

This is a repository copy of *Processing of Arabic Diacritical Marks::Phonological-Syntactic Disambiguation of Homographic Verbs and Visual Crowding Effects*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/99470/>

Version: Accepted Version

Article:

Hermena, Ehab, Drieghe, Denis, Hellmuth, Sam orcid.org/0000-0002-0062-904X et al. (1 more author) (2015) *Processing of Arabic Diacritical Marks::Phonological-Syntactic Disambiguation of Homographic Verbs and Visual Crowding Effects*. *Journal of Experimental Psychology: Human Perception and Performance*. pp. 494-507. ISSN 1939-1277

<https://doi.org/10.1037/xhp0000032>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Processing of Arabic Diacritical Marks: Phonological-Syntactic Disambiguation of
Homographic Verbs and Visual Crowding Effects

Ehab W. Hermena¹, Denis Drieghe¹, Sam Hellmuth² and Simon P. Liversedge¹

¹ University of Southampton, ² University of York

Author Note

Ehab W. Hermena, Denis Drieghe, and Simon P. Liversedge, Center for Vision and Cognition, Psychology, University of Southampton, Southampton, England. Sam Hellmuth, Department of Language and Linguistic Science, University of York, York, England.

This research was supported by Research Grant RPG-2013-205 from the Leverhulme Trust. We thank Sachiko Kinoshita, Jon Andoni Duñabeita and two anonymous reviewers for their helpful suggestions on an earlier version of this article.

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

Abstract

Diacritics convey vowel sounds in Arabic, allowing accurate word pronunciation. Mostly, modern Arabic is printed non-diacritised. Otherwise, diacritics appear either only on homographic words when not disambiguated by surrounding text or on all words as in religious or educational texts. In an eye tracking experiment we examined sentence processing in the absence of diacritics, and when diacritics were presented in either modes. Heterophonic-homographic target verbs that have different pronunciations in active and passive (e.g., ضرب /d^ar^ab^a/, *hit*; ضرب /d^orⁱb^a/, *was hit*) were embedded in temporarily ambiguous sentences where in the absence of diacritics, readers cannot be certain whether the verb was active or passive. Passive sentences were disambiguated by an extra word (e.g., بيد /bⁱj^ad/, *by the hand of*). Our results show that readers benefitted from the disambiguating diacritics when present only on the homographic verb. When disambiguating diacritics were absent, Arabic readers followed their parsing preference for active verb analysis, and garden path effects were observed. When reading fully diacritised sentences, readers incurred only a small cost, likely due to increased visual crowding, but did not extensively process the (mostly superfluous) diacritics, thus resulting in a lack of benefit from the disambiguating diacritics on the passive verb.

Keywords: diacritics, reading Arabic, eye movements, garden path, parsing preference

Reading in a number of the world's languages has been studied using the methodology of tracking readers' eye movements (see e.g. Rayner, 1998; 2009). Eye tracking research in Arabic sentence reading has explored basic properties of the perceptual and oculomotor systems (e.g., the size of readers' effective visual field during reading Arabic text—the perceptual span, Jordan, Almabruk, Gadalla, McGowan, White, Abedipour, & Paterson, 2014) and questions relating to single word processing (e.g. Farid & Grainger, 1996; Jordan, Almabruk, McGowan, & Paterson, 2011; Jordan, Paterson, & Almabruk, 2010). Eye movement investigations exploring sentence and text processing, beyond the level of the isolated word (e.g., Roman & Pavard, 1987; Roman, Pavard, & Asselah, 1985) can help us to better understand the influence of the unique linguistic properties of Arabic on reading (e.g. the role of vowel phonology in Arabic in the computation of syntactic structure, and the construction of semantic representations). In addition, these unique properties allow us to pose novel questions concerning written language processing in general. Eye tracking is a non-intrusive way of studying the cognitive processes associated with reading since readers' eye movements are tightly linked with these processes (e.g., Liversedge & Findlay, 2000; Rayner & Liversedge, 2004; 2011; Rayner, 1998; 2009). The research reported here used eye-tracking methodology to explore readers' processing of Arabic vowel phonology, and whether, and how, this phonological processing interacts with syntactic processing.

Arabic is an alphabetic language, which, like Hebrew, is written and read from right to left. Also like Hebrew, letters mainly denote consonant sounds, whereas short vowels are denoted by diacritical marks (diacritics hereafter, see Abu-Rabia, 2002; Haywood & Nahmad, 1965; Schulz, 2004), which are added to the written letters. This vowel information has the potential to disambiguate the pronunciation of words

with identical orthography (homographs) but which have a number of possible pronunciations (i.e., heterophones). An example of such a word in English would be the word *lead*, which denotes either a verb, or a noun. Heterophonic-homographs of this type are very common in Arabic: One estimate is that every second or third word in Arabic text is a heterophonic-homograph (e.g. Abu-Rabia 1997a; 1998). For instance, the string كُتِبَ /ktb/ can be pronounced $k^u t^u b$, meaning *books* (noun); $k^a t^a b^a$, meaning *he wrote* (past tense active verb); or $k^u t^i b^a$, meaning *was written* (past tense passive verb). In these examples, the vowels presented in ^{superscript} are used to illustrate how the pronunciation of each consonant is altered depending on the added ^{vowel} sounds: T^u , T^a , or T^i (e.g., the specific examples above denote respectively كُتِبَ, كَتَبَ, and كَتِبَ each with its intended pronunciation). This is in contrast to the non-diacritised form كُتِبَ. Diacritics then can inform the reader as to the pronunciations of these identical letter strings that are different for each syntactic case: Noun, active verb, or passive verb respectively. In other words, it is clear from this example that in addition to providing phonological information, diacritics also provide information about the syntactic status of a letter string, which can potentially serve as (syntactically) disambiguating information if such strings are encountered in a temporarily structurally ambiguous context.

Modern Standard Arabic (MSA) is typically printed without diacritics (Schulz, 2004). Thus, when heterophonic-homographs are presented in text, skilled readers use grammatical and semantic cues to disambiguate them (Abu-Rabia, 1997a; 1997b; 1998). If presented as single words, readers make use of the word diacritisation, if present, to ensure they pronounce the word accurately. In printed MSA, heterophonic-homographs may be diacritised if the surrounding text does not disambiguate them (Schulz, 2004). Some texts are printed fully-diacritised, that is,

with diacritics added to every word. This is the least common form of use of diacritics but is far from being rare or unusual. Fully-diacritised texts are typically: (a) Children's educational texts (up to 9-10 years of age and thereafter diacritics are removed, Abu-Rabia, 1999; similar to Hebrew texts for learners, e.g. Shany, Bar-on, & Katzir, 2012); and, (b) Texts where great precision of pronunciation is required, such as religious texts and poetry, (Haywood & Nahmad, 1965; Schulz, 2004), which are encountered regularly by skilled and novice readers. Whether skilled readers process diacritics in a similar way when dealing with these two modes of diacritisation (added only to the heterophonic-homographs, or to the whole text) during natural reading, is one issue the current research aimed to explore.

Until now, most investigations focusing on processing diacritics in both Arabic and Hebrew has been limited to isolated word tasks (e.g. lexical decision tasks, LDTs; and naming), with or without priming (e.g. Bentin, Bargai, & Katz, 1984; Bentin & Frost, 1987; Koriat, 1984); and offline measures in text reading tasks (e.g. reading speed and accuracy, or comprehension measures, e.g. Abu-Rabia, 1997a; 1997b; 1998; 1999). Some studies have suggested that the presence of diacritics contributes to improved reading accuracy and comprehension. Abu-Rabia (1997a), for instance, presented readers with single words, sentences and paragraphs, which were diacritised, or non-diacritised. The readers were 10th grade (about 16 years old) native Arabic speakers, who were split into high and low ability on the basis of their performance on a fully-diacritised, single word, reading aloud task. With reading (aloud) accuracy as the dependent measure, Abu-Rabia reported that both groups of readers benefited from the presence of diacritics, particularly in single-word reading. Similar findings were reported for similar tasks in other investigations where reading involved different types of Arabic texts (Quranic, poetry, as well as newspaper and

informative prose, e.g., Abu-Rabia, 1997b; 1998; 1999; 2001). Abu-Rabia concluded that diacritics facilitate reading Arabic aloud. Additionally, the improved silent reading comprehension findings for reading diacritised Arabic and Hebrew texts, in Abu-Rabia (2001), replicate those reported by Shimron and Sivan (1994), where the presence of diacritics improved native Hebrew readers' performance.

Other investigations of processing diacritised and non-diacritised single Hebrew words (e.g., Bentin & Frost, 1987; Koriat, 1984; 1985a; 1985b) highlighted, to some extent, the interaction between the presence of diacritics and task demands. In all of these investigations, the presence of diacritics in Hebrew had no, or a negative (slowing), influence on response time in lexical decision times (LDTs). Indeed, performance on these tasks was less affected by the presence of diacritics than in word naming tasks, where performance was significantly slower when the words were diacritised (though, c.f. Roman & Pavard, 1987, Experiment 3). As an explanation for this pattern of findings, Bentin and Frost suggested that full and complete phonological access, including diacritics-based phonological information, is necessary in naming tasks, compared to LDTs. However, Bentin and Frost also suggested that if diacritics are present in LDTs, readers do not ignore them completely, for example, they documented a word frequency effect whereby homographic words diacritised as the high-frequency pronunciation were responded to faster than those diacritised as the low-frequency pronunciation.

