Parafoveal Processing of Arabic Diacritical Marks

Ehab W. Hermena, Simon P. Liversedge, and Denis Drieghe

University of Southampton

Author Note

Ehab W. Hermena, Denis Drieghe, and Simon P. Liversedge, Center for Vision and Cognition, Psychology, University of Southampton, Southampton, England.

Correspondence regarding this article should be addressed to Ehab W. Hermena, Centre for Vision and Cognition, School of Psychology, Shackleton Building, University of Southampton, Highfield Campus, Southampton, SO17 1BJ, UK. Tel: +44(0)2380 595078, E-mail: ewh1g09@soton.ac.uk

This research was supported by Leverhulme Trust Research Grant RPG-2013-205.

Abstract

Diacritics are glyph-like marks on letters that convey vowel information in Arabic, thus allowing for accurate pronunciation and disambiguation of homographs. For skilled readers, diacritics are usually omitted except when their omission causes ambiguity. Undiacritized homographs are very common in Arabic and are predominantly heterophones (where each meaning sounds different), with one version more common (dominant) than the others (subordinate). In this study we investigated parafoveal processing of diacritics during reading. We presented native readers with heterophonic homographs embedded in sentences with diacritization that instantiated either dominant or subordinate pronunciations of the homographs. Using the boundary paradigm, we presented previews of these words carrying either: identical diacritization to the target; inaccurate diacritization, such that if the target had dominant diacritization, the preview contained subordinate diacritization, and vice versa; or no diacritics. The results showed that readers processed the identity of diacritics parafoveally, such that inaccurate previews of the diacritics resulted in inflated fixation durations, particularly for fixations originating at close launch sites. Moreover, our results clearly indicate that readers' expectation for dominant or subordinate diacritization patterns influences their parafoveal and foveal processing of diacritics. Specifically, a perceived absence of diacritics (either in no-diacritics previews, or because the eyes were too far away to process the presence of diacritics) induced an expectation for the dominant pronunciation, whereas the perceived presence of diacritics induced an expectation for the subordinate meaning.

Keywords: diacritics; reading Arabic; eye movements; parafoveal processing

Public Significance Statement

We investigated skilled readers' processing of diacritics (symbols that convey word pronunciation information) while reading Arabic sentences. It has been argued that for scientists and educators to achieve a better understanding of skilled reading, universally, studying reading in more world languages is necessary. We used an advanced method to record where readers' eyes fixated, and for how long, during sentence reading. Readers' eye movements are known to be directly related to the cognitive processes underlying written language comprehension. The results suggested that, subject to the distance between the previous fixation and the diacritized word, readers were able to process the diacritics prior to actually fixating them. Also, subject to the partial information available about the diacritics from that distance, readers' expectation for a particular pattern of diacritics to be present influenced the speed with which they processed diacritized words. The study provided new important insight into phonological processing during reading.

(150 Words)

Substantial evidence from eye movement investigations in reading has established that during a fixation, readers process the fixated word as well as pre-processing the upcoming word. Given that typically upcoming words fall outside foveal vision, pre-processing of such words is referred to as *parafoveal processing* (for reviews see Rayner, 1998; 2009; Schotter, Angele, & Rayner, 2012). Investigations of parafoveal processing have utilized the influential boundary paradigm (Rayner, 1975), where an invisible boundary is inserted into the text typically immediately before a target word. Prior to crossing this boundary, the reader is presented with a preview of the upcoming word that may, or may not, be identical to the target word, or that may share certain linguistic characteristics with the target word (e.g., phonological, *beech* – *beach*). The display changes while the reader's eyes move across the invisible boundary towards the target word, and the target word is displayed correctly when the reader fixates it. Importantly, the reader is typically unaware of the display change because of the suppression of vision during saccades (Matin, 1974). Experiments clearly show that when readers are given a valid (i.e., identical) parafoveal preview of the upcoming word (e.g., *beach* as preview of *beach*), fixation durations on this word, once it is fixated, are reduced—the so-called *preview benefit*, compared to when the previews are not valid (e.g., the string *dmaeb* as a preview of *beach*, e.g., Rayner, 1975; Rayner, Well, Pollatsek, & Bertera, 1982).

The boundary paradigm allows researchers to investigate the types of information readers extract from parafoveal words prior to their fixation. Indeed, investigations in many languages have shown that giving readers parafoveal previews which share orthographic and/or phonological information with the target results in preview benefits, relative to when previews lack such information (e.g., Ashby, Treiman, Kessler, & Rayner, 2006; Henderson, Dixon, Petersen, Twilley, & Ferreira, 1995; Miellet & Sparrow, 2004; Pollatsek, Lesch, Morris, & Rayner, 1992; Rayner, McConkie, & Ehrlich, 1978). Other investigations have shown that world languages differ in the extent to which semantic or syntactic information can be accessed parafoveally. For instance, some investigations have reported preview benefits when the preview shared semantic information with the target in Chinese (e.g., Yan, Richter, Shu, & Kliegl, 2009; Yan, Zhou, Shu, & Kliegl, 2012; Yang, Wang, Tong, & Rayner, 2010; Yang, Wang, Xu, & Rayner, 2009), and in German (e.g., Hohenstein, & Kliegl, 2014; Hohenstein, Laubrock, & Kliegl, 2010). In the investigations conducted in German, for instance, previews that were semantically related to the target words resulted in preview benefits (e.g., *Schädel*, meaning *skull*, as a preview for *Knochen*, meaning *bones*), relative to previews that were not semantically related although orthographically similar (e.g., *Stiefel*, meaning *boots*, see Hohenstein, & Kliegl, 2014). In English however, Schotter (2013) found semantic preview benefits only when the preview and target were synonymous words (e.g., *video* as preview of *movie*), not when the preview was merely related semantically to the target (e.g., *audio* as preview of *movie*). In this respect, Schotter's findings are in line with previous investigations that reported no semantic preview benefit for semantically related words in English (e.g., *ocean – river* in Rayner, Balota, & Pollatsek, 1986; see also Rayner, Schotter, & Drieghe, 2014).

As for syntactic processing in the parafovea, recent work in Korean reported preview benefits when the preview was a correct syntactic match of the target (Korean contains orthographic markers that convey whether the word is the subject or object of the sentence), compared to when the preview was a syntactic mismatch (Kim, Radach, & Vorstius, 2012). However, the very limited number of investigations conducted on syntactic parafoveal processing in English have indicated that readers do not use syntactically disambiguating parafoveal information, at least for reduced relative clause sentences (Clifton, Traxler, Mohamed, Williams, Morris, & Rayner, 2003).

What such studies investigating alphabetic language processing have in common is that the parafoveal preview manipulations typically involved changes in the letters of the preview relative to the target. Parafoveal processing of other linguistic units, such as diacritical marks (diacritics hereafter for short), has not been studied (with the exception of the Korean study examining orthographic markers indicating syntax, Kim et al, 2012). The study we report here is the first to investigate parafoveal processing of diacritics using the boundary paradigm. Diacritics are glyph-like marks that mainly add vowel sound information for instance in Hebrew and Arabic. In both these Semitic languages the vast majority of words consist of consonants only (see Abu-Rabia, 1999; 2001; Shany, Bar-on, & Katzir, 2012). Diacritics can also modify the pronunciation of vowel sounds in other languages (e.g., the umlaut in German, e.g., *fallen* vs. *fällen*; and also in English words from other origins such as *naïve* from French). Here we report an investigation of parafoveal processing of Arabic diacritical marks.

As mentioned above, Arabic words are predominantly composed of consonants (Haywood & Nahmad, 1965; Schulz, 2004). Although the letter-sound translations for Arabic consonants are transparent, that is, each consonant is associated with the same sound all the time (e.g., ك = /k/, and ت = /t/), the exact pronunciation of a consonant string depends on how each consonant is vowelized (e.g., /k^{a}/ vs. /k^{o}/; or /t^{a}/ vs. /t^{o}/ vs. /t^{i}/ in the Arabic string كتب /ktb/ which can be pronounced as: /k^{o}t^{o}b/, or /k^{a}t^{a}b^{a}/, or /k^{o}t^{i}b^{a}/, etc., where ^{superscript} indicates the vowel information added by diacritics). Thus, the letter string كتب /ktb/ would carry the following diacritization patterns to make the three different pronunciations above, respectively: كُتُبْ, كَتَبَ, كُتِبَ. In other words, the undiacritized string كتب /ktb/, as a single word, is an ambiguous homograph. Such words in Arabic can have multiple pronunciations (i.e., they are heterophones), and different semantic and syntactic representations associated with each pronunciation (e.g., for the three versions of كتب above, respectively: كُتُبْ /k^{o}t^{o}b/ means *books*, plural noun; كَتَبَ /k^{a}t^{a}b^{a}/ means [*he*] *wrote*, past masculine verb; and كُتِبَ /k^{u}t^{i}b^{a}/ means [*was*] *written*, past passive verb). This type of homographic word is very common in Arabic: Every second or third word in ordinary text is such a homograph (e.g., Abu-Rabia 1997a; 1998).

Fully diacritized Arabic texts ordinarily appear in religious works, educational books (for learners up to 9-10 years old), and in texts where spoken accuracy is important (e.g., poetry, Haywood & Nahmad, 1965; Schulz, 2004). In these texts, all words are diacritized. However, diacritics are, predominantly, not printed in other day-to-day modern Arabic texts. Rather, readers become skilled in using the text's context and syntactic structure to disambiguate homographs (Abu-Rabia, 1997a; 1997b; 1998). The exception is that diacritics are added to some individual ambiguous words in the text, if the surrounding text does not adequately disambiguate them (see Hermena, Drieghe, Hellmuth, & Liversedge, 2015; Schulz, 2004).

Surveying ambiguous homographic words in Arabic and the use of diacritics in print, Hermena et al. (2015) indicated that the vast majority of Arabic ambiguous homographic words are *biased homographs* (see e.g., Rayner & Duffy, 1986; Sereno, O'Donnell, & Rayner, 2006). Essentially, the multiple pronunciations of the Arabic homographs are not *equally* commonly encountered, or produced, by readers. Also, note that each of the multiple pronunciations of Arabic homographs (more than seven different pronunciations in some instances) can be associated with different semantic and syntactic representations (e.g., the different meanings and grammatical cases associated with the different pronunciations of the string كتب /ktb/ mentioned above). An experimental pre-screen procedure conducted as part of the experimentation reported below (see stimuli norming section), confirmed that some word pronunciations were more frequently encountered in print than others. Additionally, these pronunciations were more frequently generated by readers when asked to add diacritics to an ambiguous single word, and when asked to place the ambiguous word in a sentence that clarifies its pronunciation and meaning. We refer to these more frequent pronunciations as *dominant*, whereas the less frequently encountered or generated pronunciations as *subordinate*. We also refer to the diacritization patterns that represent these pronunciations as dominant or subordinate diacritization patterns, respectively. To illustrate, the string قدر /qdr/ has five common pronunciations (i.e., pronunciations that are used in modern language; not obsolete or archaic). Of these pronunciations, the version قَدَرٌ /q^{a}d^{a}r^{un}/ (noun, singular, masculine, meaning *fate*) occurs more frequently in text, and is generated considerably more by producers than, for instance, the pronunciations قَدْرٌ /q^{a}dr^{un}/ (noun, singular, masculine, meaning *amount* or *value*), or قِدْرٌ /q^{i}dr^{un}/ (noun, singular, masculine, meaning *vessel* or *container*). Of these three pronunciations, the final one is the least often encountered and produced by the readers sampled in our pre-screening.

As mentioned above, diacritics are added to some individual ambiguous words in printed text, in principle, only if the surrounding text does not adequately disambiguate them, regardless of whether the dominant or subordinate pronunciations are instantiated by the text (see Hermena et al., 2015; also Schulz, 2004). However, our surveys clearly indicated that in printed modern Arabic text diacritics are mostly added to the homograph to point the reader towards one of its subordinate pronunciations in a non-constraining context (Hermena et al., 2015). Thus, in printed modern Arabic text readers encounter: (i) non-diacritized homographs that are clearly disambiguated by the surrounding text as the dominant version; (ii) non-diacritized homographs that are clearly disambiguated by the surrounding text as the subordinate version; or (iii) diacritized homographs that are not disambiguated by the surrounding text as the subordinate version. The fourth possibility—diacritized homographs that are not disambiguated by the surrounding text as the dominant version, is encountered very close to never.

