non-verbal IQ, the larger the font size effect. This result suggested that children with higher non-verbal IQ may have developed more font-size dependent perceptual representations for words during learning.

Interestingly, when children pronounced words in a sentence context as in our sentence reading task, their pronunciation accuracy was not affected by font size change. This result may be because pronunciation of words in a sentence context could be facilitated by contextual information from neighboring words and sentence semantics, in contrast to reading out isolated single words. However, consistent with our hypothesis, enlarging the font size from a regular size interfered with their eye movement planning, increasing the entropy (decreasing the consistency) of their gaze transition behavior among words when reading out the sentence. The font size change affected gaze transition consistency particularly among the first few fixations during sentence reading, as the effect was observed in the entropies of the first 3 fixations, but not in overall entropy. Our exploratory analyses showed that the font size effect in the marginal entropy of the first fixation was associated with visual selective attention and verbal working memory: the better the visual selective attention and verbal working memory, the larger the font size effect. In contrast, a larger font size effect in the conditional entropy of the third fixation given the second was associated with poorer executive planning ability as assessed in the Tower of London task. These results suggested different mechanisms underlying the font size effect in entropy at different stages during sentence reading. The first fixation was typically used to locate the beginning of the sentence for reading. Those with better visual selective attention and verbal working memory may have developed more font-size dependent perceptual representations for locating the sentence beginning, resulting in a larger font size effect. Indeed, the neural mechanisms important for attention orienting within the frontal-parietal network have been shown to be involved in selective attention and working memory functions, suggesting a close relationship among working memory, selective attention, and attention orienting (Ku, 2018). In contrast, gaze transition consistency of the third fixation given the second fixation reflected how well children planned the target of the next word given the current context. Those with better executive planning ability may be able to better concentrate on the contextual information for sentence reading and thus were less affected by font size difference. Consistent with this speculation, higher executive planning ability has been shown to be associated with better adaptive reading behavior to achieve the reading goal (e.g., Micai et al., 2021; Kwok & Hsiao, 2022). Together these results suggested that young children learning to read may initially develop font-size dependent perceptual representations of words for pronunciation and for locating where to start reading and where to read next in a sentence. These perceptual representations may gradually become less font-size dependent with increasing reading experience with different font sizes. Consistent with this speculation, font size effects were not typically observed in adult readers in the literature (Beymer et al., 2022; Darroch et al., 2005; but see Liu et al., 2018).

In conclusion, here we showed that making the font size larger than what young children were typically exposed to during English reading did not facilitate their reading performance in pronunciation accuracy. Instead, it may impair their word pronunciation accuracy. This impairment was associated with children's non-verbal IQ but not changes in eye movement behavior due to font size change, suggesting that it may be related to font-size dependent perceptual representations of words instead of information extraction strategy. The font size change also decreased children's gaze transition consistency among words during sentence reading, suggesting that it could interfere with eye movement planning for locating where to start reading and where to read next given the current context. These results suggested that young children may gradually develop font-size dependent perceptual representations for words and for eye movement planning in a sentence before forming font-size independent reading behavior as observed in adults through increasing experiences in reading with different font sizes. Thus, presenting reading materials in a font size that is suitable for the development of font-size independent perceptual representations may better facilitate reading development than simply using a larger font.

Acknowledgements

We are grateful to Research Grant Council of Hong Kong (GRF #17608621 to Hsiao). We thank Karen H. Lau for her help in collecting part of the data.

References

An, J., & Hsiao, J. H. (2021). Modulation of mood on eye movement and face recognition performance. *Emotion*, *21*(3), 617-630.

Arnell, K. M., Joanisse, M. F., Klein, R. M., Busseri, M. A., & Tannock, R. (2009). Decomposing the relation between Rapid Automatized Naming (RAN) and reading ability. *Can. J. Exp. Psychol.*, *63*(3), 173-184.

Beymer, D., Russell, D., & Orton, P. (2022). An eye tracking study of how font size and type influence online reading. *Proceedings of the 22nd British HCI Group Annual Conference on HCI 2008: People and Computers XXII: Culture, Creativity, Interaction*. BCS HCI, Liverpool, United Kingdom.

Bishop, C.M. (2006). *Pattern Recognition and Machine Learning.* New York: Springer.

Chan, F. H. F., Barry, T. J., Chan, A. B., & Hsiao, J. H. (2020). Understanding visual attention to face emotions in social anxiety using hidden Markov models. *Cogn. Emot.*, *34*(8), 1704-1710.

