Are alternative meanings of an Arabic homograph activated even when it is disambiguated by vowel diacritics?

Maryse Maroun and J. Richard Hanley

Department of Psychology, University of Essex, Colchester, UK

Accepted for publication *16th July 2020* in Writing Systems Research

Corresponding Author

J. Richard Hanley

Department of Psychology,

University of Essex,

Colchester CO4 3SQ,

UK

rhanley@essex.ac.uk

Are alternative meanings of an Arabic homograph activated even when it is disambiguated by vowel diacritics?

Abstract

The diacritical markers that represent most of the vowels in the Arabic orthography are typically omitted from written texts, thereby making many Arabic words phonologically and semantically ambiguous. Such words are known as heterophonic homographs and are associated with different pronunciations and meanings. The three experiments reported in this paper were conducted to investigate how Arabic readers understand diacritized heterophonic homographs. The results suggested that even when diacritics were added to disambiguate a heterophonic homograph, it was still initially processed as if it was ambiguous and its alternative meaning(s) were activated.

Keywords Arabic reading; vowelization; semantic categorization

There are several aspects of the Arabic writing system that distinguish it from European orthographies and make it particularly interesting to investigate. In common with other Semitic scripts such as Hebrew, Arabic is primarily a consonantal system in which most letters represent consonants. Although there are three letters in the Arabic writing system that can represent long vowels, short vowels are written as diacritical marks, of which two stand above (/a/ and/u/), and one stands below the body of the word (/i/). With, however, the exception of children's books, poetry, and liturgical texts, printed material in the Arab world rarely includes diacritics. It follows that many written words in Arabic texts appear as sequences of consonants. Because readers of Arabic typically encounter the unvowelized script in their daily life, these consonantal structures are likely to become highly familiar to them. Findings from Arabic (Abu-Leil, Share & Ibrahim, 2014; Asadi, 2017; Bourisley, Haynes, Bourisley & Mody, 2013; Roman & Pavard, 1987; Saiegh-Haddad, & Schiff, 2014) and Hebrew (e.g. Frost & Bentin, 1987, 1992) have shown that word recognition is significantly quicker when words were presented as consonantal sequences than when they are vowelized. These results suggest that lexical access in Arabic and Hebrew is based on identification of these consonantal structures. According to Frost (1998), word recognition would therefore rely on a mental representation of a sequence of letters in which the identity of the vowels is not specified. Using English as an illustration, the orthographic representation of the word *pint* would comprise a CVCC segment /p-nt/ with no entry for the vowel (Frost, 1998). It is only after this consonantal representation has been activated that it is possible for the reader to retrieve the semantic and phonological information with which the word is associated.

The consequence of omitting short vowels from Semitic texts and presenting words as sequences of consonants is increased ambiguity. Because many words share a consonantal structure, heterophonic homographs are commonplace in Arabic and Hebrew. In English, a heterophonic homograph would be a word such as *tear* where a single spelling maps onto two different meanings and pronunciations, and readers must rely upon the context to determine which meaning is appropriate. An example of a consonantal structure that is associated with two different meanings in Arabic is presented in Figure 1. Homographs are so common in unvowelized Arabic that it has been estimated that approximately every third word in a passage is ambiguous (Abu-Rabia, 1997). The majority of these homographs are biased in the sense that they have a dominant and a subordinate meaning (e.g. Hermena, Drieghe, Hellmuth & Liversedge, 2015).

Insert Figure 1 about here

Some research has investigated the way in which ambiguity is resolved by contextual priming in Hebrew (e.g. Peleg & Aviatar, 2009) and Arabic (Hayadre, Kurzon, Peleg, & Zohar, 2015). The goal of the current research, however, is to investigate the role of vowelization in the comprehension of ambiguous words. Abu-Rabia (1996, 1998, 1999) showed that vowelization increased the accuracy with which single words and paragraphs were read aloud by both skilled and less skilled readers of Arabic, and improved the ability of school students to answer comprehension questions about passages that they had read. Maroun and Hanley (2017) extended this line of research by investigating whether the presence of vowel diacritics was particularly significant for the comprehension of heterophonic homographs. They asked adult readers of Arabic to decide whether written words had a *living* meaning. The materials included heterophonic homographs that had one *living* and one *non-living* meaning. Results showed that the presence of diacritics significantly increased the accuracy of semantic decisions about ambiguous words but had no effect on the accuracy of decisions about unambiguous words. It therefore appears that the presence of diacritics can allow readers of Arabic to overcome difficulties in accessing the appropriate meaning of ambiguous consonantal sequences.

