

10-10-2021

Effects of word predictability on eye movements during Arabic reading

Maryam A. AlJassmi
Zayed University

Kayleigh L. Warrington
University of Leicester; Nottingham Trent University

Victoria A. McGowan
University of Leicester

Sarah J. White
University of Leicester

Kevin B. Paterson
University of Leicester

Follow this and additional works at: <https://zuscholars.zu.ac.ae/works>



Part of the [Social and Behavioral Sciences Commons](#)

Recommended Citation

AlJassmi, Maryam A.; Warrington, Kayleigh L.; McGowan, Victoria A.; White, Sarah J.; and Paterson, Kevin B., "Effects of word predictability on eye movements during Arabic reading" (2021). *All Works*. 4616.
<https://zuscholars.zu.ac.ae/works/4616>

This Article is brought to you for free and open access by ZU Scholars. It has been accepted for inclusion in All Works by an authorized administrator of ZU Scholars. For more information, please contact scholars@zu.ac.ae.



Effects of word predictability on eye movements during Arabic reading

Maryam A. AlJassmi^{1,2} · Kayleigh L. Warrington^{2,3} · Victoria A. McGowan² · Sarah J. White² · Kevin B. Paterson²

Accepted: 31 August 2021
© Crown 2021

Abstract

Contextual predictability influences both the probability and duration of eye fixations on words when reading Latinate alphabetic scripts like English and German. However, it is unknown whether word predictability influences eye movements in reading similarly for Semitic languages like Arabic, which are alphabetic languages with very different visual and linguistic characteristics. Such knowledge is nevertheless important for establishing the generality of mechanisms of eye-movement control across different alphabetic writing systems. Accordingly, we investigated word predictability effects in Arabic in two eye-movement experiments. Both produced shorter fixation times for words with high compared to low predictability, consistent with previous findings. Predictability did not influence skipping probabilities for (four- to eight-letter) words of varying length and morphological complexity (Experiment 1). However, it did for short (three- to four-letter) words with simpler structures (Experiment 2). We suggest that word-skipping is reduced, and affected less by contextual predictability, in Arabic compared to Latinate alphabetic reading, because of specific orthographic and morphological characteristics of the Arabic script.

Keywords Arabic · Eye movements during reading · Word predictability · Word-skipping

Introduction

During reading, the eyes move in a series of rapid, ballistic movements (saccades) that carry the reader's gaze across lines of text, separated by brief fixational pauses during which individual words are processed in high acuity (Liversedge & Findlay, 2000; Rayner, 1998, 2009). Substantial research has focused on understanding this eye-movement behaviour and, in particular, whether it is under cognitive control and so influenced by aspects of online language processing, such as lexical access (see, e.g., Rayner et al., 1996).

Findings from numerous studies suggest that forward-directed eye movements in reading are influenced by a combination of low-level visual information about the length and location of words, and higher-level knowledge about a word's

frequency of written usage and its predictability from the prior sentence context (for reviews, see Rayner, 1998, 2009). These factors have been shown to influence the probability of a word being fixated (and therefore whether it is skipped), as well as how long the eyes dwell on fixated words. Specifically, studies show that words that are shorter, of higher lexical frequency, or more contextually predictable have lower fixation probabilities (and so are skipped more often). Moreover, when fixated, these words have shorter fixation times (e.g., Altarriba et al., 1996; Choi et al., 2017; Gollan et al., 2011; Hand et al., 2010; Inhoff & Rayner, 1986; Joseph et al., 2009; Kliegl et al., 2004; Paterson et al., 2013a, 2013b; Rayner, 1979; Rayner et al., 1996, 2011; Rayner et al., 2004; Rayner & McConkie, 1976; Staub & Benatar, 2013). Such findings are important as they demonstrate that decisions about when and where to move the eyes in reading are influenced by factors affecting the process of word identification. Moreover, such research has been fundamental to the development of sophisticated computational models of eye-movement control in reading, including the E-Z Reader (e.g., Reichle et al., 1998) and SWIFT (e.g., Engbert et al., 2005) models (for a recent review, see Reichle, 2021).

While both types of models can accurately simulate forward-directed eye movements in reading, there are crucial

✉ Maryam A. AlJassmi
Maryam.AlJassmi@zu.ac.ae

¹ Department of Psychology, Zayed University, Dubai, UAE

² Department of Neuroscience, Psychology and Behaviour, University of Leicester, Leicester, UK

³ Department of Psychology, Nottingham Trent University, Trent, UK

differences in their underlying assumptions. One approach, exemplified by the E-Z Reader model, assumes that word identification is the engine that drives the forward movement of the eyes. Within this model, attention is allocated to words strictly serially, so that they are identified one at a time. In an initial stage of word identification, called the *familiarity check* (or L1), the system computes whether a word is likely to be identified imminently, which triggers the programming of a saccade (i.e., eye movement). In a second stage (L2), lexical access is completed. Both stages are assumed to be influenced by lexical variables, including word length, frequency and predictability. In the case of LI, these variables are thought to influence the probability and duration of fixations on words. In particular, it is assumed that the time taken to complete the familiarity check for the currently fixated word (word n) is a function of these variables. Once this is completed, covert attention is assumed to shift from this word to the following word (word $n+1$), initiating a new familiarity check for $n+1$ and the programming of a saccade to that word. Crucially, if the $n+1$ familiarity check can be completed before the saccade program is ready, the system may revoke this saccade program and initiate a new one that skips $n+1$ and targets a saccade towards the next word along.

By comparison with E-Z Reader, SWIFT is based on the assumption that multiple words are processed in parallel within a region around each fixation location. Within this model, it is assumed that saccade timing is primarily regulated by an autonomous timer that maintains a preferred reading speed, and that saccades are targeted towards words based on their patterns of lexical activation. As with E-Z Reader, this process can be influenced by lexical variables, including word length, frequency and predictability. However, word frequency and predictability have a lesser role in the SWIFT model and are assumed to influence oculomotor processes only occasionally. For instance, the initiation of a saccade may be interrupted, so that the reader dwells on a word for longer if this word proves difficult to identify because of its low lexical frequency or low predictability. Similarly, while word-skipping is assumed to be driven primarily by word length, knowledge about a word's frequency and predictability can influence the selection of saccade targets, producing increased skipping rates for words that are of high lexical frequency or highly predictable. Accordingly, despite differing in their emphasis on cognitive factors, both models allow for factors affecting the identification of words to play a role in eye guidance.

These influences on eye guidance have been extensively investigated in writing systems based on the Latin alphabet, such as English and German. However, comparatively little research has investigated effects in Semitic scripts such as Arabic and Hebrew, despite these having very different visual and linguistic characteristics (see, e.g., AlJassmi et al., 2021). Such knowledge is nevertheless crucial for establishing the generality of mechanisms of eye guidance across different

alphabetic writing systems. Of these languages, Hebrew has received most attention to date, although eye-movement studies in this script largely focus on morphological and syntactic processing (e.g., Deutsch, 1998; Deutsch et al., 2000; Deutsch et al., 2003; Deutsch et al., 2005; Deutsch et al., 2018; Deutsch et al., 2021; Deutsch & Bentin, 2001; Kuperman & Deutsch, 2020), with few studies investigating lexical influences on eye movements. A key study, by Deutsch and Rayner (1999), showed that word length has a similar influence on eye movements in Hebrew reading as observed for Latin-based scripts like English, despite its different reading direction. Currently, there are few eye-movement studies in Arabic, although several recent studies have investigated effects of word length and frequency (Hermena et al., 2017; Hermena et al., 2019; Paterson et al., 2015), showing that these variables can influence fixation times for words but have only small or non-significant influences on word-skipping. Accordingly, the present research aimed to extend our understanding of lexical influences on eye guidance in Arabic reading by examining word-predictability effects.

As noted above, research with Latin-based languages suggests that, when reading a sentence beginning “The manager tends to be in a very bad mood every morning before he has his...”, whether the next word in the sentence is highly predictable (e.g., “coffee”) or less predictable (e.g., “drink”) can affect eye movements. In particular, even though both words are possible continuations, the more predictable word will have a higher skipping probability and receive shorter fixations. Our goal was to establish whether similar predictability effects are observed in Arabic reading. While we might expect to observe such effects in fixation times on words, there were reasons to doubt whether predictability would influence word-skipping. It was of particular concern that words are skipped rarely in Arabic reading. Previous Arabic studies report low word-skipping rates (< 10% prevalence; Hermena et al., 2017; Hermena et al., 2019; Paterson et al., 2015), as compared to 20–30% prevalence for skilled readers of Latin-based languages (Brysbaert et al., 2005; Rayner, 1998, Rayner, 2009). This high rate of word-skipping for Latin-based scripts is partially attributable to the frequent skipping of short, predictable words in these scripts (see Brysbaert et al., 2005; Rayner & McConkie, 1976; Rayner et al., 1996). While Arabic words are generally short, with four- to eight-letter words accounting for 90% of items in the Aralex word database (Boudelaa & Marslen-Wilson, 2010), this is because vowels are omitted in most forms of Arabic text unless required to disambiguate a word (with the exception of religious texts and texts designed for beginning readers). Moreover, although words tend to include multiple affixes (prefixes, suffixes and infixes) to convey grammatical information, these are represented by few letters. Importantly, however, predictability has been shown to influence fixation probability and duration for both short and long words in scripts like English (Rayner et al., 2011),

raising the question of whether similar effects might be observed for Arabic.