In light of the findings discussed above, we can consider two questions. First, to what extent are diacritics processed automatically, that is, involuntarily, in a reflexive and mandatory manner during normal text reading? And second, does the mode of the diacritisation (diacritics only on critical heterophonic-homographs, vs. fully diacritised text) affect the manner in which text is processed during normal

reading? Roman and Pavard (1987) tracked native Arabic readers' eye movements while silently reading passages of text (about 95 words) which were either fully- or non-diacritised. They reported that the presence of full text diacritics resulted in a significant reduction in reading speed, as well as significant increases in the number of fixations made and the duration of fixations. They also reported a 75 ms increase in gaze duration per word in the fully-diacritised condition. The authors suggested that these effects are attributable to one of two causes whereby the presence of full text diacritics could result in slower reading. First, diacritics might increase what they termed *perceptual noise*, whereby the additional visual information (the diacritics) may interfere with adjacent visual materials, which may result in delayed word identification—an issue which we return to below. Second, the presence of diacritics might induce additional diacritics-based syntactic and semantic information processing that would be more costly in terms of processing time. However, the results do not clarify which, or whether both, possible causes produce the findings. In addition, it is not clear to what extent the results reflect online processes associated solely with reading given that the participants in this investigation were told that when they finished reading each passage they would be required to provide a verbal summary of its content. The reading times, thus, may reflect the costs of processes other than reading (e.g., memorization and rehearsal).

Note also, that when only an ambiguous word is diacritised in a sentence, and not the entire sentence as per the Roman and Pavard (1987) study, and participants were required to read sentences aloud, readers were very accurate (almost at ceiling level) in their pronunciations compared with when these verbs are not diacritised (Roman et al., 1985). However, again it is not clear whether processing of diacritics in this situation is influenced by the task demand of reading aloud (see also Abu-

Rabia, 1997b; 1998; 1999; 2001; or in Hebrew word naming tasks e.g., Bentin & Frost, 1987) compared to normal silent reading.

To further explore processing of diacritics to disambiguate heterophonic-homographs during natural silent reading, Roman et al. (1985) reported an eye tracking experiment where they used verbs that were heterophonic-homographs whose active and passive voice pronunciations differ, while letter orthography remains identical. This is an ideal means of exploring the processing of diacritics that can be used to disambiguate the active and passive pronunciations of such verbs. Arabic active and passive sentences have very similar structure: Both contain a verb followed by a noun phrase (representing the *agent* in the active form, and the *patient* in the passive form). Roman et al. (1985) presented these verbs either diacritised as active or passive (disambiguated) or non-diacritised (ambiguous). The verbs were embedded in frame sentences such that the passive sentences were disambiguated as passive downstream by the presence of a prepositional phrase (PP) followed by a noun phrase representing the *agent*. Active sentences, on the other hand, ended with a noun phrase that represented the *patient*. Roman et al. reported increased reading times of diacritised passive verbs, compared to non-diacritised passive, and no difference between two active conditions (diacritised and non-diacritised).

However, due to major methodological shortcomings, the findings reported by Roman et al. (1985) are not interpretable with regards to the theoretical issues our investigation deals with. For instance, Roman et al. required the readers to repeat sentences as accurately as possible, with eyes closed, after the initial reading. It is possible that the memorisation demands of this additional task may have altered readers' eye movement patterns collected during reading. More importantly, when analysing fixation durations, the authors did not differentiate between first pass

(initial reading) measures, which index early processes (e.g. lexical) and second pass (re-reading) measures, which index later processes (e.g., reanalyses or integration). This severely limits our ability to interpret their results and learn about the time course of linguistic processing in this experiment. Furthermore, the pattern of findings reported at the passive disambiguating region (PP) is unexpected and suggests that readers did not detect their misanalysis of non-diacritised passive sentences as active. This deviates from established findings that readers do detect their own misanalyses at disambiguating regions, and considerable disruption to processing is typically recorded at these regions (see e.g. Frazier & Rayner, 1982; Lipka, 2002; Liversedge, Paterson, & Clayes, 2002; Murray & Liversedge, 1994). Finally, the extremely low power of this investigation (only 5 participants were tested) may have compounded the problems with this investigation.

In the current investigation we aimed to address a number of important theoretical questions that, as yet, have not been addressed satisfactorily in the literature. The first of these relates to how readers process diacritics in normal Arabic reading, and whether this processing is influenced by the mode of text diacritisation. The second, tightly linked, question relates to whether Arabic readers have an initial preference for processing verb-noun pairs as actives over passives, and whether, and how, they make use of disambiguating diacritics to resolve the ambiguity of the verb voice (active vs. passive). To this end, we used verbs that are heterophonic-homographs whose active and passive voice pronunciations differ, while letter orthography remains identical (as per Roman et al., 1985). Similarly, we embedded these verbs in carefully constructed sentence frames to examine the processing of disambiguating diacritics during normal sentence reading. In other languages, English for instance, readers have a strong and well-documented preference for

analysing *noun-verb-noun* structures as *subject-verb-object* (SVO, i.e., the active analysis, e.g., Fodor, Bever & Garrett, 1974; see also Ferreira, 2003; Rayner, Carlson & Frazier, 1983; Liversedge et al., 2002), and it is possible that a similar preference is at work in Arabic. A preliminary indication that this may be the case was obtained from a survey that we conducted of over 5,000 sentences from Arabic texts (books from various genres and literary appendices of newspapers), from a number of Arabic-speaking countries (Egypt, Kuwait, Jordan, Lebanon, Saudi Arabia and Qatar), which were published over the last 30 years. In these texts we explored those occasions on which words were diacritised, and we found that active verbs were never assigned diacritics when they appeared in non-diacritised text. Typically, passive verbs were diacritised, but only when they were not disambiguated as passives by the surrounding sentential context (100% of instances encountered in our survey). Assuming that diacritics are used to overtly demark verbs as passives when they are ambiguous and must be processed in a non-preferred form, then it seems likely that the non-diacritised versions of such verbs are typically analysed as the active.

If Arabic readers do have a parsing preference for an active over a passive analysis, then they are likely to pursue such an analysis when they encounter a non-diacritised verb followed by a noun (typical word order in Arabic), e.g. *دفعت الطالبة* (*the student pushed*, see Figure 1) embedded in a neutral context. Whilst this analysis would be accurate in the non-diacritised active condition (Act-Non-D), in the non-diacritised passive condition (Pass-Non-D), this analysis would turn out to be incorrect. Hence, when the reader processes the text at the disambiguating PP region (region ii, *Figure 1*), they should be garden pathed (e.g., Frazier & Rayner, 1982) and disruption to processing should be observed at this point in the sentence.

<Figure 1 about here>

Specifically, when the sentences contain a non-diacritised passive verb, then at disambiguation, we would expect inflated first pass reading times (first fixation: the first fixation a reader makes on a word during first reading; single fixation: the first and only fixation made by the reader on a word during first pass reading; and gaze duration: the sum of the durations of all fixations made by the reader on a word from entering the region of interest until they exit this region going forward or backwards in the text). Such effects would be consistent with previous findings in English, for instance, Rayner et al. (1983), who reported that readers experienced disruption in the disambiguating region of relative clause sentences that were temporarily ambiguous between a passive and a simple active reading. Rayner et al. argued that readers were garden pathed because they initially processed such ambiguous sentences according to the principle of *minimal attachment* (attaching newly encountered material in a sentence to existing representational structures for the sentence, using the fewest nodes possible). We might also observe disruption to processing in subsequent sentence regions (e.g., region iii the spillover region, and region iv, the end of sentence region; see *Figure 1*). Such effects might reflect not only disruption associated with the detection of the initial syntactic misanalysis, but also processing associated with the recovery from that initial misanalysis. Furthermore, we would expect this disruption to persist beyond first pass reading times into later reading measures in the disambiguating and possibly also the subsequent regions. Following the work of Rayner, Frazier and their colleagues, there have been numerous investigations that have documented the presence of such effects in readers' eye movement records when their initial parsing preferences were challenged (e.g.,

Clifton, Traxler, Mohamed, Williams, Morris, & Rayner, 2003; Lipka, 2002; Liversedge et al., 2002; Murray & Liversedge, 1994; Paterson, Liversedge, & Underwood, 1999; Rayner & Frazier, 1987). However, to date, there has been no eye movement research to investigate whether comparable garden path effects occur in Arabic reading. Thus, if readers process the disambiguating diacritics on the homographic verb (in the Pass-Verb-D condition), and they use this on-line to guide syntactic processing, then no garden path effect should occur at the disambiguating PP. Thus, comparing reading times at the PP for the Pass-Non-D and the Pass-Verb-D conditions (*Figure 1*) should allow us to learn (1) whether readers' have a initial parsing preference for an active analysis, and (2), whether the presence of disambiguating diacritics on the ambiguous verb causes readers not to be garden pathed.

Given the characteristics of homography and the use of diacritisation in Arabic discussed above, we can use the same experimental conditions (Pass-Verb-D and Pass-Full-D) to answer our original theoretical question of whether diacritics-based parsing guidance is contingent on the mode of diacritisation (i.e., only on the homograph vs. full sentence). It is an open question whether diacritics on the main verb of the sentence will provide as effective a cue to sentential syntactic structure when the sentence is fully diacritised, relative to when the verb alone is diacritised. In the fully-diacritised passive condition (Pass-Full-D), if readers process the phono-syntactic information provided by the diacritics at the verb, then, similar to the prediction made for the Pass-Verb-D condition, garden path effects should be absent. However, given that adult readers have long experience of processing non-diacritised text, then their approach to processing diacritics in full diacritisation mode may not necessarily be similar to that when diacritics are added only to homographs that

require disambiguation. Furthermore, for these readers, fully-diacritised texts are, typically, regularly encountered, and well-learnt texts (arguably over-learnt in some cases, e.g. religious texts or poetic verses), which require little disambiguation through diacritics-based phonological analysis. Consequently, to skilled readers, diacritics in full sentence diacritisation mode may be to a certain extent redundant, and, therefore, potentially ignored as a cue to the appropriate syntactic analysis during normal reading. Thus, if processing of diacritics is indeed influenced by the mode of diacritisation, a different pattern of findings may be plausibly predicted. In its most extreme form, for fully diacritised text in the Pass-Full-D condition, readers may fail to process the disambiguating diacritics at the passive verb, and consequently experience garden path effects upon arrival at the disambiguating, PP, region, similar to the non-diacritised passive condition.

