Psicológica (2015), 36, 123-142.

Language-specific and language-general factors in text reading in Arabic: Evidence from the Missing Letter Effect

Amany Saady, Raphiq Ibrahim* and Zohar Eviatar

University of Haifa, Israel

The goal of the present study was to extend the models explaining the missing-letter effect (MLE) to an additional language and orthography, and to test the role of phonology in silent reading in Arabic. We also examined orthographic effects such as letter position and letter shape, morphological effects such as pseudo-prefixes, and phonological effects such as pronounceability. The results showed that readers miss letters more often in function words and prefixes than in content words, more in second position than in first position, more often when the letters are silent than pronounced, and less often when the letter shape is more symmetric and stable. The results show that these aspects of the missing letter effect can be generalized over writing systems that are not alphabetic, suggesting that the models proposed to explain the MLE in all the orthographies tested may reflect a universal aspect of reading.

The study of language is in an interesting state of flux. The major paradigm which proposed that all languages are essentially the same (e.g., Chomsky, 2002) has been challenged by the typological proposition that the major characteristic of human languages is their diversity (Evans & Levinson, 2009). Obviously, languages have common factors, as they all occur in the context of human brains interacting with each other. Also obviously, languages differ on almost every level of analysis that has been proposed. This is the general context in which the study of cross-language effects becomes crucial: which language characteristics are universal, because all occur in human brains, and which language characteristics are specific to a culture? One of the places in which languages seem to be most

* Correspondence to: Raphiq Ibrahim. Department of Learning Disabilities. University of Haifa. Mount Carmel, Haifa. 31905 Israel. E-mail: raphiq@psy.haifa.ac.il

different from each other is in the way that they are transcribed in writing systems. In this paper, we examine specific characteristics of the Arabic writing system and the manner in which they affect the missing-letter effect.

The 'missing-letter effect' (MLE) occurs when participants are asked to read texts and to circle a particular letter every time they encounter it. Participants tend to miss the target letter when it appears in smaller and more frequent words (e.g., Corcoran, 1966).In English, it has been found that letters are missed more often in function words than in content words, as these are both shorter and more frequent (e.g., Healy, 1976). Two major hypotheses have been proposed to account for this. The first is *the unitization model*(Healy, 1976; Healy & Drewnowski, 1983; Healy, Oliver, & Macnamara, 1987), which attributes importance to the size of the processed units over the course of reading. These researchers suggest a hierarchy of the levels of perceptive processing, where the degree of automaticity of processing written materials is the crucial factor in determining the level of detail with which a stimulus is processed. The assumption is that reading is performed on parallel levels of analyses: word - syllable- letter-feature, and that when processing on a higher level is completed, processing at the lower levels is aborted or truncated. Thus, when a person reads a very familiar word and it is easily encoded, processing will not continue to the level of separate letters. Because less effort is invested in processing familiar words such as function words, target letters in these words are missed more often.

The second hypothesis is the *structural model* (Koriat & Greenberg, 1994; Koriat, Greenberg & Goldshmid, 1991), which suggests that processing of text requires encoding of both structure and meaning. The interpretation of the structure sets the basis for the interpretation of the meaning, while at the same time, a displacement of attention from structure to content occurs in the course of the process. Because function words have a structural role, they are not attended, and thus target letters in function words are missed more often. Letters are missed as a result of the syntactic role of the words in which they occur, not as a result of the length or frequency of the words. Koriat, Greenberg and Goldshmid, (1991) used the characteristics of Hebrew to dissociate syntactic role from length and frequency. In Hebrew, many function morphemes can appear as separate letters which are attached as prefixes to content words. Although words with prefixes werenot more frequent than content words, participants missed the target letter more often when it served as a function prefix than when it was part of a content word. These findings were interpreted as indicating that the missing letter effect reflects *post-lexical* processing in

which the prominence of the function morpheme is weakened during the transition from structure to meaning.

The factors that affect letter detection during reading have been studied and debated for a number of decades. Among these, are effects of the phonemic manifestation that the letter represents (e.g., Corcoran, 1966; Schneider & Healy,1993, cf. Drewnowsky & Healy, 1982) as well as effects of the visual 'gestalt goodness' of the letter shape (Schneider & Healy, 1993, Saint-Aubin & Piorier, 1997). In addition, the position of the letter in the word also affects its detection, where the probability of missing a letter in the beginning of the word is lower than missing it elsewhere, and this is independent of pronounceability (Corcoran, 1966; Schneider & Healy,1993).