The decision to skip a predictable word is necessarily based on both information obtained from outside of foveal vision (i.e., in parafoveal vision) and predictions derived from the prior sentence context. Such decisions will therefore depend on the strength of the predictions and the quality of parafoveal information, including available information about word length and orthography (e.g., Balota et al., 1985; Chang, Hao, et al., 2020a; Chang, Zhang, et al., 2020b; Choi et al., 2017; Juhasz et al., 2008; Schotter et al., 2015; Staub, 2020; Staub & Goddard, 2019; Veldre & Andrews, 2018; White et al., 2005). One possibility is that the available parafoveal information does not support predictability effects on word-skipping in Arabic. There are several reasons this might be the case. A first is that Arabic uses a semi-cursive script in which letters are connected by ligatures (small lines), while the orthography includes letters of varying width and groups of letters that are differentiated by only minor visual differences (e.g., ت, ب and ث, differ only by the presence and location of small dots, while ف and ق differ slightly in shape and number of small dots). The semi-cursive nature of the script and variation in letter width may create a situation in which a word's physical length does not help constrain word identities, as words with the same number of letters can differ in physical length, while words with different numbers of letters can have the same physical length (Hermena et al., 2017). A further possibility is that similarity in the shape of groups of letters may impair the discriminability of letters in parafoveal words. Studies using masked priming techniques suggest that the influence of this letter similarity is relatively small when words are in foveal vision (Perea et al., 2016). However, other research using lateralized displays suggests greater confusability of letter identities outside of foveal vision (e.g., Eviatar et al., 2004). Eviatar et al. further argued that this difficulty is compounded by reading direction. Specifically, they argue that, because Arabic is read from right to left, parafoveal words will be perceived in the left visual field during reading. Information about these words will therefore first project to a reader's right cerebral hemisphere (because of contralateral retinal projections, e.g., Jordan & Paterson, 2009), which for most (i.e., right-handed) readers has less efficient word-recognition capabilities. Following this account, orthography and reading direction may conspire to limit parafoveal processing in Arabic reading.

Secondly, whereas words in Latinate languages often have informative word beginnings that can help constrain parafoveal word identities (e.g., Farid & Grainger, 1996; Hand et al., 2012; Pagán et al., 2016), Arabic (like Hebrew) uses a non-concatenative morphology in which triples of letters that express a word's core meaning intermingle with other letters to convey its inflectional meaning. For instance, the root consonants ك ت ب (often transcribed as "ktb") express

the general meaning of "writing" in Arabic and these consonants intermingle with other letters (as indicated using underlining in the following examples) to create specific meanings, including يَكْتُبُ ("he writes") كَاتِبٌ ("writer"), and كِتَابَةٌ ("writing"). Consequently, letters that express a word's core meaning are not always contiguous with each other, and can be distributed throughout a word rather than located towards its beginning (Boudelaa, 2013; Ratcliffe, 2013). Readers therefore might not benefit from a parafoveal preview of the beginning letters in Arabic words, as compared to in scripts like English (e.g., Hand et al., 2012; Pagán et al., 2016), as the most informative elements seldom appear at the word beginning (Farid & Grainger, 1996). Finally, because of the script's agglutinative (or fusional) nature, Arabic words can be highly morphologically complex, using multiple affixes (i.e., prefixes, suffixes and infixes) to convey grammatical information. Words can therefore be informationally dense, making them more difficult to identify parafoveally. This morphological complexity may also incur a foveal processing cost for fixated words that reduces resources available for parafoveal processing (e.g., Henderson & Ferreira, 1990; Payne et al., 2016; Roman & Pavard, 1987).

Given this potential for impoverishment of parafoveal processing, it will be important to establish the influence of word predictability when reading Arabic. This may yield similar effects to Latinate scripts, such that fixation times are shorter and skipping rates higher for more predictable words. Alternatively, effects may be more limited, such that predictability influences reading times but not word-skipping. We conducted two experiments to investigate this. In both, we recorded the eye movements of fluent native Arabic readers while reading sentences containing a target word with high or low predictability from the prior sentence context. Experiment 1 used four- to eight-letter target words varying in morphological complexity, allowing us to assess predictability effects on reading times and word-skipping for typical Arabic words. With Experiment 2, we intentionally selected only three- to four-letter morphologically simple words to test effects under conditions that might promote parafoveal processing and maximise the likelihood of observing a skipping effect.

General method

Ethics statement Ethical approval was received from ethics committees at the University of Leicester and Zayed University.

Participants In each experiment, 40 fluent native Arabic readers aged 18–28 years ($M = 21$ years; all female) were recruited from Zayed University (13 participated in both experiments). All had normal or corrected-to-normal vision, assessed using a standard eye chart. Three participants from

Experiment 1 and one from Experiment 2 were excluded due to poor performance on questions used to test comprehension. To assess the power of our design, we performed a power simulation using the *simr* package (Green & MacLeod, 2016) in R (R version 4.0.2, R Core Team, 2016). This was conducted for three critical eye-movement variables, word-skipping probability, first-fixation duration, and gaze duration, which are informative about the initial (i.e., first-pass) processing of words. No previous studies have examined word predictability effects on eye movements in Arabic reading. The simulation therefore used effect sizes computed using means and standard deviations from three previous word-predictability studies conducted in English (Frisson et al., 2017; Sereno et al., 2018; Staub, 2020). Power curves based on these effect sizes are shown in Fig. 1. The results indicate that a sample size of 40 participants should be sufficient to detect a word predictability effect of the size reported in previous studies for all three critical variables with at least 80% power (Brysbaert & Stevens, 2018).

Stimuli and design Stimuli were two sets of 72 sentence frames that included one of two interchangeable target words (see Fig. 2). These had either high or low predictability from the prior sentence context. Experiment 1 used a variety of four- to eight-letter words and Experiment 2 used three- to four-letter nouns as targets. As Arabic conventionally uses a semi-cursive proportional script, high- and low-predictability words were matched for letter length across high- and low-predictability conditions (Experiment 1, $M = 6.1$, $SD = 1.1$; Experiment 2, $M = 3.5$, $SD = .5$; $t(143) = -24.48$, $p < .001$), and closely for spatial width (Experiment 1, high-predictability = 1.31° , low-predictability = 1.29° , $t(71) = 1.29$, $p = .20$; Experiment 2, high-predictability = $.89^\circ$, low-predictability = 1.00° , $t(71) = 1.23$, $p = .22$). The spatial width of words was shorter in Experiment 2 compared to Experiment 1 as a result of selecting words with fewer letters for this experiment ($t(143) = -14.32$, $p < .001$). The high- and low-frequency target words in each experiment were also matched for lexical frequency (in counts per million; Experiment 1, high-predictability = 68.1, $SD = 152.5$, low-predictability = 68.4, $SD = 190.8$, $t(71) = .01$, $p = .99$; Experiment 2, high-predictability = 81.9, $SD = 84.4$, low-predictability = 80.7, $SD = 169.5$, $t(70) = .01$, $p = .99$), using the *Aralex* database (Boudelaa & Marslen-Wilson, 2010). While target words had numerically higher lexical frequencies in Experiment 2 than in Experiment 1, this difference was not statistically significant ($t(143) = .68$, $p = .50$). Sentences were 10–20 words long and the target word always appeared near the sentence middle. Arabic uses small marks (diacritics) to indicate vowels, primarily in formal texts or to avoid ambiguity in everyday texts (Ratcliffe, 2013). Accordingly, these were added only when words were ambiguous (0.7% of words), and never for target words.

