



OPEN

DATA DESCRIPTOR

The Eye Movement Database of Passage Reading in Vertically Written Traditional Mongolian

Yaqian Borogjoon Bao¹✉, Xingshan Li^{2,3} & Victor Kuperman¹

This paper introduces an eye-tracking corpus of passage reading data in the vertical writing system of traditional Mongolian. This corpus extends the Multilingual Eye Movement Corpus (MECO) database and includes data from 66 native readers of traditional Mongolian script reading 12 texts comprising 99 sentences and 2,592 words. This traditional Mongolian MECO corpus aims to address the research gap in reading studies on understudied languages. As one of the very few actively used vertical writing systems, these data offer unique insights into the cognitive and visual processing demands of vertical reading. The paper provides reliability estimates for the data and reports lexical benchmark effects of word frequency and length. Additionally, the corpus provides a valuable opportunity for cross-linguistic comparisons of eye movement data, especially with horizontal writing systems, contributing to a better understanding of how reading direction influences cognitive processing.

Background & Summary

Reading is the process of constructing mental representations from printed text. It encompasses decoding visual information from written words, extracting their meanings, applying syntactic rules to organize these words into coherent grammatical units, and integrating information both across sentences and with external knowledge to develop a unified and coherent mental representation of the content¹. An intriguing aspect that influences both learning to read and reading as a daily practiced activity is the immense natural variation in human languages, scripts, and orthographic principles. Given this diversity, a critical question arises: Which aspects of the reading mechanism are shaped by the properties of different writing systems? Consequently, the primary objective of reading research is to formulate theories that elucidate both the universal and script-specific phenomena of the reading process across various writing systems². These theories aim to capture the common cognitive operations that underlie the interpretation of printed language, regardless of the script^{3,4}.

To achieve this goal, research on reading requires a robust supply of cross-linguistic data. Such data is essential to examine how reading mechanisms operate across different languages and writing systems. In one of the first efforts to meet this need, Siegelman *et al.*⁵ released the first wave of the Multilingual Eye-movement Corpus (MECO), a comprehensive dataset containing eye-tracking data from readers of 13 different languages: Dutch, English, Estonian, Finnish, German, Greek, Hebrew, Italian, Korean, Norwegian, Russian, Spanish, and Turkish. The eye-tracking data were collected using similar methods and apparatus, while participants read the highly comparable content in each language. Thus, MECO data enable standardized comparisons across different written languages. These data offer a valuable resource for exploring which mechanisms of reading are universal across different written languages and which are driven by specific facets of individual languages or writing systems. To give a few examples of utilizing the MECO dataset, researchers demonstrated a remarkable uniformity in the magnitude of benchmark lexical effects on word reading across alphabetic languages. Longer and less frequent words⁶ and words less predictable in their linguistic context⁷ take longer to process, and the size of each of the effects is nearly identical across languages under consideration.

While the MECO dataset covers a variety of alphabetic languages, vertical writing systems such as traditional Chinese, Japanese, Korean, and traditional Mongolian remain relatively underrepresented. Many East Asian scripts, including Chinese characters, Korean hangul, and Japanese kana, can be written either horizontally or vertically. This flexibility arises from their structure, as these scripts are primarily composed of discrete

¹McMaster University, Hamilton, Canada. ²Institute of Psychology, Chinese Academy of Sciences, Beijing, China.

³Department of Psychology, University of Chinese Academy of Sciences, Beijing, China. ✉e-mail: baoy47@mcmaster.ca

logographic or syllabic units, each fitting within a uniform square block, making them adaptable to different writing orientations. Although horizontal writing has become increasingly common in modern times, vertical writing remains frequently used in regions such as Hong Kong, Japan, Macau, Taiwan, Mongolia and Inner Mongolian Autonomous Region of China. Research comparing horizontal and vertical reading has highlighted the influence of reading direction and experience on eye-movement patterns. Osaka and Oda (1991)⁸ found that the perceptual span for vertically written Japanese was 5–6 character spaces, slightly smaller than the 7-character span for horizontally written Japanese reported in their later study⁹. Similarly, studies on vertical word identification revealed that 4–5 character spaces can be processed per fixation in vertical lists, compared to 10 in horizontal lists, with vertical text requiring longer fixations¹⁰. Yan *et al.*¹¹ showed that while reading speeds were similar for horizontal and vertical traditional Chinese reading, eye-movement patterns differed significantly. Vertical reading involved longer fixations and shorter saccades but demonstrated better saccade targeting accuracy. Yan *et al.*¹² further explored the role of reading experience in shaping perceptual spans for horizontal and vertical orientations in Traditional Chinese. They revealed that the perceptual span for vertical reading (3 characters below fixation and 1 above) was smaller than that for horizontal reading (4 characters to the right and 1 to the left). Importantly, participants with greater experience in vertical reading exhibited a larger perceptual span for vertical text, emphasizing the effect of direction-specific reading experience in adapting eye-movement strategies to the demands of different text orientations. These findings underscore the effect of reading direction on eye-movement patterns and cognitive processes and highlight the importance of including more vertical scripts in comparative reading research.

This study advances the agenda of supplying the reading research with behavioral datasets from typologically diverse writing systems. Specifically, this study collected eye-movement data from traditional Mongolian reading, which has a unique vertical reading direction. Mongolian is an Altaic language spoken mainly in Mongolia and the Inner Mongolian Autonomous Region of China. Mongolian is written in two scripts: Cyrillic and the traditional Mongolian script. The traditional Mongolian script, used by Mongolians since the 13th century, remains the primary writing system in Inner Mongolia. In contrast, the Cyrillic script became the official writing system in Mongolia in the 1940s.

The traditional Mongolian script is not the only vertical writing system in the world, see above. However, it differs significantly from other vertical writing systems. First, the traditional Mongolian script is unique in that it is written vertically from top to bottom, but the lines progress from left to right. This is the opposite of other vertical writing systems, such as traditional Chinese and Japanese, where the lines progress from right to left. Second, traditional Mongolian is written with spaces between words, but the letters within each word are inseparably connected, with letters cursively joining to form syllables, and syllables having initial, medial, and final forms that adapt to different positions within a word. In contrast, other logographic scripts like Chinese, the syllabic script like Japanese kana as well as most alphabetic writing systems, consist of distinct, separate letters, characters or blocks that are not connected. Moreover, traditional (and simplified) Chinese uses no spaces to demarcate words. Third, the traditional Mongolian script is strictly vertical in its native form. It can only be presented horizontally by rotating the words 90 degrees counterclockwise, which occurs in very restricted conditions, such as temporarily appearing horizontally when typing on a phone before being displayed vertically. In contrast, other vertical writing systems, such as Chinese and Japanese, can easily be adapted to horizontal writing without rotation, with text naturally flowing from left to right in modern usage. These distinct properties of the traditional Mongolian script may affect eye movement patterns, cognitive load, and overall reading fluency in ways that differ from those seen in other vertical or horizontal scripts. For example, the strict vertical reading direction might create distinct visual and cognitive patterns for readers, while the cursive nature of traditional Mongolian may impose unique demands on visual processing, as readers must recognize connected letterforms that vary depending on their position in a word. This makes traditional Mongolian a particularly interesting case for reading research.

