Interdisciplinary Journal of Linguistics
Volume [16] 2023, Pp. 1-12

LINGUISTIC ACTIVATION AND BI-DIRECTIONAL READING IN URDU: AN ELECTROPHYSIOLOGICAL INVESTIGATION

Azizuddin Khan*

ABSTRACT

Urdu orthography is bidirectional in which words are read from right to left and the numerals from left to right direction. The changes in reading directions of words and numerals pose challenges for the readers with regard to their lexical recognition strategies, inhibition of control for planning saccadic eye movements, and shifting of attention. The present study examined the effect of change in reading direction on brain activities with the help of Electroencephalogram (EEG). The study was performed on nineteen adult native Urdu speakers (M=12, F=7; Avg. Age= 37.4 years). It was observed that upon the change in reading direction in articulating numerals, an event-related potential (ERP) occurs with a latency of 200ms in the left anterior temporal and posterior middle temporal region. However, the polarity of the peak was observed positive for changing the reading direction of numerals from left-to-right to the direction of words from right-to-left. The same peak was observed to be negative when the direction changed vice versa. It was also observed that the areas responsible for recognition of information (right anterior superior temporal region) show early ERPs when the change in direction occurs from right-to-left to left-to-right as compared to the change in reading direction from left-to-right to right-to-left.

Keywords: Bidirectional Reading, Lexical Access, Morphological Processing, Orthography, Shifting of Attention

1. Introduction

Reading in Urdu is characteristically performed from right to left however, in case of numerals the reading direction changes from left to right. Such a change in reading direction may necessitate for extra efforts by imposing cognitive load. The phenomenon of bidirectional reading arises in orthographies of some alphabetic languages due to certain linguistic reasons particularly morphological, and syntactic reasons. Such a phenomenon can be intrinsic to an orthography or can exhibit extrinsically as well. In Urdu, the typical bidirectional reading pattern is a consequence of an intrinsic characteristic feature of its orthography.

* Indian Institute of Technology Bombay, Mumbai, India

Therefore, while reading a sentence which contain numbers written in numerals, the reader assumes a right to left reading pattern for words and phrases but changes the direction from left to right for reading numerals. Similarly, in a multilingual situation the phenomena of code-switching and code-mixing give rise to bidirectional orthographies. In such cases, the matrix language determines the typical direction of reading while as the lexical and grammatical elements of an embedded language can display a change in direction e.g. in an inter-sentential code-switching of Urdu to English, Urdu assumes to the matrix language position where the elements of English perform as an embedded language. Therefore, upon reading such a sentence the reading direction starts from right to left in reading Urdu and turns left to right for English. Such a type of bidirectional reading pattern occurs extrinsic to an orthography. Nonetheless, the presence of images and emojis in a text can technically exhibit bidirectional reading pattern however, such a pattern is not considered principally bidirectional because images and emojis are not part of alphabetic orthographies.

Syntactically, Urdu has an SOV (subject-object-verb) word order in which specifiers and quantifiers such as numbers usually assume pre-object position to complete a sentence, and morphologically it undergoes particularly multifaceted processes for nouns verbs, thereby changing their shapes and occurrences in a sentence. Such morphological processes are observed as in derivational and inflectional affixations, formation of compound nouns, and causative verb formations. Interestingly, Urdu is written in Devanagri and Roman scripts also which form its exographs but as a standard endograph it has a complex *Abjad* class of Persio-Arabic orthography, written principally in connected and cursive *Nasta'liq* style where morpheme boundaries do not express overtly. The complex orthography of Urdu remarkably regulates its syntactic procedures which encompasses all the morphological processes to exhibit a well expressed right-left writing pattern. As a part of orthographic system to disambiguate between otherwise similar lexemes and morphemes, the use of diacritic marks in Urdu plays an essential role in comprehension of words and phrases. Basically, such diacritic marks determine the characteristics of vowel sounds but also show the consonant connections in a word or phrase. The distinctively occurring feature of Indo-Aryan languages i.e. the consonant gemination, expressed by a diacritic mark in Urdu. Moreover, the Urdu orthography expresses the changing shapes of consonants in word-initial, word-medial, and word-final positions by the virtue of Allography. The use of diacritic marks in Urdu orthography as an essential tool for orthographically regulated structural or lexical disambiguation, as a result solves the issue of grapheme-phoneme correspondences. When Urdu is written without diacritic marks, it exhibits an opaque orthography and there is no one-one correspondence between graphemes and the phonemes. Such an opaque orthography poses challenges for the reader to articulate the words and phrases correctly in a sentence. Also, the readers derive on the contextual information for comprehending challenging orthographical structures. The studies carried out in bidirectional reading focus on eye movements and perceptual span during change in reading directions. The perceptual span in English and other languages read from left to right has received much attention (see Rayner, 2009); while the perceptual span in languages read from right to left has received significantly less. The findings for the importance of reading direction in determining asymmetry in the perceptual span are based on findings from Pollatsek, Bolozky, Well, and Rayner's (1981) single study on the perceptual span in Hebrew (read from right-left). Native Hebrew and English readers in Israel saw sentences with a window of regular text stretching 14 characters to the left of fixation and 4 characters to the right, or 4 characters to the left of fixation and 14 characters to the right. The results showed that Hebrew reading performance was better when the windows were asymmetric to the left, whereas English reading performance was better when the windows were asymmetric to the right. As a result, according to Pollatsek et al., the overall direction of reading appears to impact the asymmetry of the perceptual span. Khan, Loberg, and Hautala (2017) conducted a study on twenty-one Urdu speaking subjects and found that high cognitive resources were used for reading long numerals read in left-right direction in Urdu sentences, and found that the change against the default reading direction was preceded by highly inflated fixation duration, pinpointing the oculomotor, attentional and cognitive demands in executing sudden changes in reading direction.