Moreover, if the word has multiple subordinate pronunciations (such as the current example قدر /qdr), printed diacritics in text would typically point the reader towards the correct pronunciation, that is most likely to be the subordinate pronunciation that best fits the text context and structure of the sentence. For instance, in addition to the three pronunciations presented above for the string قدر /qdr/, other subordinate pronunciations include: قَدَّرَ /q^{a}d^{da}r^{a}/ which is a past tense, masculine, active voice verb, meaning *[he] estimated/destined*; and قُدِّرَ /q^{o}d^{di}r^{a}/ a past tense, masculine, passive voice verb, meaning *[was] estimated / destined*. The actual subordinate diacritization pattern that would appear on the string قدر /qdr/ in a sentence (e.g., the noun version meaning *vessel*, or the verb version *estimated/destined*), will be the one which best fits the syntactic structure and context of the sentence. Indeed, constructing a comprehensible Arabic sentence where structure and context do not constrain the reader towards a smaller number of possible alternative pronunciations to choose from would be nearly impossible. In the example قدر /qdr/, the sentence structure would ordinarily rule out either the verb, or noun interpretations. Thus the ambiguity of the homograph is reduced somewhat given that the number of plausible representations becomes limited (e.g., the three noun pronunciations, with /$q^ad^ar^{un}$/ being the dominant; or the two verb versions, with the active voice pronunciation being the dominant, see Hermena et al., 2015; Schulz, 2004).

Processing of Arabic diacritics has been studied in text reading aloud, silent reading comprehension, and single word naming tasks (e.g., Abu-Rabia, 1997a; 1997b; 1998; 1999; 2001). Abu-Rabia (1997a; 2001), for instance, reported that the presence of diacritics in text resulted in improved accuracy of reading aloud, as well as reading comprehension. Additionally, a small number of eye movement investigations have examined processing diacritics during silent reading (Hermena et al., 2015; Roman & Pavard, 1987). Hermena et al. investigated the processing of diacritics that disambiguated homographic verbs as either active or passive. Their findings clearly showed that readers are sensitive to the presence of diacritics prior to fixating the diacritized word such that they skipped the upcoming word significantly less when it was diacritized, compared to when it was not. Furthermore, processing the diacritics on a target verb during first pass reading did not increase fixation durations on those verbs compared to their non-diacritized form. Hermena et al. also found that readers were successful in making use of diacritics to disambiguate the target verb as passive, however this was contingent on the mode of diacritization. Essentially, when the homographic verb was the only diacritized word in the sentence, the readers successfully disambiguated the target verb as passive. However, when diacritics were added to all words in the passive sentence, a relatively uncommon situation for normal reading, as indicated above, the readers failed to make use of the disambiguating diacritics on the verb. The results suggested that skilled readers do not process (mostly-redundant) full sentence diacritics, and in this situation opt to rely on sentence context and structure to disambiguate any present homographs. Additionally, in fully-diacritized active sentences, the only cost found for the presence of the full sentence diacritization was a small (6ms) increase in average fixation duration, relative to the non-diacritized active sentences. This small effect was statistically significant and was attributed to the increased visual and/or informational density in the fully diacritized condition. The absence of any evidence that readers engaged in detailed phonological processing of full sentence diacritics was interpreted as a cognitive resource-saving strategy.

As mentioned, parafoveal processing of diacritics remains understudied. Apart from the finding discussed above that diacritized parafoveal target words were skipped less than non-diacritized words (Hermena et al., 2015), the extent to which readers process upcoming diacritics remains unknown. The study reported here investigated parafoveal processing of Arabic diacritics, that is, prior to fixating the diacritized target homographic word. All the target homographs were embedded in sentences where the preceding text constrained the readers towards a small number of plausible alternative versions of the target homograph, but did not completely disambiguate which version of the homograph was present (i.e., the dominant or a subordinate version). Thus we ensured that the use of diacritics in all sentences was ecologically valid according to the principle that diacritics are added to disambiguate homographs that are embedded in text that does not fully disambiguate them. The target homographs were given diacritics of either dominant or subordinate pronunciations. As is detailed below, we employed pre-screening procedures to allow us to learn the dominant and subordinate representations for each of the target homographs. These procedures included production of possible representations of the homographs (indicating lexical availability), and frequency of occurrence in text. Our approach was thus pragmatic, and did not follow any particular theoretical rationale as to how access to dominant and subordinate representations of homographs occurs lexically during processing in Arabic. The pattern of diacritization corresponding to the most encountered and produced pronunciation of the homograph was designated as the dominant diacritization pattern, and the pattern of diacritization corresponding to the least encountered and produced pronunciation of the homograph was designated as the subordinate diacritization pattern (i.e., we chose the most, and the least available representations associated with the word, in an attempt to maximise the effectiveness of our experimental manipulation).

With previous evidence suggesting that readers are sensitive to the presence of diacritics in the parafovea as was apparent in the decreased skipping rates of diacritized versus undiacritized words in our previous study (Hermena et al., 2015), we aimed to expand these findings in the current investigation. We aimed to establish whether, besides being sensitive to the presence or absence of diacritics, readers actually identify the diacritics parafoveally. If readers do identify diacritics parafoveally, then we would expect processing benefits, manifesting as reduced fixation durations on the target homographs, when the readers have an identical parafoveal preview of the diacritics, relative to when the preview is inaccurate.

In addition, we aimed to learn whether processing of dominant parafoveal diacritization patterns might result in greater facilitation (or potentially, cost), relative to processing of subordinate parafoveal diacritization patterns. It seems reasonable to hypothesize that if readers identify patterns of parafoveal diacritization, then the presence of a dominant pattern might well result in processing facilitation, relative to a subordinate pattern. This would be in line with the widely accepted findings for frequency-mediated processing of semantically ambiguous words, for example, where processing facilitation is obtained for more frequently occurring meanings (see reviews in Hyönä, 2011; Juhasz & Pollatsek, 2011; Rayner, 1998; 2009). To be clear, findings from non-reading tasks (e.g., cross-modal priming) show that, for biased homographic words, such as our targets, with multiple semantic representations, these representations are accessed in the order of frequency (e.g., Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982; see also Simpson & Burgess, 1985). Additionally, multiple researchers have argued that during text reading, readers treat subordinate versions of ambiguous words as low frequency words (e.g., Sereno et al., 2006; see also Reichle, Rayner, & Polatsek, 2003; Sereno, Brewer, & O'Donell, 2003; Sereno, Pact, & Rayner, 1992), that is, they are more costly to access and process, and the subordinate versions are activated later than the dominant version of the same word[1].

Alternatively, it is possible that the presence of diacritics in the parafovea, *per se*, may alter readers' performance. Recall that: (i) the target homographic words are placed in a partially-constraining context which supports both the dominant and the subordinate version of the homograph, and (ii) the presence of diacritics in print, as discussed above, mostly guides the readers towards one of the subordinate, pronunciations of the word (Hermena et al., 2015). As such, the presence of diacritics in the parafovea might plausibly alert the reader to expect that the upcoming word would have a subordinate pronunciation. In other words, the mere presence of diacritics in the parafovea may guide the readers towards *expecting* subordinate diacritization to be present. If this is the case, we could expect processing facilitation for the *expected* subordinate diacritization patterns, relative to the dominant. Such results would be theoretically very interesting because they would suggest that parafoveal (and foveal) processing of diacritics is not frequency-mediated. Rather, when diacritics are perceived in the parafovea, frequency-mediated processing is suspended, or overridden, by an expectation for a subordinate interpretation of the word.

Additionally, we can predict that inaccurate previews of diacritics may result in processing costs if readers do identify diacritics parafoveally. Costs of inaccurate previews should be observed for both dominant and subordinate diacritics. Furthermore, these costs should reduce, or completely mask, any processing benefits observed for: (i) dominant diacritics, if processing of diacritics is frequency-mediated whereby dominant diacritics are easier and faster to process than subordinate diacritics; and (ii) subordinate diacritics, if the processing of diacritics is influenced by sensitivity to the presence of diacritics and there is an expectation for a subordinate pronunciation to be present.

To investigate these hypotheses concerning parafoveal processing of diacritics, we presented readers with target words that either carried the dominant or a subordinate diacritization pattern. These target words were embedded in frame sentences, and we used the boundary paradigm (Rayner, 1975) to manipulate the parafoveal preview of the diacritics available to the readers prior to fixating the target word. Specifically, we presented the readers with parafoveal previews of the diacritics which were either identical; inaccurate previews; or previews which contained no diacritics (see sample stimuli in Figure 1). Thus, we manipulated two independent variables; target word diacritization (dominant, subordinate) and preview availability of diacritics (identical, inaccurate, and no-diacritics).

Another variable which we decided to include *a priori* in our analyses was the launch site, or the distance between the location of the fixation prior to fixating the diacritized target word and the beginning of the region which contained the target word. Our main reason for including launch site in our analyses, as we detail below, is the fact that prior literature suggests that the quality of parafoveal processing is modulated by launch site, with better parafoveal processing for closer launch sites (e.g., Fitzsimmons & Drieghe, 2011). We predicted that this would apply particularly to parafoveal processing of diacritics given their smaller size compared to letters.

<Insert Figure 1 about here>

To be explicit, our hypotheses were: (i) We expected that skipping would be reduced following parafoveal previews which contained diacritics (i.e., previews containing either identical or inaccurate diacritics), compared to when the previews contained no diacritics. This would be in line with our previous findings (Hermena et al., 2015) and further supports the suggestion that readers are sensitive to the presence of diacritics parafoveally. (ii) If readers initiate pre-processing of the identity of diacritics parafoveally, then, in line with previous literature, identical previews of the diacritics would result in preview benefit once the diacritized word was fixated. By contrast, inaccurate previews would result in processing costs (increased fixation durations). (iii) As for the no-diacritics previews, we predicted that readers may expect the pronunciation of the upcoming word to be the dominant one, as is the case most of the time in their reading experience. Thus, we predicted that if the subordinate diacritics were present on the target following a no-diacritics preview, readers may have initially misprocessed the target word to some degree and this would result in processing costs. (iv) The exact pattern of results obtained, that is, whether there was a processing facilitation for dominant over subordinate diacritization, or vice versa would depend on the nature of processing of parafoveal diacritics. Specifically, the direction of the effect would depend on whether readers do identify diacritics parafoveally or not, and whether processing of the diacritics is frequency-mediated. Potentially, the presence of parafoveal diacritics might result in frequency-mediated processing being overridden, thereby signalling to the reader to expect a subordinate diacritization.

We can also make explicit hypotheses regarding the role of how launch site may influence processing of diacritics. (v) We expect that any effects obtained will be amplified for closer launch sites. To be specific, it is plausible that if readers do identify diacritics parafoveally, their influence will be greatest at close launch sites given visual acuity limitations. Similarly, readers' expectations about upcoming diacritics may be altered depending on launch site. This is because at far launch sites, readers may have no clear preview, or a highly degraded one, of the upcoming diacritics. Under such conditions, readers may expect that the upcoming word is not diacritized, and thus expect the word to have a dominant pronunciation (similar to our predictions about the no-diacritics preview condition). On the other hand, at closer launch sites, where preview permits perception of the presence of upcoming diacritics, readers' expectations may shift towards a subordinate analysis of the upcoming word, given their experience of printed diacritics predominantly pointing towards subordinate pronunciations of homographs. (vi) Finally, if readers are able to not only perceive the presence or absence of diacritics but also to process their identity (again, this would be more likely at close launch sites), one more issue can be investigated. Namely, it remains to be seen whether any effect of expectation (for the subordinate diacritics) would remain, or would be undone by identifying the diacritics at close launch site. If the latter scenario is the case, then identical previews should result in comparable facilitation for both dominant (not expected) and subordinate (expected) diacritics.