In addition to these differences that set the traditional Mongolian script apart from other vertical (and even more so horizontal) writing systems, traditional Mongolian is one of the most understudied written languages in reading research. Only few studies have used eye-tracking techniques to investigate the perceptual span in traditional Mongolian reading. For instance, Borjigin *et al.*¹³ conducted an eye-tracking experiment on readers of traditional Mongolian to measure the size of their perceptual span during reading. They found that the perceptual span in traditional Mongolian extended one syllable above the fixation and three syllables below the fixation, providing consistent evidence that the asymmetry toward the reading direction of the perceptual span is universal across different languages. Furthermore, Su *et al.*¹⁴ conducted a gaze-contingent eye-tracking experiment with traditional Mongolian readers and demonstrated that perceptual mechanisms during reading are highly flexible, effectively adjusting to align with changes in reading direction. These findings notwithstanding, many other aspects of traditional Mongolian reading behavior remain largely unexplored. In this context, developing a substantial eye-tracking corpus for traditional Mongolian reading is crucial to addressing this research gap and providing a valuable dataset for cross-linguistic comparisons.

In this study, we are publishing MECO-Traditional Mongolian, an eye-tracking corpus that includes data from 66 readers of vertically oriented traditional Mongolian reading 12 texts, comprising 99 sentences and 2,592 words. To enable cross-linguistic comparisons, the experimental procedures and text stimuli creation strictly followed the guidelines and procedures of the MECO project⁵.

Methods

Participants. 68 participants (53 females, age = 20.82 ± 2.15 years, range from 18 to 27) from Inner Mongolia Normal University in Hohhot, China, participated in this eye-tracking experiment. Before the experimental session, each participant provided written consent in accordance with the guidelines and regulations approved by the ethics committee (20230314) of the Institute of Psychology, Chinese Academy of Sciences, and reviewed by the

Text #	Topic	# sent	# word	Similarity
1*	Janus	10	218	0.86
2	Shaka	9	188	0.75
3*	Doping	9	263	0.95
4	Thylacine	10	207	0.52
5	World Environment Day	5	214	0.77
6	Monocle	10	201	0.89
7*	Wine tasting	8	245	0.87
8	Orange juice	6	157	0.75
9	Beekeeping	7	212	0.67
10	National flag	9	228	0.93
11*	International Union for Conservation of Nature	8	249	0.90
12*	Vehicle registration plate	8	210	0.82

Table 1. Detailed information on each text. *Note:* An asterisk (*) indicates translated matched texts, while the absence of an asterisk indicates non-matched texts. # sent and # word represent the number of sentences and words in each text, respectively. Similarity refers to the semantic similarity with the original English texts, ranging from −1 (opposite meanings) to 1 (perfect similarity), with 0 indicating orthogonal meanings.

ethics committee of Inner Mongolia Normal University, where the study was conducted. The Ethics Committee reviewed and approved the study protocol, including participant recruitment, data collection, and the use of the data for publication and sharing. All participants are native speakers of Mongolian and have attended Mongolian schools from kindergarten through university, where they studied all subjects in Mongolian. This extensive education has made them proficient in reading vertically written traditional Mongolian. Their self-ratings of Mongolian proficiency (on a scale of 0–10) are as follows: speaking (mean = 9.03 ± 0.69), oral comprehension (mean = 9.10 ± 0.76), and reading (mean = 9.16 ± 0.64). In addition to their native language, all participants are proficient in Chinese, which they began learning in the third year of elementary school. Each participant has normal or corrected-to-normal eyesight. Basic demographic information is available on the project’s OSF page (see Data records below).

Materials. As with all other partner sites participating in the MECO project, this experiment used a set of 12 texts in the first language of the readers, i.e., traditional Mongolian. In the MECO project, the texts were initially created in English as Wikipedia-style encyclopedic entries. Five of the 12 texts were selected as “matched texts” and directly translated from English into other languages. The remaining 7 texts were chosen as “non-matched texts,” and were created by each language site following guidelines to maintain the same topic, prosaic genre, similar length, and comparable level of difficulty as in the original English texts (see details in Siegelman *et al.*⁵). Following these guidelines, this study translated 5 texts directly from the English texts used in the MECO project to traditional Mongolian (matched texts). The remaining 7 non-matched texts were originally crafted in traditional Mongolian but adhered to the same topics, prosaic genre, length, and difficulty as their English counterparts. Detailed information on each resulting text stimulus is shown in Table 1.

This study employed the same methods outlined in the supplementary materials of Siegelman *et al.*⁵ to assess the quality of the traditional Mongolian translation and to evaluate the readability and complexity of traditional Mongolian texts compared to other languages from the MECO project. First, we measured the cosine semantic similarity between the English back-translation of traditional Mongolian texts and the original English texts using pretrained latent semantic analysis vectors (LSA, TASA corpus, downloaded from https://sites.google.com/site/fritzgntr/software-resources/semantic_spaces), see more details about LSA in Günther *et al.*¹⁵. The cosine semantic similarity of each text is shown in Table 1. Overall, the semantic content of the traditional Mongolian texts is highly similar to the original English texts (Mean cosine = 0.81, SD = 0.12). Matched texts were numerically more similar to the English originals than non-matched texts ($M_{\text{matched}} = 0.88$, $SD_{\text{matched}} = 0.04$; $M_{\text{non-matched}} = 0.76$, $SD_{\text{non-matched}} = 0.13$; Welch- $t_{(7.99)} = -2.24$, $p = 0.06$).

Second, we used the Cohmetrix 3.0 web tool (www.cohmetrix.com)^{16,17} to obtain scores for text readability and complexity. Text readability was assessed using the Flesch-Kincaid readability and L2 readability scores, while text complexity was quantified by the metrics of narrativity, simplicity, concreteness, cohesion, deep cohesion, verb cohesion, connectivity, and temporality. We then combined the scores of traditional Mongolian texts with those from other 13 languages in the first wave of the MECO project to determine if there were any significant cross-linguistic differences in text readability and complexity. We set the two readability measures and the eight complexity measures as dependent variables, with the original language of the text (the 14-level categorical predictor) as the independent variable in a series of regression models. *P*-values of the variance *F*-test were corrected using the Bonferroni method to avoid family-wise inflation of Type I error¹⁸.