The present study investigates further the way in which a phonologically and semantically disambiguated word (heterophonic homograph with diacritics) is comprehended. Does this written word's orthography activate only the meaning that is consistent with the vowelization, or does it also activate the alternative semantic representation(s) that share its consonantal structure? Maroun and Hanley (2017) showed that diacritics allow the reader to access semantic representations that are inaccessible when the word is unvowelized. One possibility is that when a written word is disambiguated by vowelization, it is processed in the same way as an unambiguous word. That is, only the diacritized meaning of the homograph is activated. Another possibility is that, to at least some degree, words that are homographs when unvowelized are processed as if they are ambiguous even in the presence of diacritics. In other words, it may be that even when an ambiguous consonantal sequence is disambiguated by diacritics, the meaning that corresponds to the alternative vowelized form of the sequence is also activated. If it is only the diacritized meaning that is accessed, then it would follow that readers rely upon the diacritics when identifying vowelized words. If, however, the alternative meaning is also activated, then it would appear reasonable to conclude that, even when a word is accompanied by vowel diacritics, the consonantal structure is sometimes used to identify the word.

To address this issue, the following experiments investigated the processes that are involved in accessing the meaning of ambiguous words in their disambiguated (diacritized) form. Consider as an example an undiacritized heterophone such as حداد*/ħda:d/*. This is a balanced heterophone that is associated with two approximately equally frequent pronunciationsحِداد*/ħida:d/ mourning* and حَدّاد*/ħadda:d/ blacksmith*. In everyday life, a reader of unvowelized Arabic will use the consonantal sequence together with the context in which it appears to access the word's intended meaning. When, however, he or she processes a diacritized version such as حَدّاد*/ħadda:d/ blacksmith*, there are at least two possibilities. One possibility is that the reader focuses on the letters and diacritics and accesses exclusively the *blacksmith* meaning of the word. This is likely to require the use of sub-lexical letter/diacritic > phoneme correspondences. Another possibility is that the reader, being accustomed to reading the word in its undiacritized form, first activates the orthographic representation associated with the consonantal sequence حداد */ħda:d/* and accesses all of the meanings that are linked to it. A meaning that is inconsistent with the vowelization might therefore be generated by a reader even when the word has been disambiguated by diacritics.

To this end, performance on two types of diacritized word was compared in Experiment 1a. *Related* words comprised words with consonantal structures that are associated with two non-living meanings when undiacritized (e.g. شَعْر*/ʃaʕr/ hair,* شِعْر*/ʃiʕr/ poetry*). *Unrelated* words comprised words with consonantal structures that are associated with one non-living and one living meaning (e.g. عالَم /ʕa:lam/ *world* and عالِم /ʕa:lim/ *scientist*). The experimental task was to decide whether a target word had a non-living meaning. Since both meanings of *related* words had a non-living meaning, no effect on the speed or accuracy of semantic decisions about related words should be observed if the meaning that is associated with the alternative vowelization of the consonantal structure was activated. This is because both meanings would elicit the same response ("*yes*"). With *unrelated* words, however, the meaning associated with the alternative vowleization of the homograph would be *living*. If the alternative meaning of the word was activated, there might then be a reduction in the speed and/or accuracy with which the semantic decision could be made. This is because the alternative *living* meaning might generate conflict by suggesting to the participant that the appropriate response should be "*no*" rather than the correct answer ("yes"). If the *living* meaning were to be activated but discounted, then there should be an effect of *relatedness* on decision latencies. If the *living* meaning is actually selected as the word's meaning, then there should be an effect of *relatedness* on accuracy. If, conversely, the diacritized version of the homograph is processed as if it was unambiguous, then *related* words and *unrelated* words should be processed equally quickly and accurately.