Word predictability was measured using a cloze procedure (Rayner & Well, 1996; Schwanenflugel, 1986; Taylor, 1953). Stimuli for each experiment were truncated immediately prior to the target word and undergraduate participants from Zayed University (24 per experiment), who did not participate in the experiments, provided written completions for each fragment. Target words were considered highly predictable if produced on more than 60% of completions, and of low predictability if produced on less than 15% of completions (Experiment 1, high-predictability, $M = 89\%$, $SD = 10$, low-predictability, $M = 0.4\%$, $SD = 2$, $t(71) = 78.52$, $p < .001$; Experiment 2, high-predictability, $M = 90\%$, $SD = 9$, low-predictability, $M = 1\%$, $SD = 3$, $t(71) = 77.43$, $p < .001$). Note that these cut-offs are consistent with cut-offs used in other word predictability studies (Frisson et al., 2017; Miell et al., 2007; Rayner et al., 2011; Staub, 2020; see Staub, 2015, for a discussion). For each sentence, 24 additional participants (per experiment) rated sentence plausibility on a 5-point scale (1 = highly implausible, 5 = highly plausible). This was high for all sentences, but higher for sentences with highly predictable words than less-predictable words (Experiment 1, high-predictability, $M = 4.6$, $SD = 0.3$, low-predictability, $M = 3.6$, $SD = 0.6$, $t(71) = 11.78$, $p < .001$; Experiment 2, high-predictability, $M = 4.7$, $SD = 0.2$, low-predictability, $M = 3.4$, $SD = 0.7$, $t(71) = 15.87$, $p < .001$), consistent with other research (Frisson et al., 2017).

For each experiment, sentence frame and target word combinations were split into two lists, each containing all 72 frames and an equal number of high- and low-predictability target words. Participants were pseudo-randomly assigned to each list. Sentences in each list were presented in random order to each participant, preceded by eight practice sentences. The experiments had a within-participants design with the factor word predictability (high, low).

Apparatus and procedure Eye movements were recorded at Zayed University using an EyeLink 1000 plus (SR Research) tower-mounted eye-tracker. Right eye-gaze location was recorded each millisecond during binocular viewing. This was in line with other research on reading in right-to-left languages, such as Arabic, Hebrew and Uyghur, that have recorded eye movements from the right eye only (Deutsch et al., 2005, 2021; Deutsch & Rayner, 1999; Hermena et al., 2015, 2016, 2017, 2019; Kuperman & Deutsch, 2020; Paterson et al., 2015; Yan et al., 2014; Zhou et al., 2021). Sentences were displayed on a high-definition 24-in. BenQ monitor at 80-cm viewing distance. Text was presented in 14-point using a commonly used proportional font (Arial).

Participants were tested individually. Before the experiment began, participants were instructed to read normally and for comprehension. The eye-tracker was then calibrated using a 3-point horizontal calibration (ensuring spatial error $< .30^\circ$). Calibration accuracy was assessed before each trial and the eye-tracker re-calibrated as necessary to maintain this low

spatial error. At the beginning of each trial, a fixation square equal in size to a character space was presented on the right of the screen (as Arabic is read from right-to-left). Once the participant fixated this location, a sentence was presented with its first letter replacing the square. Once the participant finished reading, they pressed a response button. The sentence then disappeared, replaced on 50% of trials by a comprehension question requiring a yes/no button-press response. The experiment lasted about 30 min for each participant.

Data analysis Following standard procedures, fixations < 80 ms or > 1,200 ms were discarded (affecting 3.6% fixations in Experiment 1, and 3.2% of fixations in Experiment 2). We also removed trials in which: (1) track loss occurred (affecting < 0.01% of trials in both Experiment 1 and Experiment 2), (2) a blink occurred on the target word or during an immediately adjacent fixation (affecting 11% of trials in Experiment 1 and 6% of trials in Experiment 2), or (3) when saccades to the target word were longer than ten characters (i.e., > 140 pixels, with each character subtending 14 pixels on average; affecting 0.01% of trials in Experiment 1 and 0.02% in Experiment 2). This resulted in the exclusion of 12% of trials in Experiment 1, with 2,542 trials remaining, and 7% of trials in Experiment 2, with 2,671 trials, remaining. Finally, we removed observations more than 2.5 standard deviations from the mean of each participant for each fixation measure (FFD: affecting 2.3% of data in Experiment 1 and 1.7% in Experiment 2; GD: affecting 2.1% of data in Experiment 1 and 1.9% in Experiment 2).

The remaining data were analysed using the lme4 package (version 1.1-27; Bates et al., 2015) in R (R Core Team, 2016). Linear mixed-effects models (LMMs) were used for continuous measures and generalized linear mixed models (GLMMs) for binominal measures. For each measure, a maximal random effects structure was used (Barr et al., 2013), with participants and stimuli as crossed random effects and word predictability as a fixed effect, with all models converging successfully. Contrasts were defined using the `contr.sdif` function in the MASS package (Venables & Ripley, 2002), with two levels of predictability (high, low). Analyses for untransformed and log-transformed reading times data produced the same patterns of results, so only results for untransformed data are reported for transparency.

Sentence-level analyses were computed to provide normative data for skilled Arabic reading. These comprised reading rate (words per minute), average fixation duration, average number of fixations per word, number of regressions (backward eye movements), and progressive saccade length (average length of forward eye movements, in average character length in the stimulus set). We also report target word-level eye movement measures used to test hypotheses concerning word predictability effects. To decrease the likelihood of Type I error due to testing multiple eye movement variables (von der Malsburg & Angele, 2017), we focused analyses on three

first-pass measures that are widely reported in word predictability studies (Balota et al., 1985; Ehrlich & Rayner, 1981; Rayner et al., 2005; Rayner & Well, 1996). First-pass reading refers to processing that takes place during an initial encounter with a word, prior to a fixation on the next words or a regression to re-inspect earlier text. To assess first-pass effects of word predictability, we examined: (1) word-skipping (SKIP), which is the probability of not fixating a word during first-pass reading, (2) first-fixation duration (FFD), which is the length of the first fixation on a word during first-pass reading; and (3) gaze duration (GD), which is the sum of all first-pass fixations on a word. To further reduce the likelihood of a Type I error, we used a Bonferroni correction to adjust for multiple comparisons. To do so, we divided the alpha threshold (0.05) by the number of critical dependent measures (3) to yield an alpha of .02 (t/z values = 2.39) for word-level measures of primary interest. We report several additional eye-movement variables as exploratory analyses that include t/z values without commenting on their statistical significance. These were single-fixation duration (SFD; length of the fixation on a word receiving only one first-pass fixation), regressions-out (RO; probability of a first-pass regression from a word), total reading time (TRT; sum of all fixations on a word) and regressions-in (RI; probability of a regression back to a word).

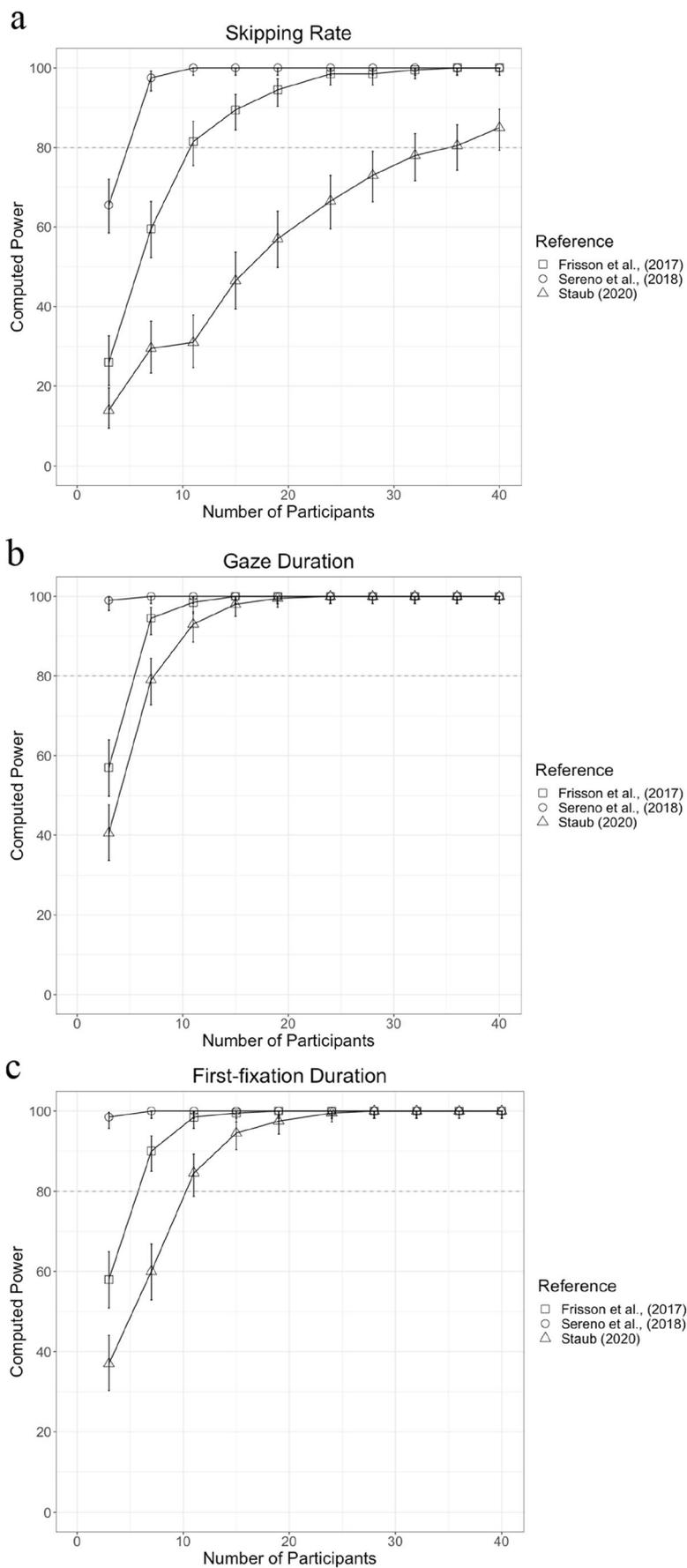
Experiment 1: Results and discussion

Data files and related resources for Experiments 1 and 2 are available from the University of Leicester online Figshare repository: <https://figshare.com/s/4382467c9a132fd2e15c>

Accuracy for responses to comprehension questions was high ($M = 96\%$), above 90% for all participants, indicating good comprehension.