Another unavoidable aspect of processing diacritics during reading is the fact that full sentence diacritics, when present, add a considerable amount of visual information to the text. This is clear from comparing the sentences in the non-diacritised conditions (Act- and Pass-Non-D) with the fully-diacritised ones (Act- and Pass-Full-D) in *Figure 1*. In other writing systems, English for instance, increasing the amount of visual information in the same space is sometimes referred to as *visual crowding*, which results in lateral inhibition, whereby interference of adjacent visual materials (c.f., adjacent letters within a word) slows the identification of that word (see e.g., Paterson & Jordan, 2010; Slattery & Rayner, 2013). The effects of crowding or lateral inhibition have also been explored via investigations of word spacing in alphabetic (Bouma, 1970; 1973; Drieghe, Brysbaert & Desmet, 2005, see also Pelli, Tillman, Freeman, Su, Berger, & Majaj, 2007) and other writing systems (e.g., Chinese: Bai, Yan, Liversedge, Zang, & Rayner, 2008; Zang, Liang, Bai, Yan,

& Liversedge, 2013). In all these investigations, visual crowding was mainly described in terms of the effects of letter-, character-, and word-spacing, and was shown to reduce the speed of reading (e.g., increased numbers and durations of fixations; as well as decreased word-per-minute reading counts). Reading speed typically recovered when crowding was reduced. Slattery and Rayner (2013), for instance, found that as the within-word letter space decreased from normal size to half-normal size, average fixation duration increased by about 6 ms. Somewhat similarly, Liversedge, Zang, Zhang, Bai, Yan, and Drieghe (2014) manipulated the visual properties of Chinese characters in terms of the number of strokes comprising a character. Chinese characters vary in their visual and linguistic complexity while occupying the same amount of space (e.g. Zang, Liversedge, Bai, & Yan, 2011), and an increased number of strokes in a character leads to an increase in the time required for its identification (Yang & McConkie, 1999). In their experiment, Liversedge, et al. orthogonally manipulated Chinese character frequency and visual complexity, where the number of strokes in target characters was considered a proxy of visual complexity. They reported interactive effects for first and single fixation and gaze durations such that readers spent about 9-11 ms longer fixating low frequency characters of high complexity relative to characters in all the other conditions. Arguably, increased visual complexity due to the presence of diacritics could result in similar effects for Arabic sentence reading; this is what Roman and Pavard (1987) termed *perceptual noise*. The amount of visual information in fully diacritised sentences is clearly increased relative to the same sentence in its non-diacritised form. As yet, whether full-diacritisation results in similar visual crowding effects during natural reading have not been explored. Obviously, reducing letter-spacing in Latinate languages, increasing stroke count in Chinese characters, and the presence of

full sentence diacritics in Arabic are very different phenomena, but perhaps like the earlier two, full sentence diacritisation in Arabic is a source of visual crowding effects during reading. No investigation of the influence of diacritics on eye movements during reading can ignore this issue given the striking visual differences between diacritised and non-diacritised sentential forms, and the fact that diacritised Arabic text is used by default in both language teaching materials and is regularly encountered in significant pieces of Arabic literature. Therefore, we also considered this aspect of processing in the current study by comparing global measures of eye movement behaviour for the diacritised and non-diacritised active conditions (Act-Non-D vs. Act-Full-D, *Figure 1*). These two conditions are ideal to examine how full diacritisation may result in crowding effects as no garden path effects are expected for them, and any increase in reading times may be attributed solely to the increased visual complexity of the diacritised text. The findings discussed above (e.g. Drieghe et al., 2005; Liversedge et al., 2014; Slattery & Rayner, 2013) suggest that crowding effects are subtle and manifest as small, but significant, increases in fixation durations. Thus, we anticipated that average fixation durations, total number of fixations and total sentence reading times would be increased for the sentences in the Act-Full-D condition relative to those in the Act-Non-D condition.

To summarise, in this experiment we explored the theoretical questions pertaining to processing of diacritics, in both modes of diacritisation (on the homograph alone, or on the full sentence), to disambiguate heterophonic-homographic verbs. Furthermore, our sentence manipulation allowed us to simultaneously answer another linked theoretical question relating to whether Arabic readers have a parsing preference for simple active analysis over passive. If readers do exhibit a simple active parsing preference, then they should be garden pathed at the

disambiguating PP in passive sentences without diacritics (the Pass-Non-D condition), relative to those with diacritics on the verb (Pass-Verb-D condition), and potentially, fully diacritised passive sentences (Pass-Full-D). Any such disruption may also spill over into later sentence regions (e.g., the end of sentence region), and into later reading time measures on the disambiguating and subsequent regions (e.g., go past time: the sum of all fixation durations made from entering the region of interest until exiting this region forwards), reflecting sentence re-analysis associated with recovery from the initial garden path. The absence of garden path effects in the Pass-Verb-D and Pass-Full-D conditions would provide strong evidence for the automatic use of information conveyed by diacritics in relation to initial syntactic commitments, regardless of the mode of diacritisation. Finally, we explored whether full sentence diacritisation produced visual crowding effects, that we anticipated would manifest as small but significant increases in fixation times for the diacritised compared with the non-diacritised active sentences (Act-Full-D vs. Act-Non-D sentences respectively). It is important to reiterate that, although less common, fully-diacritised texts are regularly encountered by Arabic readers, and for this reason, it is very important to consider reading behaviour in response to this mode of text diacritisation.

Method

Participants

Twenty five adult native Arabic speakers were paid £10 for participation. All participants were UK residents or visitors (e.g., international students). The

participants (12 females) ranged in age between 18 and 61 (mean = 34.5, SD = 10.5, 8 participants > 36 years, 3 participants > 50 years, 1 participant > 60 years¹). All participants had normal or corrected-to-normal vision. The majority of participants spoke and read English as a second language. Participants were required to provide additional information about their Arabic reading habits and experience, and the time they spent away from their native Arabic-speaking countries. 19 of the participants (76%) indicated that they read Arabic daily, and the remainder indicated that they read Arabic at least once a week. All participants had normal or lens-corrected vision, and all reported being able to clearly see the words and diacritics on the screen during a practice block. Their high comprehension scores (see Results) clearly indicate that they were able to see and read the materials on the screen with ease.

Stimuli

Forty target sentences were constructed, with 5 versions of each, making up the 5 experimental conditions (full- and non-diacritised active and full-, non- and verb-only-diacritised passive). *Figure 1* above contains a sample stimulus sentence. As explained above, the main interest regions, in canonical order, are the verb (region i), which in each sentence was a heterophonic-homograph, followed by a noun which was the agent in active sentences and the patient in passive sentences. These were followed by a 5-7 words long interjectory phrase, which was always of neutral meaning and provided no disambiguating information to the readers about whether the homographic verb was active or passive. This ensured that any disambiguating material was outside of the perceptual span whilst fixating the verb. Subsequently the sentences contain a second noun phrase which represented the patient in active, whereas in passive sentences it represented the agent (region iii), or the spill-over

region, given that it was preceded by the optional prepositional phrase (PP, region ii) which acted as the passive-disambiguating region.

Sixty filler sentences of similar length and complexity to the target sentences, which also contained active or passive verbs, were also presented to the participants. These sentences were presented, like the targets, with full-, verb-only, or no-diacritics. Ten additional sentences made up a practice block, thus each participant read 110 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 cursive non-monospaced font (Traditional Arabic, size 18, roughly equivalent to English text in Times New Roman font size 14). Our choice to use a non-monospaced font was deliberate to ascertain that the reading experience was as natural as possible (e.g., Boudelaa & Marslen-Wilson, 2001; 2005; Ibrahim & Eviatar, 2009), given that monospaced fonts are typically not used in Arabic print. The reason for using this particular non-monospaced font is for the clarity with which it displays the diacritical marks.

All sentences were normed whereby 10 ratings of verb commonness and 10 ratings of sentence structure naturalness were obtained on 5-point scales (1 = the verb is rare/the sentence structure is highly unnatural, 5 = the verb is very common/the sentence structure is perfectly natural). The additional participants who performed the norming did not take part in the eye tracking experiment and were recruited using Amazon Mechanical Turkers (AMT). In addition to the required ratings, the Turkers were required to perform other linguistic tasks (e.g., construct original new sentences) to demonstrate that they were native and skilled Arabic speakers. Participants were paid \$10 if their work was accepted. No participant took part in the task more than

once. The ratings showed that the target verbs were rated as highly common (average = 4.46, SD = 0.35, range = 3.3 – 5), and the naturalness of the sentences was rated as high (average = 4.37, SD = 0.40, range = 3.2 – 5). The target verbs were on average 3.7 letters long (SD = 0.73, range = 3-6), with the majority of the verbs used being either 3 or 4 letters long. The target verbs had an average orthographic frequency of 28.6 per million (SD = 43.1, range = 0.13 – 168.12) in the Aralex corpus (Boudelaa & Marslen-Wilson, 2010).

Prior to taking part in the eye movement experiment each participant took part in a paper-based screening task. Each participant was presented with a printed passage extracted from an Arabic newspaper (146 words, the topic was the Kuwaiti Stock Market), and was required to read this aloud to allow the experimenter for screening to check their accuracy (percentage of words read accurately). This resulted in the exclusion of 1 participant (mean text reading accuracy = 98.86%, SD = 0.96, range = 97.26 – 100%).