The critical experimental task that was employed in this study was semantic categorization (Experiment 1a). However, the same set of stimuli was also used in two additional experimental tasks (Experiment 1b and 1c). Experiment 1b required participants to make orthographic lexical decisions about the diacritized homographs. By definition, semantic categorization involves semantic access. Lexical decision, however, involves the recognition of the orthographic form of the written word but does not necessarily require that the semantic representation of the word is activated before a decision can be made (e.g., Balota & Chumbley, 1984; Bentin & Frost, 1987; Coltheart et al., 2001). Consequently, the expectation was that no effect of relatedness would be observed on the lexical decision task. The lexical decision task was therefore used as a control task. This was because the sets of *related* and *unrelated* words were not matched for every dimension that might have affected performance. That is, the sets of related and unrelated words were equated for length, familiarity and dominance, but other variables such as neighborhood density could not be controlled. If the speed and accuracy with which the *related* and *unrelated* words are processed were to differ in the lexical decision task, then this would represent evidence that any observed effects of relatedness in the semantic task (Experiment 1a) were caused by uncontrolled variables.

Experiment 1c involved a speeded naming task in which participants were asked to read each diacritized word aloud as quickly and as accurately as possible. The naming task was used as another control task designed to assist the interpretation of performance on the semantic task. As with lexical decision in Experiment 1b, it was anticipated that a naming task should not require access to the meanings of the target words. Consequently, there should be no effect of semantic relatedness on response times. A further aim was to investigate whether the participants could consistently recognize and read out the diacritized written words correctly, or whether they sometimes disregarded the diacritics and simply read aloud the more frequent form of its consonantal sequence. The results of Maroun and Hanley (2017) suggest that readers of Arabic do not ignore diacritics during comprehension when they accompany the word. Nevertheless, it remains possible that, being more familiar with the undiacritized script, participants were unable to use the diacritics consistently to generate the correct pronunciation of a vowelized word (e.g., Hermena et al., 2015).

Experiment 1a

Method

Participants

The sample size in the three experiments was determined with reference to the authors' previous work (Maroun and Hanley, 2017). The participants were 30 undergraduate students, all Arabic native speakers, who were volunteers. They all had a Lebanese high school degree (Baccalaureate), and could therefore be considered as proficient in Arabic reading, as Arabic language is widely covered in the Lebanese baccalaureate curriculum. Like the majority of Lebanese university students, they had all pursued their studies in a second language (English or French). They had not completed any university degree, and were enrolled in a school of public health at the Lebanese University. Their ages ranged between 18 and 25 years. None of the participants had school difficulties or suffered from neurological, emotional, attentional, or learning disorders. Their vision was normal or corrected to normal. All participants signed a consent form approved by the University of Essex prior to performing the experimental tasks. Ethical approval was granted by the University of Essex Science and Health Faculty Ethics committee.

Materials

Insert table 1 about here

Two lists were created containing 21 words that were between three and six letters long. All of the words were nouns preceded by the copula 'al' (equivalent to the English copula 'the') to prevent confusion with verbs or adjectives. The first set (referred to below as the *unrelated* word list) was taken from words chosen by Maroun and Hanley (2017). They contained consonantal sequences that were heterophonic homographs when unvowelized. Each consonantal sequence was associated with two different pronunciations that corresponded to a *living* and a *nonliving* meaning. All the words were presented with diacritics, and the *nonliving* form of the ambiguous word was always presented regardless of whether or not it was the dominant form of the word. Each word had previously been graded for subjective familiarity and for dominance of meaning by 10 independent raters.