Sentence-level measures Table 1 shows mean sentence-level measures. These are similar to those in previous Arabic research (Hermena et al., 2017, 2019; Jordan et al., 2014; Paterson et al., 2015), and so appear typical for Arabic reading.

Hypothesis-testing analyses for word-level measures Mean target word-level measures are shown in Table 2 and statistical effects for hypothesis-testing variables are summarized in Table 3. No predictability effects were observed for word-skipping rates ($|z| < 1$), which were low for high- and low-predictability words (7.2% vs. 7.8%). We further explored possible effects of word predictability on target word-skipping, by including the launch site of the saccade as an additional variable. This is based on the observation that predictability effects on word-skipping are more likely to be observed when saccades are launched from locations closer to the target word (e.g., Fitzsimmons & Drieghe, 2011) see Fig. 3. Launch site was calculated as the distance (in degrees of



◀ **Fig. 1** Power Estimates for Experiments 1 and 2 for (a) Word-Skipping Probability, (b) Gaze Duration for Target Words and (c) First-Fixation Duration

visual angle) from the right border of the target word, which we converted into number of letters for transparency (based on average letter width in the stimulus set). Launch sites more than six letters from the target word were excluded, as readers were unlikely to obtain a useful preview of the target word from a fixation this distant (affecting 5.0% of saccades). With saccade launch site included as a continuous variable, there was no main effect of word predictability ($b = 0.21$, $SE = 0.32$, $z = -0.65$). While skipping rates were higher for launch sites closer to the target word ($b = 0.82$, $SE = 0.08$, $z = -10.20$), there was no interaction between word predictability and launch site ($b = 0.20$, $SE = 0.16$, $z = 1.25$). The indication, therefore, is that word-skipping was a low-frequency event that was not influenced by target-word predictability.

We observed clear effects of word predictability in both first fixations and gaze durations for target words that were not skipped (first-fixation duration, high-predictability, $M = 248$ ms, low-predictability, $M = 260$ ms; gaze duration, high-predictability, $M = 287$, low-predictability, $M = 305$ ms; $|t|s > 3.08$). Consistent with findings from Latin-based languages, such as English and German, fixation times were shorter when words were highly predictable, as compared to less predictable, from the prior sentence context.

Exploratory word-level analyses Additional word-level analyses indicated that target words had shorter single-fixation durations and total reading times when they were highly predictable rather than less predictable from the prior sentence context. In addition, there was a lower probability of a regression back to target words that were highly predictable rather than less predictable, with no indication of an effect of predictability on the probability of a regression from the target word.

The absence of a predictability effect in word-skipping is divergent from findings for Latin-based languages (Balota et al., 1985; Ehrlich & Rayner, 1981; Rayner & Well, 1996). This is a likely consequence of reduced word-skipping in Arabic reading, possibly due to difficulty identifying words parafoveally. Indeed, word-skipping effects emerged only for close launch sites that allow better parafoveal processing. In the *Introduction*, we proposed such effects might be attributable to specific characteristics of the Arabic script, including impoverished parafoveal processing of orthographic and morphological information, as well as the greater morphological complexity of Arabic words. While we do not systematically assess the contribution of these factors, Experiment 2 addressed this issue by examining effects for short words with a simple morphological structure. This was achieved by creating a new stimulus set (see Fig. 2) comprising only three- to four-letter target words that were morphologically simple

(composed wholly or mostly of root consonants, without prefixes, such as the bigram $_l$, which preceded 80% of target words in Experiment 1). The crucial question was whether this would reveal a predictability effect in word-skipping.

Experiment 2: Results and discussion

As in Experiment 1, comprehension accuracy was high ($M = 97\%$) and above 90% for all participants.

Sentence-level measures Table 1 reports sentence-level means for Experiment 2. These were in line with eye-movement parameters in Experiment 1 and other Arabic studies (Hermena et al., 2017, 2019; Jordan et al., 2014; Paterson et al., 2015).

Hypothesis-testing analyses for word-level measures Mean target word-level measures are shown in Table 2 and statistical effects for variables used to test hypotheses concerning word predictability effects are summarized in Table 3. Word-skipping appeared to be higher in Experiment 2 than in Experiment 1 (Experiment 1, 7.5%; Experiment 2, 24.5%). Moreover, there was an effect of word predictability due to higher skipping of target words that were highly predictable (26.8%) rather than less predictable (22.1%, $z = -2.64$). As in Experiment 1, we explored this effect further by including saccade launch site as a continuous variable and excluding launch sites from beyond six letter spaces from the right edge of the target word (affecting 5.3% of saccades). This produced a main effect of word predictability ($b = -0.57$, $SE = 0.23$, $z = -2.52$), due to higher skipping probabilities for highly predictable compared to less predictable words. There was also a main effect of saccade launch site ($b = -0.13$, $SE = 0.04$, $z = -3.94$), due to higher skipping probabilities for launch sites closer to the target word. However, no significant interaction between these variables was observed ($b = 0.07$, $SE = 0.07$, $z = 1.06$).

As in Experiment 1, first-fixation durations and gaze durations for target words that were not skipped in first-pass reading yielded effects of word predictability, with shorter reading times for high- than low-predictability words (FFD = 234 ms vs. 252 ms, GD = 255 ms vs. 281 ms, $|t|s > 3.38$). Experiment 2 therefore revealed that predictability can influence both skipping probability and fixations times for words in Arabic reading, at least for words that are short and morphologically simple.

Exploratory word-level analyses Similar to Experiment 1, exploratory analyses of additional word-level measures suggested predictability effects in single-fixation durations and total reading times for words, with shorter reading times for highly predictable compared to less predictable words. As in

Experiment 1	
High Predictability	← يكون مزاج المدير سيئاً جداً قبل شربه <u>القهوة</u> لذلك يتجنبه الجميع في الصباح.
Low Predictability	← يكون مزاج المدير سيئاً جداً قبل شربه <u>الشراب</u> لذلك يتجنبه الجميع في الصباح.
<i>Translation</i>	The manager tends to be in a very bad mood every morning before he has his <u>coffee/ drink</u> , this is why everyone avoids him.
Experiment 2	
High Predictability	← قامت والدتي بقلع ضرسي بقوة وخفة، حيث أنني لم أشعر بأي ألم أثناء قيامها بذلك.
Low Predictability	← قامت والدتي بقلع ضرسي بقوة وخفة، حيث أنني لم أشعر بأي تعب أثناء قيامها بذلك.
<i>Translation</i>	My mother gently extracted my tooth, and I did not feel any <u>pain/ discomfort</u> .

Fig. 2 Example sentences in Experiments 1 and 2. Target words are underlined but shown normally in the experiment. Arrows indicate reading direction

Experiment 1, there was a higher probability of a regression back to highly predictable than less predictable target words, but no indication that predictability influenced the probability of a first-pass regression from these words.

General discussion

We report two experiments investigating effects of word predictability in Arabic reading. Both show clear predictability effects in fixations times for words, consistent with findings for other scripts, including Latin-based scripts like English (e.g., Rayner & Well, 1996). However, a predictability effect on word-skipping for short, morphologically simple words in Experiment 2 was not observed for longer words with more varied structures in Experiment 1. Additional analyses that examined these effects as a function of saccade launch site (measured from up to six letter spaces to the right of the target word) showed no indication of a word predictability effect in Experiment 1, and no difference in the size of the word

predictability effect as a function of launch site in Experiment 2. Our findings suggest that, while it is possible to observe predictability effects on word-skipping in Arabic reading, such events are likely to be rare, and possibly restricted to situations in which words are short, morphologically simple, and highly predictable from the prior sentence context.