Following the participation in the eye tracking procedure, participants were presented with a printed list of single words (target words were 36 diacritised words and 24 undiacritised words) to read aloud from a card. The diacritisation patterns on the target words were similar to those of the target verbs in the target sentences and allowed us to test their fluency in the use of diacritics. Participants were fluent decoders of diacritical marks (mean word reading accuracy = 92.78%, SD = 8.52, range = 72.2 – 100%).

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 words were in black on a light grey background. The display was 70cm from the participants, and at this distance, an average of 3.8 characters equalled about 1° of visual angle.

Participants used a Microsoft gaming button box to enter their responses to comprehension questions and to terminate trials after reading the sentences.

When participants read screening materials aloud their voices were recorded using a standard digital voice recorder.

Design

The stimuli in the 5 conditions (Act-Non-D, Act-Full-D, Pass-Non-D, Pass-Verb-D, & Pass-Full-D, see *Figure 1*), were counterbalanced using a Latin square and presented in random order such that all participants saw each sentence only once, and they saw an equal number of target stimuli from all conditions.

Procedure

This experiment was approved by the University of Southampton Ethics Committee. Upon arrival, participants were given a description of the apparatus and instructions for the experiment.

The participants started by signing consent forms, then they read aloud the first paper-test (the paragraph) while being audio-recorded. This was followed by the eye tracking procedure, then finally the single word reading task, again while being audio-recorded.

The eye tracker was calibrated, following a horizontal 3-point calibration at the beginning of the experiment and the calibration was validated. Calibration accuracy was always $< 0.25^\circ$, otherwise calibration and validation were repeated. The participants were told to read silently, and that they would periodically be required to answer yes/no questions about the sentences. Participants read the 10 practice sentences followed by the 100 experimental trials. Drift measurement was performed at the beginning of each trial with the fixation circle ($1^\circ \times 1^\circ$) appearing at the location of the first character of the sentence. 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 35-45 minutes.

Results

We used the *lmer* package (*lme4*, Bates, Maechler, & Bolker, 2011) within the R environment for statistical computing (R-Core Development Team, 2013) to run Linear Mixed Models. We report the output for the specific contrasts that allow us to address the theoretical questions we set out above. In all reported contrasts, the subjects and items were classed as random factors. For each contrast we report beta

values (b), standard error (SE), and t statistics for reading time and fixation count data. For each contrast we started with models containing maximal random effects structure (Barr, Levy, Scheepers & Tily, 2013) that was trimmed if failure to converge occurred (first by removing correlations between random effects, and if necessary also by removing interactions). All findings reported here are from successfully converging models.

The reported contrasts were carried out on log transformed fixation time data to reduce distribution skewing (Baayen, Davidson, & Bates, 2008). For skipping and regression data we performed logistic LMMs and thus report z statistics.

In all analyses, fixation times shorter than 80ms, or longer than 800ms were removed. However, fixations shorter than 80ms that were located within 10 pixels or less (about 0.33° of visual angle approximately) from another longer fixation, were merged into the longer fixations. Furthermore, for each of the fixation duration measures, we removed data points ± 2.5 standard deviations away from the mean fixation duration per participant.

Direct comparison between active and passive conditions was not possible given the inclusion of an extra word in the passive conditions (the disambiguating prepositional phrase), and also given the difference in the verb voice. Therefore, our analyses of the active and passive conditions will be carried out separately.

Specifically, in the passive conditions we focus on local analyses of specific sentence regions to examine readers' use of verb diacritics to avoid potential garden path effects in the Pass-Verb-D and Pass-Full-D conditions, whereas in active, sentences we focus on global measures to examine the processing of diacritics in the full diacritisation mode, in simple active text, and in the absence of potential garden path effects.

Comprehension questions followed 25% of all sentences. Participants responded accurately on average 84% of the time ($SD = 7.6$, range = 70.37 – 100%) indicating that participants read and understood the sentences. There were no differences between the accuracy scores across the conditions.

Examining the numerical trends for the older adult participants (4 who were over 50 years of age), we can confirm that the performance of this group was almost identical with the remaining sample in all eye movement measures reported below. This is an important point given that previous research has shown that older adults may adopt different text scanning strategies (Rayner, Reichle, Stroud, Williams, & Pollatsek, 2006) and respond differently to visual properties of text (Jordan, McGowan, & Paterson, 2014). Furthermore, the performance of the participants in the offline measure (sentence comprehension question) was not influenced by their chronological age ($r = 0.1$).

Garden Path Effects

To explore these effects, we ran two contrasts: First, between the Pass-Verb-D and the Pass-Non-D conditions; and second, between the Pass-Verb-D and the Pass-Full-D conditions. In both contrasts, the Pass-Verb-D was coded as the baseline condition, which is ecologically appropriate given the prevalence of diacritising single homographic words for disambiguation in most Modern Standard Arabic texts. We report analyses of eye movement measures at the regions outlined in Figure 1 (the verb, the disambiguating PP, the spillover, and end-of-sentence regions).

The Verb Region

First pass measures for this region included *skipping rate*, *first fixation duration*, *single fixation* and *gaze duration*. We also examined the extent to which readers made regressive saccades to revisit the verb region from subsequent sentence regions (*regressions in*). Descriptive statistics for all of the regions and all of the measures we report here and below are available in *Table 1*, and inferential statistics in *Table 2*.

<Table 1 about here>

<Table 2 about here>

Our analyses revealed that the verb was skipped significantly more in the Pass-Non-D condition compared to the Pass-Verb-D condition. However, there was no difference in skipping between the two diacritised (verb only and full sentence) conditions. This suggests a sensitivity to the presence of diacritics on the upcoming word that increases the likelihood of its being fixated. As for reading time measures: *first fixation*, *single fixation* and *gaze durations*, the removal of outliers resulted in removing 6.4%, 3.9%, and 4.9% of the data points from the analyses of these measures respectively. The results indicated that there were no significant differences between the passive conditions in any of the contrasts. There were also no significant differences between the conditions for the *regression in* measure. Thus, at the verb region, the only reliable effect occurred for word skipping, and this effect was entirely driven by the visual complexity of the verb. When it was diacritised and visually complex it was more likely to be fixated than when it was not, regardless of whether just the verb, or alternatively, the whole sentence itself was diacritised. It is

extremely unlikely that this effect has anything to do with the syntactic commitments that readers were making on line at the verb since at the point when the decision is made regarding whether to skip or fixate the verb, the eyes are positioned at a word preceding it. Thus, if syntactic considerations were to be associated with this decision, it would suggest that readers had parafoveally processed the word, fully identified it, accessed its syntactic category and made a parsing commitment all on the basis of a parafoveal view. We consider this possibility to be extremely unlikely, and if we are correct, then it implies that the effect is entirely driven by the visual complexity of the diacritised compared to the non-diacritised verb. This finding is in line with results from Liversedge et al. (2014) who also found visual complexity effects for word skipping in Chinese.

The Disambiguating (PP) Region

In addition to skipping, first fixation duration, single fixation duration and gaze duration, we will report the later processing measure of regressions out of the disambiguating region. The regressions out measure is defined as the proportion of regressive, rightward in Arabic, saccades from the PP region, and is known to be a measure sensitive to disruption associated with garden path effects (Frazier & Rayner, 1982; Murray & Liversedge, 1994).

There were no significant differences between the three passive conditions in any of the contrasts for skipping rates. We did, however, obtain reliable and consistent effects for the early reading time measures: *first fixation*, *single fixation*, and *gaze duration*, the removal of outliers resulted in removing 6.9%, 4.5%, and 7.1% of the data points from the analyses of these measures respectively. The results

indicated that readers spent significantly less time reading the disambiguating region in the Pass-Verb-D condition, compared to the Pass-Non-D condition. This difference provides evidence that readers were garden pathed when reading non-diacritised passive sentences relative to sentences in which the verb only was marked with diacritics as a passive.

The contrast between the Pass-Verb-D and the Pass-Full-D conditions, however, yielded a pattern of results that might initially be considered somewhat surprising. Recall that we predicted that if readers used the diacritics in fully diacritised sentences to guide initial parsing decisions, then garden path effects should be absent at the disambiguating PP. Recall also, however, that we considered an alternative possibility, namely, that to skilled readers full sentence diacritics may be redundant, and therefore, they may not engage in processing of information conveyed by the diacritics to the same degree that they do when diacritics are assigned exclusively to specific homographs for disambiguation (i.e., in the Pass-Verb-D condition). Our results showed that at the disambiguating region in the Pass-Full-D condition, readers made longer first and single fixations and had longer gaze durations compared to the Pass-Verb-D condition. This difference suggests that readers experienced garden path effects in the fully diacritised condition relative to the verb only diacritised condition. This occurred despite the presence of the disambiguating diacritics on the passive verb in both conditions. This pattern of effects indicates that when reading fully diacritised sentences, our sample of adult, skilled, readers did not make use of the diacritical cues at the verb to inform their initial syntactic commitments. This suggests that readers disengage from processing of diacritical information during silent reading to guide syntactic processing when diacritics are applied uniformly over the entire sentence, relative to when they are used more

discriminately to specifically disambiguate individual homographs.

We next considered regressions from the disambiguating PP, and found that there were no differences across conditions. Readers were no more likely to make a regression to re-read the text when they were garden pathed than when they were not. Furthermore, since the regression rate was very low (on average 7.3%), we did not analyse go past (regression path duration) reading times.