Moreover, the numbers are written either in words or in Arabic numerals and thus the format of writing numbers is either verbal (words) or Arabic Numerals (digits). Some studies (e.g., Colvin et al., 2005; Cohen Kadosh et al., 2005; Pinel et al., 2003; Szûcs and Csépe, 2004) have looked at how numerical information is processed according to number format and found that specific neural activities are dedicated to processing specific number formats, while other cortical areas represent conceptual information. In Urdu, the numerals are written in Eastern-Arabic numeral system or Arabic–Indic numerals. These numerals are orthographically different (e.g. ۲-۱-۴-۳-than usual Arabi (۵c numeral system however, they are read and written from left to right direction similarly as that of the Arabic numeral system (i.e. 1-2-3-4-5). Hence, when a number is read in a given Urdu sentence it is read as that of the universal Arabic numeral system and follows all the characteristic properties for place values of the digits. Pinel et al. (2001) discovered that numerical input notation selectively activated different cortical areas (verbal or Arabic notation). During verbal processing, the bilateral extrastriate cortices and a left precentral region were more activated than during digit processing, whereas during digit processing, the right fusiform gyrus (FG) and bilateral inferior parietal and frontal areas were more activated. As a result, it has been shown that numeral symbolic ability is reliant on parieto-temporal connectivity, which facilitates communication between intraparietal quantity processing areas and occipito-temporal areas engaged in symbol recognition (Dehaene et.al., 2003). The present study aims to understand brain responses when direction change occurs in reading bidirectional languages. It is observed that the change in reading direction (i.e., words to numerals, numerals to words) activates the left posterior middle temporal region with a latency of 160 -220 ms (see. Szûcs and Csépe, 2004; Dehaene, 1996). It corresponds to the N1 response predominantly left-sided to the verbal numbers and more bilateral to the digits. It is also observed that the polarity of the ERPs changes according to the change in reading direction (e.g., Pinel et al., 2003). The ERPs analysis performed in the current work coincides with the earlier studies.

1.1 Protocol

The protocol was conducted following the Declaration of Helsinki and the protocol was approved by the ethical committee of the Indian Institute of Technology Bombay, India.

1.1.1 Participants' Selection

Nineteen adult participants (12 Male and 7 Female) were selected for the experiment, who were proficient in Urdu reading. All participants were school teachers and their mother tongue was also Urdu. To grade their proficiency level in the language, a sentence reading task was conducted which were including words as well as numerals. The proficiency in Urdu reading was graded according to the time taken by the participant in the reading of sentences. The participants were aged in the range of 22 to 53 years with a mean of 37.4 years. None of the subjects had a history of any type of neurological disorder and they were not having any medical treatment at the time of the experiment. Subjects gave their informed consent before participating in the experiment.