As shown in Table 2, no significant differences were observed in either the matched or non-matched texts for any of the dependent variables related to text readability or complexity, both before and after the family-wide correction for multiple comparisons. By comparing semantic similarity, text readability, and text complexity, we suggest that the text stimuli used in this study are well-controlled in line with the MECO project and are valid for conducting cross-language analysis without confounding factors arising from the texts themselves.

Readability/ complexity metrics	Matched texts: F-value	Matched texts: p-value (corrected)	Non-matched texts: F-value	Non-matched texts: p-value (corrected)
flesch_readability	0.209	0.998 (1)	1.799	0.056 (0.564)
l2_readability	0.92	0.539 (1)	0.728	0.731 (1)
narrativity	0.344	0.981 (1)	1.044	0.418 (1)
simplicity	2.409	0.012 (0.118)	1.963	0.034 (0.341)
concreteness	0.159	1 (1)	0.44	0.95 (1)
cohesion	1.042	0.427 (1)	0.539	0.894 (1)
deep_cohesion	0.65	0.801 (1)	0.661	0.795 (1)
verb_cohesion	1.42	0.18 (1)	0.61	0.839 (1)
connectivity	0.119	1 (1)	1.374	0.189 (1)
temporality	0.193	0.999 (1)	1.079	0.388 (1)

Table 2. Regression analysis of readability and complexity measures. *Note:* The *F*-values and corresponding *p*-values are reported for models run separately for matched and non-matched texts. *P*-values in parentheses are Bonferroni-corrected for multiple comparisons.

Traditional Mongolian texts were printed in the commonly used Menk Qagan Tig (proportional font, not monospaced), size 19 points, with 1.5 spacing. Due to programming issues with correctly displaying and processing vertically oriented texts in traditional Mongolian, all passages were converted to the image format for use in the presentation software Experiment Builder (version 2.4.1, SR Research, Kanata, Ontario, Canada), rather than being displayed as text.

To make sure participants paid full attention while reading the texts during the experiment session, each text was followed by four yes/no comprehension questions. All the materials are available on the project’s OSF page (see Data records below). Participants also completed an additional task alongside the reading task: a nonverbal IQ test using the short version of the Culture Fair Test-3 (CFT20), subset 3 Matrices, Form A¹⁹, to provide a standardized measure of nonverbal intelligence.

Apparatus and procedure. Eye movements were recorded using the Tower Mount EyeLink 1000 eye tracker (SR Research, Kanata, Ontario, Canada) at a sampling rate of 1000 Hz. The stimuli were displayed on a 20-inch Lenovo L2021 monitor with a resolution of 1024 × 768 pixels. Participants were seated 54 cm away from the monitor and used a chin rest and head restraint to minimize head movements. They read sentences binocularly, but only their right eye was monitored. Before each trial, a fixation dot appeared on the monitor slightly above the first word of the passage. The trial began once the participant fixated on the dot. This drift check and correction procedure occurred at the beginning of each trial, with calibration monitored by the experimenter and redone as needed. Each of the 12 texts was presented on a separate screen. Participants were instructed to read the passages silently for comprehension and press the space bar once they finished reading each passage.

Before starting the eye-tracking experiment session, each participant completed a basic demographic and language proficiency questionnaire, as well as the nonverbal CFT20 test. Afterward, the experimenter guided them through setting up the eye-tracking session and conducted a nine-point calibration and validation. Participants were then introduced to the experiment with one practice text, followed by four yes/no questions displayed one at a time. Participants pressed “1” on the keyboard for “yes” and “0” for “no” to answer the questions. Once it was confirmed that the participants understood the task, the formal experiment began in following the same procedure.

Data preprocessing. Using the eye-tracking data analysis software Data Viewer (version 4.4.1, SR Research, Kanata, Ontario, Canada), two research assistants visually inspected the quality of the eye-tracking data. They individually labeled trials with fixations consistently misaligned with text lines—likely due to poor calibration—as well as blank trials caused by software interruptions, marking them as “bad” trials. A third reviewer verified the labels, and – upon discussion with the original annotators – excluded problematic trials. As a result, 11 trials were removed from the analysis. Participants with fewer than five usable trials were also excluded, leading to the removal of one participant. Additionally, one participant’s data file was corrupted and could not be analyzed. To further refine the dataset, we filtered out fixations shorter than 80 ms and excluded the top 1% of total fixation durations and number of fixations at the word level within each participant’s specific distribution. After these procedures, the final dataset included 66 participants, 781 trials, and 165,174 out of 171,072 data points.

Since the stimuli were presented as images, the areas of interest were manually created at the word and sentence levels in Data Viewer (version 4.4.1, SR Research, Kanata, Ontario, Canada). The eye-tracking data were then reported as interest area reports corresponding to these words and sentences. In 12 trials, research assistants manually corrected fixations on the first and last text lines that were misaligned with the vertical text lines, likely due to poor calibration near the screen boundaries. The correction procedure involved moving fixations only horizontally—either to the right or left—without altering their vertical position. This approach was applied when the fixations for an entire line displayed a consistent misalignment pattern. Research assistants would review and assess the overall fixation patterns for one or two lines of text before making adjustments to ensure the fixations were aligned with the words. Importantly, the corrections were constrained to horizontal movements only, ensuring the fixation remained aligned with the correct word, without shifting vertically and