In addition to investigating parafoveal processing of diacritics at the target word region, we will also, for the sake of completeness, explore whether previews of the diacritics influence processing of the pre-target word (so-called parafoveal-on-foveal effects reported in investigations not involving diacritics manipulations, see Inhoff, Starr, & Shindler, 2000; Pynte, Kennedy, & Ducrot, 2004; Rayner, Warren, Juhasz, & Liversedge, 2004; Starr & Inhoff, 2004; also Drieghe, 2011 for review). Additionally, we also explore whether effects of processing the disambiguating diacritics on the target word spill over into the post-target region as has been reported in previous investigations that, again, did not involve manipulations of diacritics (e.g., Frazier & Rayner, 1987; 1990; Pickering & Frisson, 2001; Rayner et al., 2006).

Method

Participants

Thirty-six adult native Arabic speakers were paid £15 for participation. All participants were UK residents or visitors (e.g., international students). The participants (23 females) ranged in age between 18 and 47 (mean = 32.5, SD = 8.7). All participants had normal or corrected vision, and all reported being able to clearly see the words and diacritics on the screen during a practice block. The majority of participants spoke and read English as a second language. All participants read Arabic text regularly (on daily or weekly basis). Although the participants knew that we were investigating reading in Arabic, they were naïve as to the exact purpose of the experiment.

Stimuli

Fifty-four sets of target sentences were constructed, each of which contained two frame sentences, one with the target word carrying the dominant diacritization, and the other carrying a subordinate pattern (see Figure 1). In all stimuli sets, the frame sentences were identical until the target word, after which the sentences differed to suit either the dominant or subordinate versions of the target. In 13 of the 54 sets, both dominant and subordinate versions of the target word were nouns; in the remaining sets they were verbs. The target words had an average of 4 different pronunciations (SD = 1.4, range = 2 – 7, mode = 4). Given the partial sentential (syntactic) constraint of the sentence up to the target word, each of the target words had one dominant pronunciation, and on average 2 plausible subordinate pronunciations (mean = 1.5, SD = 0.6, range = 1 – 3, mode = 1). Note that the preceding sentential context did not constrain towards the dominant or any of the plausible subordinate pronunciations of the target. The process of selecting the dominant and subordinate diacritization patterns for each of the target words is detailed below in the stimuli norming section.

In all experimental sentences, the invisible boundary (dashed line in Fig. 1) was placed immediately before the space preceding the target word. Prior to crossing this boundary, the readers had access to a parafoveal preview of the target word with identical diacritics, inaccurate diacritics, or no-diacritics. The inaccurate preview was basically the opposite diacritization pattern, that is, for targets with the dominant diacritization, the inaccurate preview corresponded to the subordinate diacritization pattern, and vice versa. Following crossing the boundary, the target word was always displayed with its correct diacritization pattern.

Eighty-five filler sentences of similar length and complexity to the target sentences were also presented to the participants. Eleven additional sentences made up a practice block, thus each participant read 150 sentences in total.

All sentences were written and displayed on a single line and in natural cursive script. We used a commonly available and widely used proportional font (Traditional Arabic, size 18, which is comparable in size to English text in Times New Roman font size 14).

Stimuli norming.

The target words in the sentences had a mean orthographic frequency of 124.9 per million (SD = 217.9, range = 0.18 – 1130.05) in the Aralex corpus (Boudelaa & Marslen-Wilson, 2010). However, this corpus does not contain any information as to the dominant or subordinate word pronunciations (i.e., diacritization patterns). To determine the dominant and the subordinate patterns of diacritization for each of the 54 target words used, we adopted 3 norming steps. In the first step we presented a set of single ambiguous words (135) to native Arabic readers (Amazon Mechanical Turkers, AMTs), who did not take part in the eye tracking experiment, and we asked them to place diacritics on these words. We obtained 15 different responses for each of the words. The pronunciation designated as dominant was always the one that was used in the majority of the AMTs' responses, with the proviso that it should be used no less than twice as much as the version selected as subordinate. The pronunciation designated as the subordinate was always the least used in the AMTs responses, and from the same syntactic class (verb or noun) as the dominant pronunciation.

In the second step, we asked another set of AMTs to create sentences, each containing one of the words. Given that in sentences these ambiguous words would be disambiguated towards the meaning intended by the writer, we took this as an index of the dominant and subordinate pronunciations of these words. We obtained 15 different sentences for each target word. Similar to the first step, a pronunciation was designated as dominant when it was used in the absolute majority of the AMTs' responses, at least twice as much as the version selected as subordinate. The subordinate pronunciation was also the least used by the AMTs, from the same syntactic class as the dominant pronunciation. At the end of this stage we obtained 79 words where both norming steps were in agreement.

In the final step, we used the first 100 hits from a Google search for each one of the 79 words. The number of times, out of a 100, a certain pronunciation of each word was present in the Google hits was taken as an additional index as to which pronunciation was dominant, and which subordinate. The dominant pronunciation appeared at least twice as frequently in the Google hits as the subordinate pronunciation, and both versions were from the same syntactic class. The 54 words used in the current experiment were the ones where all three norming procedures were in agreement as to which pronunciation was dominant, and which was subordinate. For the final 54 target words selected in the norming procedure described above, dominant diacritization patterns were given in the single word diacritization step, on average, 69% of the time (SD = 15.4, range = 53 – 87%), compared to subordinate pronunciations which appeared only in 21% of the time (SD = 9.7, range = 7 – 40%). In the sentence generation step, dominant pronunciations were used in sentences, on average, 67% of the time (SD = 15.9, range = 40 – 67%), compared to subordinate pronunciations which appeared only in 15% of the time (SD = 8.4, range = 7 – 33%).

Finally, in the Google 100 hits, the dominant pronunciation was present, on average, in 71% of the first 100 hits (SD = 23.6, range = 23 – 99%) compared to the subordinate pronunciation which was present, on average, in 8% of the first 100 hits (SD = 9, range = 1 – 38%).

In addition, we obtained 10 cloze predictability ratings for the target word within each sentence. In this procedure, 10 participants were given sentences up to, but not including, the target word, and were asked to complete the sentence. If participants produced any of the target words to continue the sentence, this was taken as an indication that the target word was predictable given the context of the sentence. With the exception of one sentence, none of the target words, in either of their dominant or subordinate versions, were produced by the AMT raters. The sentence where one version was predictable was changed, and re-norming revealed that the target word was no longer predictable. Finally, we obtained 10 ratings from 10 AMTs as to the naturalness of the sentence structure of all target sentences with both dominant and subordinate target diacritization. On a 5-point scale (1 = structure is highly unusual, 5 = structure is highly natural), overall sentence structure naturalness ratings for all stimuli were high (mean = 4.3, SD = 0.82, range = 3 – 5). Structure naturalness ratings for sentences containing the dominant and subordinate versions of the target were very similar (dominant: mean = 4.26, SD = 0.73, range = 3 – 5; subordinate: mean = 4.28, SD = 0.69, range 3 – 5; dominant vs. subordinate structure naturalness ratings: $t < 1$).

Apparatus

An SR Research Eyelink 1000 tracker was used to record participants' eye movements while they read the sentences. Viewing was binocular, but eye movements were recorded from the right eye only. The eye tracker was interfaced with a Dell Precision 390 computer, with all sentences presented on a 20 inch ViewSonic Professional Series *P227f* CRT monitor. The participants leaned on a headrest, which supported their chin and forehead during reading to reduce head movements. The text was displayed in black on a light grey background. The display was 81 cm away from the participants, and at this distance, on average, 3.2 characters equalled 1° of visual angle.

The CRT monitor was programmed to run at a refresh rate of 140 Hz, however due to an error not detected until the completion of data collection, the monitor was actually running at 60 Hz. This is a somewhat slow refresh rate for boundary paradigm experiments. We thus adopted a thorough data cleaning procedure (see Results) to remove all trials where the display change did not take place during the readers' saccade towards the target word.

The participants used a VPixx RESPONSEPixx VP-BB-1 button box to enter their responses to comprehension questions and to terminate trials after reading the sentences. Finally, a standard digital voice recorder was used to record participants reading aloud of the materials used for reading skill screening (details below).

Design

We manipulated two independent, within-participants, variables: (i) diacritics preview (identical, inaccurate, or no-diacritics previews); and (ii) diacritization pattern on target word (dominant or subordinate). These variables were counterbalanced using a Latin square design (see example in Figure 1), and presented in a random order such that participants saw each sentence only once in any condition, and they saw an equal number of target stimuli from all conditions.

Another variable that we included in our analyses, was the launch site for the saccade into the target word. We measured launch site as the distance between the location of the pre-target fixation and the beginning (or right boundary - because Arabic is read from right to left) of the interest area containing the target word. In our statistical models we treated launch site distance as a fixed, continuous variable (e.g., Slattery, Staub, & Rayner, 2012).

Procedure

This experiment was approved by the University of Southampton Ethics Committee. Upon arrival at the lab, participants were given a description of the apparatus and instructions for the experiment. After signing the consent forms, participants read aloud the reading-skill screening text (346 words, which provided, in Arabic, a general introduction to the research) while being audio-recorded. This was followed by the eye tracking procedure. Finally, to assess readers' accuracy in decoding diacritics, we presented them with a single word reading aloud task (target words were 36 diacritized words, as well as 24 filler, non-diacritized, words), again while being audio-recorded.

Prior to collecting eye movement data, the eye tracker was calibrated using a horizontal 3-point calibration, and the calibration was validated. Maximum error of calibration accuracy was always $< 0.25°$, otherwise calibration and validation were repeated. Prior to the onset of each sentence, a circular fixation target ($1° \times 1°$)

appeared on the screen in the location of the first character of the sentence. If a stable fixation was detected on the target (fixation trigger), the display changed and the sentence was displayed. Recalibration was performed if a stable fixation was not detected on the circular target.

The participants were told to read silently, and that they would periodically be required to use the button box to provide a yes/no answer to the questions that followed some sentences. Participants read the 11 practice sentences followed by the 139 experimental and filler sentences. In addition to the fixation trigger, and to ensure the accuracy of eye tracking, drift measurement was performed at the beginning of each trial with the circular fixation target (1°×1°). Re-calibration was performed if necessary. Participants were allowed to take breaks whenever they needed, and following any breaks the tracker was re-calibrated. The testing session lasted 60-80 minutes depending on how many breaks were taken.

Following the collection of eye movement data in each session the experimenter asked each participant if they noticed any changes or flicker on the screen. Only one participant reported noticing some flickering around the middle of sentences on 5-6 occasions. This participant was replaced and the data discarded.

Results

For all reported analyses, fixation times shorter than 80ms, or longer than 800 ms were removed. However, fixations shorter than 80 ms that were located within 10 pixels or less (0.31° of visual angle approximately) from another longer fixation, were merged into the longer fixations. Along with removing trials where blinks or track loss occurred, this resulted in removing approximately 5.4% of all data points (1839 fixations remained). Furthermore, for each of the fixation duration measures, we removed data points ±2.5 standard deviations away from the mean fixation duration per participant and condition as outliers. The resulting percentages of data loss for outlier trimming per measure are reported in Table 1 below.

We furthermore removed data points relating to fixations on the target word where the display change was inaccurate. We removed data points when display changes happened prior to readers initiating a saccade towards the target (4.2% of data points). Subsequently, we removed data points for instances where display changes happened late, that is, after the reader crossed the invisible boundary and began fixating the target word. Removing data points for changes where the delay in display change was > 0 ms resulted in removing 5.6% of data points[2]. Finally, removing data points where readers crossed the boundary very briefly and then returned to the pre-target region resulted in removing no data points. The data cleaning procedures affected all experimental conditions equally (mean number of observations per condition = 290, SD = 5.6, range = 283 – 297).

Following a preliminary analysis, we removed all observations where launch site into the target word was farther than 80 pixels (or average of 10.4 characters; average character size = 7.7 pixels), given the scarcity of observations where launch site was farther. This resulted in removing a further 0.8% of data points (1,632 data points remaining).

Our screening and comprehension monitoring tasks revealed that the participating readers were highly skilled and had no difficulty comprehending target stimuli. In the screening procedure where participants read text aloud, with the exception of one participant who was replaced, all 36 participants were highly accurate in reading (mean percentage of words read accurately = 97.3%, SD = 0.98, range = 95.6 – 100%). Additionally, comprehension questions followed about 30% of all target sentences in the eye tracking part of the study. Participants responded accurately on average 90% of the time (SD = 5.3, range = 82 – 100%) indicating that participants read and understood the sentences. There were no differences between the accuracy scores across the conditions. Finally, for the single word reading aloud task that we used to investigate readers' accuracy of decoding diacritics, all 36 participants were highly accurate (mean word reading accuracy = 93.5%, SD = 7.3, range = 84.2 – 100%).