The spillover (PP+1) region

We computed *skipping rate*, *first* and *single fixation durations* and *gaze duration*. For the fixation duration measures, the removal of outliers resulted in removing 3.1%, 2.7%, and 6.2% of the data points from the analyses of the *first fixation duration*, *single fixation duration* and *gaze duration* on the spillover region and there were no significant effects for any measures.

End-of-sentence region

For the final region we computed first pass reading times. Removing outliers resulted in removing 4.6% of the data points from the analysis for this measure. We also computed the probability that readers made a regression to examine an earlier portion of the sentence. Finally, we computed the time participants spent re-reading the sentence. We defined re-reading time as the sum of all the fixations the reader made after making a regression from the final region until they pressed the button to indicate they had understood the sentence. Removing outliers resulted in removing 5.2% of the data points from the analysis for this measure.

This region received significantly longer first pass reading time in the Pass-Non-D condition compared to the Pass-Verb-D. This suggests that garden path effects maintained through this region: The increased first pass reading time suggests that readers needed more time to process the final sentence region and perhaps to perform sentence integration in the absence of the disambiguating verb diacritics (Pass-Non-D condition), relative to when they were present (Pass-Verb-D condition). A similar effect was observed in the contrast between the Pass-Verb-D and the Pass-Full-D conditions, with longer first pass reading time for the latter. This difference approached significance and also highlights the persistence of garden path effects in this region in the fully-diacritised passive condition, and of course, this is in line with the findings reported at the disambiguating PP region.

There were no significant differences between the passive conditions in any of the contrasts for the regression out measure. Finally, there were also no differences between the three conditions in the sentence re-reading measure.

The absence of effects in the regression out measure and for re-reading time in both the disambiguating and end of sentence regions can be taken to indicate that the disruption to processing caused by the garden path manipulation was short-lived.

Processing Diacritics in the Active Full Sentence Diacritisation Condition

We compared the Active-Full-D and Active-Non-D conditions to explore the effects of the presence of the additional visual information conveyed by the diacritics, in the absence of other effects (e.g. of processing passive sentence analysis, or related potential garden path effects) on global eye movement measures: Average fixation duration (Removing outliers resulted in removing 1.2% of the data points from the

analysis for this measure), number of fixations, and total sentence reading time² (Removing outliers resulted in removing 2.6% of the data points from the analysis of *total sentence reading time*). We coded the contrasts between the two active conditions such that the non-diacritised condition was the baseline. This is ecologically appropriate given the prevalence of non-diacritised text in adult native readers' experience with printed Modern Standard Arabic.

The descriptive statistics for these measures are reported in *Table 3* and inferential statistics are presented in *Table 4*.

<Table 3 about here>

<Table 4 about here>

The presence of full sentence diacritics resulted in a small (7 ms), but reliable, increase in average fixation duration. Fully-diacritised sentences attracted a slightly lower total number of fixations, but the difference was not significant between the two conditions. Finally, there was no difference in total sentence reading time between the two conditions.

The small but significant increase in average fixation duration in the Act-Full-D condition is similar to the costs of visual crowding effects reported by Liversedge et al. (2014) and Slattery and Rayner (2013).

Additional Analyses

We also investigated whether the reported effects of diacritisation varied as a function of longitudinal effects during the experiment (e.g. increased familiarity with stimuli, particularly the passive condition; fatigue; whether garden path effects, or the

effects of full-diacritisation were stronger earlier compared to later within the testing session). Thus, we used trial rank as a predictive variable of the reported eye movement measures, for all contrasts. We coded this in a Linear Mixed Effects model, with participants and items as random effects. Our findings showed that trial rank had no effect on any of the eye movement measures, in any of the contrasts (all $t_s < 1.6$).

Similarly, we investigated whether the reported patterns of findings were influenced by participants' chronological age. We used participants' age as a predictive variable in Linear Mixed Effects models³, with participants and items as random effects. The analyses revealed that participants' age had no effect on any of the eye movement measures, in any of the contrasts (all $t_s < 1$).

Discussion

Our investigation aimed to address a number of important theoretical issues relating to the processing of diacritics during natural silent reading in Arabic. We will discuss each of these in turn shortly. However, prior to discussing the details and theoretical significance of our findings, we note a general aspect of our results, namely, that in common with other investigations of eye movements during reading Arabic (e.g., Jordan et al., 2014), the fixations durations we report (first, single and average) are somewhat elevated compared with those observed during reading of Latinate languages (see Rayner, 1998; 2009). We are not entirely sure why this is the case, but in all likelihood reflects a general cross-linguistic difference in reading behaviour (rather like that observed between Chinese and other Latinate languages). We suspect that this may be due to the morphological complexity of Arabic words,

though this suggestion is clearly speculative at this stage. Another factor that may be taken into consideration with regards to this aspect of our results is that the participants were all bilingual (Arabic-English), and it may be useful to replicate these patterns with a monolingual sample in the future. However, we think it is unlikely that this factor limits how generalizable our findings are, given that (1), the participants were demonstrably fluent in reading Arabic (and read it on a regular basis), and (2), because it is unlikely that their knowledge of English affected their ability to read Arabic given the extensive differences between the two languages (e.g. in typography, phonology, etc.), compared to, for instance, languages with more similar characteristics such as Dutch and English (e.g. Brysbaert & Duyck, 2010).

Our selection of heterophonic-homographic verbs which can be disambiguated as active or passive through diacritisation allowed us to explore whether Arabic readers have a preference for simple active verb analysis, and whether they will use diacritics as phono-syntactic disambiguation cues for diacritised passive verbs. We predicted that if readers have this preference for active analysis, they would experience a garden path effect in the absence of diacritics which disambiguate the verb as passive (Pass-Non-D condition), compared to when these diacritics were present (the Pass-Verb-D condition).

A clear picture of the time course and costs of processing diacritics during reading began to emerge at the ambiguous verb region. The effect of diacritisation was limited to skipping rates, and did not influence fixation time measures. Readers were significantly less likely to skip the verb in the Pass-Verb-D condition. As we suggested above, this skipping effect is likely to be related to visual complexity (see e.g. Liversedge et al., 2014), which increases if the upcoming word is diacritised, and whereby non-diacritised words, being more visually simple, are skipped more often

than diacritised words. This is more plausible than suggesting that processing of the upcoming diacritised verb was performed parafoveally to such an extent that a decision to skip it, or not, was made on the basis of such deep processing.

Furthermore, our patterns of effects for the early fixation time measures on the verb suggest that disambiguating a critical homograph using the provided diacritics does not entail additional processing demands which significantly increase word reading time. Our data, therefore, also suggest that processing passive verbs does not pose that much of a cognitive demand on readers given that passive is a common, albeit less preferred, analysis (c.f. Roman et al., 1985).

The remaining findings of the contrasts performed on the passive conditions provided unequivocal answers to our research questions. At the disambiguating PP region, the most striking finding was the clear garden path effect: Readers experienced immediate disruption to processing upon arrival at this region in the Pass-Non-D condition, relative to the Pass-Verb-D condition. Compared to the Pass-Verb-D, first and single fixation and gaze durations were significantly inflated in the Pass-Non-D condition. This pattern of results clearly shows that Arabic readers have a parsing preference for simple active over passive analysis of the verb: When verb-disambiguating diacritics were not available, the readers followed this active parsing preference until arriving at the passive disambiguating region. Our results replicate previous research findings which show that when readers' initial parsing preferences are violated, the resulting disruption to processing is manifest at the disambiguating region in the violating conditions (e.g. Frazier & Rayner, 1982; Lipka, 2002; Liversedge et al., 2002; Paterson et al., 1999; Rayner et al., 1983; Rayner & Frazier, 1987). Our results document, for the first time, that readers in Arabic have a parsing

preference for simple active over passive, as is the case in other languages (e.g. English, Fodor et al., 1974; Ferreira, 2003 etc.).

The contrast between the Pass-Verb-D and Pass-Non-D conditions in the spillover region showed no significant difference between these two conditions. However, we obtained more evidence for garden path effects in the end-of-sentence region. Reading time measures of early (first pass) processing clearly show that readers were slower in the Pass-Non-D condition, compared to the Pass-Verb-D condition. This pattern of results replicates previous findings (e.g. Frazier & Rayner, 1982; Lipka, 2002).

The disruption our readers experienced appeared in early, first pass, reading measures in the disambiguating PP and end-of-sentence regions. The fact that garden path effects in the disambiguating PP region were limited to first pass measures can be seen as somewhat deviant from previous findings (e.g., Frazier & Rayner, 1982; Lipka, 2002). In these investigations garden path effects persisted into later reading measures (e.g. regression out, and go past reading time). Yet we do not find this surprising, for two reasons: (a) Passive is a natural sentence structure in Arabic, as our sentence naturalness ratings suggest, albeit it is the less preferred analysis of the verb-noun pair; and (b) Good semantic matching of the agent and patient of the sentence with its verb meant that the disambiguating region did not contain any semantic implausibility (e.g. the verb *pushed* in the provided example (see *Figure 1*) could equally be performed by the student, or by her colleague). This can be contrasted with, for example, the materials used by Rayner, Warren, Juhasz, and Liversedge (2004) where significant differences were obtained at go past reading times when the disambiguating region contained semantically implausible information.

The results thus far suggest that readers processed the available passive verb diacritics in the Pass-Verb-D condition, and thus avoided the garden path seen in the Pass-Non-D condition. Contrastingly, when the whole passive sentence was diacritised (Pass-Full-D condition), and despite the presence of the verb disambiguating diacritics, readers showed evidence of disruption to processing of a similar magnitude to when the passive verb was not diacritised. In the Pass-Full-D condition the readers showed significantly longer first and single fixation and gaze durations in the disambiguating region, as well as significantly longer first pass in the end-of-sentence region, compared to the Pass-Verb-D condition.