Cho, V., Hsiao, J. H., Chan, A. B., Ngo, H., King, N. M., & Anthonappa, R. (2022a). Understanding children's attention to dental caries through eye-tracking. *Carries Res.*

Cho, V., Hsiao, J. H., Chan, A. B., Ngo, H., King, N., & Anthonappa, R. (2022b). Eye movement analysis of children's attention for midline diastema. *Sci. Rep.*, *12*, 7462.

Cho, V., Hsiao, J. H., Chan, A. B., Ngo, H., King, N. M., & Anthonappa, R. (2022c). Understanding children's attention to traumatic dental injuries using eye-tracking. *Dent. Traumatol.*

Chuk, T., Chan, A. B., & Hsiao, J. H. (2014). Understanding eye movements in face recognition using hidden Markov models. *J. Vision*, *14*(11), 8–8.

Chuk, T., Chan, A. B., & Hsiao, J. H. (2017). Is having similar eye movement patterns during face learning and recognition beneficial for recognition performance? Evidence from hidden Markov modeling. *Vision Res.*, *141*, 204-216.

Chuk, T., Chan, A. B., Shimojo, S., & Hsiao, J. H. (2020). Eye movement analysis with switching hidden Markov models. *Behav. Res. Methods*, *52*(3), 1026-1043.

Chuk, T., Crookes, K., Hayward, W. G., Chan, A. B., & Hsiao, J. H. (2017). Hidden Markov model analysis reveals the advantage of analytic eye movement patterns in face recognition across cultures. *Cognition*, *169*, 102–117.

Darroch, I., Goodman, J., Brewster, S. A., & Gray, P. D. (2005). The effect of age and font size on reading text on handheld computers. *Interact 2005*, *3585*, 253-266.

Hage, D. S., & Stroud, J. B. (1959). Reading proficiency and intelligence scores, verbal and nonverbal. *J. Educ. Res.*, *52*(7), 258-262.

Ho, C. S. H., Chan, D., Chung, K. K. H., Tsang, S. M., Lee, S. H., & Cheng, R. W. Y. (2007). *The Hong Kong Test of Specific Learning Difficulties in Reading and Writing for Primary School Students-[HKT-P(II)]*. Hong Kong: Hong Kong Specific Learning Difficulties Research Team.

Hsiao, J. H. (2011). Visual field differences can emerge purely from perceptual learning: Evidence from modeling Chinese character pronunciation. *Brain & Language*, *119(2)*, 89-98

Hsiao, J. H., An, J., & Chan, A. B. (2020). The role of eye movement consistency in learning to recognise faces: Computational and experimental examinations. *Proceedings of the 42nd Annual Conference of the Cognitive Science Society* (pp. 1072-1078). Cognitive Science Society.

Hsiao, J. H., An, J., Hui, V. K. S., Zheng, Y., & Chan, A. B. (2022). Understanding the role of eye movement consistency in face recognition and autism through integrating deep neural networks and hidden Markov models. *NPJ Sci. Learn.*, *7*, 28.

Hsiao, J. H., An, J., Zheng, Y., & Chan, A. B. (2021). Do portrait artists have enhanced face processing abilities? Evidence from hidden Markov modeling of eye movements. *Cognition*, *211*, 104616.

Hsiao, J. H., & Cheng, L. (2013). The modulation of stimulus structure on visual field asymmetry effects: the case of Chinese character recognition. *Quarterly Journal of Experimental Psychology*, *66*(9), 1739-1755.

Hsiao, J. H., Liao, W., & Tso, R. V. Y. (2022). Impact of mask use on face recognition: An eye tracking study. *Cogn. Res.: Princ. Implic.*, *7*, 32.

Hughes, L. E., & Wilkins, A. J. (2000). Typography in children's reading schemes may be suboptimal: evidence from measures of reading rate. *J. Res. Read.*, *23*(3), 314-324.

Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., & Rayner, K. (2009). Word length and landing position effects during reading in children and adults. *Vision Research*, *49*(16), 2078-2086.

Katzir, T., Hershko, S., & Halamish, V. (2013). The effect of font size on reading comprehension on second and fifth grade children: bigger is not always better. *PLoS One*, *8*(9), e74061.

Kiiveri, K., & Maatta, K. (2012). Children's opinions about learning to read. *Early Child Dev. Care*, *182*(6), 755-769.

Ku, Y. X. (2018). Selective attention on representation in working memory: cognitive and neural mechanisms. *PeerJ*, *6*, e4585.