We used the *lmer* package (*lme4*, version 1.1-8, Bates, Maechler, & Bolker, 2011) within the R environment for statistical computing (R-Core Development Team, 2013) to run linear mixed models (LMMs). We report $|t|$ statistics for the LMMs where effects approximately twice as large as their standard error (i.e., $|t| \geq 1.96$) are interpreted as significant. The fixed variables of all models were the experimentally manipulated preview conditions (identical, inaccurate, no diacritics) and pattern of target diacritization (dominant, or subordinate), as well as launch site (a continuous variable)[3]. Subjects and items were treated as the random variables. We always began our analyses with full models (e.g., Barr, Levy, Scheepers, & Tily, 2013) that included the main effects and their interactions, as well as maximal random effects structure. These models were systematically trimmed when failure to converge occurred, first by removing correlations between random effects, and if necessary also by removing their interactions. All findings reported here are from successfully converging models. For each contrast we report beta values (*b*), standard error (*SE*), and *t* statistics for fixation duration measures. We performed log transformation of fixation duration data to reduce distribution skewing (Baayen, Davidson, & Bates, 2008). Prior to running the models, we used the contr.sdif function in the MASS package (Venables & Ripley, 2002) to pre-specify the contrasts between the levels of the fixed factors (preview availability, and target diacritization). Following running the model, we used the Effects package (Fox, 2003; Fox & Hong, 2009) to generate visual representations of the obtained effects (Figures 2 - 6). For all analyses reported, Table 1 contains the descriptive statistics, and Table 2 contains the outputs of the LMMs.

<Insert Table 1 about here>

<Insert Table 2 about here>

Skipping Rate

Even when the random structure of the generalised linear mixed models (GLMMs) for the skipping data was reduced to a single intercept for subjects, the model did not converge. In all likelihood, this is due to the very small differences between all conditions, indeed the means were very similar (see Table 1). Thus, we only report descriptive statistics for skipping rates (Table 1). This is a somewhat surprising outcome: Based on our previous findings (Hermena et al., 2015), we anticipated that the no-diacritics previews would result in more skipping compared to previews containing diacritics.

First Fixation Duration

Whereas the means (Table 1) suggest a pattern such that first fixation durations were longest following inaccurate previews of the diacritics, and shortest following previews with no diacritics (see Table 1), the mixed linear models indicated that the only significant differences were between the preview of the inaccurate diacritics and the other preview conditions[4].

The effect of preview availability was furthermore qualified by two 2-way interactions with launch site. As Figure 2 shows, fixation durations were increased for closer launch sites when the readers were given inaccurate previews of the diacritics, and this pattern was absent from the other preview conditions (identical, and no-diacritics). There was no main effect of target diacritization, or significant interactions between target diacritization and preview availability or launch site.

<Insert Figure 2 about here>

The pattern of results obtained in first fixation suggests that processing diacritics began early, that is, parafoveally. Moreover, this processing includes identifying the diacritics such that inaccurate previews of the diacritics were costly to processing, particularly at close launch sites, where, presumably, better pre-processing of the previews occurred. Interestingly, the no-diacritics preview condition did not come with an additional cost compared to the identical preview condition.

Single Fixation Duration

Similar to first fixation, there was a significant effect of preview availability, such that single fixation durations were longer following inaccurate previews of the target compared to the other two preview conditions. The effect of preview availability was qualified by two 2-way interactions with launch site (the interaction was significant in the inaccurate preview vs. identical × launch site, and closely approached significance in the no-diacritics vs. inaccurate previews × launch site, see Table 2). Similar to first fixation, and as Figure 3 shows, fixation durations were increased for closer launch sites when the readers were given inaccurate previews of the diacritics, and this pattern was absent from the other preview conditions (identical, and no-diacritics). Thus, the single fixation data provide further evidence to suggest that processing diacritics to full identification begins parafoveally, specifically at close launch sites. Inaccurate previews of the diacritics at close launch sites resulted in increased fixation durations. Of course, it is important to note that the single fixation data form a significant proportion of the first fixation data set (and therefore commonality in patterns of effects is highly likely). Note also that in both first and single fixation measures, no difference was observed between the two diacritization patterns (dominant vs. subordinate) with regards to the influence of inaccurate previews on fixation duration.

<Insert Figure 3 about here>

Additionally, we obtained another 2-way interaction between target diacritization and launch site. As Figure 4 shows, overall, single fixation durations increased for far launch sites for the subordinate pattern, compared to the dominant pattern. This is the first set of results that shows a difference between the two diacritization patterns (dominant vs. subordinate). Note that the overlapping of the grey confidence interval bands in Figure 4 indicates that there were no significant differences as a function of target diacritization at close launch sites. Figure 4 also clearly shows that the data points at far launch sites were sparser than for closer launch sites. Note also that this interaction collapses across all three preview conditions (identical, inaccurate, and no-diacritics previews). For these reasons, we adopt due caution in interpreting this interaction. It is possible that this pattern supports our earlier suggestion that at far launch sites information about diacritics from the preview is so visually degraded that readers may not even have a clear indication of whether or not diacritics are present on the upcoming word. In the absence of clear information concerning the presence or absence of diacritics on the upcoming word, readers may have an expectation for the dominant pronunciation of that word. This perhaps explains why, for target words with the dominant diacritization pattern, single fixation durations originating from far launch sites were shorter relative to fixation durations on words with the subordinate diacritization pattern.

<Insert Figure 4 about here>

Gaze Duration

Besides a main effect of launch site, there was an effect of preview availability on gaze duration (the contrast was significant in the no-diacritics vs. inaccurate preview and marginally significant in the inaccurate vs. identical preview). The effect of preview availability was qualified by two 2-way interactions, the first one between preview availability (inaccurate vs. identical) × target diacritization, and the second one between preview availability (inaccurate vs. identical) × launch site. However, these 2-way interactions were again qualified by a 3-way interaction between preview availability (inaccurate vs. identical) × target diacritization (dominant vs. subordinate)

× launch site (see Table 2). Combined, a rather complex data pattern emerged which is made comprehensible by the visualisation of this 3-way interaction in Figure 5.

<Insert Figure 5 about here>

As can be seen in Figure 5, the patterns obtained for dominant diacritization are very similar for the different preview conditions. By contrast, a different pattern is seen for the subordinate target diacritization, where the identical preview condition clearly differs from the inaccurate preview and the no-preview conditions. For the identical preview condition, when the target word has the subordinate diacritization pattern, a standard preview benefit is observed with bigger preview benefit at closer launch sites. Importantly, this facilitation was not observed for identical previews of the dominant diacritization pattern. This pattern of results suggests that when readers had a parafoveal preview from a close launch site that clearly indicated an upcoming word with diacritics, they expected a subordinate diacritization pattern. As explained above, this is likely due to the readers' long experience with diacritics assigned to homographs in text pointing them towards one of the plausible subordinate pronunciations. When the target word was indeed carrying subordinate diacritization, facilitation (reduced gaze duration) was observed. Note that a similar facilitation was not observed for the dominant diacritization. At the outset of this experiment, we considered (hypothesis vi) whether in instances where readers would be able not only to perceive the presence or absence of diacritics, but also to extract the identity of the diacritics parafoveally, expectation for the subordinate pattern would still play a role or whether the identity of the diacritics would exclusively influence processing. We see here that readers' expectation for the subordinate pattern when diacritics are present also modulates processing of diacritics when the identity of the diacritics was processed. Therefore, at close launch sites, only the subordinate diacritics (expected based on the presence of diacritics) showed the standard preview benefit because the subordinate diacritics were both expected and identified, compared to the dominant diacritics that were identified but not expected.

For identical previews from a far launch site (Figure 5), on the other hand, we see a similar pattern to that observed in single fixation duration (Figure 4): Gaze durations are inflated for the subordinate diacritization pattern, compared to the dominant pattern, when the initial fixation on the target word originated from a far launch site. As explained above, with highly degraded previews of diacritics (or possibly none) at far launch sites, readers presumably assumed that the upcoming word was not diacritized. As such, they would have expected that the upcoming target word would conform to the dominant pronunciation. Hence, processing demands, and therefore fixation durations, were inflated when the word was directly fixated and turned out to carry the subordinate diacritization pattern instead. Coupled with results reported above, our findings indicate that readers' expectation for dominant or subordinate diacritization to be present on the target is influenced by whether or not the eyes were close enough on the preceding fixation (i.e., the parafoveal word fell on an area of the retina that delivered sufficiently high visual acuity information) to allow for a sufficiently clear preview of the diacritics in the parafovea.

A final point can be made, with regard to the gaze duration findings as illustrated in Figure 5, concerning the similarity in the patterns of effects obtained for the target words with dominant and subordinate diacritization when there had been a no-diacritics preview. Recall that we suggested that the presence and quality of the preview of the diacritics that was available to a reader would influence their expectations for either the dominant or the subordinate pattern to be present. This expectation relative to the diacritics that were present when the target was fixated, in turn, should have influenced fixation durations. If this were the case, then no-diacritics previews should have resulted in clear facilitation for the dominant pattern over the subordinate pattern. However, the pattern of results in the no-diacritics preview condition deviates from our predictions, showing a great deal of similarity between dominant and subordinate diacritics (see Table 1 & Figure 5).

These are somewhat surprising results and we can only offer a speculative explanation for this pattern. We suggest that the reason for the similarity between the results obtained for the dominant and subordinate diacritization patterns in the no-diacritics preview condition is that both patterns surprised the readers. Specifically, for the dominant pattern, although the dominant reading of the word was expected (given the absence of diacritics in the preview), the presence of the dominant diacritics upon fixation of the target would have been unexpected since readers are used to encountering subordinate diacritization patterns when they appear in print. Thus, any benefit arising due to an expectation for the dominant reading of the word (based on an absence of parafoveal diacritics) was reduced due to the onset of an unexpected diacritical form at fixation onset. This account is clearly speculative, and of course, more experimentation is necessary to better understand how diacritics are processed both parafoveally as well as foveally.

Additional Analyses

Finally, we investigated whether any processing effects related to our experimental manipulations were observable in the regions containing the pre- and post-target words. To do this, we first compared readers' last fixation durations (first pass) on the pre-target word in all experimental conditions to explore whether the parafoveal previews of the diacritics had any influence on pre-target word processing. If fixation durations on the pre-target words were influenced by the parafoveal previews of the upcoming word, we would have evidence of parafoveal-on-foveal effects (Inhoff et al., 2000; Rayner et al., 2004). We had no *a priori* expectations as to possible parafoveal-on-foveal effects resulting from parafoveal processing of diacritics. The results were unequivocal: No significant differences between the conditions were recorded at the pre-target word (all ts < 1.4). Similarly, no significant differences between the conditions were found at the post-target word (all ts < 1.3), suggesting that the influence of processing the diacritics in the various conditions did not spill over into the following region. Clearly, the effects were quite immediate and short lived.

Discussion

In this study we investigated parafoveal processing of Arabic diacritics by presenting adult native Arabic readers with homographic words which carried either the dominant or subordinate diacritization pattern. Using the boundary paradigm (Rayner, 1975), we manipulated the parafoveal preview of this diacritization pattern available to readers: Readers had access to an identical, an inaccurate (opposite pattern), or a no-diacritics preview. In our analyses, we also examined the influence of launch site on parafoveal processing.