A second set of words (referred to below as the *related* word list) was created for the purpose of this experiment. A preliminary set of 37 ambiguous words that had two non-living meanings was formed. Two pilot experiments were conducted to estimate the subjective familiarity and the meaning dominance of this initial pool of words. Ten participants having the same characteristics as the main experiment's participants were asked to rate from 1 to 5 the familiarity of the two diacritized versions of the word. Ten additional participants were asked to define the undiacritized form of the words twice giving a different meaning each time. The results were classified as the first and the second availability of the word, and were considered an indicator of the relative dominance of each meaning of the word. This set also contained 21 words and is referred to below as *related*. The words from the related and unrelated lists were matched for length, familiarity, and dominance. The related and unrelated lists had similar mean familiarity (related *mean* = 4.51, *SD* = 0.47; unrelated mean = 4.69, *SD* = 0.34), similar length (related *mean* = 3.81, *SD* = 0.68; unrelated *mean* =3.57, *SD* = 0 .60), and dominance (related *mean* = 5.62, *SD* = 3.40; unrelated *mean* = 5.48, *SD* = 3.34). Independent t-tests revealed no difference in the level of familiarity (t (40) = 1.34, p > .05), length, (t (40) = 1.21, p >.05), or dominance, (t (40) =.14, p > .05) between the 2 lists. All words were diacritized. Examples of the words are shown in table 1. Of note, words that are dominant tend to also be familiar and vice versa (r = -.67, p < .0001). Forty-two diacritized unambiguous words having one living meaning were used as fillers.

Apparatus and Procedure

Superlab software was used to present the words in a random order. Words were presented in Arial size-66 font on a white screen using an HP Pavilion g6 laptop. All stimuli were fully diacritized except for their ending. In a written sentence, the ending of the word is diacritized according to its grammatical function which is beyond the scope of this study. The diacritization was similar to that found in a widely used dictionary. Response latencies were measured to the nearest millisecond. The participants were tested individually. They sat about 50 cm from the monitor and used both their hands to answer. At the beginning of each trial, a fixation cross was presented in the center of the screen for 2 seconds. Then the first word appeared in the center of the screen. The word remained on the screen until the participant pressed one of two keys on the computer keyboard. Participants were tested individually. They were instructed to look at a cross in the middle of the screen between stimuli, and to press a yes (F) key if the word they see represents a non-living thing, or to press a no key (J) if it cannot represent a living thing.

Experiment 1b

Participants, Materials and Procedure

The participants were 30 undergraduate students drawn from the same population as Experiment 1. None of them had participated in Experiment 1a. The experimental stimuli were the same as Experiment 1a. The fillers were non-words that were created by changing the first letters of the fillers with a living meaning used for experiment 1a. Participants were tested individually and were instructed to press a *yes* key on the keyboard if the word that appeared on the screen was a real Arabic word, or to press a *no* key if it was not a real word.

Experiment 1c

Participants, Materials and Procedure

The participants were 25 undergraduate students drawn from the same population as before and who had participated in neither Experiment 1a nor 1b. The experimental stimuli were the same as in Experiment 1a and 1b. Participants were tested individually and were instructed to read each presented word aloud as quickly and accurately as soon as it appeared. Inquisit software was used to present the words in a random order and record the participants' responses. Naming latency was measured from the onset of the word stimulus to the onset of the vocal response.

Results and Discussion

Experiment 1a: Semantic Categorization.

Insert Table 2 about here

Mean semantic categorization decision latencies for correct responses and mean correct answers were calculated across participants. Significant effects reported in all experiments were based on a .05 alpha level. Accuracy was a measure of word comprehension; it revealed how well a single written word was understood. Reaction times (RTs) to recognize words, measured the speed of word comprehension.

A paired-samples t-test was conducted to compare the number of correct responses. It compared the number of recognized unrelated words (one of their meanings is non-living and one is living) with the number of recognized related words (both of their meanings are non-living). Another paired-sample t-test was conducted to compare RTs to correctly recognized unrelated and related words. Bayes factors were also calculated using the JASP (2014) program. There was a main interfering effect of the competing meaning of the ambiguous word on accuracy scores, $t(29) = 6.62$, $p < .0001$, $d = 0.92$, Bayes Factor = 54112) and on reaction times, $t(29) = 4.76$, $p < .001$, $d = 0.41$, Bayes Factor = 495). On average, participants had significantly higher mean accuracy scores on the 21 diacritized words that had two non-living versions/meanings when undiacritized (M = 17.67/21, SD = 2.79) than on the 21 words diacritized words that had one living and one non-living version/meaning when undiacritized (M = 14.70/21, SD = 3.47). Bayesian analysis revealed that such a difference was over 50,000 times more likely under the experimental hypothesis than under the null hypothesis. Reading speed was also significantly slower when the word had one non-living and one living meaning (M = 2215, SD = 800), than when it had two non-living meanings (M =

1920, *SD* = 628). Bayesian analysis revealed that such a difference was 495 times more likely under the experimental hypothesis than under the null hypothesis.