This finding is in line with the more general observation that word-skipping occurs relatively infrequently in Arabic reading. In the present case, we observed low skipping rates, averaging about 7.5% for target words of variable length and complexity in Experiment 1, and around only 24.5% in Experiment 2 even though target words in this experiment were short and morphologically simple. We note also that the higher skipping rate in Experiment 2 relative to Experiment 1 could relate to differences in the spatial extent of target words across the two experiments. The average skipping rate for Experiment 1 was broadly similar to that reported in other Arabic studies (Hermena et al., 2017, 2019; Paterson et al., 2015). This contrasts with skipping rates for Latin-based scripts like English, which averages around 20–30%, rising to 60% or more for short words. How Arabic word-skipping compares with other Semitic languages, like Hebrew, currently is difficult to establish. One issue is that many Hebrew studies do not report skipping rates (e.g., Dank et al., 2015; Deutsch, 1998; Deutsch et al., 2000, 2003, 2005, 2021; Deutsch & Bentin, 2001; Kuperman & Deutsch, 2020; Nazir et al., 2004; Velan et al., 2013; Yablonski et al., 2017). However, data from one study, which examined word-length effects, showed quite high skipping rates, similar to that for Latin-based scripts, with approximately 30% skipping for short (three-letter) words, falling to about 18% for longer (seven-letter) words (Deutsch & Rayner, 1999). A further

Table 1 Sentence-level measures for Experiments 1 and 2

Measure	Experiment 1	Experiment 2
Reading rate (wpm)	203 (26)	208 (31)
Average fixation duration (ms)	249 (4)	242 (4)
Average number of fixations per word	1.2 (.01)	1.2 (.01)
Number of regressions (%)	13.8 (.9)	13.6 (.9)
Progressive saccade length (chars)	7.3 (.03)	7.2 (.03)

Note. The standard error of the mean is shown in parentheses

Table 2 Target word-level measures for Experiments 1 and 2

Eye-movement variables	Experiment 1		Experiment 2	
	High Predictability	Low Predictability	High Predictability	Low Predictability
Hypothesis-testing variables				
Word-skipping probability (%)	7.2 (1)	7.8 (1)	26.8 (1)	22.1 (1)
First-fixation duration (ms)	248 (2)	260 (2)	234 (2)	252 (3)
Gaze duration (ms)	287 (3)	305 (3)	255 (3)	281 (4)
Exploratory variables				
Single-fixation duration (ms)	258 (3)	277 (3)	240 (3)	263 (3)
Total reading time (ms)	322 (5)	378 (6)	289 (4)	350 (6)
Regressions-out (%)	5.4 (1)	6.6 (1)	8.0 (1)	8.5 (1)
Regressions-in (%)	3.3 (1)	9.2 (1)	4.8 (1)	12.4 (1)

Note. The standard error of the mean is shown in parentheses

issue is that few studies have been conducted in other comparable languages. For example, it could be informative to assess word-skipping for Uyghur, which is an agglutinative Turkic language that uses the Arabic script but does not omit vowel information in words or use the Semitic morphology. However, such comparisons currently are difficult as relatively few studies have been reported for this script, and existing studies report variable skipping rates, ranging from about 8% in a study by Yan et al. (2014) to around 29% in a recent study by Zhou et al. (2021). It nevertheless seems clear from research to date that, as compared with readers of Latin-based languages like English, Arabic readers do not gain efficiency by skipping words, and fixate most words in a sentence at least once.

Table 3 Word-level statistical effects for Experiments 1 and 2

Source		SKIP	FFD	GD
Experiment 1				
(Intercept)	Estimate	-3.12	255.48	297.61
	SE	0.24	5.44	7.37
	t/z	-12.83	46.97	40.41
Predictability	Estimate	0.11	12.66	18.42
	SE	0.26	4.10	5.93
	t/z	0.41	3.09*	3.10*
Experiment 2				
(Intercept)	Estimate	-1.17	243.90	262.52
	SE	0.14	4.61	5.71
	t/z	-8.36	52.93	46.83
Predictability	Estimate	-0.39	17.87	25.30
	SE	0.15	4.58	7.47
	t/z	-2.64*	3.91*	3.39*

*Denotes statistical significance ($t/z > 2.39$)

We argued in the *Introduction* that this low rate of word-skipping might be a consequence of specific orthographic and morphological characteristics of the Arabic script (see also AlJassmi et al., 2021). In particular, we noted that the use of a semi-cursive script with variable letter widths might limit the usefulness of parafoveal length information for constraining word identities. In addition, words might be more difficult to identify parafoveally because of close similarity in the shape of certain groups of letters (e.g., Eviatar et al., 2004). We also noted that readers might not benefit from parafoveal preview of beginning letters in words, as compared to in scripts like English, because Arabic uses a non-concatenative morphology in which the most informative elements (the consonant triple that conveys core word information) seldom appears at the word beginning (Farid & Grainger, 1996). Finally, the agglutinative (or fusional) nature of the Arabic morphology produces complex, informationally dense words that may be difficult to process parafoveally, and may also impose foveal processing costs for fixated words that limit resources for parafoveal processing (e.g., Henderson & Ferreira, 1990; Payne et al., 2016; Roman & Pavard, 1987). Such costs may explain why word-skipping is infrequent and why re-fixation probabilities (whereby words receive multiple fixations) are high (e.g., Paterson et al., 2015).

This raises the question of what information can be processed parafoveally in Arabic reading. This question has received relatively limited attention to date, although there is evidence that readers are sensitive to the use of diacritics to disambiguate words (Hermena et al., 2016). These are glyph-like marks placed above or below Arabic letters to convey vowel information. Diacritics are seldom used unless required to disambiguate words (but are always present in religious texts and widely used in texts designed for beginning readers). The study by Hermena et al. used the boundary paradigm (Rayner, 1975) to make surreptitious changes to parafoveal previews of diacritics ahead of the reader fixating a specific

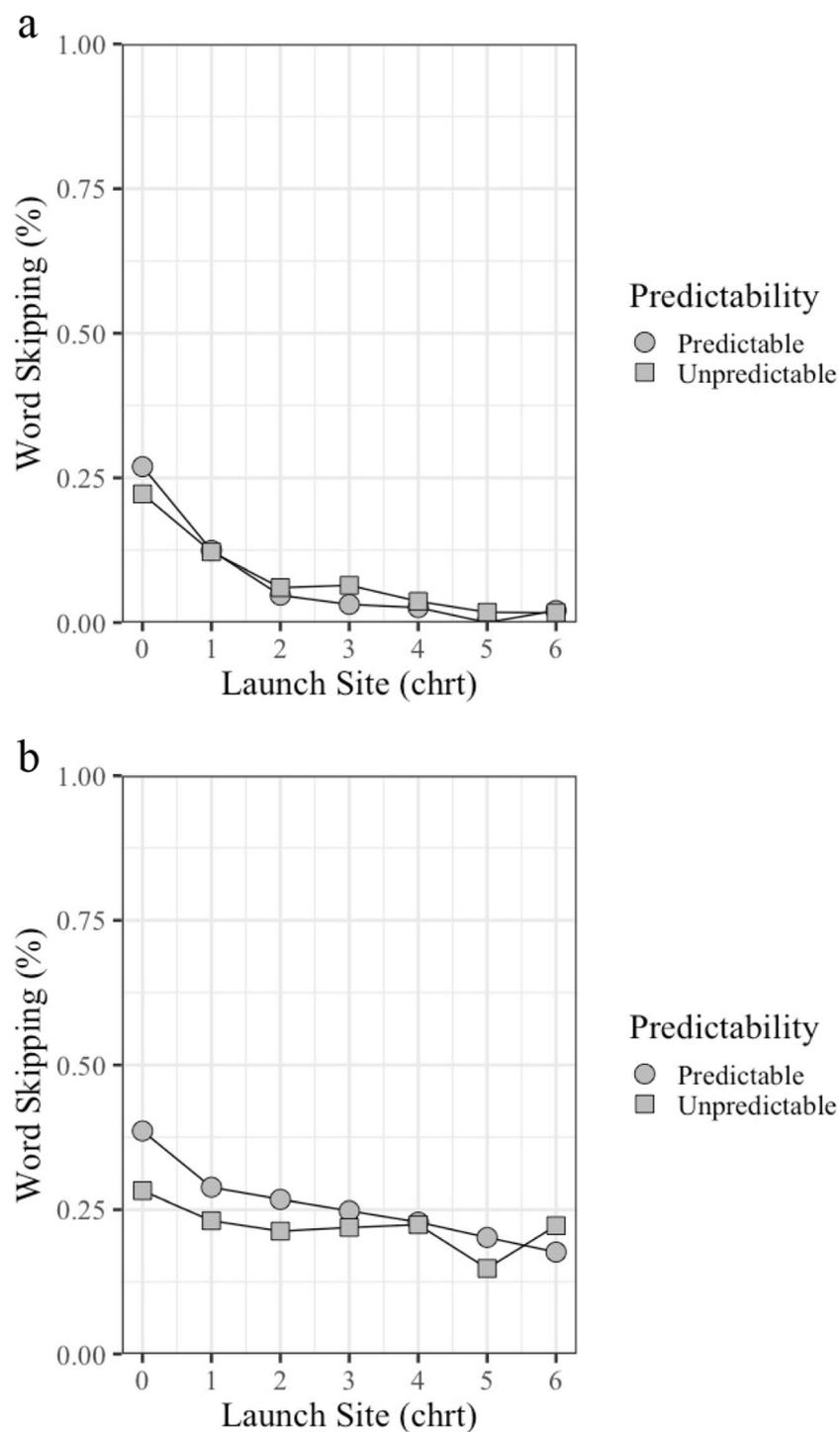


Fig. 3 Effects of Saccade Launch Site on Word-Skipping Probability for (a) Experiment 1 and (b) Experiment 2