1.1.2 Sentence Types

Four types of sentences were selected for the experiment, in which two types of sentences were consisting bi-directionality in reading (i.e. words as well as numerals) and two types of sentences were consisting uni-directionality in reading (i.e. only words). The two types of bi-directional sentences were consisting either short or long numerals. Similar sentences were created for uni-directional reading by replacing numerals by words. Sentences were named as Long Numerals, Short Numerals, Long words, and Short Words shown in Figure 1.

- **Long Numerals:**

کلام نے اسکول کے بچوں کے لیے ۷۵۰۹۹۰ روپے دیے۔

(Kalam ne school ke bacchon ke liye **750990** rupye diye.)

- **Long Word:**

کلام نے اسکول کے بچوں کے لیے دو کروڑ روپے دیے۔

(Kalam ne school ke bacchon ke liye **do crore** rupye diye.)

- **Short Numerals:**

میرے گھر کی لمبائی کم سے کم ۱۴۳ فٹ ہے۔

(Mere ghar ki lambayi kam se kam **143** foot hai.)

- **Short Word:**

میرے گھر کی لمبائی کم از کم بیس فٹ ہے۔

(Mere ghar ki lambayi kam se kam **bees** foot hai.)

Figure 1: Types of sentences displayed for reading. The sentences consisting the long or short numerals were having the bi-directionality. The numerals were replaced by the number in words to read sentence in uni-direction.

2. Data Collection Technique

Participants were asked to wash their hair and dry it completely before starting data collection. The participants were comfortably seated in a reclining armchair in the dimly lit recording room. All personal electronics gadgets (mobile, watch, etc.) were taken out from the data recording room. A 24-inch monitor was placed 2 ft. away from the participant to display the cue.

2.1 EEG Recording Setup

A 64-channel Bio Semi Active Two system with active Ag/AgCl electrodes with CMS and DRL as the ground was used to record

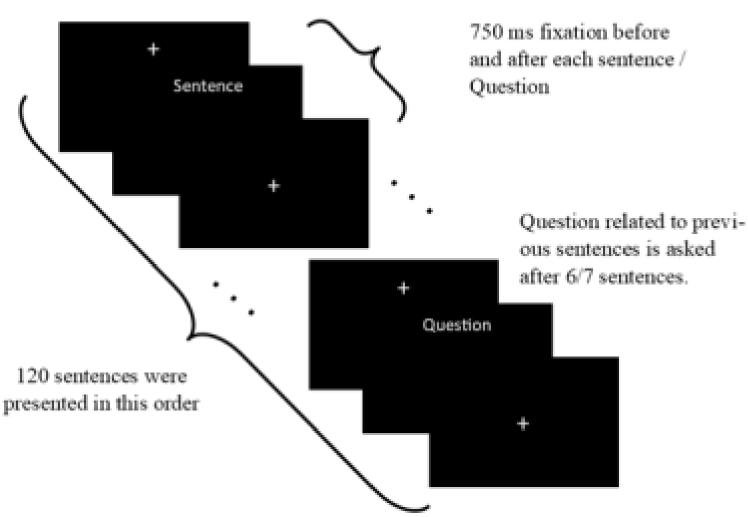


Figure 2: **Experiment Paradigm**

EEG at a sampling rate of 2048 Hz with a 24-bit analog to digital converter. Bio Semi head cap (according to the subject's head size) was used to position the 64 EEG electrodes on the scalp, according to the international 10-20 system. All the markers were provided through E-Prime 2.0 software via parallel port and recorded with EEG signals through Active Two software. The head circumference was measured using a measuring tape. Headcap was chosen according to the head size to hold the EEG electrodes. The head center was marked by measuring the center between nasion to inion and left earlobe to right earlobe. Headcap was worn to the participant's head by keeping the Cz channel at the center. A conductive gel was filled in all the holes of the headcap which was used to make proper connections between electrodes and scalp. 64 channels electrodes were connected in the headcap holes by matching the positions of the electrodes. Electrodes were connected with the Biosemi analog to digital converter which was connected with the computer for recording the EEG data. After connecting all the electrodes, impedances for all electrode sites were checked if they are less than 10 KΩ. If high impedance was found, the connection of the respective electrode was checked.

2.2 Experiment Paradigm

All sentences were displayed to the subject for 5 seconds with a gap of 750 msec fixation time as shown in Figure 2. After the display of 6/7 sentences, a multiple-choice question was displayed to the subject based on the last displayed 6/7 sentences. A participant was required to respond right answer by pressing numeral keys within 5 seconds. The question was placed to make sure whether the participant is properly reading the sentences. After the response to the question, 1 second fixation time was given before the display of the next set of sentences. The markers for the sentence appearance on the screen was provided from E-Prime to Biosemi Actve Two software.