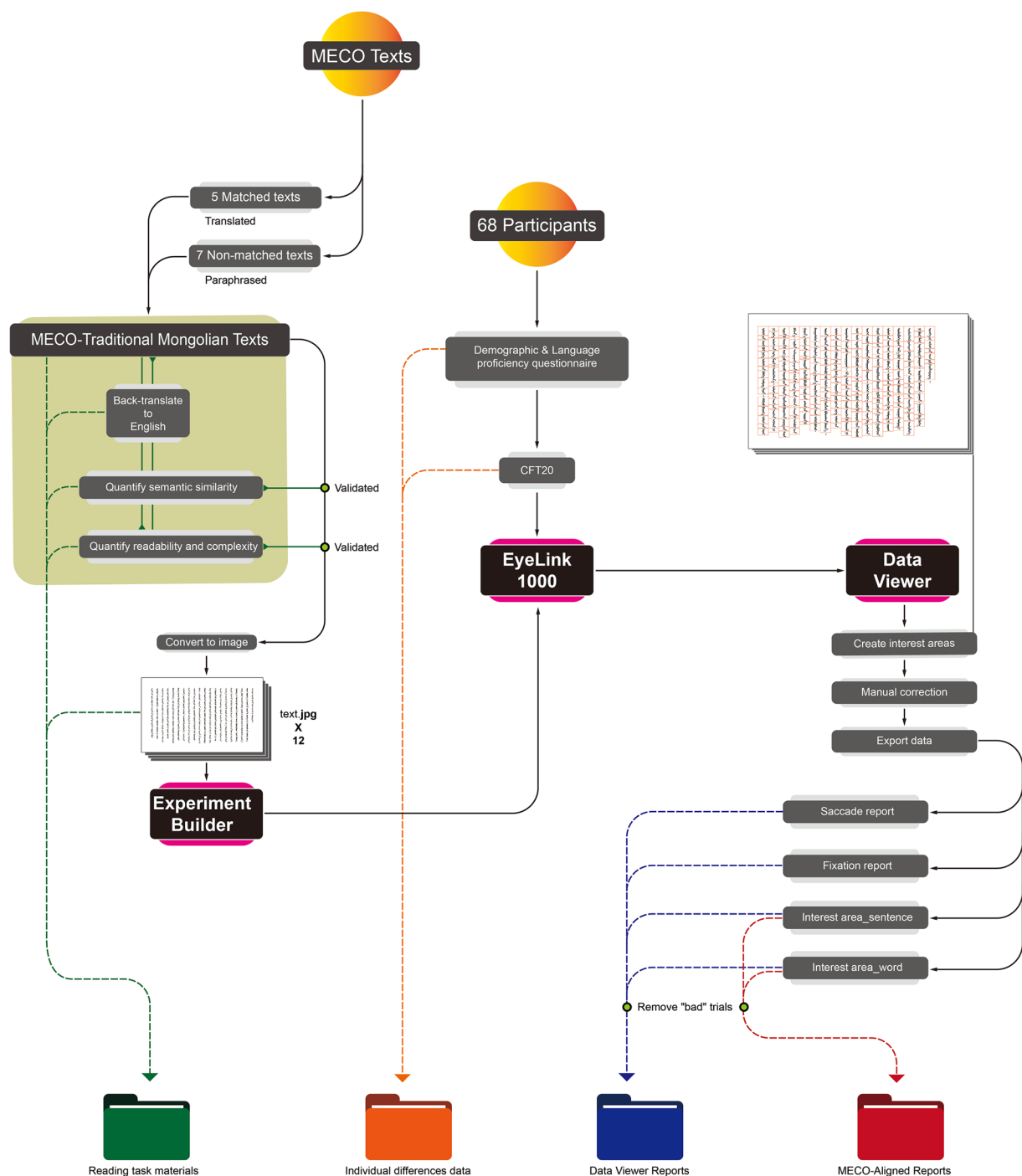


Fig. 1 Schematic illustration of the experimental workflow.

potentially changing the word position. In this regard, the Traditional Mongolian data is different from all other language samples of the MECO project available so far: The latter underwent automatic correction and assignment of fixations using the popEye software (implemented in R, version 0.6.4)²⁰, an integrated environment to pre-process and analyze eye-tracking data from reading experiments. The reason for the discrepancy is that popEye does not work with image files. The schematic illustration of the experimental workflow is shown in Fig. 1.

Data Records

The MECO-Traditional Mongolian data is available for free access in the Open Science Framework (OSF) repository²¹ under the CC BY 4.0 license. The dataset includes eye-tracking and reading comprehension data processed using the standard Data Viewer software (version 4.4.1). The eye-tracking data is represented by standard outputs of Data Viewer, i.e., saccade reports, fixation reports, and interest area reports, see details below. To align with the MECO project structure and to follow its conventions for variable naming, the interest area reports

Variable	n	mean	sd	se	median	min	max
firstFixationDuration	66	257.04	28.44	3.50	254.06	201.78	326.34
gazeDuration	66	307.67	42.27	5.20	305.47	232.97	430.06
nFixationsFirstRun	66	1.20	0.08	0.01	1.20	1.08	1.45
nFixationsTotal	66	1.26	0.36	0.04	1.25	0.68	2.17
readingRate	66	211.32	66.20	8.15	194.29	109.10	364.58
regressionIn	66	0.22	0.07	0.01	0.21	0.10	0.37
rereading	66	0.24	0.13	0.02	0.22	0.04	0.56
skipping	66	0.27	0.07	0.01	0.26	0.15	0.44
totalFixationDuration	66	426.68	91.97	11.32	426.06	276.99	661.74
accuracy	66	0.78	0.10	0.01	0.76	0.62	0.98
accuracyMatched	66	0.73	0.15	0.02	0.75	0.45	1
cft	66	8.73	2.24	0.28	9	0	12

Table 3. Descriptive statistics of the dataset.

have also been recalculated and relabelled to match the MECO variables. Therefore, the storage is organized into two main folders: Data Viewer Reports and MECO-Aligned Reports.

Data Viewer Reports contains the original reports generated directly by the Data Viewer software without modifications. The files in this folder include: 'Saccade_report.csv', 'Fixation_report.csv', 'Interest_area_report_word.csv', and 'Interest_area_report_sentence.csv'. These files retain the original variables and variable names generated by the Data Viewer software. The variable names can be accessed in the Data Viewer manual²².

MECO-Aligned Reports contains eye-tracking data that has been recalculated and reformatted to match the MECO project's variable structure. These reports enable a seamless integration of traditional Mongolian data with the data available in MECO for all other languages and makes cross-linguistic comparisons that include traditional Mongolian possible. A detailed legend document is included to explain the conversion formulas applied to the original Data Viewer variables.

MECO-Aligned Reports folder is organized into several subfolders. The 'auxiliary files' comprise two sub-sets. One is 'descriptive stats,' which includes descriptive analyses of comprehension accuracy (descriptive_acc_passage.csv) and reading rate (descriptive_rate_passage.csv) for each text. Additionally, this subset provides descriptive analyses of eye movement measures reported at the word level, as presented in Table 3. We also computed correlations between behavioral measures of reading, calculated from participant-level means for word-level reports, as shown in Table 4. The other subfolder is 'reading task materials,' which includes the original texts, images used for displaying text and questions, all back-translations in English, text readability and complexity metrics, and similarity scores. In addition, the lexical properties of word frequency and word length both in pixels and letters are included in 'word_list.csv' file. The 'code' folder includes the R script used for data analysis (Main.R). The 'primary data' folder includes 'comprehension data,' 'eye-tracking data,' and 'individual differences data.' The 'IAS files' folder includes the interest area files used in Data Viewer for each text at the sentence and word levels. Detailed information is listed below:

- auxiliary files
 - descriptive stats
 - descriptive_acc_passage_mo.csv (comprehension question accuracy)
 - descriptive_eyemove_mo.csv (word-level eye movement measures)
 - descriptive_rate_passage_mo.csv (reading rate in words per minute)
 - (all files listed above provide data in a descriptive summary format)
 - reading task materials
 - back_translation_mo.xlsx
 - passage pictures
 - question pictures
 - texts_mo.xlsx
 - cohmetrix_mo.csv
 - similarity_mo.csv
 - word_list.csv (word frequency and word length information)
- code
 - Main.R (main code used for data analysis)
- primary data

	skipping	firstFixationDuration	gazeDuration	totalFixationDuration	readingRate	nFixationsFirstRun	nFixationsTotal	regressionIn	rereading	cft_score
skipping		0.01	−0.31	−0.75	0.9	−0.68	−0.92	−0.65	−0.85	−0.22
firstFixationDuration	0.927		0.87	0.47	−0.34	0.11	−0.04	−0.1	−0.07	−0.05
gazeDuration	0.011	<0.001		0.69	−0.59	0.58	0.28	0.09	0.16	0.02
totalFixationDuration	<0.001	<0.001	<0.001		−0.92	0.61	0.85	0.64	0.81	0.18
readingRate	<0.001	0.005	<0.001	<0.001		−0.66	−0.88	−0.68	−0.83	−0.18
nFixationsFirstRun	<0.001	0.361	<0.001	<0.001	<0.001		0.65	0.34	0.45	0.13
nFixationsTotal	<0.001	0.763	0.023	<0.001	<0.001	<0.001		0.77	0.97	0.24
regressionIn	<0.001	0.446	0.477	<0.001	<0.001	0.005	<0.001		0.79	0.2
rereading	<0.001	0.579	0.203	<0.001	<0.001	<0.001	<0.001	<0.001		0.24
cft_score	0.076	0.696	0.904	0.143	0.155	0.314	0.053	0.104	0.055	

Table 4. Correlation table for reading measures. *Note:* values above the diagonal show Pearson correlations; values below the diagonal show *p*-values.

- comprehension data
 - mo_all_acc.rda (answers for each question by each participant; each participant answers a total of 48 questions, with 4 questions per trial.)
 - mo_br_acc.rda (accuracy for each trial by each participant)
 - mo_acc.rda (accuracy rate for each participant)
- eye tracking data
 - eye_data_trimmed_mo.rda (eye movement measures for each word, interest area report at the word level)
 - mo_readrate.rda (reading rate for each participant in each trial)
 - passage_data.csv (eye movement measures summed by trial)
 - sentence_data.csv (eye movement measures for each sentence, interest area report at the sentence level)
- individual differences data
 - cft20.csv (cft IQ test score for each question by each participant)
 - cft_summary.csv (summed cft score for each participant)
 - language_experience_and_proficiency_questionnaire
- IAS files
 - sentence_IAS (interest area files for each trial at the sentence level)
 - word_IAS (interest area files for each trial at the word level)

Technical Validation

Quantitative reliability validation. The reliability of this data was estimated using the split-half technique at both the participant level and the word token level. First, to assess the stability of each eye movement measure within participants, we examined the correlation between mean values for “odd” and “even” words within each participant. In addition, we also estimated the reliability of the nonverbal CFT test and comprehension question accuracy among participants. The only exception to this procedure was the estimation of reliability for reading rate, which was assessed by calculating the intra-class correlation coefficients (ICC) to examine the degree of agreement in reading rates across the 12 texts. Second, to assess reliability at the word token level, we examined the correlation between means for “odd” and “even” participants within each word token for each eye movement measure. For both participant-level and word token-level reliability, we computed raw correlations and Spearman–Brown-corrected values.

The high level of correlation in eye movement measures at the participant level (mean $rs = 0.99$), as shown in Table 5, confirmed the stable reliability of these measures within participants. The ICC for reading rate was 0.99. In contrast, the reliability of estimates at the word token level was slightly lower than at the participant level but still reached a high correlation (mean $rs = 0.88$), as shown in Table 6. Additionally, both the nonverbal CFT score and comprehension accuracy achieved high levels of correlation, with values of 0.63 and 0.76, respectively.

Validation of benchmark word-length and word-frequency effect. Word frequency and word length are two fundamental lexical variables that consistently influence reading behavior across languages. Decades of eye-tracking research have established that higher-frequency words are fixated for shorter durations and are more likely to be skipped, whereas longer words tend to require longer fixation durations and are skipped

Measures	Uncorrected_reliability	Spearman_Brown_corrected
skipping	0.97	0.99
firstFixationDuration	0.98	0.99
gazeDuration	0.98	0.99
totalFixationDuration	0.99	1
nFixationsFirstRun	0.96	0.98
nFixationsTotal	0.99	1
regressionIn	0.96	0.98
rereading	0.99	1
cft_score	0.46	0.63
accuracy	0.61	0.76

Table 5. Reliability estimates at the participant level.

Measures	Raw_correlation	Spearman_Brown_corrected
skipping	0.92	0.96
firstFixationDuration	0.56	0.71
gazeDuration	0.79	0.89
totalFixationDuration	0.85	0.92
nFixationsFirstRun	0.80	0.89
nFixationsTotal	0.93	0.96
regressionIn	0.69	0.82
rereading	0.81	0.90

Table 6. Reliability estimates at the word token level.

less frequently^{23,24}. These effects are considered benchmark findings in reading research and have been replicated across multiple languages and orthographic systems^{6,25,26}. To validate whether these well-documented lexical effects hold in traditional Mongolian reading, we examined how word frequency and word length influence first fixation duration (FFD), gaze duration (GD), total fixation duration (TFD), and skipping probability (SKIP). Given the unique vertical orthography of the traditional Mongolian script, it is important to establish whether these effects manifest similarly to those observed in horizontally written languages.

Since no publicly available online corpus of traditional Mongolian exists, we relied on the *Modern Mongolian Frequency Dictionary*²⁷, published in 1998, which is written in traditional Mongolian. Each word in our experimental materials was manually checked against this dictionary, with words not found in the dictionary assigned a frequency value of 1. Word frequency (log-transformed) ranged from 0 to 3.15 ($M = 0.56$, $SD = 0.92$). Word length was measured using two approaches: (1) physical length in pixels and (2) the number of letters. In traditional Mongolian script, consonants and vowels are written in a connected, cursive manner, forming ligatures that can alter the overall visual shape and length of a word. As a result, the physical length of a word (measured in pixels) does not always correspond directly to the number of letters it contains. Depending on the research objective, either measure may be used. In our dataset, word length measured in letters ranged from 1 to 17 ($M = 4.94$, $SD = 2.71$), while word length measured in pixels ranged from 12 to 150 pixels ($M = 52.61$, $SD = 23.05$). Since the correlation between the two measures was high ($r = 0.95$), we conducted further analysis using only word length in letters.