We hypothesized that if readers identified diacritics parafoveally, most likely only at close launch sites, we would observe processing benefit for identical, compared to inaccurate, previews. As for no-diacritics previews, we anticipated that in the absence of diacritics in the parafovea, readers may have an expectation that the pronunciation of the upcoming word would be the dominant one, and thus predicted facilitation for dominant, compared to subordinate, diacritics in this condition. We also hypothesized that processing of diacritics may be frequency-mediated, with facilitation observed for the dominant diacritization pattern. Alternatively, the presence of diacritics in the parafovea may alert readers that the upcoming word is to be pronounced as one of the subordinate versions—that is, to expect subordinate diacritization pattern to be present. This is based on Arabic readers' experience with encountering the subordinate diacritization patterns in print to guide them towards the less-preferred pronunciations of homographs. If this were the case, parafoveal diacritics would produce facilitation for the subordinate relative to the dominant pattern. Additionally, we anticipated that any obtained effects would be amplified at close launch sites, given that identification of parafoveal diacritics is perhaps only possible at close launch sites. We also suggested that readers' expectations for a particular diacritization pattern to be present on the target word may be influenced by launch site. Specifically, at far launch sites with no, or a highly degraded preview of the diacritics, readers may expect the upcoming word to have a dominant pronunciation. Conversely, at a close launch site, with clear preview of the upcoming diacritics, readers may expect the upcoming word to conform to a subordinate pronunciation and to carry subordinate diacritics. Finally, at close launch sites, when the eyes perceived not only the absence or presence of diacritics, but were also able to extract the identity of the diacritics, we considered whether readers' expectation (for the subordinate pronunciation) still influenced processing of the target word, or instead whether any influence of the expectation would be undone by readers actually identifying the diacritics.

The first of our results, the skipping probabilities of the diacritized target words, challenged our expectation that previews of the target which contained diacritics would result in less skipping than previews containing no diacritics. We based our prediction on previous similar results we obtained (Hermena et al., 2015). It is hard to explain the discrepancy between the current and previous results. One hypothesis is that in our previous investigation, one of the conditions and some filler items contained fully-diacritized sentences. As such, readers' sensitivity to the presence of diacritics might have been increased relative to the current investigation where no fully-diacritized sentences were included in either the experimental or filler sentences. Of course this is currently only a hypothesis and future investigations are needed in order to better understand how diacritics on parafoveal words affects word skipping in Arabic.

Next, let us consider the fixation data on the target word itself, after the boundary change had occurred, and when the target was presented in its fully diacritized dominant or subordinate form. Early measures, namely first and single fixation duration, demonstrated clearly that readers engaged in parafoveal pre-processing of the upcoming diacritics. In both fixation duration measures, we reported 2-way interactions between preview availability and launch site, such that following inaccurate previews of the diacritics, fixation durations on the target were inflated, particularly for closer launch sites (Figures 2 & 3). Note that these effects occurred for target words with both dominant and subordinate diacritization at fixation. Furthermore, results showed that this effect was relatively short-lived, influencing only the initial fixation made on the target: This pattern of results was not observed in gaze duration. This finding strongly suggests that readers have identified the diacritics parafoveally, particularly at closer launch sites, such that inaccurate previews of the diacritics resulted in processing costs. This pattern also supported our hypothesis that observed effects for parafoveal pre-processing would be amplified at closer launch sites, given the improved quality of parafoveal processing (see Fitzsimmons & Drieghe, 2011; Kennison & Clifton, 1995; Miellet & Sparrow, 2004).

We also obtained a 2-way interaction between target diacritization and launch site in single fixation duration (Figure 4). This interaction showed that single fixations on the target word carrying the subordinate diacritization pattern were inflated, compared to the dominant pattern, when originating from a far launch site. We suggested that, in line with our hypothesis regarding the availability of diacritics in the parafovea, this pattern indicates that, at far launch sites, where parafoveal previews of the diacritics are degraded, readers expected that the upcoming word would probably have no diacritics and that the word would have the dominant pronunciation. The interaction illustrated in Figure 4 supports this suggestion: Durations of fixations originating at far launch sites were inflated for the subordinate diacritization. Furthermore, this suggestion regarding readers' expectation about the upcoming target word at far launch sites was supported by the significant 3-way interaction reported in gaze duration (Figure 5, identical preview panel). An aspect of this 3-way interaction (preview availability × target diacritization × launch site) is similar to the pattern reported in the 2-way interaction (target diacritization × launch site) in single fixation duration. Namely, gaze durations were inflated on the subordinate diacritics in the identical preview condition, at far launch sites. This was clearly not the case for the dominant diacritization.

As for the no-diacritics preview condition, our hypothesis that in this condition the presence of dominant diacritics on the target would result in facilitation, relative to the subordinate pattern, was not supported by the results. Indeed, the pattern of results (including the means) of gaze duration was very similar for both the dominant and subordinate diacritics. However, as we speculated above, the similarity of the results for the dominant and subordinate diacritics may be because both patterns were unexpected for the readers in the no-diacritics preview condition. More experimentation is perhaps necessary to fully explain the results we obtained.

To summarize thus far, our results show clearly that readers initiated parafoveal pre-processing of the diacritics whereby at close launch sites parafoveal diacritics were identified. Indeed, inaccurate previews of the diacritics resulted in inflated initial fixation durations (first and single) on the target word, particularly at close launch sites. This pattern clearly indicated that the parafoveal pre-processing of diacritics is modulated by launch site. Our results also suggest that when the parafoveal preview of the diacritics was highly degraded at far launch site, readers' expectation was for the pronunciation of the upcoming word to be the dominant one. When the subordinate diacritics were present instead, a cost to processing was recorded at far launch site in single fixation, and gaze duration (identical preview).

The remainder of our results elucidated how readers' expectations for the subordinate diacritization pattern at close launch site modulated processing of the upcoming diacritics. In gaze duration we observed clearly a traditional preview benefit for identical previews of the diacritics, but only for the subordinate pattern. Specifically, gaze durations were reduced on target words carrying the subordinate diacritization pattern when initial fixations originated from closer launch sites. This pattern of results was not observed for dominant diacritics (see Figure 5). This clearly indicates that at close launch sites identification of the parafoveal diacritics resulted in preview benefit, but only for the *expected* subordinate diacritics. In other words, at close launch sites, the benefit of identification of parafoveal diacritics is modulated by readers' expectation for the subordinate pattern to be present. As we explained earlier, readers developed the expectation for subordinate diacritization to be present in print given their long experience in reading Arabic text. This is because the printed diacritization usually directs readers towards the less frequent and less preferred versions of the homographs, whereas dominant (and preferred) pronunciations are typically left undiacritized (Hermena et al., 2015). As such, at close launch sites, the subordinate diacritics were *both* identified (from the identical preview), and expected by the readers. By contrast, and although readers also identified the dominant diacritics when they had an identical preview at close launch site, this pattern was not expected. Recall that in all experimental sentences, context prior to the target homograph did not constrain the readers towards either dominant or subordinate interpretations. The results thus suggest that even in the absence of constraining context, the presence of diacritics in the parafovea, particularly at close launch sites, alerts the readers that the upcoming homograph is likely to be disambiguated towards a subordinate analysis, and thus readers expect to see subordinate diacritization pattern on the target word once it is fixated. This expectation has subsequently modulated processing of the diacritics such that identical preview benefit was observed only for the expected subordinate diacritization pattern.

Thus, overall, our results reveal that readers' expectations as to which diacritization pattern will be present on the upcoming word depends on whether or not the fixation location of the preceding fixation allowed for a sufficiently detailed preview (in terms of visual acuity) of the diacritics. Specifically, at far launch sites, if the preview of the diacritics did not allow for even their presence to be detected, then readers expect the word to conform to the dominant pronunciation. By contrast, when the launch site is close enough to allow for a sufficiently clear parafoveal preview of the diacritics (i.e., close launch sites), readers' expectations were altered and they expected a subordinate diacritization pattern to be present. Thus, the results clearly indicate that readers' expectation for a particular diacritization pattern modulates their parafoveal and subsequent foveal processing of diacritics. This explains the fact that preview benefit for identical previews was only observed for the *expected* subordinate diacritization.

The results also indicate that our earlier suggestion that processing diacritics may be frequency-mediated may have been rather simplistic. We documented that dominant diacritization patterns do not yield the widely-reported frequency effects of facilitation of the dominant over the subordinate interpretations of homographic words (e.g., Reichle et al., 2003; Sereno et al., 2006; see also Binder, 2003; Binder & Rayner, 1998; Duffy et al., 1988). Rather, the processing benefit for dominant diacritization patterns is only observed when readers expected this pattern to be present when only a highly degraded parafoveal preview of the diacritics was available at far launch sites. Similarly, the results indicated that for parafoveal processing of diacritics, the presence of an identical preview results in preview benefit only when the target word is carrying the expected subordinate diacritization. These findings can be contrasted with previous investigations of parafoveal processing (e.g., Ashby et al., 2006; Henderson et al., 1995; Rayner, 1975; Rayner et al., 1982, etc.) where identical previews of targets always resulted in preview benefit. Our findings thus provide a clear demonstration that readers' expectations — influenced by both experience with the linguistic materials being manipulated, in this case Arabic diacritics, as well as launch site — modulate parafoveal and foveal processing of diacritics.

Finally, the additional analyses we performed showed that there was neither a parafoveal-on-foveal effect on the pre-target word, nor a spill-over effect on the post-target word. That is, the specific pattern of diacritization (dominant or subordinate), and the quality of the parafoveal preview available, do not influence processing demands prior or subsequent to fixation of the diacritized word itself. We propose that the absence of evidence for parafoveal-on-foveal effects for processing diacritics to be more consistent with eye guidance models which stipulate serial processing, namely the E-Z Reader model (e.g., Pollatsek, Reichle, & Rayner, 2006; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle et al., 2003). In E-Z Reader attention is allocated in a serial manner and word identification occurs sequentially, unlike models which propose gradient allocation of attention and parallel processing of words, as in the SWIFT model (Engbert, Longtin, & Kliegl, 2002; Engbert, Nuthmann, Richter, & Kliegl, 2005). The absence of spill-over effects, on the other hand, can be attributed to a number of factors. To begin with, our sample of Arabic readers was made up of skilled readers, who were all highly skilled at decoding diacritics. Another factor is perhaps the absence of any mismatch between the sentence context subsequent to the target word and the target word, in both target diacritization conditions. Thus, any costs associated with processing the target word did not spill-over to the subsequent word.

To summarize, this is the first investigation of parafoveal processing of diacritics in Arabic using the boundary paradigm (Rayner, 1975). Our fixation duration results show that readers begin processing diacritics parafoveally, prior to fixating the target word. Specifically, diacritics are identified at close launch sites.

Furthermore, readers' expectations for a particular pattern of diacritics to be present, the dominant pattern or a subordinate one, is influenced by the clarity of the preview of diacritics. The clarity and quality of this preview is in turn influenced by the distance between the location of the pre-target fixation and the target word—launch site. At far launch sites, readers' expectations are for a dominant pronunciation; whereas at close launch site the expectation (when diacritics are present parafoveally) is for a subordinate pronunciation. Importantly, at close launch site processing of the diacritics is influenced by readers' ability to identify the diacritics parafoveally, and this processing is also modulated by the expectation for the subordinate diacritization pattern to be present. This expands upon our previous findings about readers' sensitivity to the presence of diacritics in the parafovea (Hermena et al., 2015). Although we did not replicate this finding for the skipping measure, first and single fixation durations clearly demonstrated in the current investigation that readers initiated pre-processing and identification of diacritics parafoveally such that at close launch site inaccurate previews resulted in increased fixation durations. Our findings thus provide an insight into how processing diacritics is modulated by a number of interacting variables: (i) The pattern of diacritization present on the target word (dominant or subordinate); (ii) The type of preview available to the readers prior to fixating the diacritized word; (iii) The quality of the preview available of the diacritics, based on launch site; and (iv) Readers' expectations for a particular pattern of diacritics to be present, which is influenced in turn by the quality of the preview available to the reader, as well as their experience of encountering subordinate diacritics in print.