Clearly, then, the mode of diacritisation (only on the homograph vs. on the entire sentence) impacts how readers process diacritics. That is to say, readers do not always appear to engage in automatic (involuntarily, reflexive and mandatory) processing of diacritics: Whereas it is clear that readers processed the present disambiguating diacritics on the verb in the Pass-Verb-D condition such that they were able to extract the disambiguating information provided by these diacritics, there is no evidence that such processing occurred in Pass-Full-D condition, indicating that diacritics are not necessarily automatically processed. In the Pass-Full-D condition participants treated full sentence diacritics as if they were redundant and thus failed to utilise the verb diacritics as a phono-syntactic disambiguation cue. Failure to use the diacritical information resulted in readers experiencing garden path effects.

Comparing the active conditions (Act-Non-D and Act Full-D) to learn about readers' processing of full sentence diacritics, revealed that full sentence diacritisation resulted in a small, reliable, increase in average fixation duration. This was offset by a small (non reliable) decrease in the number of fixations made in that condition resulting in very comparable total sentence reading times in both conditions. The

presence of full sentence diacritisation thus has resulted only in a small increase in average fixation duration. We are inclined to attribute this effect to visual crowding, where the presence of the additional visual information (the diacritics) slows readers' uptake of visual information. However, there is an alternative possibility. While fully-diacritised texts are regularly encountered by native Arabic readers, these texts are typically religious or other literary works. Given that our experimental sentences were neither religious nor from formal literature, then it can be argued that the effects may be caused by readers not being familiar with reading fully-diacritised single sentences which are not of these types. One way to assess whether this might have been the case is to investigate whether readers' performance with fully-diacritised sentences changed as they encountered successive examples of such sentences throughout the experiment. We therefore examined this possibility by taking the trial number into account in relation to our dependent measures. Our data indicated that readers' processing of full sentence diacritics did not change as their familiarity with the stimuli developed across the duration of the experiment. Of course, we cannot completely rule out the possibility that any familiarity effects that may have explained the full diacritics effect could have developed over a period much longer than the duration of the experimental testing session. However, in our view, these analyses certainly weaken any such explanation. Rather, this small cost seems more reasonably attributed to the undisputable increase in visual crowding resulting from the presence of full sentence diacritics, and is in line with the small but reliable increase in average fixation duration (6 ms) reported by Slattery and Rayner (2013) as a result of increasing visual crowding. This pattern is also in line with the small (9-11 ms) increase in fixation time reported by Liversedge et al. (2014). However, unlike Slattery and Rayner's findings, diacritics-based crowding effects did not result in a

significant increase in total sentence reading time for the crowded Act-Full-D condition, compared to Act-Non-D. If a diacritics-based crowding effect is what is driving this pattern of findings, then its weakness, relative to what was reported by Slattery and Rayner, is perhaps due to Arabic readers' experience with reading fully-diacritised texts. Similarly, the subtle effect of visual crowding we obtained is weaker than the effects reported by Roman and Pavard (1987). It is hard to speculate as to what was driving the increased fixation times in their fully diacritised conditions. One possibility is that the stimuli were made up of entire paragraphs, rather than single sentences, and faster first-pass reading of paragraphs compared to single sentences may have compounded the visual crowding effects (e.g. longer gaze durations for sentence reading compared to passage reading, Radach, Huestegge, & Reilly, 2008; Wochna & Juhasz, 2013).

In our view, the theoretical implications of these results are important: The small effect, likely due to visual crowding, and the fact that there were no other differences between the two active conditions (similar total number of fixations and total sentence reading times) lend more support to the suggestion that readers do not engage automatically, or to any significant extent, in cognitive processes to access the linguistic information provided by the diacritics in the full diacritisation mode. This is in line with the findings reported above regarding the Pass-Full-D condition. If readers treat full sentence diacritisation as information that is redundant and therefore do not automatically process the phonetic and syntactic information provided by these diacritics in fully-diacritised sentences, then it might be considered that, under "normal" reading circumstances, this would be a cognitive resources-saving strategy that adult skilled Arabic readers are capable of using. Typically these readers utilise

text context and syntactic structure to effectively disambiguate homographic words (e.g. Abu-Rabia, 1997a; 1997b; 1998).

Compared to the literature reviewed above (e.g. Abu-Rabia, 1997a; 1997b; 1998; 1999; 2001), which suggested that the presence of diacritics contributed to improved reading accuracy and, or, comprehension performance, our findings bring considerable insight into the time course of processing diacritics, and the influence of the mode of diacritisation on processing during natural silent reading. Similarly, our findings clarify how readers process diacritics in the course of natural silent reading compared to in lexical decision or word naming tasks. In such single word tasks participant are typically asked to *perform the task accurately* and are thus obliged to perform complete analyses of the diacritised letter-strings, and this may explain the additional processing time costs reported when diacritics are present (see similar suggestions made by Bentin & Frost, 1987). By contrast, when readers are asked to *read sentences for comprehension*, as in our experiment, this does not entail performing lengthy additional analyses of the linguistic information provided by the diacritics when reading fully-diacritised sentences. In our experiment, it is likely that readers followed their assumption about full sentence diacritics being redundant, in both fully-diacritised conditions, active and passive, and spared cognitive resources. This, in turn, resulted in the different outcomes for the fully-diacritised passive compared to the fully-diacritised active condition.

To summarise, our findings address a number of theoretical issues in reading Arabic for which, up until now, there has been no interpretable eye movement data. By exploring processing diacritics during natural reading of Arabic, we have demonstrated that Arabic readers extract phono-syntactic information from diacritics to disambiguate heterophonic-homographic verbs. Readers did this most effectively

when the diacritics were assigned only to the homographic word. This lends support to current practices in printing most Arabic texts, where diacritics are only printed on heterophonic-homographs that are not disambiguated by the surrounding text. We also demonstrated that Arabic readers have a preference for simple active analysis of ambiguous verbs: Readers followed this preference when these verbs were embedded in a temporarily ambiguous context. To our knowledge, this is the first time that this preference has been documented in the literature on reading Arabic. When this initial parsing preference for active was violated, readers experienced garden path effects at the disambiguating PP region in the non-diacritised passive condition. Finally, our data suggest that skilled readers show slightly different eye movement behaviour when reading fully-diacritised sentences, compared to non-diacritised sentences, when both contain no violation of their syntactic parsing preferences: A modest increase in average fixation duration, and a relative decrease in number of fixations—a trade-off resulting in comparable sentence reading times. These results are likely to reflect visual crowding effects.

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
- Abu-Rabia, S. (2002). Reading in a root-based-morphology language: the case of Arabic. *Journal of Research in Reading*, 25, 299-309. doi:10.1111/1467-9817.00177

- 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
- Bai, X., Yan, G., Liversedge, S. P., Zang, C., & Rayner, K. (2008). Reading spaced and unspaced Chinese text: Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 1277-1287. doi:10.1037/0096-1523.34.5.1277
- 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 pack- age version 0.999375-42 [Computer software]. Available from <http://CRAN.R-project.org/package=lme4>
- Bentin, S., Bargai, N., & Katz, L. (1984). Orthographic and phonemic coding for lexical access: Evidence from Hebrew. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 353-368. doi:10.1037/0278-7393.10.3.353
- Bentin, S., & Frost, R. (1987). Processing lexical ambiguity and visual word recognition in a deep orthography. *Memory & Cognition*, *15*, 13-23. doi:10.3758/BF03197708
- Boudelaa, S., & Marslen-Wilson, W. D. (2001). Morphological units in the Arabic mental lexicon. *Cognition*, *81*, 65-92. doi:10.1016/S0010-0277(01)00119-6
- Boudelaa, S., & Marslen-Wilson, W. D. (2005). Discontinuous morphology in time: Incremental masked priming in Arabic. *Language and Cognitive Processes*, *20*, 207-260. doi:10.1080/01690960444000106

- 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
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*, 177-178. doi:10.1038/226177a0
- Bouma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research*, *13*, 767-782. doi:10.1016/0042-6989(73)90041-2
- Brybaert, M., & Duyck, W. (2010). Is it time to leave behind the Revised Hierarchical Model of bilingual language after fifteen years of service? *Bilingualism: Language and Cognition*, *13*, 359-371.
doi:10.1017/S1366728909990344
- Clifton, C., Jr., 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 and Language*, *49*, 317-334.
doi:10.1016/S0749-596X(03)00070-6
- Drieghe, D., Brybaert, M., & Desmet, T. (2005). Parafoveal-on-foveal effects on eye movements in text reading: Does an extra space make a difference? *Vision Research*, *45*, 1693-1706. doi:10.1016/j.visres.2005.01.010
- Farid, M., & Grainger, J. (1996). How initial fixation position influences visual word recognition: A comparison of French and Arabic. *Brain and Language*, *53*, 351-368. doi:10.1006/brln.1996.0053
- Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*, 164-203. doi:10.1016/S0010-0285(03)00005-7

- Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The psychology of language: An introduction to psycholinguistics and generative grammar*. NY: McGraw-Hill.
- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14*, 178-210. doi:10.1016/0010-0285(82)90008-1
- Haywood, J. A., & Nahmad, H. M. (1965). *A new Arabic grammar of the written language*. Surrey: Lund Humphries.
- Ibrahim, R., & Eviatar, Z. (2009). Language status and hemispheric involvement in reading: Evidence from trilingual Arabic speakers tested in Arabic, Hebrew, and English. *Neuropsychology*, *2*, 240-254. doi:10.1037/a0014193
- Jordan, T. R., Almabruk, A. A., Gadalla, E. A., McGowan, V. A., White, S. J., Abedipour, L., & Paterson, K. B. (2014). Reading direction and the central perceptual span: Evidence from Arabic and English. *Psychonomic Bulletin & Review*, *21*, 202-211. doi:10.3758/s13423-013-0510-4
- Jordan, T. R., Almabruk, A. A., McGowan, V. A., & Paterson, K. B. (2011). Evaluating hemispheric divisions in processing fixated words: The evidence from Arabic. *Cortex*, *47*, 992-997. doi:10.1016/j.cortex.2011.02.012
- Jordan, T. R., McGowan, V. A., & Paterson, K. B. (2014). Reading with filtered fixations: Adult age differences in the effectiveness of low-level properties of text within central vision. *Psychology and Aging*, *29*, 229-235. doi:10.1037/a0035948
- Jordan, T. R., Paterson, K. B., & Almabruk, A. A. (2010). Revealing the superior predictability of words in Arabic. *Perception*, *39*, 426-428. doi:10.1068/p6637

- Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S., & Rayner, K. (2009). Word length and landing position effects during reading in children and adults. *Vision Research*, *49*, 2078-2086. doi:10.1016/j.visres.2009.05.015
- Koriat, A. (1984). Reading without vowels: lexical access in Hebrew. In H. Bauma, & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 227-242). Hillsdale, NJ: Earlbaum. Retrieved from http://iipdm.haifa.ac.il/images/Articles/reading_without_vowels1.pdf
- Koriat, A. (1985a). Lexical access for low- and high-frequency words in Hebrew. *Memory & Cognition*, *13*, 37-44. doi:10.3758/BF03198441
- Koriat, A. (1985b). Lateralization effects in reading pointed and unpointed Hebrew. *British Journal of Psychology*, *76*, 161-173. doi:10.1111/j.2044-8295.1985.tb01939.x
- Lipka, S. (2002). Reading sentences with a late closure ambiguity: Does semantic information help? *Language and Cognitive Processes*, *17*, 271-298. doi:10.1080/01690960143000029
- Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Science*, *4*, 6-14. doi:10.1016/S1364-6613(99)01418-7
- Liversedge, S. P., Paterson, K. B., & Clayes, E. L. (2002). The influence of only on syntactic processing of “long” relative clause sentences. *The Quarterly Journal of Experimental Psychology*, *55A*, 225-240. doi:10.1080/02724980143000253
- Liversedge, S. P., Zang, C., Zhang, M., Bai, X., Yan, G., & Drieghe, D. (2014). The effect of visual complexity and word frequency on eye movements during reading Chinese. *Visual Cognition*, *22*, 441-457. doi:10.1080/13506285.2014.889260
- Murray W. S., & Liversedge, S. P. (1994). Referential context effect on syntactic processing. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on*

sentence processing (pp. 359-388). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

- Paterson, K. B., & Jordan, T. R. (2010). Effects of increased letter spacing on word identification and eye guidance during reading. *Memory & Cognition*, *38*, 502-5012. doi:10.3758/MC.38.4.502
- Paterson, K. B., Liversedge, S. P., & Underwood, G. (1999). The influence of focus operators on syntactic processing of short relative clause sentences. *The Quarterly Journal of Experimental Psychology*, *52A*(3), 717-737. doi:10.1080/713755827
- Pelli, D. G., Tillman, K. A., Freeman, J., Su, M., Berger, T. D., and Majaj, N. J. (2007). Crowding and eccentricity determine reading rate. *Journal of Vision*, *7*, 20-36. doi:10.1167/7.2.20
- Radach, R., Huestegge, L., & Reilly, R. (2008). The role of global top-down factors in local eye movement control in reading. *Psychological Research*, *72*, 675-688. doi:10.1007/s00426-008-0173-3
- Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology*, *41*, 211-236. doi:10.1016/0022-0965(86)90037-8
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372-422. doi:10.1037/0033-2909.124.3.372
- 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., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing – eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*, 22, 358-374. doi:10.1016/S0022-5371(83)90236-0
- Rayner, K., & Frazier, L. (1978). Parsing temporarily ambiguous complements. *The Quarterly Journal of Experimental Psychology*, 39, 657-673.
doi:10.1080/14640748708401808
- Rayner, K., & Liversedge, S. P. (2004). Visual and linguistic processing during eye fixations in reading. In F. Ferreira and J. Henderson (Eds.) *The interface of language, vision and action: Eye movements and the visual world* (pp. 59-104). New York: Psychology Press.
- Rayner, K., & Liversedge, S. (2011). Linguistic and cognitive influences on eye movements during reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye Movements* (pp. 751-766). Oxford, England: OUP.
- Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging*, 21, 448-465.
doi:10.1037/0882-7974.21.3.448
- Rayner, K., Warren, T., Juhasz, B. J., & Liversedge, S. P. (2004). The effect of plausibility on eye movements in reading. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 30, 1290-1301. doi:10.1037/0278-7393.30.6.1290

- Roman, G., & Pavard, B. (1987). A comparative study: How we read in Arabic and French. In J. K. O'Regan & A. Levy-Schoen (Eds.), *Eye Movements from Physiology to Cognition* (pp. 431-440). Amsterdam: Elsevier
- Roman, G., Pavard, B., & Asselah, B. (1985). Traitement perceptif des phrases ambiguës en Arabe. *Cahiers de Psychologie Cognitive*, 5, 5-22.
- Schultz, E. (2004). *A student grammar of Modern Standard Arabic*. Cambridge: CUP.
- 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
- Shimron, J., & Sivan, T. (1994). Reading proficiency and orthography: Evidence from Hebrew. *Language Learning*, 44, 5-27. doi:10.1111/j.1467-1770.1994.tb01447.x
- Slattery, T. J., & Rayner, K. (2013). Effects of intraword and interword spacing on eye movements during reading: Exploring the optimal use of space in a line of text. *Attention, Perception, and Psychophysics*, 75, 1275-1292. doi:10.3758/s13414-013-0463-8
- Wochna, K. L., & Juhasz, B. J. (2012). Context length and reading novel words: An eye-movement investigation. *British Journal of Psychology*, 104, 347 – 363. doi:10.1111/j.2044-8295.2012.02127.x.
- Yang, H. M., & McConkie, G. W. (1999). Reading Chinese: Some basic eye-movement characteristics. In H.-C. Chen (Ed.), *Reading Chinese script: A cognitive analysis* (pp. 207–222). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

- Zang, C., Liang, F., Bai, X., & Liversedge, S. P. (2013). Interword spacing and landing position effects during Chinese reading in children and adults. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 720-734. doi:10.1037/a0030097
- Zang, C., Liversedge, S. P., Bai, X., & Yan, G. (2011). Eye movements during Chinese reading. In S. P. Liversedge, I. Gilchrist, & S. Everling. (Eds.), *The Oxford handbook of eye movements* (pp. 961–978). Oxford: Oxford University Press.

Footnotes

¹ Although this participant would qualify as an older adult, the data were used in the analyses given that his performance did not suggest that he was underperforming or an outlier: a) the participant's screening text accuracy (see Stimuli section) level was 99.3%; b) their eye movement data showed identical patterns to other participants on all reported measures; c) their sentence comprehension score was 85.2%; and finally, d) their diacritics decoding accuracy on the single word reading task (see Stimuli section) was 97.2%.

² Although saccade amplitude is often reported in global analyses as an additional indicator of ease of processing (with smaller saccade amplitude denoting reading of difficult linguistic material, lack of skill or cautious reading, Frazier & Rayner, 1982; Joseph, Liversedge, Blythe, White, & Rayner, 2009; Rayner, 1986), we do not report saccade amplitudes here. This is mainly because of our decision to use a non-monospaced font to display the sentences. Letters in monospaced fonts make Arabic text look extremely unnatural. The use of a non-monospaced font makes the text look far more natural, however, it also compromises the possibility of directly comparing saccade lengths across conditions.

³ We would like to thank Jon Andoni Duñabeitia for making this suggestion.

PROCESSING DIACRITICS IN READING ARABIC

Table 1.

Descriptive Statistics of Eye Movement Measures of Sentence Regions in the Passive Conditions

	<i>Verb Region</i>			<i>Disambiguating (PP) Region</i>			<i>Spillover (PP + I) Region.</i>			<i>End-of-Sentence Region</i>		
	Pass- Verb-D	Pass- Full-D	Pass- Non-D	Pass- Verb-D	Pass- Full-D	Pass- Non-D	Pass- Verb-D	Pass- Full-D	Pass- Non-D	Pass- Verb-D	Pass- Full-D	Pass- Non-D
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Skipping Rate	0.125 (0.21)	0.135 (0.19)	0.19 (0.13)	0.19 (0.17)	0.15 (0.19)	0.195 (0.13)	0.16 (0.17)	0.17 (0.21)	0.14 (0.22)	-	-	-
First fixation (ms)	281 (19.51)	292 (30.73)	299 (44.04)	259 (16.40)	291 (26.17)	295 (34.45)	281 (30.52)	284 (25.41)	268 (20.49)	-	-	-
Single fixation (ms)	287 (31.82)	305 (40.8)	298 (45.44)	259 (22.21)	308 (33.36)	301 (35.10)	278 (35.03)	292 (26.10)	274 (24.14)	-	-	-
Gaze duration / First pass (ms)	350 (40.26)	361 (50.81)	334 (47.30)	312 (36.28)	353 (34.70)	347 (52.47)	324 (40.77)	360 (49.30)	297 (31.40)	1271 (160.4)	1369 (205)	1444 (171.3)
Regression in	0.35 (0.18)	0.40 (0.17)	0.35 (0.14)	-	-	-	-	-	-	-	-	-
Regression out	-	-	-	0.07 (0.20)	0.09 (0.21)	0.06 (0.19)	0.19 (0.23)	0.16 (0.2)	0.2 (0.22)	0.64 (0.19)	0.65 (0.17)	0.65 (0.19)
Sentence re-reading	-	-	-	-	-	-	-	-	-	1877 (1310.9)	1969 (1414.2)	1503 (991.1)

Note. Pass-Verb-D = passive with verb-only diacritics, Pass-Non-D = non-diacritised passive, Pass-Full-D = fully-diacritised passive

Table 2.