Descriptive statistics (Fig. 2) revealed that higher-frequency words were processed more quickly and skipped more often, while longer words required more processing time and were skipped less frequently. These effects were confirmed using linear mixed-effects models for fixation durations and a generalized linear mixed-effects model for skip probability (Table 7). Regression model results showed a significant negative effect of word frequency on fixation durations (FFD: $b = -7.032$, $t = -9.049$, $p < 0.001$; GD: $b = -5.354$, $t = -4.901$, $p < 0.001$; TFD: $b = -18.092$, $t = -8.385$, $p < 0.001$), indicating that frequent words are processed more efficiently. In contrast, word length had a strong positive effect on fixation durations (FFD: $b = 1.938$, $t = 8.083$, $p < 0.001$; GD: $b = 11.542$, $t = 33.745$, $p < 0.001$; TFD: $b = 22.773$, $t = 33.427$, $p < 0.001$), confirming that longer words require increased processing effort. For skipping probability, the logistic mixed-effects model indicated that higher-frequency words were more likely to be skipped ($b = 0.26$, $z = 12.378$, $p < 0.001$), while longer words were less likely to be skipped ($b = -0.437$, $z = -53.376$, $p < 0.001$).

These findings confirm that the benchmark word-frequency and word-length effects observed in horizontally written languages extend to the vertically written traditional Mongolian script. This validation strengthens the cross-linguistic comparability of the dataset and provides a basis for future analyses examining the influence of reading direction on cognitive processing.

Cross-language comparisons. The primary aim of MECO-Traditional Mongolian was to address the lack of eye movement data from understudied languages and to make a significant contribution to the theoretical understanding of cognitive processing across different languages. Additionally, since the MECO-Traditional

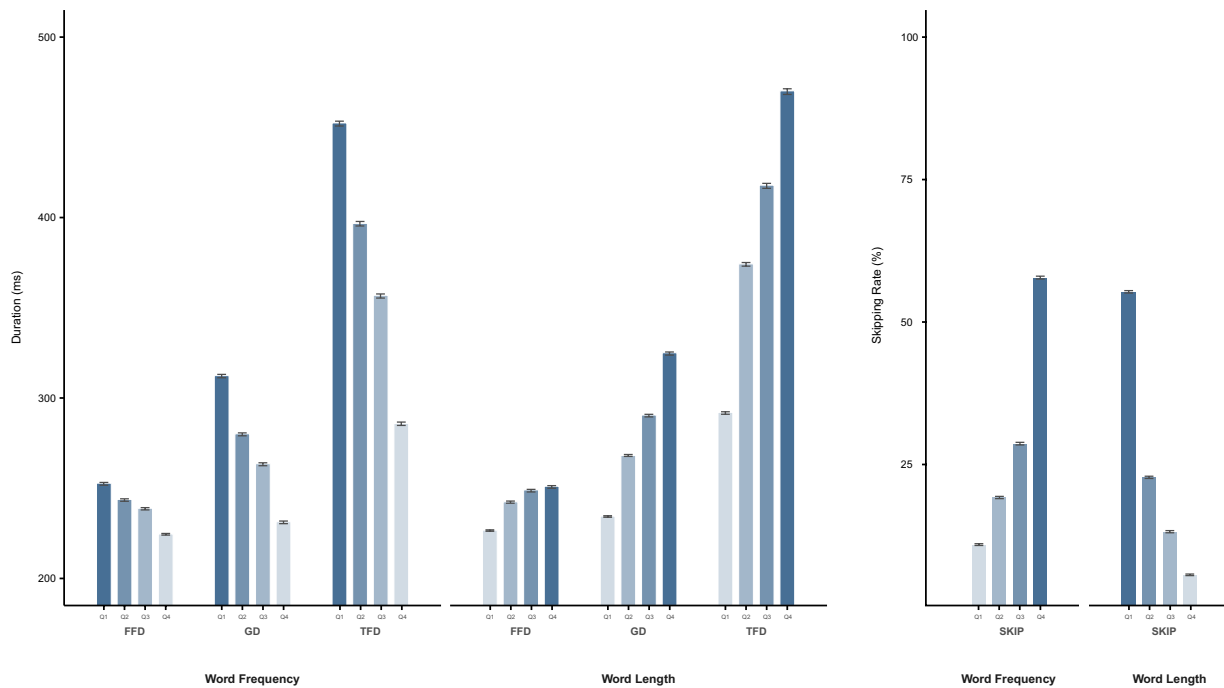


Fig. 2 Descriptive statistics of word frequency and word length effects on eye-movement measures. *Note:* Q1, Q2, Q3, and Q4 represent quartiles of word frequency and word length. Q1 corresponds to the lowest word frequency and shortest word length, while Q4 represents the highest word frequency and longest word length. FFD (*first fixation duration*), GD (*gaze duration*), TFD (*total fixation duration*), and SKIP (*skipping probability*) are shown with means and standard errors.

Measure	Predictor	<i>b</i> value	Cohen's <i>d</i> /OR	<i>t</i> / <i>z</i> value
FFD	Word Frequency	−7.032	−0.053	−9.049***
	Word Length	1.938	0.053	8.083***
GD	Word Frequency	−5.354	−0.032	−4.901***
	Word Length	11.542	0.248	33.745***
TFD	Word Frequency	−18.092	−0.058	−8.385***
	Word Length	22.773	0.265	33.427***
SKIP	Word Frequency	0.26	1.297	12.378***
	Word Length	−0.437	0.646	−53.376***

Table 7. Results for the effects of word frequency and word length on the main eye-movement measures. *Note:* For SKIP, the Cohen's *d* column represents Odds Ratios (OR), and the *t*/*z* column represents Z-values. **p* < 0.05, ***p* < 0.01, ****p* < 0.001. Abbreviations: FFD: *first fixation duration*; GD: *gaze duration*; TFD: *total fixation duration*; SKIP: *skipping probability*.

Mongolian dataset was processed under the guidelines and procedures of the MECO project, it is essential to conduct a cross-language comparison with the MECO project datasets to validate the comparative utility of the MECO-Traditional Mongolian dataset.

First, we merged the MECO-Traditional Mongolian data with data from 13 other languages in Wave 1 of the MECO project⁵. We then calculated the mean values and standard errors for nine eye-movement measures, the CFT20 nonverbal IQ test, and comprehension accuracy, grouped by language and participant. As shown in Fig. 3, the reading strategies employed by traditional Mongolian readers stand out from those of other languages; they spend more time on the first fixation duration than readers of any other language and have longer gaze durations and total fixation durations compared to reading in most other languages. Mongolian reading also shows a lower rate of rereading and a smaller number of fixations. Comprehension accuracy and the CFT-20 scores in the non-verbal intelligence test are close to the cross-linguistic average.