Recently, Ibrahim (2011) examined the MLE in Arabic, and showed that miss rates were higher in function words and function prefixes than in content words, supporting the structural model of the MLE. Musseler, Koriat, and Nißlein (2000) showed that the miss rate is sensitive to sentence structure in German, where more typical sentence structure resulted in higher miss rates than less frequent structures, also supporting the structural model. Musseler, Nißlein, and Koriat (2005) showed that orthographic cues to structure (such as capitalization in German, which indicates nouns) also affect the MLE. Saint-Aubin and Poirier (1997) report data consistent with the structural account in French, but also report effects of specific letters, which they explain as resulting from the visual properties of the letters. The structural model does not predict that phonological or orthographic characteristics of letters should affect the MLE, whereas the unitization model does accommodate this finding. Greenberg, Healy, Koriat, and Kreiner (2004) presented a model which integrates the unitization and structural models, the Guidance Organization model, (GO). The GO model posits that unitization processes (which are subject to the visual characteristics of the letters in words) underlie the efficient recognition of function words that guide the computation of the structure of the sentence, thus incorporating both prelexical and postlexical sources for the missing letter effect.

In the present study, we continue our exploration of the process of reading in Arabic (Eviatar & Ibrahim, 2004; 2007; Eviatar, Ibrahim & Ganayim, 2004;Ibrahim & Eviatar, 2009; 2012), which, we have suggested, poses unique challenges to readers. To do this, we use the unique characteristics of Arabic to examine more closely the effects of phonology, visual shape, and letter position on the missing letter effect. As detailed below, the phonological manifestation of the letter 'l' depends on the

identity of the letter that it precedes. Given that the determiner 'al' (the) is added as a prefix to nouns, we can examine the effects of pronounceability on the missing letter effect, holding letter position and function constant. Specifically, we will examine whether the processes of building morphological and text structure during text processing in Arabic, integrate pre- and post-lexical phonological processes.

The Challenge of Reading in Arabic

Arabic orthography is an abjad (Daniels, 1990), wherein the consonants are represented by letters, and the vowels by optional diacritics.All verbs and most nouns are written primarily as roots that are differently affixed and vowelled to form the words of the lexicon (Prunet, Beland, & Idrissi, 2000). Most written texts do not include short vowels. When vowels are included (in poetry, childrens' books and liturgical texts), they are signified by diacritical marks above or below the letters. Inclusion of these diacritical marks completely specifies the phonological form of the orthographic string, making it completely transparent in terms of orthography/phonology relations. Thus, voweled Arabic words are orthographically transparent, in the sense that all of the phonological information necessary for identification is represented. Unvowelled Arabic texts are orthographically opaque, and include many heterophonic as well as homophonic homographs. Information about the phonological form of words must be inferred from the morphological, the contextual and the lexical cues present in the text. Arabic is read from right-to-left.

Previous research has shown that skilled reading of single words and texts in Arabic is slower than in other languages (Azzam, 1984; Eviatar & Ibrahim, 2004; Abu-Rabia, 2001). In addition, reading acquisition by beginning readers seems to be more challenging than in other languages (Saiegh-Haddad, 2003; Asaad & Eviatar, 2013; 2014). Three major reasons have been proposed for this (e.g., Eviatar & Ibrahim, 2013): The diglossic situation in most Arab countries (where the written language is different from the spoken language) (Saiegh-Haddad, 2004, Eviatar & Ibrahim, 2000); the visual complexity of the letters (Ibrahim, Eviatar, & Aharon-Perez, 2002); and the specific disability of the right hemisphere to distinguish between very similar graphemes that denote very different phonemes (Eviatar, Ibrahim, & Ganayim, 2004).