3. Data Analysis

EEG data was imported into MATLB environment. The raw EEG signal from each electrode was filtered using a zero-phase band-pass filter with a cutoff frequency of 0.01-250 Hz to remove low-frequency trends and high-frequency noises. Each EEG epoch corresponding to the duration of reading was extracted with the help of sentence onset marker given by E-Prime. Sentences were categorized by the position of number appeared in the sentences. The horizontal eye movement while reading the sentences was monitored using EEG activities acquired from channels AF7 and AF8. EEG epochs at the time of changing reading direction was extracted using the step change observed in the channels AF7 and AF8 around the timing of number reading in the sentences. Timestamp during the number reading in the sentences was approximately measured by the position of number in the sentence. The EEG segment for the duration of change in reading direction was extracted in two conditions, i.e. when direction of reading changed from right-to-left to left-to-right (reading numerals to words) and when direction of reading change from left-to-right to right-to-left (reading words to numerals). The EEG segments were extracted with the block size -200 to 600 ms, centered at the time of change in reading direction. The grand average of ERPs in both conditions (changing direction of reading in either way) was obtained from the data of all subjects. Topographic plots were also obtained for the change in reading directions in either side using the average of amplitudes from -10ms to 10 ms of changing reading direction.

3.1 Coherence Analysis

Coherence analysis is performed to understand the simultaneous brain activation at different areas for a particular activity. Coherence is a mathematical factor that measures the linear dependence between two signals coming from two brain regions at a noticeable distance. Coherence can be calculated as per Equation 1.

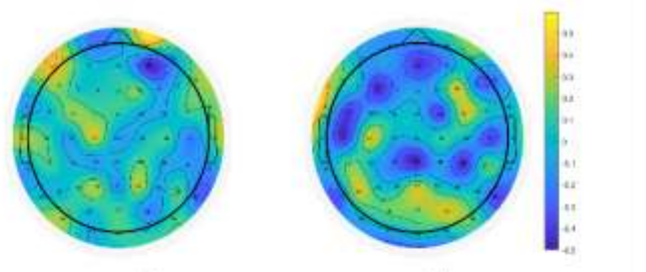
$$coh_{xy}(f) = \left|S_{xy}(f)\right|^2 / |S_{xx}(f)|\left|S_{yy}(f)\right| \quad (1)$$

Where $S_{xy}(f)$ is the cross-spectral density between signals x and y, and $S_{xx}(f)$, $S_{yy}(f)$ are the auto power spectral densities of signals x and y respectively.

Subject's proficiency in Language reading v/s cognitive load and saccade effect: Proficiency help readers to not consider higher cognitive load during reading. For an approximate evaluation, the amplitude changes in EEG activity when a large saccade happens (changing the direction of reading from word to digit and vice-versa) was measured. The subject-wise average amplitude change in the EEG activity during the change of reading direction is calculated by measuring the amplitude change around the reading direction change from the EEG recorded at channel AF7 and AF8 for each sentence. The amplitude change is measured by normalizing the signal in the range of -1 to 1.

3.2 Results

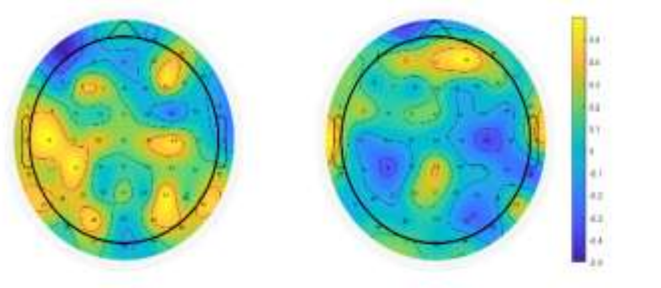
Topographic plots were created during the change in reading direction (average of EEG during -10ms to 10ms) in either cases (changing the direction of reading from word to digit and vice-versa). Topographic plots show that the potential change happened only near the right anterior frontal lobe (AF8) for both cases and the other brain region was not much active shown in Figure 3.