target word in each sentence. When this information was inaccurate, it disrupted subsequent processing of the target word, suggesting that this information was available during parafoveal processing. This finding therefore reveals sensitivity to subtle orthographic cues during parafoveal processing.

Moreover, it establishes the boundary paradigm as a method for investigating other aspects of parafoveal orthographic processing in Arabic reading, including the use of cues to word length and the precision with which letter identities can be processed. Moreover, the technique has the capacity to reveal

how contextual constraints and visual and orthographic factors act together to guide eye movements (e.g., Balota et al., 1985; Choi et al., 2017; White et al., 2005).

Similar considerations may apply to morphological factors. Evidence from research in Hebrew suggests readers can acquire quite detailed morphological information during parafoveal processing. For example, several studies using the boundary paradigm show that readers are sensitive to the presence of letter sequences that convey the core meaning of a word (Deutsch et al., 2003, 2005). In Hebrew, as in Arabic and other Semitic languages, words are constructed using a non-concatenative derivational morphology in which a sequence of letters that form the word's root, which captures its core meaning, interleave with other letters in a word (see Deutsch et al., 1998). Studies using the boundary paradigm show that preview of a word containing the same root as a target word, as compared with other shared letters, results in shorter target word fixations (Deutsch et al., 2003), and that the use of morphological preview information can be mediated by context (Deutsch et al., 2005). The indication, therefore, is that detailed morphological information can be acquired parafoveally during Hebrew reading. This raises the question of whether similar information is obtained in Arabic reading. While evidence exists to suggest this might be the case (e.g., Hermena et al., 2019; Hermena et al., 2021), this issue requires further investigation, including needing to establish whether parafoveal processing of morphological information is influenced by contextual constraints. Moreover, this approach might also be used to investigate whether the location of root information can influence processing, such that parafoveal preview effects might be stronger when this information is closer to the beginning of words.

The present findings might also be considered in the context of a debate concerning whether lexical prediction can be observed in the absence of parafoveal information. We argued in the *Introduction* that the decision to skip a word is based on both predictions derived from prior context and parafoveal preview information. This is in line with the view that at least partial parafoveal information is required to support lexical prediction (see Staub, 2015; Staub & Goddard, 2019). Evidence for this view comes from studies using the boundary paradigm to compare effects of valid and invalid previews. In these experiments, an invisible boundary is placed in front of a target word in a sentence. Before the reader's gaze crosses this boundary, the target word is shown as a normal (i.e., a valid) preview or masked (e.g., by replacing the target word's letters with either visually similar or dissimilar letters) to create an invalid preview. As soon as the reader's gaze crosses the boundary, the target word quickly reverts to normal, allowing the experimenter to compare effects of valid and invalid previews on subsequent processing. Crucially, studies using this approach show that predictability effects are obtained only following valid previews or visually similar invalid previews

(e.g., Balota et al., 1985; Staub & Goddard, 2019; also Chang et al., 2020a,b).

This observation has led researchers to argue that lexical prediction facilitates the early processing of a word by pre-activating its features and letters, and that this depends crucially on the availability of parafoveal information (Staub, 2015; Staub & Goddard, 2019). However, an alternative account holds that readers can use lexical prediction to guess upcoming words even in the absence of parafoveal information, and that effects in the boundary paradigm may arise from specific processing costs for invalid previews. Support for this alternative view comes from research showing that predictability effects can be obtained even for words appears at the beginning of a line of text, and therefore in the absence of parafoveal preview (Parker et al., 2017). We have argued that parafoveal processing of Arabic words frequently is impoverished because of specific orthographic and morphological characteristics of the script. Moreover, we showed in Experiment 1, using words varying in length and morphological complexity, that this can eliminate predictability effects in word-skipping. Might these results therefore provide further evidence that lexical prediction can take place in the absence of parafoveal information? We think not as we observed clear predictability effects in fixation times for words. What instead seems likely is that the typically more impoverished parafoveal processing of words in Arabic reading limits the influence of lexical prediction so that it seldom affects word-skipping.

This raises the question of how the present findings might be accommodated by models of eye-movement control. The currently dominant E-Z Reader and SWIFT models were designed to simulate eye movement for Latin-based languages like English and German. It is therefore important to establish the generality of these models by considering their capacity to explain effects in other scripts, including Arabic. The models can accurately simulate effects of key lexical variables (word length, frequency and predictability) on the probability and duration of fixations on words in English and German (e.g., Engbert et al., 2005; Reichle, 2021; Reichle et al., 1998). It seems likely that the models can also account for word length and frequency effects in Arabic reading (Hermena et al., 2017, 2019; Paterson et al., 2015), as well as the word predictability effects in the present study. The E-Z Reader models might account for predictability effects on fixation times by assuming that this influences the familiarity check (i.e., L1 processing) for these words. The model might also account for the limited influence of predictability on word-skipping by assuming that a familiarity check for parafoveal words proceeds more quickly for shorter words that are morphologically simple, such as the target words in Experiment 2 rather than the more varied words used in Experiment 1. In a similar vein, SWIFT might accommodate word predictability effects in fixation times for both types of words, and allow for skipping of

the shorter and morphologically simpler target words in Experiment 2 to be more affected by contextual constraints, as compared to the more varied items in Experiment 1.

Beyond these findings, it should be clear that we currently lack a model of eye-movement control that can account for more specific influences of Arabic orthography and morphology (see Hermena & Reichle, 2020, for further discussion). This necessarily will include mechanisms that can account for effects of the locations of core (i.e., root) information and number of morphological units on the processing of words (AlJassmi et al., 2021; Tibi et al., 2020; Tibi & Kirby, 2017). Consequently, while existing models might account for the lexical effects reported in the present experiments, more sophisticated mechanisms will be required to account for the influence of Arabic orthography and morphology on eye-movement control in reading.

Acknowledgements This research was supported by a PhD research studentship from the UAE Ministry of Education awarded to Maryam AlJassmi.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, 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 changes were made. 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/4.0/>.