Second, we employed a hierarchical cluster analysis, following the methodology used in the MECO project, to investigate whether linguistic similarities between languages are reflected in the oculomotor patterns of their readers. For this analysis, we used three key eye-movement measures—skipping rate, gaze duration, and total number of fixations—to construct a vector representing each Mongolian participant. The Euclidean distances between all pairs of scaled participant vectors were computed, and these distances were averaged by language to generate a measure of dissimilarity between languages⁵. This language-level distance data was then analyzed

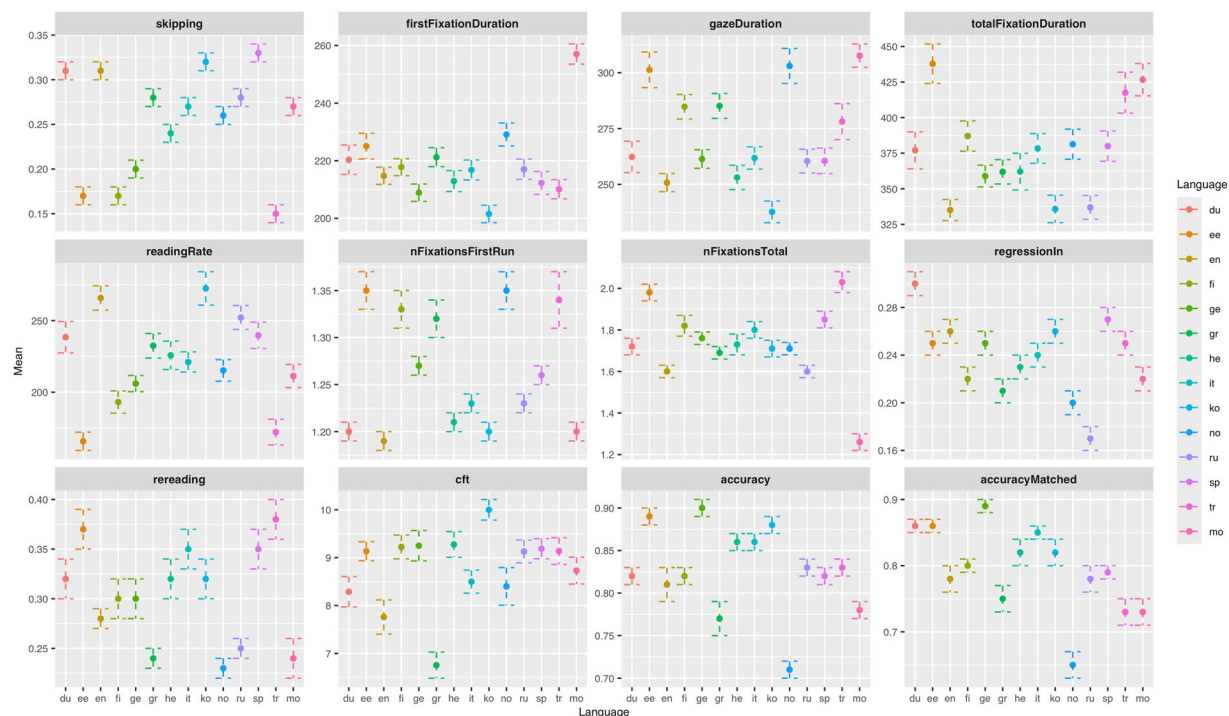


Fig. 3 Cross-language comparison across languages on mean value. *Note:* Language codes used in this figure are as follows: **du**: Dutch; **ee**: Estonian; **en**: English; **fi**: Finnish; **ge**: German; **gr**: Greek; **he**: Hebrew; **it**: Italian; **ko**: Korean; **no**: Norwegian; **ru**: Russian; **sp**: Spanish; **tr**: Turkish; **mo**: traditional Mongolian.

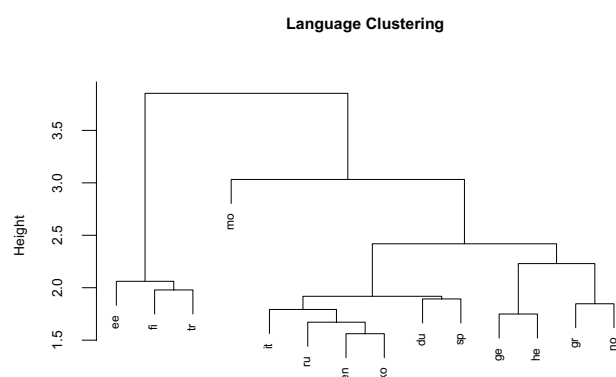


Fig. 4 Hierarchical cluster analysis of eye-movement measures across languages from MECO Wave 1 data. *Note:* Language codes used in this figure are as follows: **du**: Dutch; **ee**: Estonian; **en**: English; **fi**: Finnish; **ge**: German; **gr**: Greek; **he**: Hebrew; **it**: Italian; **ko**: Korean; **no**: Norwegian; **ru**: Russian; **sp**: Spanish; **tr**: Turkish; **mo**: traditional Mongolian.

through hierarchical cluster analysis, using the Ward clustering criterion, implemented via the *hclust* function in R²⁸. The resulting cluster solution (see Fig. 4) shows the most distinctive split to be between agglutinative alphabetic non-Indo-European languages (Estonian, Finnish, and Turkish) and the remainder of the language samples. The next split shows that traditional Mongolian reading behavior forms a distinct branch, separating it from all other, mostly Indo-European languages, in the cluster. This distinct placement highlights the unique oculomotor patterns of traditional Mongolian readers, likely driven by the language's agglutinative structure and vertical orthography. The divergence from alphabetic (mostly Indo-European) languages suggests that traditional Mongolian reading strategies differ significantly from those employed by readers of other scripts or writing systems.

The cross-linguistic analysis, as well as analyses of lexical benchmark effects, indicate that the MECO-Traditional Mongolian eye-movement corpus provides valuable data from an understudied language. These data offer insight into how traditional Mongolian readers process text, contributing to a broader understanding of cognitive processing across diverse linguistic systems. The distinct clustering of

traditional Mongolian underscores the potential for identifying new reading patterns in languages that have previously lacked empirical data.

Code availability

The code used in this study is provided at project's OSF page (<https://osf.io/3j9ut/>) with detailed readme.txt files.