The visual complexity of the Arabic orthography is accentuated by the fact that all letters in words are connected to each other from both sides, except 6 letters which connect only to the preceding letter. Some ligatures are not linear, in that one letter may be written above another. Additional

complexity has to do with the fact that most of the letters change shape depending on their placement in the word. For example, the letter ع (circled in the following examples) is written differently in the beginning (علي),, the middle (لعب), and the end (مع) of words. Taouk and Coltheart (2004, Experiment 2) took advantage of this property to test the manner in which skilled and novice readers read single words. The participants were presented with words in Arabic written either correctly, or graphemically distorted, with letters in the right places, but in the wrong form for their placement. The results showed that skilled readers were more affected by the graphemic distortion than novice readers. The authors concluded that skilled readers of Arabic use a whole-word encoding strategy, whereas novice readers use a phonological encoding strategy. This hypothesis was supported by Simon, Bernard, Lalonde, and Rebai, (2006) who examined French and Arabic readers, and found electrophysiological evidence that French readers utilize a grapheme-phoneme conversion strategy whereas Arabic speakers do not. Contradictory evidence was presented by Bentin and Ibrahim (1996), who showed that in a lexical decision task, pseudo-words that were phonologically similar to words in spoken Arabic (which is not written) took longer to reject than those that were not. They concluded that phonological encoding is a mandatory stage in reading Arabic. These studies used single words either in lexical decision or naming tasks. In the present study, the stimuli were sentences, which were read silently, and no pseudo-words were presented. Thus, the task was as close to normal reading as we could make it.

In order to explore the role of phonology in the reading of Arabic and in the MLE, we took advantage of another special characteristic of Modern Standard Arabic: The 28 letters of the alphabet are divided into two groups of 14: sun and moon letters, which differ in the way that they affect the pronunciation of the determiner 'al'(the). This is a bound morpheme (ألـ) that connects to the beginning of the noun. When the determiner 'al' appears before words beginning with moon letters, it is fully pronounced. However, when it appears before sun letters, the /l/ is silent. Thus when the word 'the-chair' is written (the determiner is underlined- الكرسي), it is pronounced /alkursi/, because the letter ك which denotes /k/, is a moon letter. However, when 'al' appears in front of a sun letter, such as ش which denotes the sound /ʃ/in the word 'the-sun'(the determiner is underlined - ألشمس), it is pronounced /ashams/, with the /l/ sound becoming silent. Thus, if the letter ل,which denotes the sound /l/, is missed more often when it precedes sun letters (when it is silent), than when it precedes moon letters

(when it is pronounced), this would support the hypothesis that even during silent reading of text, there is mandatory phonological encoding in Arabic.

In the present study we report the results of three experiments. In the first, we tested the generalizability of the structural model to Arabic. We used two function words: 'to'ela-ألى, which is a free morpheme, and 'the' - ألـ, which is a bound morpheme, added as a prefix). The rate of missing the letter ل was compared in these words and in matched content words in which the letter appeared in the second position in the word. In addition, the function word ألى('to') can also be represented as a bound morpheme, where the letter ل is attached to the beginning of a content word as in (للسوق – to the market). In all of the stimuli used in this experiment, the target letter ل preceded moon letters, so that it always had an acoustic manifestation when the words were read aloud. Thus, different rates of misses in these words as compared to matched content words would support the structural model of reading, suggesting that the syntactic role of the words was the defining factor in its depth of processing, not length and frequency.

In the second experiment we used the function prefix ألـ, but in half of the stimuli, it preceded a content word beginning with a sun letter (which would result in its being unpronounced) and in half it preceded a moon letter, (which would result in its being pronounced). The major question here was whether this characteristic would affect letter detection in silent reading. Such an effect would suggest that there is a mandatory phonological stage during silent reading of Arabic.

In the third experiment we examined the effects of the visual properties of the letters and of letter position on the rate of misses. Are asymmetric (less 'Gestalt good') letters missed more often in different positions irrespective of the syntactic role they play?

METHOD

Participants. Sixty six students from the University of Haifa (15 males, mean age 23.5, range 20-30) participated in all three experiments. All had normal or corrected-to-normal vision and no reported learning disability. All were native Arabic speakers who had completed Arabic-language high school and had succeeded in at least level 3 (out of 5) of the matriculation exams in the Arabic language. Thus, all of them can be regarded as skilled readers of the Arabic language. The participants were recruited via advertisements on campus. Of the participants, 39 received payment (30 NIS, approximately $7.00), and 28 received course credit

points. The participants come from Arab villages and cities in the north of Israel. All participants reported that the spoken language with family members is Arabic.