References

- AlJassmi, M. A., Hermena, E. W., & Paterson, K. B. (2021). Eye movements in Arabic reading. In D. Nteliheos and T. Leung (Eds.), *Experimental Arabic Linguistics*, 10, 85–108. <https://doi.org/10.1075/sal.10.03alj>
- Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition*, 24, 477–492. <https://doi.org/10.3758/bf03200936>
- Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology*, 17, 364–390. [https://doi.org/10.1016/0010-0285\(85\)90013-1](https://doi.org/10.1016/0010-0285(85)90013-1)
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48.
- Boudelaa, S. (2013). Psycholinguistics. In J. Owens (Ed.), *The Oxford handbook of Arabic linguistics* (pp. 369–391). Oxford University Press.
- Boudelaa, S., & Marslen-Wilson, W. D. (2010). Aralex: A lexical database for Modern Standard Arabic. *Behavior Research Methods*, 42, 481–487. <https://doi.org/10.3758/BRM.42.2.481>
- Brysbaert, M., & Stevens, M. (2018). Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *Journal of Cognition*, 1, 1–20. <https://doi.org/10.5334/joc.10>
- Brysbaert, M., Drieghe, D., & Vitu, F. (2005). Word skipping: Implications for theories of eye movement control in reading. In G. Underwood (Ed.), *Cognitive processes in eye guidance* (pp. 53–78). Oxford, UK: Oxford University Press
- Chang, M., Hao, L., Zhao, S., Li, L., Paterson, K. B., & Wang, J. (2020a). Flexible parafoveal encoding of character order supports word predictability effects in Chinese reading: Evidence from eye movements. *Attention, Perception, & Psychophysics* <https://doi.org/10.3758/s13414-020-02050-x>
- Chang, M., Zhang, K., Hao, L., Zhao, S., McGowan, V. A., Warrington, K. L., ... Gunn, S. C. (2020b). Word predictability depends on parafoveal preview validity in Chinese reading. *Visual Cognition*, 28, 33–40. <https://doi.org/10.1080/13506285.2020.1714825>
- Choi, W., Lowder, M. W., Ferreira, F., Swaab, T. Y., & Henderson, J. M. (2017). Effects of word predictability and preview lexicality on eye movements during reading: A comparison between young and older adults. *Psychology and Aging*, 32, 232–242. <https://doi.org/10.1037/pag0000160>
- Dank, M., Deutsch, A., & Bock, K. (2015). Resolving conflicts in natural and grammatical gender agreement: Evidence from eye movements. *Journal of Psycholinguistic Research*, 44, 435–467. <https://doi.org/10.1007/s10936-014-9291-9>
- Deutsch, A. (1998). Subject-predicate agreement in Hebrew: Interrelations with semantic processes. *Language & Cognitive Processes*, 13, 575–597. <https://doi.org/10.1080/016909698386447>
- Deutsch, A., & Bentin, S. (2001). Syntactic and semantic factors in processing gender agreement in Hebrew: Evidence from ERPs and eye movements. *Journal of Memory and Language*, 45, 200–224. <https://doi.org/10.1006/jmla.2000.2768>
- Deutsch, A., & Rayner, K. (1999). Initial fixation location effects in reading Hebrew words. *Language & Cognitive Processes*, 14, 393–421. <https://doi.org/10.1080/016909699386284>
- Deutsch, A., Frost, R., & Forster, K. (1998). Verbs and nouns are organized and accessed differently in the mental lexicon: Evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 1238–1255. <https://doi.org/10.1037/0278-7393.24.5.1238>
- Deutsch, A., Frost, R., Pollatsek, A., & Rayner, K. (2000). Early morphological effects in word recognition in Hebrew: Evidence from parafoveal preview benefit. *Language & Cognitive Processes*, 15, 487–506. <https://doi.org/10.1080/01690960050119670>
- Deutsch, A., Frost, R., Pelleg, S., Pollatsek, A., & Rayner, K. (2003). Early morphological effects in reading: Evidence from parafoveal preview benefit in Hebrew. *Psychonomic Bulletin & Review*, 10, 415–422. <https://doi.org/10.3758/BF03196500>
- Deutsch, A., Frost, R., Pollatsek, A., & Rayner, K. (2005). Morphological parafoveal preview benefit effects in reading: Evidence from Hebrew. *Language & Cognitive Processes*, 20, 341–371. <https://doi.org/10.1080/01690960444000115>
- Deutsch, A., Velan, H., & Michaly, T. (2018). Decomposition in a non-concatenated morphological structure involves more than just the roots: Evidence from fast priming. *Quarterly Journal of Experimental Psychology*, 71, 85–92. <https://doi.org/10.1080/17470218.2016.1250788>
- Deutsch, A., Velan, H., Merzbach, Y., & Michaly, T. (2021). The dependence of root extraction in a non-concatenated morphology on the

- word-specific orthographic context. *Journal of Memory and Language*, 116. <https://doi.org/10.1016/j.jml.2020.104182>
- Ehrlich, S. F., & Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior*, 20, 641-655. [https://doi.org/10.1016/S0022-5371\(81\)90220-6](https://doi.org/10.1016/S0022-5371(81)90220-6)
- Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, 112, 777-813. <https://doi.org/10.1037/0033-295X.112.4.777>
- Eviatar, Z., Ibrahim, R., & Ganayim, D. (2004). Orthography and the Hemispheres: Visual and Linguistic Aspects of Letter Processing. *Neuropsychology*, 18, 174-184. <https://doi.org/10.1037/0894-4105.18.1.174>
- Farid, M., & Grainger, J. (1996). How initial fixation position influences visual word recognition: A comparison of French and Arabic. *Brain and Language*, 53, 351-368. <https://doi.org/10.1006/brln.1996.0053>
- Fitzsimmons, G., & Drieghe, D. (2011). The influence of number of syllables on word skipping during reading. *Psychonomic Bulletin & Review*, 18, 736-741. <https://doi.org/10.3758/s13423-011-0105-x>
- Frisson, S., Harvey, D. R., & Staub, A. (2017). No prediction error cost in reading: Evidence from eye movements. *Journal of Memory and Language*, 95, 200-214. <https://doi.org/10.1016/j.jml.2017.04.007>
- Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, 140, 186-209. <https://doi.org/10.1037/a0022256>
- Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, 7, 493-498. <https://doi.org/10.1111/2041-210X.12504>
- Hand, C. J., Miellet, S., O'Donnell, P. J., & Sereno, S. C. (2010). The frequency-predictability interaction in reading: It depends where you're coming from. *Journal of Experimental Psychology: Human Perception and Performance*, 36, 1294-1313. <https://doi.org/10.1037/a0020363>
- Hand, C. J., O'Donnell, P. J., & Sereno, S. C. (2012). Word-initial letters influence fixation durations during fluent reading. *Frontiers in Psychology: Language Sciences*, 3, 1-19. <https://doi.org/10.3389/fpsyg.2012.00085>
- Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 417-429. <https://doi.org/10.1037/0278-7393.16.3.417>
- Hermena, E. W., & Reichle, E. D. (2020). Insights from the study of Arabic reading. *Lang & Ling Compass*, 14, e12400. <https://doi.org/10.1111/lnc3.12400>
- Hermena, E. W., Drieghe, D., Hellmuth, S., & Liversedge, S. P. (2015). Processing of Arabic Diacritical Marks: Phonological-Syntactic Disambiguation of Homographic Verbs and Visual Crowding Effects. *Journal of Experimental Psychology: Human Perception and Performance*, 41, 494-507. <https://doi.org/10.1037/xhp0000032>
- Hermena, E. W., Liversedge, S. P., & Drieghe, D. (2016). Parafoveal Processing of Arabic Diacritical Marks. *Journal of Experimental Psychology: Human Perception and Performance*, 42, 2021-2038. <https://doi.org/10.1037/xhp0000294>
- Hermena, E. W., Liversedge, S. P., & Drieghe, D. (2017). The influence of a word's number of letters, spatial extent, and initial bigram characteristics on eye movement control during reading: Evidence from Arabic. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43, 451-471. <https://doi.org/10.1037/xlm0000319>
- Hermena, E. W., Liversedge, S. P., Bouamama, S., & Drieghe, D. (2019). Orthographic and root frequency effects in Arabic: Evidence from eye movements and lexical decision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45, 934-954. <https://doi.org/10.1037/xlm0000626>
- Hermena, E. W., Juma, E. J., & AlJassmi, M. (2021). Parafoveal processing of orthographic, morphological, and semantic information during reading Arabic: A boundary paradigm investigation. *PLoS One*, 16, e0254745. <https://doi.org/10.1371/journal.pone.0254745>
- Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics*, 40, 431-439. <https://doi.org/10.3758/BF03208203>
- Jordan, T. R., & Paterson, K. B. (2009). Re-evaluating split-fovea processing in word recognition: A critical assessment of recent research. *Neuropsychologia*, 47, 2341-2353. <https://doi.org/10.1016/j.neuropsychologia.2008.07.020>
- Jordan, T. R., Almabruk, A. A., Gadalla, E. A., McGowan, V. A., White, S. J., Abedipour, L., & Paterson, K. B. (2014). Reading direction and the central perceptual span: Evidence from Arabic and English. *Psychonomic Bulletin & Review*, 21, 505-511. <https://doi.org/10.3758/s13423-013-0510-4>
- Joseph, H. S., 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, 2078-2086. <https://doi.org/10.1016/j.visres.2009.05.015>
- Juhász, B. J., White, S. J., Liversedge, S. P., & Rayner, K. (2008). Eye movements and the use of parafoveal word length information in reading. *Journal of Experimental Psychology: Human Perception and Performance*, 34, 1560-1579. <https://doi.org/10.1037/a0012319>
- Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, 16, 262-284. <https://doi.org/10.1080/09541440340000213>
- Kuperman, V., & Deutsch, A. (2020). Morphological and visual cues in compound word reading: Eye-tracking evidence from Hebrew. *The Quarterly Journal of Experimental Psychology* <https://doi.org/10.1177/1747021820940297>
- Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences*, 4, 6-14. [https://doi.org/10.1016/s1364-6613\(99\)01418-7](https://doi.org/10.1016/s1364-6613(99)01418-7)
- Miellet, S., Sparrow, L., & Sereno, S. C. (2007). Word frequency and predictability effects in reading French: An evaluation of the EZ Reader model. *Psychonomic Bulletin & Review*, 14, 762-769. <https://doi.org/10.3758/BF03196834>
- Nazir, T. A., Ben-Boutayab, N., Decoppet, N., Deutsch, A., & Frost, R. (2004). Reading habits, perceptual learning, and recognition of printed words. *Brain and Language*, 88, 294-311. [https://doi.org/10.1016/S0093-934X\(03\)00168-8](https://doi.org/10.1016/S0093-934X(03)00168-8)
- Pagán, A., Blythe, H. I., & Liversedge, S. P. (2016). Parafoveal preprocessing of word initial trigrams during reading in adults and children. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42, 411-432. <https://doi.org/10.1037/xlm0000175>
- Parker, A. J., Kirkby, J. A., & Slattery, T. J. (2017). Predictability effects during reading in the absence of parafoveal preview. *Journal of Cognitive Psychology*, 29, 902-911. <https://doi.org/10.1080/20445911.2017.1340303>
- Paterson, K. B., McGowan, V. A., & Jordan, T. R. (2013a). Aging and the control of binocular eye movements during reading. *Psychology and Aging*, 28, 789-795. <https://doi.org/10.1037/a0033328>
- Paterson, K. B., McGowan, V. A., & Jordan, T. R. (2013b). Filtered text reveals adult age differences in reading: Evidence from eye movements. *Psychology and Aging*, 28, 352-364. <https://doi.org/10.1037/a0030350>
- Paterson, K. B., Almabruk, A. A., McGowan, V. A., White, S. J., & Jordan, T. R. (2015). Effects of word length on eye movement control: The evidence from Arabic. *Psychonomic Bulletin & Review*, 22, 1443-1450. <https://doi.org/10.3758/s13423-015-0809-4>