(a) (b)

Figure 3: Topographic plot during change in direction of reading occurs for sentences with long numbers in digits. (a) Saccade in the left direction when changing the reading of words to digit. (b) Saccade in the left direction when changing the reading of digit to words

The reason behind the higher activation in the right anterior frontal lobe in both cases was that the reader's sudden eye movement happens in the same direction i.e. right to left in both cases. The brain activity during the reading of words (i.e. right to left) and the brain activity during the reading of digits (i.e. left to right) was also analysed. The topographic plots were taken for the average amplitude of EEG signals during the reading of words and digits as shown in Figure 4.



(a) (b)

Figure 4: Topographic plot during (a) reading of words and (b) the reading of digit

It is clear that Wernicke's area is active during the reading of words as well as the digits. However, the polarity of activation is different in both cases. During words reading, bilateral extrastriate cortices also activate which supports the Pinel et al. (2001) study. It is also observed that the frontal lobe and inferior parietal regions are more activated during digit reading.

The ERPs were extracted during the change in direction of reading occurred i.e. reading numerals to words (direction of reading change from right-to-left to left-to-right) and reading words to numerals (direction of reading change from left-to-right to right-to-left) with the block size -200 to 600ms, centered at the time of change in reading direction. The grand average of ERPs was obtained from the data of all subjects as shown in Figure 5.

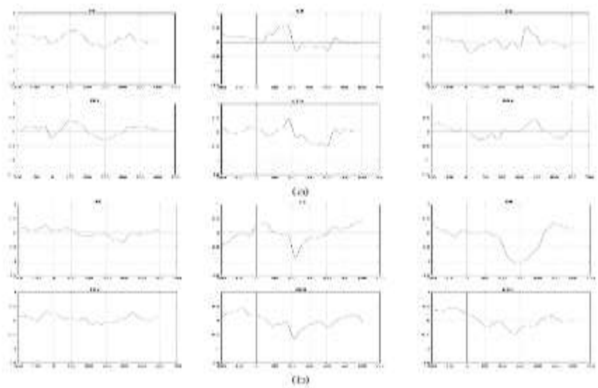


Figure 5: Event related potentials during change in direction of reading occurs i.e. (a) reading numerals to words and (b) reading words to numerals

Major activations in ERPs were seen over the left anterior temporal and posterior middle temporal region, and right anterior superior temporal and parietal-occipital region. In the left hemisphere, P200 activity occurs in channels C5 and CP3, while Channels T7 and TP7 showed the P100 and N300 activities for changing the reading activity from numerals to words. The reason behind activation in the left temporal region is that the reading of words requires the understanding of language, remembering verbal information, and formation of speech corresponding to the displayed word. For the same event, late positive deflection i.e. P400 occurs over channels C6 and PO4 (right hemisphere), which are also responsible for the information gathering.

When changing the reading activity from numerals to words, N200 activity occurs in channels C5 and CP3, while no specific ERP occurs over the left anterior inferior temporal lobe i.e. T7 and TP7. It is obvious the reading of digits does not require the memorization of verbal information or formation of speech. Reading of digits requires the memorization of non-verbal information, recognition of the information, etc. which can be seen over the right hemisphere as N300 activity occurs in the channels C6 and PO4.

After comparison between the ERPs during the change of reading direction in, either way, it was clearly observed that the change in direction of reading activated the left posterior middle temporal region with a latency of 200ms. However, the polarity of the potential changes according to the direction of reading. The coherence analysis is performed for all 64 channels. For reading words from right to left direction, maximum coherence was obtained between electrode pairs FPz-FT7-F7, FC1-C3-FPz, FPz-CP2-P2, AF8-C4-P2-CP2, and FC4-AF4-FC6-P5.

The subject-wise average amplitude change in the EEG activity during the change of reading direction shows that the readers with very high proficiency had lesser change in brain activities at anterior frontal lobe (AF7 and AF8), which make the difficulties in identification of the onset of the change in reading direction. It can be possible due to high parafoveal vision of such readers as shown in Table 1 and Table 2.