Experiment 1b: Lexical decision.

Mean RTs and accuracy scores are shown in Table 2. Results of the paired-samples t-tests showed that there was no longer any interfering effect of the meaning of the competing version of the ambiguous word on either accuracy scores, $t(29) = 1.89$, $p = .07$, $d = 0.35$, Bayes factor = .927), or reaction times, $t(29) = 1.51$, $p = .14$, $d = 0.19$, Bayes factor = .538). The number of correct responses to lexical decisions for words having an unrelated (living) competing version (*M* = 16.77/21, *SD* = 4.65) did not differ from that for words having a related (non-living) competing meaning (*M* = 17.30/21, *SD* = 5.23). On average, time to decide on lexicality did not differ between words having a different (i.e., living) competing meaning (*M* = 1780, *SD* = 840) and those having a related (i.e., non-living) meaning (*M* = 1637, *SD* = 657). Bayesian analysis revealed that the differences between the means for both accuracy and decision latencies were more likely to have occurred under the null hypothesis than under the experimental hypothesis.

Experiment 1c: Speeded Naming

Mean RTs and accuracy scores are shown in Table 2. Results of the paired-samples t-tests showed that there was no significant difference in the accuracy scores of words with two non-living meanings (*M* = 20.20/21, *SD* = 0.82) and those with one living and one non-living meaning meanings (*M* = 20.44/21, *SD* = 0.72), $t(24) = 1.30$, $p = .21$, $d = 0.31$, Bayes factor = .445). Likewise, there was no significant difference, $t(24) = 1.77$, $p = .09$, $d = 0.17$, Bayes factor = .814), between the reaction times to words with related meanings: (*M* = 881, *SD* = 125) and words with unrelated meanings: (*M* = 860, *SD* = 106). Bayesian analysis revealed that the differences between the means for both accuracy and decision latencies were more likely to have occurred under the null hypothesis than under the experimental hypothesis.

Across-Experiments Comparisons

The results of the t-tests revealed significant differences in the semantic decision task but not the lexical decision nor speeded naming tasks. Nevertheless, the conclusions that can be drawn from the data would be more secure if it could further demonstrated that differences between the two types of word were significantly greater in the semantic categorization task than in the other two tasks. The number of accurate responses and the RTs in Experiments 1a, 1b and 1c were therefore subjected to a two-way mixed analysis of variance. Task (semantic categorization, lexical decision, naming) was a between-subjects variable, and relatedness (related, unrelated) was the within-subjects variable. The main effect of relatedness was significant for accuracy scores F (1, 82) = 30.8, $p< .001$ and for reaction times F (1, 82) = 11.9, $p = .001$. Crucially, the interaction between task and relatedness was highly significant for both accuracy, F (2, 82) = 24.5, $p \le .001$ and reaction times, F (2, 82) = 4.9, $p = .009$. These significant interactions support the view that the effect of relatedness was different across tasks, with larger effects of relatedness on the semantic decision task than on the lexical decision or naming tasks.

Discussion

The results of Experiment 1a have provided important information about the way in which readers of Arabic process vowelized words. Earlier research by Maroun and Hanley (2017) revealed that participants could use the diacritized form of an ambiguous word to access its intended pronunciation and meaning. Maroun and

Hanley claimed that, on some occasions, vowelization allowed participants to activate a semantic representation that they would not have been able access from a word's consonantal structure alone.

Experiment 1a investigated whether readers of Arabic *consistently* accessed the meaning of a vowelized word by processing the diacritics that it contained. If so, semantic representations that were inconsistent with the word's vowelization should not have been activated even if they were consistent with the word's undiacritized consonantal sequence. There should therefore have been no effect of *relatedness* on semantic decision-making in Experiment 1a because the diacritized form of the word should have led participants directly to its *non-living* meaning. The fact that the consonantal structure of the unvowelized words from the *unrelated* set were also associated with a *living* meaning should have had no effect on the speed or accuracy with which a *non-living* decision could be made. The finding that there was an effect of relatedness on both the speed and accuracy with which semantic decisions were made in Experiment 1a is therefore an important finding. It shows that participants sometimes accessed a diacritized word's meaning by ignoring the diacritics and focusing on its consonantal structure.