Design. As mentioned above, we carried out three experiments in order to test different hypotheses:

Experiment 1. The goal of this experiment was replication of the effects of syntactic role (function or content word) in the missing-letter effect, as was found in Hebrew, and the examination of the effect of letter position in a ligatured orthography. As did Koriat et al. (1991) in Hebrew, we used both bound and free function morphemes. The dependent variable was the percentage of times participants missed the letter ل in the critical words in the sentence stimuli. There were two independent factors: the syntactic role of the word in which the target letter appeared (content word versus function word) and the position in which the letter appeared (first position or second position in the word). Examples of these sentences, together with the design, are shown in the top panel of Table 1. There were 14 instances of ل appearing in second position in a content word, and 10 instances in each of the other experimental conditions.

Experiment 2. The goal of this experiment was to see whether the phonological manifestation of the target letter affects its detection. Here again the dependent variable was the % rate of misses for the letter ل. The independent variable was the type of word in which the target letter appeared. There were three types (levels) of words: i) content words which begin with the combination ألـ (‘al’), ii) words with the determiner أل (‘al’ - the) before a moon letter (where it would have been pronounced, had it been read out loud), and iii) words with the determiner أل (‘al’ - the) before a sun letter (where it would have been silent, had it been read out loud). Thus, the orthographic context was identical in the three conditions, but the syntactic function was different between content words and function prefixes, and phonological manifestation was different between prefixes where the target letter appeared before moon and sun letters. Examples of the stimuli used in this experiment are shown in Table 2. There were 14 instances of each type of sentence.

Experiment 3. The third experiment focused on the effects of the identity of the specific letter. Here we compared rates of misses of the letter ل with rates of misses of the letter أ. This was done in three separate analyses:

a) In the first, we used two independent variables: the identity of the letter (ل or أ), and the syntactic role of the morpheme in which it occurred (as part of a function word or morpheme, or as part of a content word). Here

all the target letters appeared in the first position in the critical words. There were 10 instances each for the conditions using the target ل and 14 instances for the conditions using the target أ. The percent misses was the dependent variable. The design and examples of the items are shown in the top panel of Table 3.

Table 1: Design of Experiment 1 and examples of the stimuli used in each condition. The target letter is circled in the examples.

Experiment 1: effects of syntactic role and letter position- target letter ل			
		Syntactic role	
		Function	Content
Position in the word	Initial letter	لقد كتبنا أمس لصديقنا رسالة وداع مؤثرة	لقد كتبت صديقي لميس رسالة وداع مؤثرة
	Second letter	لقد كتبنا أمس إلى صديقنا رسالة وداع مؤثرة	يوصي اختصاصيو تغذية كثيرون بعدم تناول أطعمة تحوي ملحاً بنسبة كبيرة.

b) In the second analysis we looked at the effects of letter identity and position in the word. This analysis used only content words. The independent variables here were letter identity (ل or أ) and letter position (initial or second).There were 10 instances of the letter ل in inital position and 14 instances of the other conditions. The design and examples of the stimuli are shown in the middle panel of Table 3.

c) In the third analysis we compared the rate of misses for the target letter أ when it was word-initial in a content word that included ل in second position, resulting in a pseudo-prefix, versus when it was word-initial in a content word that has another letter in second position, so it does not look like a prefix. There were 14 instances of each type of item. Examples are shown in the bottom panel of Table 3.

Table 2: Design of experiment 2 and examples of the stimuli used. The target letter is circled in the examples.

Experiment 2- effects of syntactic role and phonology		
Syntactic role in the sentence - target letter ل		
Content Word beginning with the combination 'al' (pseudo-prefixed)	**Function prefix**	
	Function prefix preceding a moon letter (/l/ pronounced)	**Function prefix preceding a sun letter (/l/ silent)**
يوصي اختصاصيو تغذية كثيرون بتناول أطعمةً تحوي أليافاً بنسبة كبيرة	يوصي اختصاصيو تغذية كثيرون بتناول أطعمة تحوي البروتين بنسبة كبيرة.	يوصي اختصاصيو تغذية كثيرون بعدم تناول أطعمةً تحوي الدهنيات بنسبة كبيرة

Materials. Each participant received three booklets that included unconnected sentences in Arabic; Booklet 1 was three pages long and Booklets 2 and 3 were four pages long, with 15 sentences on each page. The sentences in each booklet were ordered randomly. Sentences in all of the booklets were presented in the same order. For Booklets 1 and 2 the target letter was ل, and the conditions shown in Tables 1 and 2 were distributed among these two booklets. Booklet 3 contained all the stimuli for which the target letter was أ. Booklet 1 included 15 filler sentences and booklets 2 and 3 included 12 filler sentences each. Thus, our participants read 181 sentences in all.