References

Abu-Rabia, S. (1997a). Reading in Arabic orthography: The effect of vowels and context on reading accuracy of poor and skilled native Arabic readers. *Reading and Writing: An Interdisciplinary Journal*, *9*, 65-78. doi:10.1023/A:1007962408827

Abu-Rabia, S. (1997b). Reading in Arabic orthography: The effect of vowels and context on reading accuracy of poor and skilled native Arabic readers in reading paragraphs, sentences and isolated words. *Journal of Psycholinguistic Research, 26*, 465-482. doi:10.1023/A:1025034220924

Abu-Rabia, S. (1998). Reading Arabic texts: Effects of text type, reader type, and vowelization. *Reading and Writing: An Interdisciplinary Journal,* 10, 106–119. doi:10.1023/A:1007906222227

Abu-Rabia, S. (1999). The effect of Arabic vowels on the reading comprehension of second- and sixth-grade native Arab children. *Journal of Psycholinguistic Research*, *28*, 93-101. doi:10.1023/A:1023291620997

Abu-Rabia, S. (2001). The role of vowels in reading Semitic scripts: Data from Arabic and Hebrew. *Reading and Writing: An Interdisciplinary Journal*, 14, 39-59. doi:10.1023/A:1008147606320

Altarriba, J., Kambe, G., Pollatsek, A., & Rayner, K. (2001). Semantic codes are not used in integrating information across eye fixations in reading: Evidence from fluent Spanish–English bilinguals. *Perception & Psychophysics, 63*, 875–890. doi:10.3758/BF03194444

Ashby, J., Treiman, R., Kessler, B., & Rayner, K. (2006). Vowel processing during silent reading: evidence from eye movements. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, *32*, 416-424. doi:10.1037/0278-7393.32.2.416

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modelling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412. doi:10.1016/j.jml .2007.12.005

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. doi:10.1016/j.jml .2012.11.001

Bates, D., Maechler, M., & Bolker, B. (2011). Lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-42 [Computer software]. Available from http://CRAN.R-project.org/package=lme4

Binder, K. S. (2003). Sentential and discourse topic effects on lexical ambiguity processing: An eye-movement examination. *Memory & Cognition*, *31*, 690–702. doi:10.3758/BF03196108

Binder, K. S., & Rayner, K. (1998). Contextual strength does not modulate the subordinate bias effect: Evidence from eye fixations and self-paced reading. *Psychonomic Bulletin & Review*, *5*, 271–276. doi:10.3758/BF03212950

Boudelaa, S., & Marslen-Wilson, W. D. (2010). Aralex: A lexical database for Modern Standard Arabic. *Behavior Research Methods*, *42*, 481-487. doi:10.3758/BRM.42.2.481

Clifton, C., Traxler, M. J., Mohamed, M. T., Williams, R. S., Morris, R. K., & Rayner, K. (2003). The use of thematic role information in parsing: Syntactic processing autonomy revisited. *Journal of Memory & Language*, *49*, 317-334. doi:10.1016/S0749-596X(03)00070-6

Dopkins, S., Morris, R. K., & Rayner, K. (1992). Lexical ambiguity and eye fixations in reading: A test of competing models of lexical ambiguity resolution. *Journal of Memory and Language, 31,* 461–477. doi:10.1016/0749-596X(92)90023-Q

Drieghe, D. (2011). Parafoveal-on-foveal effects on eye movements during reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye Movements* (pp. 839-855). Oxford, England: OUP.

Duffy, S. A., Morris, R. K., & Rayner, K. (1988). Lexical ambiguity and fixation times in reading. *Journal of Memory and Language, 27,* 429– 446. doi:10.1016/0749-596X(88)90066-6

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision Research, 42,* 621–636. doi:10.1016/S0042-6989(01)00301-7

Engbert, R., Nuthmann, A., Richter, E., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112,* 777–813. doi:10.1037/0033-295X.112.4.777

Fitzsimmons, G., & Drieghe, D. (2011). The influence of number of syllables on word skipping during reading. *Psychonomic Bulleting and Review*, *18*, 736-741. doi:10.3758/s13423-011-0105-x

Folk, J. R., & Morris, R. K. (2003). Effects of syntactic category assignment on lexical ambiguity resolution in reading: An eye movement analysis. *Memory & Cognition, 31,* 87–99. doi:10.3758/BF03196085

Fox, J. (2003). Effect displays in R for generalised linear models. *Journal of Statistical Software, 8,* 1-27.

Fox, J., & J. Hong (2009). Effect displays in R for multinomial and proportionalodds logit models: Extensions to the effects package. *Journal of Statistical Software*, *32*, 1-24.

Frazier, L., & Rayner, K. (1987). Resolution of syntactic category ambiguities: Eye movements in parsing lexically ambiguous sentences. *Journal of Memory and Language*, *26*, 505–526. doi:10.1016/0749-596X(87)90137-9

Frazier, L., & Rayner, K. (1990). Taking on semantic commitments: Processing multiple meanings versus multiple senses. *Journal of Memory and Language*, *29*, 181–200. doi:10.1016/0749-596X(90)90071-7

Haywood, J. A., & Nahmad, H. M. (1965). *A new Arabic grammar of the written language*. Surrey: Lund Humphries.

Henderson, J. M., Dixon, P., Petersen, A., Twilley, L. C., & Ferreira, F. (1995). Evidence for the use of phonological representations during transsaccadic word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 82–97. doi:10.1037/0096-1523.21.1.82

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 effect. *Journal of Experimental Psychology: Human Perception and Performance*, *41*, 494-507. doi:10.1037/xhp0000032

Hohenstein, S., & Kliegl, R. (2014). Semantic preview benefit during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40,* 166–190. doi:10.1037/a0033670

Hohenstein, S., Laubrock, J., & Kliegl, R. (2010). Semantic preview benefit in eye movements during reading: A parafoveal fast-priming study. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, *36*, 1150–1170. doi:10.1037/a0020233

Hyönä, J. (2011). Foceal and parafoveal processing during reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye Movements* (pp. 819-838). Oxford, England: OUP.

Hyönä, J., & Häikiö, T. (2005). Is emotional content obtained from parafoveal words during reading? An eye movement analysis. *Scandinavian Journal of Psychology*, *46*, 475–483. doi:10.1111/j.1467-9450.2005.00479.x

Inhoff, A. W., Starr, M., & Shindler, K. L. (2000). Is the processing of words during eye fixations in reading strictly serial? *Perception and Psychophysics, 62*,1474-1484. doi:10.3758/BF03212147

Juhasz, B. J., & Pollatsek, A. (2011). Lexical influences on eye movements in reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye Movements* (pp. 873-893). Oxford, England: OUP.

Kennison, S. M., & Clifton, C. (1995). Determinants of parafoveal preview benefit in high and low working memory capacity readers: Implications for eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 68-81. doi:10.1037/0278-7393.21.1.68

Kim, Y-S., Radach, R., & Vorstius, C. (2012). Eye movements and parafoveal processing during reading Korean. Reading and Writing, 25, 1053-1078. doi:10.1007/s11145-011-9349-0

Matin, E. (1974). Saccadic suppression: A review and analysis. *Psychological Bulletin*, *81*, 899-917. doi:10.1037/h0037368

Miellet, S., & Sparrow, L. (2004). Phonological codes are assembled before word fixation: Evidence from boundary paradigm in sentence reading. *Brain and Language, 90*, 299-310. doi:10.1016/S0093-934X(03)00442-5

Pacht, J. M., & Rayner, K. (1993). The processing of homophonic homographs during reading: Evidence from eye movements studies. *Journal of Psycholinguistic Research, 22,* 251–271. doi:10.1007/BF01067833

Pickering, M. J., & Frisson, S. (2001). Processing ambiguous verbs: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 556–573. doi:10.1037/0278-7393.27.2.556

Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 148-162. doi:10.1037/0096-1523.18.1.148

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology, 52*, 1–52. doi:10.1016/j.cogpsych.2005.06.001

Pynte, J., Kennedy, A., & Ducrot, S. (2004). The influence of parafoveal typological errors on eye movements in reading. *European Journal of Cognitive Psychology*, *16*, 178-202. doi:10.1080/09541440340000169

R Core Team. (2013). A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. http:// www.R-project.org/

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7*, 65-81. doi:10.1016/0010-0285(75)90005-5

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, 62, 1457-1506. doi:10.1080/17470210902816461

Rayner, K., Balota, D. A., & Pollatsek, A. (1986). Against parafoveal semantic preprocessing during eye fixations in reading. *Canadian Journal of Psychology, 40,* 473–483. doi:10.1037/h0080111

Rayner, K., Cook, A. E., Juhasz, B. J., & Frazier, L. (2006). Immediate disambiguation of lexically ambiguous words during reading: Evidence from eye movements. *British Journal of Psychology, 97,* 467-482. doi:10.1348/000712605X89363

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14,* 191-201. doi:10.3758/BF03197692

Rayner, K., & Frazier, L. (1989). Selection mechanisms in reading lexically ambiguous words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 779–790. doi:10.1037/0278-7393.15.5.779

Rayner, K., McConkie, G. W., & Ehrlich, S. (1978). Eye movements and integrating information across fixations. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 529–544. doi:10.1037/0096-1523.4.4.529

Rayner, K., Pacht, J. M., & Duffy, S. A. (1994). Effects of prior encounter and global discourse bias on the processing of lexically ambiguous words: Evidence from eye fixations. *Journal of Memory and Language, 33,* 527–544. doi:10.1006/jmla.1994.1025

Rayner, K., Schotter, E. R., & Drieghe, D. (2014). Lack of semantic parafoveal preview benefit in reading revisited. *Psychonomic Bulletin & Review*, *21*, 1067-1072. doi:10.3758/s13423-014-0582-9

Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *30*, 1290-1301. doi: 10.1037/0278-7393.30.6.1290

Rayner, K., Well, A. D., Pollatsek, A., & Bertera, J. H. (1982). The availability of useful information to the right of fixation in reading. *Perception and Psychophysics, 31*,537-550. doi:10.3758/BF03204186

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. doi:10.1037/0033-295X.105.1.125

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26,* 445–526. doi:10.1017/S0140525X03000104

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, and Psychophysics*, *74*, 5-35. doi:10.3758/s13414-011-0219-2

Schultz, E. (2004). *A student grammar of Modern Standard Arabic*. Cambridge: CUP.

Seidenberg, M. S., Tanenhaus, M. K., Leiman, J. M., & Bienkowski, M. (1982). Automatic access of the meanings of ambiguous words in context: Some limitations of knowledge-based processing. *Cognitive Psychology, 14,* 489-537. doi:10.1016/0010-0285(82)90017-2

Sereno, S. C. (1995). Resolution of lexical ambiguity: Evidence from an eye movement priming paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 582–595. doi:10.1037/0278-7393.21.3.582

Sereno, S. C., Brewer, C. C., & O'Donnell, P. J. (2003). Context effects in word recognition: Evidence for early interactive processing. *Psychological Science, 14,* 328–333. doi:10.1111/1467-9280.14471

Sereno, S. C., O'Donnell, P. J., & Rayner, K. (2006). Eye movements and lexical ambiguity resolution: Investigating the subordinate-bias effect. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 335-350. doi:10.1037/0096-1523.32.2.335

Sereno, S. C., Pacht, J. M., & Rayner, K. (1992). The effect of meaning frequency on processing lexically ambiguous words: Evidence from eye fixations. *Psychological Science, 3,* 296–300. doi:10.1111/j.1467-9280.1992.tb00676.x

Shany, M., Bar-On, A., & Katzir, T. (2012). Reading different orthographic structures in the shallow-pointed Hebrew script: a cross-grade study in elementary school. *Reading and Writing, 25,* 1217-1238. doi:10.1007/s11145-011-9314-y

Simpson, G. B., & Burgess, C. (1985). Activation and selection processes in the recognition of ambiguous words. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 28-39. doi:10.1037/0096-1523.11.1.28

Slattery, T. J., Angele, B., & Rayner, K. (2011). Eye movements and display change detection during reading. *Journal of Experimental Psychology: Human Perception and Performance*, *37,* 1924-1938. doi:10.1037/a0024322

Slattery, T. J., Staub, A., & Rayner, K. (2012). Saccade launch site as a predictor of fixation durations in reading: Comment on Hand, Miellet, O'Donnell, and Sereno (2010). *Journal of Experimental Psychology: Human Perception and Performance*, *38*, 251-261. doi:10.1037/a0025980

Starr, M., & Inhoff, A. W. (2004). Attention allocation to the right and left of a fixated word: Use of orthographic information from multiple words during reading. *European Journal of Cognitive Psychology*, *16*, 203-225. doi:10.1080/09541440340000150

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th Edition). New York: Springer.