The finding that accuracy was significantly higher in the *related* condition in Experiment 1a suggests that on some trials in the *unrelated* condition the *living* rather than the *non-living* meaning of the word was activated. Activation of a *living* meaning presumably convinced the participant that the word did not have a non-living meaning and elicited an incorrect "no" response on the semantic decision task.

Even when a correct response was made to a word from the *unrelated* set, decision latencies were significantly slower than in the *related* condition. It therefore appears that both of the semantic representations associated with a word's consonantal structure were sometimes activated. In the *unrelated* condition, the living meaning seems to have competed with non-living meaning, with extra time being required to resolve the conflict. Of course, both of a word's meanings are also likely to have been activated in the *related* condition. In the *related* condition, however, activation of both meanings would not be expected to reduce the speed or accuracy of semantic decisions. This is because the alternative meaning of the word would also have been *non-living* and would not have led to conflict with the vowelized *non-living* meaning. So, even if a meaning was activated that was inconsistent with a word's vowelization in the *related* condition, it would have nevertheless have led to an accurate response on the semantic decision task.

The effect of relatedness on RTs in Experiment 1a is consistent with the claim (Folk and Morris, 1995; Onifer and Swinney, 1981) that alternative meanings of heterophonic homographs are sometimes activated during reading. In Hebrew, however, the dominant meaning of a homophonic heterophone appears to strongly inhibit its subsidiary meaning (Bitan et al., 2017). The effect of semantic relatedness on accuracy that we observed in Experiment 1a suggests that even if alternative meanings are initially activated from the consonantal sequence, one of these meanings often inhibits the other (Gottlob, Goldinger, Stone and Van Orden, 1999). It appears that in the *unrelated* condition, the *living* meaning that was associated with the word's consonantal structure sometimes inhibited its *non-living* meaning and elicited a slow or incorrect response. It seems likely that the dominant meaning of a consonantal sequence inhibited the less dominant meaning. When the dominant meaning of a consonantal sequence was *living*, then it seems to have sometimes guided participants in the direction of an incorrect response despite the presence of vowel diacritics that were inconsistent with the *living* meaning. Our restricted set of words meant that we were unable to determine whether interference was specific to consonantal sequences where the dominant meaning was *living* or whether it also occurred when the dominant meaning was *non-living*. In future research, it would be interesting to discover whether there would be priming from the unvowelized meaning of a vowelized heterophonic homograph to words with which it is semantically associated.

The results of Experiment 1a might, of course, have had a more prosaic explanation because it was not possible to match the words in the related and unrelated sets for all of the variables that might have affected performance. If, however, performance in the unrelated set was associated with greater speed and accuracy because of these uncontrolled variables, then a similar pattern should also have been observed in lexical decision (Experiments 1b) and speeded naming (Experiment 1c). In fact, though, although there were small differences between the relevant means in Experiment 1b and 1c, none of these differences were significant, and Bayes factors showed that these differences were more likely to have occurred under the null hypothesis than the experimental hypothesis. Most important of all, a significant task x word type interaction showed that effects of relatedness were strongest when participants were performing semantic categorization.

In conclusion, Maroun and Hanley (2017) showed that, on some occasions, readers of Arabic used diacritics to disambiguate heterophonic homographs that they would otherwise have misinterpreted. The present study has shown that, on other occasions, readers of Arabic ignore diacritics and misinterpret heterophonic homographs that they would have comprehended correctly if they had paid attention to vowelization. The accuracy with which participants could read aloud the vowelized words on a separate speeded naming task makes it unlikely that these errors came about because participants were insufficiently familiar with the vowelized script to process the diacritics efficiently. This outcome provides further evidence of the key role that mental representations of consonantal sequences play in identifying written words in Arabic.

References

Abu-Leil, A.K., Share, D.L & Ibrahim, R. (2014). How does speed and accuracy in reading relate to reading comprehension in Arabic? *Psicológica, 35*, 251-276.

Abu-Rabia, S. (1997). 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 Research in Reading*, *20,* 137-147.

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.

Al-Mounjid (3rd ed). (2008). Beirut: Dar Almachreck (in Arabic).