Subject ID	1	2	3	4	5	6	7	8	9	10
Normalized value of Amplitude change	0.42	0.61	0.44	0.35	0.62	0.32	0.38	0.39	0.34	0.38
Subject ID	11	12	13	14	15	16	17	18	19	Average
Normalized value of Amplitude change	0.39	0.28	0.36	0.46	0.34	0.46	0.51	0.42	0.48	0.42

Table 1: Normalized Amplitude Changes in EEG during Saccade for "Long Numerals"

Subject ID	1	2	3	4	5	6	7	8	9	10
Normalized value of Amplitude change	0.37	0.52	0.38	0.33	0.55	0.28	0.36	0.35	0.30	0.36
Subject ID	11	12	13	14	15	16	17	18	19	Average
Normalized value of Amplitude change	0.33	0.25	0.34	0.39	0.32	0.44	0.45	0.36	0.44	0.38

Table 2: Normalized Amplitude Changes in EEG during Saccade for "Short-Numerals"

4. Discussion

The present study aimed to understand the brain activities during the change of reading direction in the Urdu language. Only a few studies were found in the literature that explored the brain responses for reading words (right-to-left) and numbers (left-to-right). However, the effect of the sudden change in reading direction while performing continuous reading is not addressed for studying brain responses during the events. In the present study, a continuous reading task was performed for the sentences consisting of the words that can be read from right to left and numerals that can be read from left to right.

The brain responses analysis begins with the onset detection of the change in reading direction. The challenge is that the time required for reading a text varies for different readers, depending on various factors like reader proficiency, word length, reading numerals, etc. For the fair study, it is required to find the exact time when the reader changes the direction of reading, which means when the reader starts reading a word after a numeral or vice-versa. To resolve such a challenge, eye tracking was performed using the EEG acquired from the anterior frontal region of the brain.

It is also essential to monitor the reader's alertness while performing the reading task. So, a question was asked to the reader after reading six or seven sentences. The questions were designed in a way; if the reader had paid attention to the previous sentences, the answer would be easy to respond to the question.

The experiment protocol discussed earlier has considered maximum precautions that can help provide accurate brain responses corresponding to the change in reading direction. The sentence design part took care of the word length and the time required for the sentence reading by the subject. Sentences were displayed to the subject for natural reading. The position of the numeral in the sentence was also noted to help in the approximation of timestamp while the event occurs. The sudden step-change in electric potential over the anterior frontal region near the timestamp obtained by the numeral position was considered the time of change in reading direction.

However, we acknowledge that latency is possible in detecting a change in reading direction due to the possible latency in a step change of electric potential over the anterior frontal region. It can be further rectified by combining the eye (pupil) tracking system with the EEG recording machine. Still, the present study successfully presents the relative comparison when the change in reading direction occurs either way. The study shows that the cognitive load, which is observed with the help of brain responses, change in both conditions as ERPs latencies and polarities are different.

In conclusion, the results validated the protocol for understanding the brain responses corresponding to the change in reading direction. The present study will be helpful in future investigations to explore more about the effects of reading bi-directional languages on the brain. It can help develop intervention tools for dyslexic, dyscalculic, and slow learners to improve reading such languages.

References

Farukh, A., & Vulchanova, M. (2014). Predictors of Reading in Urdu: Does Deep Orthography have an Impact? *Dyslexia*, *20*: 146-166.

Fruzsina Soltész, & Dénes Szűcs. (2014).Neural Adaptation to Non-Symbolic Number and Visual Shape: An Electrophysiological Study. *Biological Psychology,* 103: 203-21.

Khan, A., Loberg, O., & Hautala, J. (2017). On the Eye Movement Control of Changing Reading Direction for a Single Word: The Case of Reading Numerals in Urdu. *Journal of Psycholinguistic Research,* 46(5): 1273-1283. https://doi.org/10.1 007/s10936-017-9491-1.

Philippe Pinel, Stanislas Dehaene, Manuela Piazza, & Laurent Cohen. (2003). Three Parietal Circuits for Number Processing. *Cognitive Neuropsychology,* 20(3-6): 487-506. doi:10.1080/026432902440 00239.

Pollatsek A. Bolozky S. Well A. D., & Rayner K. (1981). Asymmetries in the Perceptual Span for Israeli Readers. *Brain and Language* 14 (1): 174-180.

Rao, Chaitra & Vaid, J., & Srinivasan, Narayanan, & Chen, Hsin- Chin. (2011). Orthographic Characteristics Speed Hindi Word Naming but Slow Urdu Naming: Evidence from Hindi/Urdu Biliterates. *Reading and Writing*. 24: 679-695. https://doi.org/10.1007/s11145-010-9256-9.

Rayner, K. (2009). The 35th Sir Frederick Bartlett Lecture: Eye Movements and Attention in Reading, Scene Perception, and Visual Search. *Quarterly Journal of Experimental Psychology*, 62(8): 1457-1506. https://doi.org/10.1080/17470210902816461.