Contrasts of Eye Movement Measures of Sentence Regions in the Passive Conditions

	Verb Region			Disambiguating (PP) Region			Spillover (PP + 1) Region.			End-of-Sentence Region		
	<i>b</i>	SE	<i>z</i> or <i>t</i>	<i>b</i>	SE	<i>z</i> or <i>t</i>	<i>b</i>	SE	<i>z</i> or <i>t</i>	<i>b</i>	SE	<i>z</i> or <i>t</i>
EM Measure	Skipping											
(Intercept)	-2.60	0.35	-7.50***	-2.30	0.39	-5.97***	-2.38	0.37	-6.49***	-	-	-
Pass-Verb-D vs. Pass-Non-D	0.63	0.31	2.02*	0.05	0.31	0.17	-0.28	0.32	-0.88	-	-	-
Pass-Verb-D vs. Pass-Full-D	0.09	0.33	0.29	-0.41	0.32	-1.26	0.06	0.31	0.20	-	-	-
EM Measure	First Fixation											
(Intercept)	5.59	0.03	164.64***	5.53	0.03	184.63***	5.60	0.03	188.58***	-	-	-
Pass-Verb-D vs. Pass-Non-D	0.05	0.04	1.29	0.11	0.03	3.50***	-0.04	0.03	-1.07	-	-	-
Pass-Verb-D vs. Pass-Full-D	0.03	0.04	0.87	0.10	0.03	3.31***	-0.01	0.03	-0.25	-	-	-
EM Measure	Single Fixation											
(Intercept)	5.63	0.04	144.03***	5.54	0.03	178.08***	5.59	0.03	168.27***	-	-	-
Pass-Verb-D vs. Pass-Non-D	-0.01	0.04	-0.20	0.12	0.03	3.53***	-0.01	0.04	-0.35	-	-	-
Pass-Verb-D vs. Pass-Full-D	0.04	0.04	0.83	0.14	0.03	4.07***	0.01	0.04	0.33	-	-	-
EM Measure	Gaze Duration						First Pass					
(Intercept)	5.79	0.04	148.24***	5.68	0.04	144.92***	5.71	0.04	145.91***	6.96	0.07	102.32***
Pass-Verb-D vs. Pass-Non-D	-0.06	0.04	-1.43	0.10	0.04	2.64**	-0.06	0.04	-1.48	0.12	0.06	1.85*
Pass-Verb-D vs. Pass-Full-D	0.01	0.04	0.18	0.12	0.04	3.20**	0.07	0.04	1.63	0.16	0.06	2.54*
EM Measure	Regression In						Regression Out					
(Intercept)	-0.73	0.25	-2.91**	-2.95	0.36	-8.17***	-1.69	0.25	-6.70***	0.70	0.25	2.86**
Pass-Verb-D vs. Pass-Non-D	-0.04	0.23	-0.18	-0.18	0.47	-0.38	0.10	0.28	0.36	0.04	0.22	0.18
Pass-Verb-D vs. Pass-Full-D	0.22	0.23	0.97	0.35	0.42	0.84	-0.14	0.29	-0.47	0.05	0.22	0.23
EM Measure	Sentence Re-reading											
(Intercept)	-	-	-	-	-	-	-	-	-	6.75	0.16	42.85***
Pass-Verb-D vs. Pass-Non-D	-	-	-	-	-	-	-	-	-	-0.05	0.14	-0.35
Pass-Verb-D vs. Pass-Full-D	-	-	-	-	-	-	-	-	-	-0.04	0.14	-0.30

Note. EM Measure = eye movement measure, Pass-Verb-D = passive with verb-only diacritics, Pass-Non-D = non-diacritised passive, Pass-Full-D

= fully-diacritised passive, *b* = regression coefficient, SE = standard error, *t* or *z* = test statistic (*b*/*SE*). *t* or *z* values higher than 1.96 denote

statistically significant difference. Significance code: *** for .001; ** for .01; * for .05; and + for marginal (< .1) significance.

Table 3.

Descriptive Statistics of Eye Movement Global Measures of the Active Conditions.

Measure	Act-Non-D	Act-Full-D
	Mean (SD)	Mean (SD)
Average fixation duration (ms)	259 (7.24)	266 (6.20)
Total number of fixations	32 (5.47)	30 (4.04)
Total reading times (ms)	8185 (1451.15)	8119 (1190.58)

Note. Act-Non-D = non-diacritised active, Act-Full-D = fully-diacritised active.

Table 4.

Inferential Statistics of Eye Movement Global Measures of the Active Conditions.

	Active Conditions		
	<i>b</i>	SE	<i>t</i>
	Average Fixation Duration		
(Intercept)	5.58	0.02	259***
AVd vs. Active-Full-D	-0.03	0.01	-3.54***
	Number of Fixations		
(Intercept)	30.44	1.61	18.86***
AVd vs. Active-Full-D	1.35	0.80	1.69 ⁺
	Total Reading Time		
(Intercept)	8.95	0.05	169.96***
AVd vs. Active-Full-D	0.00	0.02	-0.06

Note. Act-Full-D = fully-diacritised active, Act-Non-D =

non-diacritised active, *b* = regression coefficient, SE =

standard error, *t* or *z* = test statistic (*b*/*SE*). *t* or *z* values

higher than 1.96 denote statistically significant

difference. Significance code: *** for .001; ** for .01;

* for .05; and ⁺ for marginal (< .1) significance.

← ← Direction of Reading ← ←													
1. Act-Non-D	سمع الجميع صرخة مدوية حينما <u>دفعت</u> الطالبة التي كانت في طريقها إلى المعمل زميلتها فهوت مغشياً عليها.												
2. Act-Full-D	سَمِعَ الْجَمِيعُ صَرْخَةً مُدَوِيَةً حِينَمَا <u>دَفَعَتِ</u> الطَّالِبَةُ الَّتِي كَانَتْ فِي طَرِيقِهَا إِلَى الْمَعْمَلِ زَمِيلَتَهَا فَهَوَتْ مَغْشِيًا عَلَيَّهَا.												
3. Pass-Non-D	سمع الجميع صرخة مدوية حينما <u>دفعت</u> الطالبة التي كانت في طريقها إلى المعمل <u>بيدي</u> زميلتها فهوت مغشياً عليها.												
4. Pass-Verb-D	سمع الجميع صرخة مدوية حينما <u>دُفِعَت</u> الطالبة التي كانت في طريقها إلى المعمل <u>بيدي</u> زميلتها فهوت مغشياً عليها.												
5. Pass-Full-D	سَمِعَ الْجَمِيعُ صَرْخَةً مُدَوِيَةً حِينَمَا <u>دُفِعَتِ</u> الطَّالِبَةُ الَّتِي كَانَتْ فِي طَرِيقِهَا إِلَى الْمَعْمَلِ <u>بِيَدِي</u> زَمِيلَتَهَا فَهَوَتْ مَغْشِيًا عَلَيَّهَا.												
Interest Areas	<table style="width: 100%; border-collapse: collapse;"> <tr> <td style="border-right: 1px solid black; padding: 0 5px;">سمع الجميع صرخة مدوية حينما</td> <td style="border-right: 1px solid black; padding: 0 5px;"> دُفِعَتِ </td> <td style="border-right: 1px solid black; padding: 0 5px;"> الطالبة التي كانت في طريقها إلى المعمل</td> <td style="border-right: 1px solid black; padding: 0 5px;"> بيدي </td> <td style="border-right: 1px solid black; padding: 0 5px;"> زميلتها </td> <td style="border-right: 1px solid black; padding: 0 5px;"> فهوت مغشياً عليها.</td> </tr> <tr> <td style="border-right: 1px solid black; text-align: center;">iv</td> <td style="border-right: 1px solid black; text-align: center;">i</td> <td style="border-right: 1px solid black; text-align: center;">ii</td> <td style="border-right: 1px solid black; text-align: center;">iii</td> <td style="border-right: 1px solid black;"></td> <td></td> </tr> </table>	سمع الجميع صرخة مدوية حينما	دُفِعَتِ	الطالبة التي كانت في طريقها إلى المعمل	بيدي	زميلتها	فهوت مغشياً عليها.	iv	i	ii	iii		
سمع الجميع صرخة مدوية حينما	دُفِعَتِ	الطالبة التي كانت في طريقها إلى المعمل	بيدي	زميلتها	فهوت مغشياً عليها.								
iv	i	ii	iii										
Translation	Everyone heard a loud scream when the student <u>pushed</u> ([was] pushed) on her way to the lab (by the hand of) her colleague so she fell unconscious.												

Figure 1. Sample stimulus set. The 5 conditions: Non-Diacritised Active (Act-Non-D); Fully-Diacritised Active (Act-Full-D); Non-Diacritised Passive (Pass-Non-D); Verb-only-Diacritised Passive (Pass-Verb-D); and Fully-Diacritised Passive (Pass-Full-D). Stimulus translation provided. In the translation, the passive alternative is given within brackets: (verb) and (disambiguating PP region). In both the Arabic text and in the translation, the verb is underlined, and so is the prepositional phrase (PP) acting as the disambiguating region. The interest areas where eye movement data were analysed are |marked| with vertical lines and are labelled: i = the verb region; ii = the disambiguating PP region; iii = the spill-over region; and iv = the end of sentence region.