Asadi, I.A. (2017). Reading Arabic with the diacritics for short vowels: vowelised but not necessarily easy to read. *Writing Systems Research, 2,* 137-147.

Balota, R.A. & Chumbley, J.I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage*, Journal of Experimental Psychology: Human Perception and Performance 10*, 340–357.

Bentin, S., & Frost, R. (1987). Processing lexical ambiguity and visual word recognition in a deep orthography. *Memory and Cognition*, *15,* 13–23.

Bitan, T., Kaftory, A., Meiri-Leib, A., Peleg, O ., & Eviatar, Z. (2017). Phonological ambiguity modulates resolution of semantic ambiguity during reading: An fMRI study of Hebrew. *Neuropsychology, 31*, 759-777.

Bourisley, A.K., Haynes, C., Bourisley, N., & Mody, M. (2013). Neural correlates of diacritics in Arabic: An fMRI study. *Journal of Neurolinguistics, 26*, 195-206.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*, 204–256.

Feldman, L. B., & Turvey, M. T. (1983). Word recognition in Serbo- Croatian is phonologically analytic. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 288-298.

Folk, J.R., & Morris, R.K. (1995). Multiple lexical codes in reading: Evidence from eye movements, naming time, and oral reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1412-1429.

Frost, R. (1998) Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin 123*(1):71–99.

Frost, R. (2012). Towards a universal model of reading. *Behavioral and Brain Sciences, 35*, 263–329.

Frost, R. & Bentin, S. (1992a). Processing phonological and semantic ambiguity: Evidence from semantic priming at different SOAs, *Journal of Experimental Psychology: Learning, Memory and Cognition 18*, 58–68.

Frost, R., & Bentin, S. (1992b). *Reading consonants and guessing vowels: Visual word recognition in Hebrew orthography*. Amsterdam: Elsevier.

Frost, R., Katz, L., & Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison*. Journal of Experimental Psychology: Human Perception and Performance, 13,* 104-115.

Gottlob, L.R., Goldinger, S.D., Stone, G.O. & Van Orden, G.C. (1999). Reading homographs: Orthographic, phonological and semantic dynamics. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 561-574.

Hermena E.W., Drieghe D., Hellmuth S., & Liversedge S.P. (2015). Processing of Arabic diacritical marks: phonological-syntactic disambiguation of homographic verbs and visual crowding effects. *Journal of Experimental Psychology: Human Perception and Performance, 41*(2), 494-507.

Hayadre, M., Kurzon, D., Peleg, O. & Zohar, E. (2015). Ambiguity resolution in lateralized Arabic. *Reading and Writing, 28*, 395–418

Ibrahim, R., Eviatar, Z., & Aharon-Peretz, J. (2002). The Characteristics of Arabic orthography slows its processing. *Neuropsychology, 16,* 322-326.

JASP. (2014). Retrieved from http://jasp-stats.org/

Katz, L., & Feldman, L. B. (1981). Linguistic coding in word recognition. In A. M. Lesgold & C. A. Perfetti (Eds.), *Interactive processes in reading* (pp. 85-105). Hillsdale, NJ: Erlbaum.

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In L. Katz & R. Frost (Eds.), *Orthography, phonology, morphology and meaning*. Amsterdam: Elsevier.

Klima, E. S. (1972). How alphabets might reflect language. In F. Kavanagh & I. G. Mattingly (Eds.), *Language by ear and by eye* (pp. 57-80). Cambridge, MA: MIT Press.

Liberman, I. Y., Liberman, A. M., Mattingly, I. G., & Shankweiler, D. (1980). Orthography and the beginning reader. In J. F. Kavanagh & R. L. Venezky (Eds.), *Orthography, reading, and dyslexia* (pp. 137-154). Austin, TX: Pro-Ed.

Lukatela, G., Popadic, D., Ognjenovic, P., & Turvey, M. T. (1980). Lexical decision in a phonologically shallow orthography. *Memory & Cognition, 8*, 415-423.

Maroun, M., & Hanley, J.R. (2017). Diacritics improve comprehension of the Arabic script by providing access to the meanings of heterophonic homographs. *Reading and Writing: An Interdisciplinary Journal, 30 (2),* 319–335.