White, S. J., Bertram, R., & Hyönä, J. (2008). Semantic processing of previews within compound words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 988–993. doi:10.1037/0278-7393.34.4.988

Yan, M., Richter, E. M., Shu, H., & Kliegl, R. (2009). Readers of Chinese extract semantic information from parafoveal words. *Psychonomic Bulletin & Review*, *16*, 561–566. doi:10.3758/PBR.16.3.561

Yan, M., Zhou, W., Shu, H., & Kliegl, R. (2012). Lexical and sublexical semantic preview benefits in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*, 1069-1075. doi:10.1037/a0026935

Yang, J., Wang, S., Tong, X., & Rayner, K. (2010). Semantic and plausibility effects on preview benefit during eye fixations in Chinese reading. *Reading and Writing*. Advance online publication. doi:10.1007/s11145-010-9281-8

Yang, J., Wang, S., Xu, Y., & Rayner, K. (2009). Do Chinese readers obtain preview benefit from word n + 2? Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 1192–

1204.doi:10.1037/a0013554

Footnotes

[1] Note that the target homographs in the current investigation are actually disambiguated with the correct diacritics (dominant or subordinate patterns) when fixated. As such, the contribution of sentence context towards disambiguation is not being investigated. Given this, previous investigations where sentence context disambiguated the target homograph prior to encountering it (documenting the *subordinate bias effect*, e.g., Binder, 2003; Binder & Rayner, 1998; Duffy, Morris, & Rayner, 1988; Pacht & Rayner, 1993; Rayner, Cook, Juhasz, & Frazier, 2006; Rayner & Frazier, 1989; Rayner, Pacht, & Duffy, 1994), or after encountering the homograph (e.g., Dopkins, Morris, & Rayner, 1992; Folk & Morris, 2003; Rayner & Frazier, 1989, Experiment 1; Rayner et al., 1994; Sereno, 1995; Sereno et al.,1992) may be of limited relevance in relation to the current investigation. Also, for the same reason, this discussion will not deal with models of context-based disambiguation of homographic words (e.g., the *reordered access model*, Duffy et al., 1988; and the *integration model*, Rayner & Frazier, 1989). Both models would predict that the dominant version of the word becomes available before the subordinate one (see Sereno et al., 2006).

[2] Typically trials in which the boundary change completed more than 10 ms after fixation onset are excluded in accordance with Slattery, Angele and Rayner (2011). By adopting the stricter criterion of 0 ms, we have taken into account the impact our relatively slower refresh rate of the monitor might have had on our results (compared to experiments implementing a faster refresh rate). As a reviewer pointed out, this still allows for a very few instances where the display change can happen late, although rare enough that it is unlikely to influence the results.

[3] Allowing successive contrasts to interact can lead to increased multicollinearity of the fixed factors and this may result in loss of statistical power. An anonymous reviewer suggested that we check the variance inflation factors (VIFs) for each of the predictors to determine the extent to which the results are affected by this. For all the reported eye

movement measures, the maximum VIF value obtained was 10.1 for the preview availability factor and its interactions. Although high, this VIF is still around the limits of what is considered acceptable, with the cut-off point typically adopted is VIF = 10.0. VIF values were considerably less for all the other variables (maximum values for diacritization pattern VIF = 3.3, and for launch site VIF = 1.0). We would like to thank the reviewer for pointing this issue out.

[4] Besides the theoretically more interesting contrasts that compared the inaccurate preview conditions with the other preview conditions, a separately run, additional contrast directly comparing the identical and no-diacritics preview conditions for all measures showed consistently no significant differences between these two conditions (first and single fixation durations *t*s < 1; gaze duration b = 0.13, SE = 0.079, $t = 1.65$). An anonymous reviewer suggested that, in light of running this extra contrast, it would be advisable to adopt a stricter criterion for statistical significance ($|t| \geq 2.24$, $|p| \leq .025$) when examining the contrasts we report in this manuscript. As can be seen in Table 2, the majority of our significant results have $|t|$ values > 2.24. However, it is important to note that we were conducting theoretically-motivated planned contrasts, to answer the specific experimental questions we posed, and not series of post hoc tests. We wish to thank the reviewer for the suggestion.

Table 1

Table 1

Descriptive Statistics for Skipping Rates and Fixation Durations for the Target Word

		Dominant Diacritization Target			Subordinate Diacritization Target		
		Identical Preview	Inaccurate Preview	No Diacritics Preview	Identical Preview	Inaccurate Preview	No Diacritics Preview
	Number of Observations Included (% Removed as Outliers)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Skipping Rate	1,632 (NA)	0.11 (0.3)	0.10 (0.3)	0.11 (0.3)	0.10 (0.3)	0.10 (0.3)	0.10 (0.3)
First Fixation	1.626 (0.3%)	310 (177.8)	316 (170.6)	301 (170.2)	319 (167.9)	328 (202.2)	314 (162.9)
Single Fixation	1,113 (no outliers)	330 (205.9)	343 (189.5)	305 (130.1)	334 (175.2)	335 (183.8)	328 (176.0)
Gaze Duration	1,632 (no outliers)	444 (290.9)	442 (277.4)	440 (334.7)	460 (323.7)	502 (419.2)	439 (296.0)

Table 2

Table 2

Linear Mixed Model Analyses on the Target Word

		First Fixation			Single Fixation			Gaze Duration		
		b	SE	*t*	b	SE	*t*	b	SE	*t*
	(Intercept)	5.6744	0.0349	**162.80**	5.7301	0.0448	**127.96**	5.8493	0.0605	**96.74**
Preview Availability	Inaccurate vs. Identical	0.0986	0.0496	**1.99**	0.1157	0.0584	**1.98**	0.1063	0.0567	1.88
	No-Diacritics vs. Inaccurate	-0.1214	0.0498	**-2.44**	-0.1451	0.0576	**-2.52**	-0.1367	0.0570	**-2.40**
Target Diacritization	Subordinate vs. Dominant	-0.0126	0.0409	-0.31	-0.0677	0.0477	-1.42	-0.0105	0.0468	-0.22
Launch Site Distance	Launch Site (Continuous Var.)	-0.0013	0.0007	-1.72	-0.0003	0.0009	-0.37	0.0033	0.0009	**3.84**
Preview Availability × Target Diacritization	(Inaccurate vs. Identical) × (Subordinate vs. Dominant)	0.0738	0.0993	0.74	0.1330	0.1170	1.14	0.3011	0.1136	**2.65**
	(No-Diacritics vs. Inaccurate) × (Subordinate vs. Dominant)	-0.1054	0.0997	-1.06	-0.0419	0.1151	-0.36	-0.1016	0.1140	-0.89
Preview Availability × Launch Site Distance	(Inaccurate vs. Identical) × Launch Site	-0.0039	0.0017	**-2.33**	-0.0045	0.0022	**-2.09**	-0.0041	0.0019	**-2.12**
	(No-Diacritics vs. Inaccurate) × Launch Site	0.0039	0.0017	**2.28**	0.0041	0.0021	1.93	0.0037	0.0020	1.90
Target Diacritization × Launch Site Distance	(Subordinate vs. Dominant) × Launch Site	0.0016	0.0014	1.11	0.0041	0.0018	**2.31**	0.0016	0.0016	0.99
Preview Availability × Target Diacritization × Launch Site Distance	(Inaccurate vs. Identical) × (Subordinate vs. Dominant) × Launch Site	-0.0028	0.0034	-0.84	-0.0067	0.0043	-1.55	-0.0110	0.0039	**-2.84**
	(No-Diacritics vs. Inaccurate) × (Subordinate vs. Dominant) × Launch Site	0.0042	0.0034	1.24	0.0023	0.0042	0.55	0.0028	0.0039	0.71

Figure 1

Figure 1. Sample stimulus set. The target words appeared following parafoveal previews which were either identical, inaccurate (of the opposite pattern), or non-diacritized. The target words appeared with either the dominant or subordinate diacritization pattern. Target word (and preview) location is marked by underlining. The dashed line represents the location of the invisible boundary, always immediately before the white space preceding the target word. Translation of the two frame sentences is provided. The italicised words separated by slash in the translation refer to the meaning of the parafoveal preview (or to phonological representation in case of the no-diacritics preview), in the following order: Identical, Inaccurate, No-Diacritics previews.

Figure 2

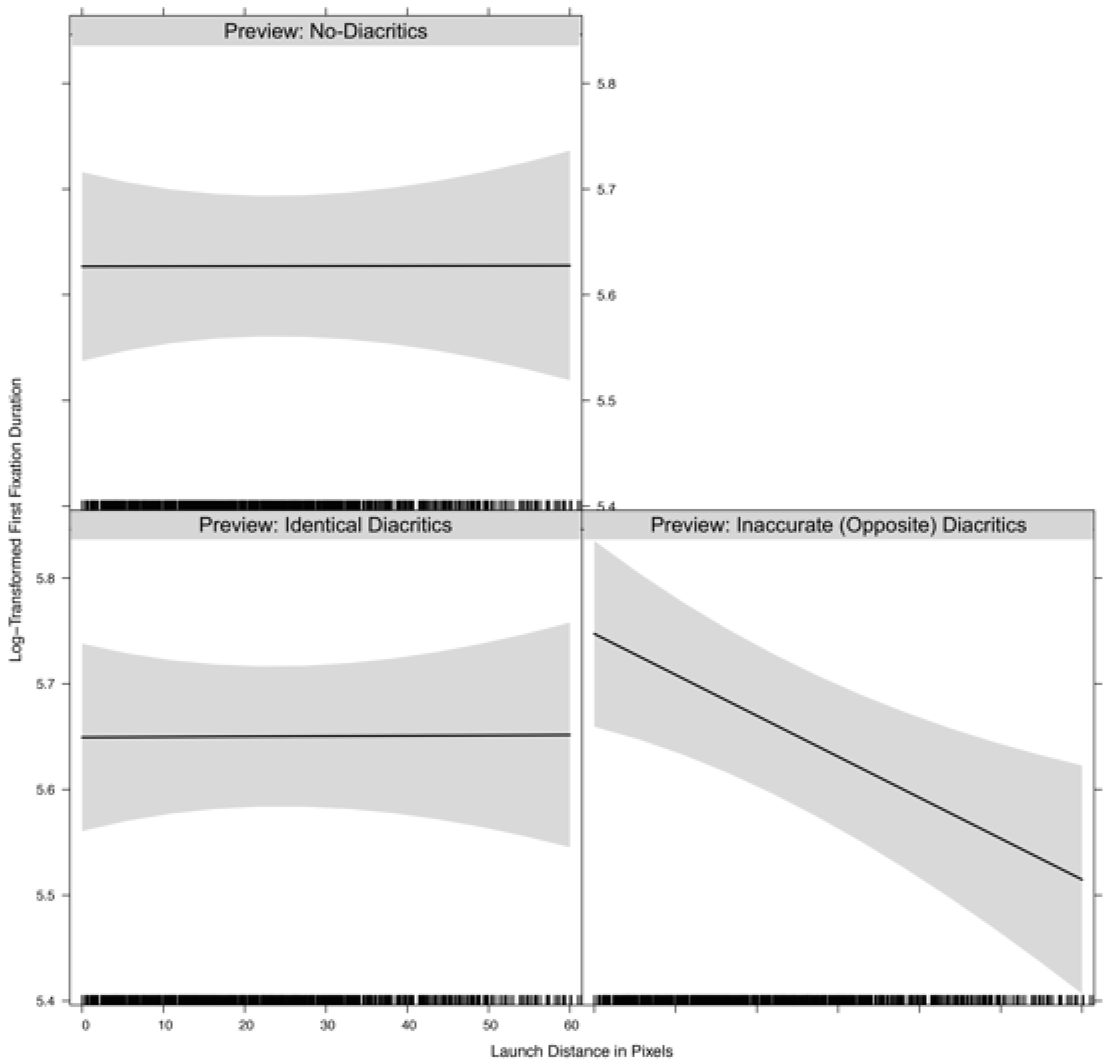

Figure 2. The interactions between preview availability × launch site in first fixation duration. The x-axis plots launch site in pixels (one character was on average 7.7 pixels wide). Launch sites are closer to the left. The y-axis plots log-transformed fixation duration. The grey bands represent 95% confidence intervals.