Received: 15 October 2024; Accepted: 6 March 2025;

Published online: 25 March 2025

References

1. Cain, K. in *The Science of Reading: A Handbook* (2nd ed., pp. 298–322. Ch14 2022).
2. Li, X., Huang, L., Yao, P. & Hyönä, J. Universal and specific reading mechanisms across different writing systems. *Nature Reviews Psychology* **1**, 133–144, <https://doi.org/10.1038/s44159-022-00022-6> (2022).
3. Frost, R. Towards a universal model of reading. *Behavioral and Brain Sciences* **35**, 263–279, <https://doi.org/10.1017/S0140525X11001841> (2012).
4. Verhoeven, L. & Perfetti, C. Universals in learning to read across languages and writing systems. *Scientific Studies of Reading* **26**, 150–164, <https://doi.org/10.1080/10888438.2021.1938575> (2022).
5. Siegelman, N. *et al.* Expanding horizons of cross-linguistic research on reading: The Multilingual Eye-movement Corpus (MECO). *Behavior Research Methods* **54**, 2843–2863, <https://doi.org/10.3758/s13428-021-01772-6> (2022).
6. Kuperman, V., Schroeder, S. & Gnetov, D. Word length and frequency effects on text reading are highly similar in 12 alphabetic languages. *Journal of Memory and Language* **135**, 104497, <https://doi.org/10.1016/j.jml.2023.104497> (2024).
7. Wilcox, E. G., Pimentel, T., Meister, C., Cotterell, R. & Levy, R. P. Testing the Predictions of Surprisal Theory in 11 Languages. *Transactions of the Association for Computational Linguistics* **11**, 1451–1470, https://doi.org/10.1162/tacl_a_00612 (2023).
8. Osaka, N. & Oda, K. Effective visual field size necessary for vertical reading during Japanese text processing. *Bulletin of the Psychonomic Society* **29**, 345–347, <https://doi.org/10.3758/BF03333939> (1991).
9. Osaka, N. Size of saccade and fixation duration of eye movements during reading: Psychophysics of Japanese text processing. *Journal of the Optical Society of America, A, Optics, Image & Science* **9**, 5–13, <https://doi.org/10.1364/josaa.9.000005> (1992).
10. Ojanpää, H., Näsänen, R. & Kojo, I. Eye movements in the visual search of word lists. *Vision Research* **42**, 1499–1512, [https://doi.org/10.1016/S0042-6989\(02\)00077-9](https://doi.org/10.1016/S0042-6989(02)00077-9) (2002).
11. Yan, Pan, J., Chang, W. & Kliegl, R. Read sideways or not: vertical saccade advantage in sentence reading. *Reading and Writing* **32**, 1911–1926, <https://doi.org/10.1007/s11145-018-9930-x> (2019).
12. Yan, M., Kliegl, R. & Pan, J. Direction-specific reading experience shapes perceptual span. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **50**, 1740–1748, <https://doi.org/10.1037/xlm0001340> (2024).
13. Borjigin, B., Zhang, G., Hou, Y. & Li, X. Perceptual span in Mongolian text reading. *Current Psychology* <https://doi.org/10.1007/s12144-024-06074-6> (2024).
14. Su, J. *et al.* Flexibility in the perceptual span during reading: Evidence from Mongolian. *Attention, Perception, & Psychophysics* **82**, 1566–1572, <https://doi.org/10.3758/s13414-019-01960-9> (2020).
15. Günther, F., Dudschig, C. & Kaup, B. LSAfun - An R package for computations based on Latent Semantic Analysis. *Behavior Research Methods* **47**, 930–944, <https://doi.org/10.3758/s13428-014-0529-0> (2015).
16. Graesser, A. C., McNamara, D. S., Louwerse, M. M. & Cai, Z. Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers* **36**, 193–202, <https://doi.org/10.3758/BF03195564> (2004).
17. Graesser, A. C., McNamara, D. S. & Kulikowich, J. M. Coh-Metrix: Providing Multilevel Analyses of Text Characteristics. *Educational Researcher* **40**, 223–234, <https://doi.org/10.3102/0013189X11413260> (2011).
18. von der Malsburg, T. & Angele, B. False Positives and Other Statistical Errors in Standard Analyses of Eye Movements in Reading. *J Mem Lang* **94**, 119–133, <https://doi.org/10.1016/j.jml.2016.10.003> (2017).
19. Weiß, R. H. *Grundintelligenzskala 2 mit Wortschatztest und Zahlenfolgetest [Basic intelligence scale 2 with vocabulary knowledge test and sequential number test]*. (Göttingen, Germany: Hogrefe., 2006).
20. Schroeder, S. popEye - An R package to analyse eye movement data from reading experiments. *GitHub repository* <https://github.com/sascha2schroeder/popEye> (2019).
21. Bao, Y. B., Kuperman, V. & Li, X. The Eye Movement Database of Passage Reading in Vertically Written Traditional Mongolian. *Open Science Framework* <https://doi.org/10.17605/OSF.IO/3J9UT> (2025).
22. EyeLink DataViewer 4.4.1 [computer software manual]. Oakville, Ontario, Canada: SR Research Ltd (2024).
23. Brysbaert, M. *et al.* The Word Frequency Effect. *Experimental Psychology* **58**, 412–424, <https://doi.org/10.1027/1618-3169/a000123> (2011).
24. Barton, J. J. S., Hanif, H. M., Eklinder Björnström, L. & Hills, C. The word-length effect in reading: A review. *Cognitive Neuropsychology* **31**, 378–412, <https://doi.org/10.1080/02643294.2014.895314> (2014).
25. Pan, J., Yan, M., Richter, E. M., Shu, H. & Kliegl, R. The Beijing Sentence Corpus: A Chinese sentence corpus with eye movement data and predictability norms. *Behavior Research Methods* **54**, 1989–2000, <https://doi.org/10.3758/s13428-021-01730-2> (2022).
26. Zhang, G. *et al.* The database of eye-movement measures on words in Chinese reading. *Scientific Data* **9**, 411, <https://doi.org/10.1038/s41597-022-01464-6> (2022).
27. Da, B., & Bao, J. Modern Mongolian Frequency Dictionary. *Inner Mongolia Education Press* (1998).
28. Langfelder, P. & Horvath, S. Fast R Functions for Robust Correlations and Hierarchical Clustering. *J Stat Softw* **46** (2012).

Acknowledgements

The contributions of the first and third authors were supported by the Social Sciences and Humanities Research Council of Canada Insight Grant (435-2021-0657; Kuperman, PI). The third author's contributions was also supported by the Partnered Research Training Grant (895-2016-1008; Libben PI), and the Canada Research Chair (Tier 2; Kuperman, PI). We sincerely thank Yongchang, for his assistance with graphic design.

Author contributions

Yaqian (Borogjoon) Bao: Conceptualization, Methodology, Formal Analysis, Software, Data curation, Writing-Original draft preparation. Xingshan Li: Writing- Reviewing and Editing, Supervision. Victor Kuperman: Writing- Reviewing and Editing, Formal, Analysis, Validation, Supervision.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Y.B.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025