Navon, D, & Shimron, J. (1981). Does word naming involve grapheme- to-phoneme translation? Evidence from Hebrew. *Journal of Verbal Learning and Verbal Behavior*, 20, 97-109.

Navon, D., & Shimron, J. (1985). *Source and function of phonological codes in processing written Hebrew: Final report to the US–Israel Binational Science Foundation*. Haifa, Israel: University of Haifa Press.

Onifer, W., & Swinney, D. A. (1981). Accessing lexical ambiguities during sentence comprehension: Effects of frequency of meaning and contextual bias. *Memory & Cognition, 9*, 225-236.

Peleg, O., & Eviatar, Z. (2009). Semantic asymmetries are modulated by phonological asymmetries: Evidence from the disambiguation of heterophonic versus homophonic homographs. *Brain and Cognition, 70*, 154–162.

Roman, G., & Pavard, B. (1987). A comparative study: How we read Arabic and French. In J. K. O'Regan, & A. Levy-Schoen (Eds.), *Eye movements: From physiology to cognition* (pp. 431-440). Amsterdam: North Holland Elsevier.

Saiegh-Haddad, E. (2003). Linguistic distance and initial reading acquisition: the case of Arabic diglossia. *Applied Psycholinguistics, 24*, 431-451.

Saiegh-Haddad, E. (2004). The impact of phonemic and lexical distance on the phonological analysis of word and pseudowords in a diglossic context. *Applied Psycholinguistics*, 25,495-512.

Saiegh-Haddad, E., & Schiff, R. (2014). The impact of diglossia on voweled and unvoweled word reading in arabic: a developmental study from childhood to adolescence. *Scientific Studies of Reading*, 25,495-512.

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

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed develop- mental model of word recognition and naming. *Psychological Review, 96*, 523-568.

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an "outlier" orthography. *Psychological Bulletin, 134*(4), 584-615.

Shimron, J. (2006). *Reading Hebrew: the language and the psychology of reading it.* Mahwah, New Jersey: Lawrence Erlbaum Associates.

Van Orden, G. C., Pennington, B. R, & Stone, G. O. (1990). Word identification in reading and the promise of subsymbolic psycholinguistics. *Psychological Review, 97,* 488-522.

Figure 1. An example of a heterophonic homograph in Arabic.

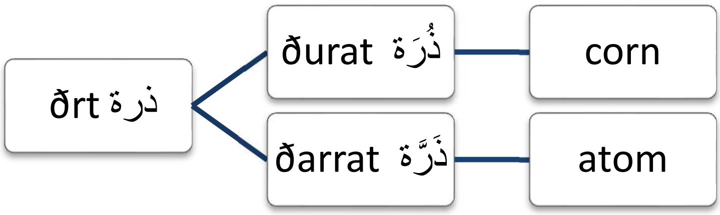

Table.1: Examples of words used in Experiment 1, 2, and 3.

Diacritized version of an ambiguous word with 2 non-living meanings	**Diacritized non-living version of an ambiguous word with 1 living and 1 non-living meaning**
السُحَاب	الخَلْف
المَرْكَب	السَّلْق
البِرْكَة	النُّور
السُكْر	الحِدَاد

Table 2: Mean and standard deviations of accuracy and latency scores in Experiments 1, 2, and 3.

	Experiment 1 **Semantic Categorization**		**Experiment 2** **Lexical Decision**		**Experiment 3** **Naming**	
	Ambiguous words with 1 living and 1 non-living meaning	Ambiguous words with 2 non-living meanings	Ambiguous words with 1 living and 1 non-living meaning	Ambiguous words with 2 non-living meanings	Ambiguous words with 1 living and 1 non-living meaning	Ambiguous words with 2 non-living meanings
Accuracy	M = 14.70 SD = 3.47	M = 17.67 SD = 2.79	M = 16.77 SD = 4.65	M = 17.30 SD = 5.23	M = 20.44 SD = 0.71	M = 20.20 SD = 0.82
ReactionTimes	M = 2215 SD = 800	M = 1920 SD = 628	M = 1780 SD = 840	M = 1637 SD = 657	M = 860 SD = 106	M = 880 SD = 125