- Payne, B. R., Stites, M. C., & Federmeier, K. D. (2016). Out of the corner of my eye: Foveal semantic load modulates parafoveal processing in reading. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 1839–1857. <https://doi.org/10.1037/xhp0000253>
- Perea, M., Abu Mallouh, R., Mohammed, A., Khalifa, B., & Carreiras, M. (2016). Do diacritical marks play a role at the early stages of word recognition in Arabic? *Frontiers in Psychology*, *7*, 1255. <https://doi.org/10.3389/fpsyg.2016.01255>
- R Core Team (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Ratcliffe, R. R. (2013). Morphology. In J. Owens (Ed.), *The Oxford Handbook of Arabic Linguistics* (pp. 71–91). Oxford University Press.
- Rayner, K. (1975). Parafoveal identification during a fixation in reading. *Acta Psychologica*, *39*, 271–281. [https://doi.org/10.1016/0001-6918\(75\)90011-6](https://doi.org/10.1016/0001-6918(75)90011-6)
- Rayner K. (1979). Eye guidance in reading: fixation locations within words. *Perception*, *8*, 21–30. <https://doi.org/10.1068/p080021>
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422. <https://doi.org/10.1037/0033-2909.124.3.372>
- Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*, *62*, 1457–1506. <https://doi.org/10.1080/17470210902816461>
- Rayner, K., & McConkie, G. W. (1976). What guides a reader's eye movements? *Vision Research*, *16*, 829–837. [https://doi.org/10.1016/0042-6989\(76\)90143-7](https://doi.org/10.1016/0042-6989(76)90143-7)
- Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review*, *3*, 504–509. <https://doi.org/10.3758/BF03214555>
- Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: a comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 1188–1200. <https://doi.org/10.1037/0096-1523.22.5.1188>
- Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 1290–1301. <https://doi.org/10.1037/0278-7393.30.6.1290>
- Rayner, K., Li, X., Juhasz, B. J., & Yan, G. (2005). The effect of word predictability on the eye movements of Chinese readers. *Psychonomic Bulletin & Review*, *12*, 1089–1093. <https://doi.org/10.3758/BF03206448>
- Rayner, K., Slattery, T., Drieghe, D., & Liversedge, S. (2011). Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psychology: Human Perception and Performance*, *37*, 514–528. <https://doi.org/10.1037/a0020990>
- Reichle, E. D. (2021). *Computational Models of Reading: A Handbook*. Oxford University Press. <https://doi.org/10.1093/oso/9780195370669.001.0001>
- Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*, 125–157. <https://doi.org/10.1037/0033-295X.105.1.125>
- Roman, G., & Pavard, B. (1987). A comparative study: How we read in Arabic and French. In J. K. O'Regan & A. Levy-Schoen (Eds.), *Eye movements from Physiology to Cognition* (pp. 431–440). Elsevier. <https://doi.org/10.1016/B978-0-444-70113-8.50064-3>
- Schotter, E. R., Lee, M., Reiderman, M., & Rayner, K. (2015). The effect of contextual constraint on parafoveal processing in reading. *Journal of Memory and Language*, *83*, 118–139. <https://doi.org/10.1016/j.jml.2015.04.005>
- Schwanenflugel, P. J. (1986). Completion norms for final words of sentences using a multiple production measure. *Behavior Research Methods, Instruments, & Computers*, *18*, 363–371. <https://doi.org/10.3758/BF03204419>
- Sereno, S. C., Hand, C. J., Shahid, A., Yao, B., & O'Donnell, P. J. (2018). Testing the limits of contextual constraint: Interactions with word frequency and parafoveal preview during fluent reading. *Quarterly Journal of Experimental Psychology*, *71*, 302–313. <https://doi.org/10.1080/17470218.2017.1327981>
- Staub, A. (2015). The Effect of Lexical Predictability on Eye Movements in Reading: Critical Review and Theoretical Interpretation. *Lang & Ling Compass*, *9*, 311–327. <https://doi.org/10.1111/lnc3.12151>
- Staub, A. (2020). Do effects of visual contrast and font difficulty on readers' eye movements interact with effects of word frequency or predictability? *Journal of Experimental Psychology: Human Perception and Performance*, *46*, 1235–1251. <https://doi.org/10.1037/xhp0000853>
- Staub, A., & Benatar, A. (2013). Individual differences in fixation duration distributions in reading. *Psychonomic Bulletin & Review*, *20*, 1304–1311. <https://doi.org/10.3758/s13423-013-0444-x>
- Staub, A., & Goddard, K. (2019). The role of preview validity in predictability and frequency effects on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*, 110–127. <https://doi.org/10.1037/xlm0000561>
- Taylor, W. L. (1953). "Cloze procedure": A new tool for measuring readability. *Journalism Quarterly*, *30*, 415–433. <https://doi.org/10.1177/107769905303000401>
- Tibi, S., & Kirby, J. R. (2017). Morphological awareness: Construct and predictive validity in Arabic. *Applied PsychoLinguistics*, *38*, 1019–1043. <https://doi.org/10.1017/S0142716417000029>
- Tibi, S., Edwards, A. A., Schatschneider, C., & Kirby, J. R. (2020). Predicting Arabic word reading: A cross-classified generalized random-effects analysis showing the critical role of morphology. *Annals of Dyslexia*, *70*, 200–219. <https://doi.org/10.1007/s11881-020-00193-y>
- Velan, H., Deutsch, A., & Frost, R. (2013). The flexibility of letter-position flexibility: Evidence from eye movements in reading Hebrew. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 1143–1152. <https://doi.org/10.1037/a0031075>
- Veldre, A., & Andrews, S. (2018). How does foveal processing difficulty affect parafoveal processing during reading?. *Journal of Memory and Language*, *103*, 74–90. <https://doi.org/10.1016/j.jml.2018.08.001>
- Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). Springer-Verlag.
- von der Malsburg, T., & Angele, B. (2017). False positives and other statistical error in standard analyses of eye movements in reading. *Journal of Memory and Language*, *94*, 119–133. <https://doi.org/10.1016/j.jml.2016.10.003>
- White, S.J., Rayner, K. & Liversedge, S.P. (2005). The influence of parafoveal word length and contextual constraint on fixation durations and word skipping in reading. *Psychonomic Bulletin & Review* *12*, 466–471. <https://doi.org/10.3758/BF03193789>
- Yablonski, M., Polat, U., Bonneh, Y. S., & Ben-Shachar, M. (2017). Microsaccades are sensitive to word structure: A novel approach

- to study language processing. *Scientific Reports*, 7, 1-11. <https://doi.org/10.1038/s41598-017-04391-4>
- Yan, M., Zhou, W., Shu, H., Yusupu, R., Miao, D., Krügel, A., & Kliegl, R. (2014). Eye movements guided by morphological structure: Evidence from the Uighur language. *Cognition*, 132, 181-215. <https://doi.org/10.1016/j.cognition.2014.03.008>
- Zhou, W., Wang, A., & Yan, M. (2021). Eye movements and the perceptual span among skilled Uighur readers. *Vision Research*, 182, 20-26. <https://doi.org/10.1016/j.visres.2021.01.005>

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