Figure 3

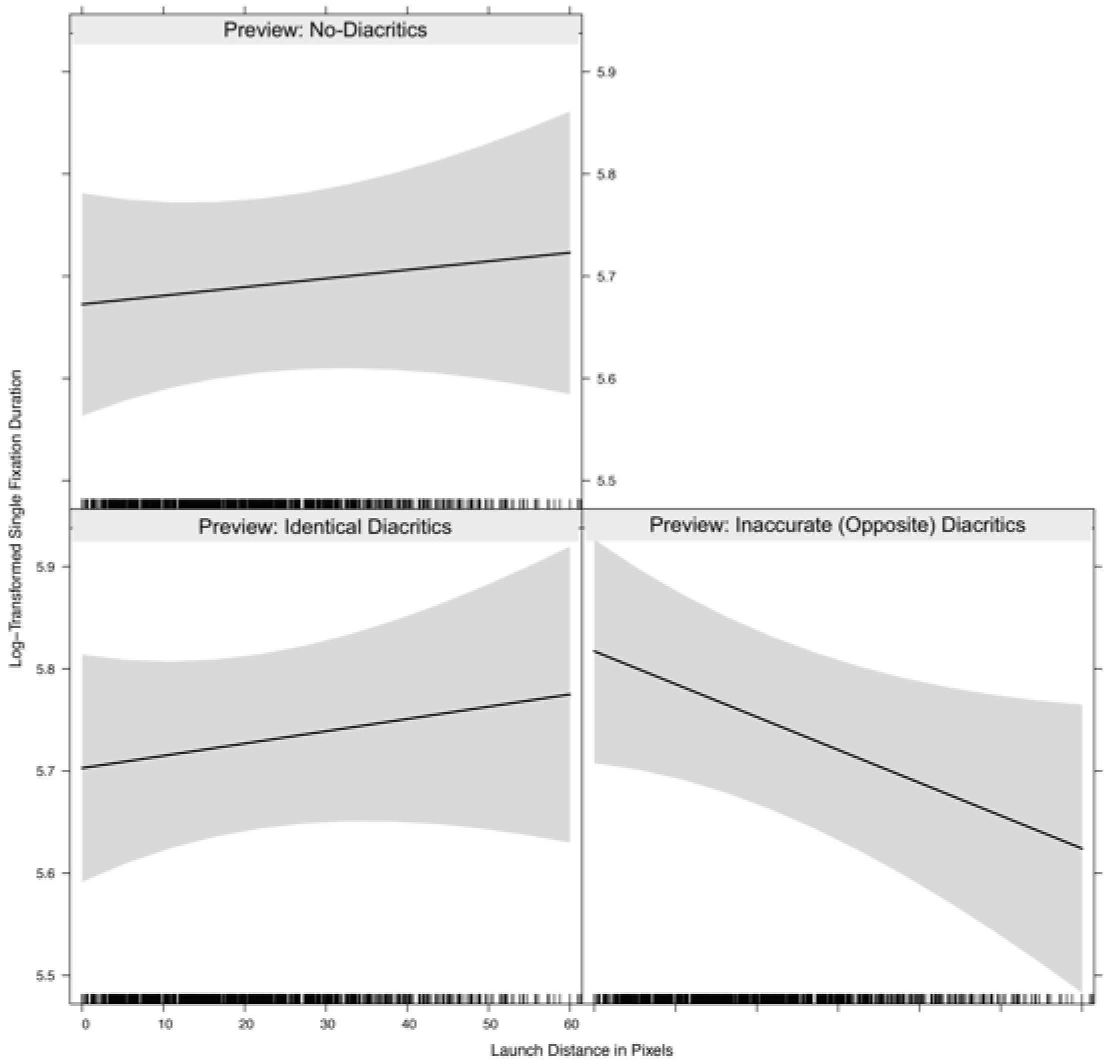

Figure 3. The interactions between preview availability × launch site in single fixation duration. The x-axis plots launch site in pixels (one character was on average 7.7 pixels wide). Launch sites are closer to the left. The y-axis plots log-transformed fixation duration. The grey bands represent 95% confidence intervals.

Figure 4

Figure 4

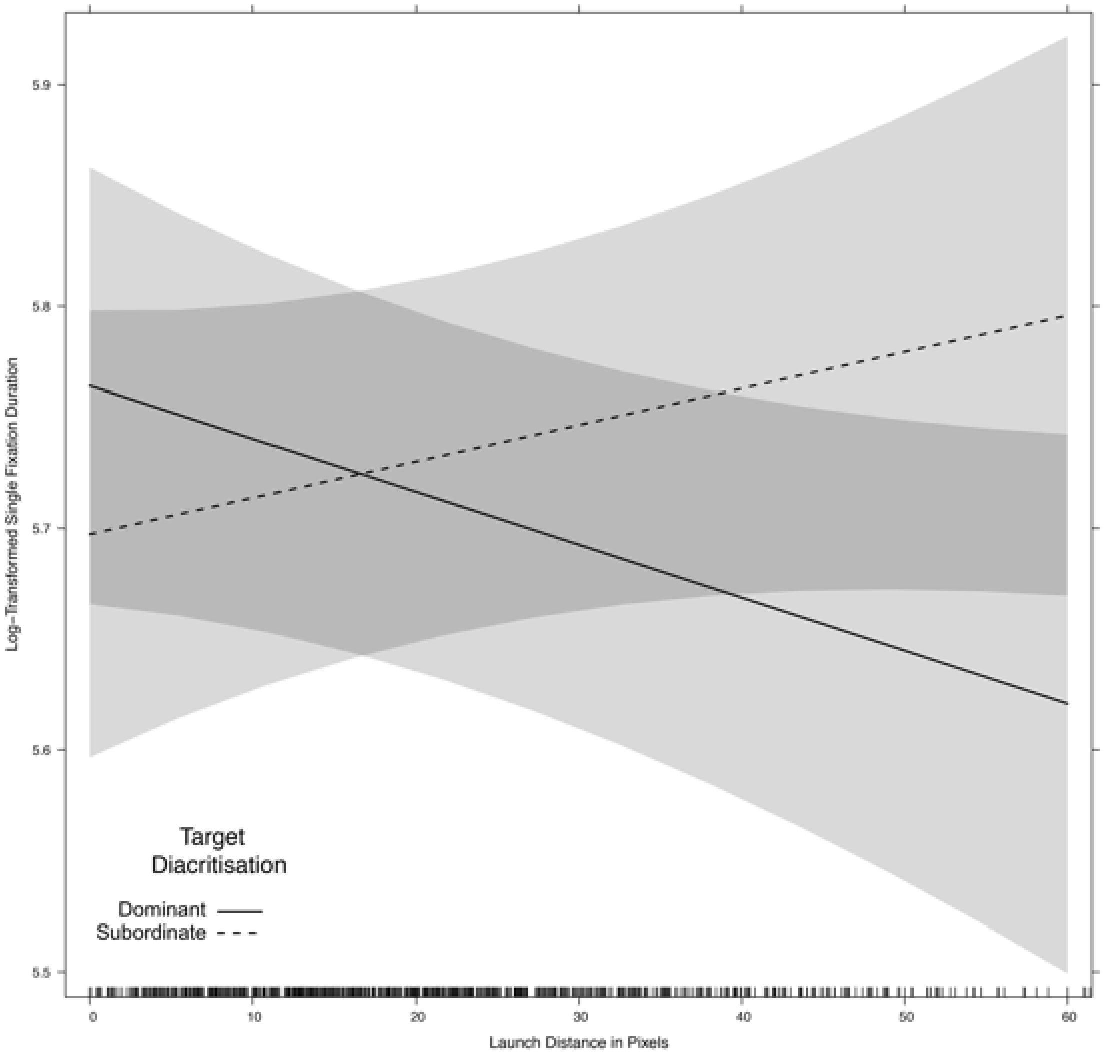

Figure 4. The interaction between target diacritization (dominant vs. subordinate) × launch site in single fixation duration. The x-axis plots launch site in pixels (one character was on average 7.7 pixels wide). Launch sites are closer to the left. The y-axis plots log-transformed fixation duration. The grey bands represent 95% confidence intervals.

Figure 5

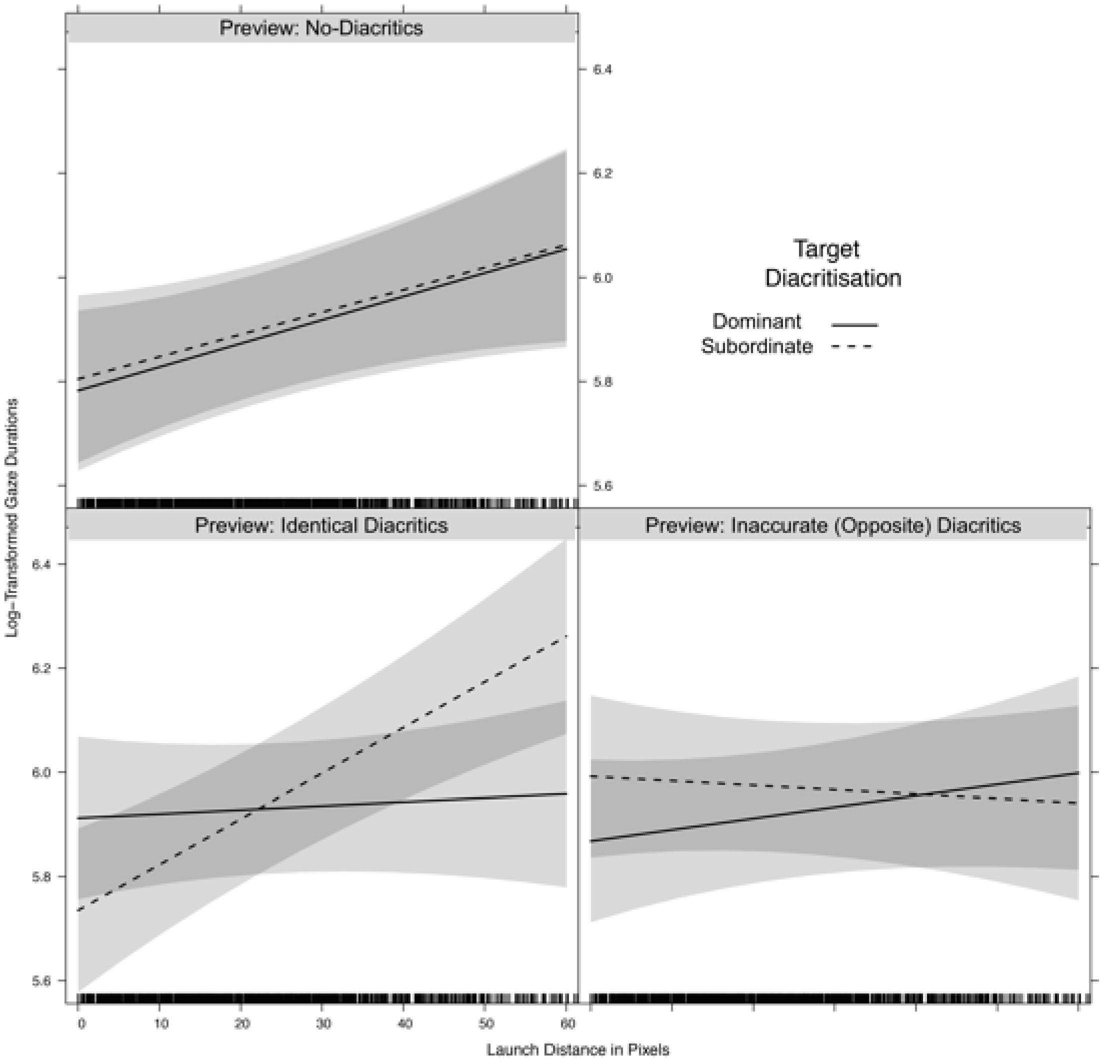

Figure 5. The 3-way interaction between preview availability (inaccurate vs. identical) × target diacritization (dominant vs. subordinate) × launch site in gaze duration. The x-axis plots launch site in pixels (one character was on average 7.7 pixels wide). Launch sites are closer to the left. The y-axis plots log-transformed fixation duration. The grey bands represent 95% confidence intervals.