Kwok, J., & Hsiao, J. H. (2022). The role of reading strategy in reading comprehension performance: An eye-movement study. *Proceedings of the 44th Annual Conference of the Cognitive Science Society* (pp. 1985-1992). Cognitive Science Society.

Lau, E. Y. Y., Eskes, G. A., Morrison, D. L., Rajda, M., & Spurr, K. F. (2010). Executive function in patients with obstructive sleep apnea treated with continuous positive airway pressure. *J. Int. Neuropsychol. Soc.*, *16*(6), 1077-1088.

Liao, W., Chong, W. C., & Hsiao, J. H. (2022). Does word boundary information facilitate Chinese sentence reading among beginning readers? *Proceedings of the 44th Annual Conference of the Cognitive Science Society* (pp. 2014-2021). Cognitive Science Society.

Liao, W., & Hsiao, J. H. (2021). The role of eye movement pattern and global-local information processing abilities in isolated English word reading. *Proceedings of the 43nd Annual Conference of the Cognitive Science Society* (pp. 2253-2259). Cognitive Science Society.

Liao, W., Li, S. T. K., & Hsiao, J. H. (2022). Music reading experience modulates eye movement pattern in English reading but not in Chinese reading. *Sci. Rep.*, *12*, 9144.

Liu, T., Yeh, S. L., & Hsiao, J. H. (2018). Transfer of the left-side bias effect in perceptual expertise: The case of simplified and traditional Chinese character recognition. *PLOS ONE*, *13*(4): e0194405.

Message. (2014, February 22). *Ideal Reading Distances*. https://messagengo.wordpress.com/eyes/ideal-reading-distances/

Micai, M., Vulchanova, M., & Saldana, D. (2021). Reading goals and executive function in autism: an eye-tracking study. *Ausism Res.*, *14*(5), 1007-1024.

Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press.

Phillips, L. H., Wynn, V. E., McPherson, S., & Gilhooly, K. J. (2001). Mental planning and the Tower of London task. *Q. J. Exp. Psychol.-A*, *54*(2), 579-597.

Raven, J. C. (1958). *Standard Progressive Matrices: Sets A, B, C, D & E*. London: HK Lewis.

Ridderinkhof, K. R., Band, G., & Logan, G. (1999). A study of adaptive behavior: effects of age and irrelevant information on the ability to inhibit one's actions. *Acta Psychol*, *101*(1999), 315-337.

Scaltritti, M., Miniukovich, A., Venuti, P., Job, R., De Angeli, A., & Sulpizio, S. (2019). Investigating effects of typographic variables on webpage reading through eye movements. *Sci. Rep.*, *9*, 12711.

Sohrabi, A. (2019). Age of acquisition effect: evidence from single-word reading and neural networks. *Basic Clin. Neurosci.*, *10*(2), 137-146.

Swicker, A. (2022, February 10). When do kids start reading? *Book Riot*. https://bookriot.com/when-do-kids-start-reading/

Tinker, M. A. (1968). Suitable typography for beginners in reading. *Education*, *88*(4), 317–320.

Tinker, M. A. (1959). Print for children's textbooks. *Education*, *80*(1), 37-40.

Wang, L., Hsiao, J. H., Chan, A. B., Cheung, J., Hung, S., & Au, T. K. (2023). On Becoming Socially Anxious: Toddlers' Attention Bias for Fearful Faces. *Developmental Psychology*, 59(2), 353-363.

Wilkins, A., Cleave, R., Grayson, N., & Wilson, L. (2009). Typography for children may be inappropriately designed. *J. Res. Read.*, *32*(4), 402-412.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock– Johnson III Tests of Cognitive Abilities*. Itasca, IL: Riverside.

Yan, M., Pan, J., & Kliegl, R. (2019). Eye movement control in Chinese reading: A cross-sectional study. *Dev. Psychol.*, *55*(11), 2275-2285.

Zheng, Y., Ye, X., & Hsiao, J. H. (2022). Does adding video and subtitles to an audio lesson facilitate its comprehension? *Learn. Instr.*, *77*, 101542.

Zhou, Y. L., McBride-Change, C., Law, A. B. Y., Li, T., Cheung, A. C. Y., Wong, A. M. Y., & Shu, H. (2014). Development of reading-related skills in Chinese and English among Hong Kong Chinese children with and without dyslexia. *J. Exp. Child Psychol.*, *122*, 75-91.