Procedure. The experiment was run in small groups. Each participant received the three small booklets. The instructions to circle the target letter in the booklets appeared at the beginning of each booklet, including a short sentence as an example. The instruction indicated that the participant must read the paragraphs at their regular reading speed. Every time he/she encountered the target letter which appeared at the top part of the page, it

should be circled. Participants were told not to go back and circle letters they thought they had missed, and not to change their reading speed. The order in which booklets were given to the participants was counterbalanced.

Table 3: Design and examples of the stimuli used in experiment 3. Analysis 1 explored the effects of letter identity and syntactic role, analysis 2 examined the effects of letter identity and position in the word (all the words used in this analysis were content words) and analysis 3 looked at the effects of orthographic context on letter detection in content words.

Experiment 3: Effects of letter identity and context			
Analysis 1: letter identity and syntactic role			
		Syntactic role	
		Function	**Content**
Target letter	ل	لقد كتبنا أمس لصديقنا رسالة وداع مؤثرة	لقد كتبت صديقتي لميس رسالة وداع مؤثرة
	أ	يوصي اختصاصيو تغذية كثيرون بعدم تناول أطعمة تحوي الهيدرات بنسبة كبيرة.	يوصي اختصاصيو تغذية كثيرون بتناول أطعمة تحوي أصنافاً كثيرة.
Analysis 2: letter identity and letter position			
		Letter position	
		initial	**second**
Target Letter	ل	لقد كتبت صديقتي لميس رسالة وداع مؤثرة	يوصي اختصاصيو تغذية كثيرون بعدم تناول أطعمةٌ تحوي ملحاً بنسبة كبيرة.
	أ	يوصي اختصاصيو تغذية كثيرون بتناول أطعمة تحوي أصنافاً كثيرة.	يوصي اختصاصيو تغذية كثيرون بتناول أطعمٌ تحوي سائلاً بنسبة كبيرة
Analysis 3: orthographic context. Target letter أ			
	Psuedo-prefix		**Not pseudo prefixed**
	يوصي اختصاصيو تغذية كثيرون بتناول أطعمة تحوي أليافاً بنسبة كبيرة.		يوصي اختصاصيو تغذية كثيرون بتناول أطعمة تحوي أصنافاً كثيرة.

RESULTS

The target letters 'ل' and 'أ' appeared in the filler sentences and in other words as well. Here we report the omission rates only of the letters in the critical words for which we controlled.

Experiment 1: Replication of the Missing-Letter Effect in Arabic and the Effects of Letter Position

In order to extend the findings of Koriat et al. (1991) to Arabic, we compared the rates of letter misses in function words and content words. We compared the number of misses of the letter ل when it was in the context of the function word 'ela' ألى, or in the second position of a content word. We were also interested in letter position, so we included conditions in which the target letter was in a function prefix position (the initial letter of the word), versus when it was in first position in a content word.

The percentage of misses that each participant made in each of these conditions was analysed with a 2-way ANOVA with syntactic role (function or content) and letter position (first or second) as within-subject factors. The analysis revealed that overall, more letters were missed in the context of function words (M=10.76%, sd=15.4) than of content words (M=3.57%, sd=6.4), $F(1,65)=33.82$, $\eta_p^2=.31$, $p<.0001$. In addition, more letters were missed in second position (M=9.86%, sd=14.6) than in first position (4.47%, sd=8.8), $F(1,65)=23.33$, $\eta_p^2=.28$. $p<.0001$. There was also a significant interaction between these factors, $F(1,65)=16.09$,$\eta_p^2=.20$, $p<.0005$,because, as shown in Figure 1, the effect of syntactic role was larger when ل was in second position than when it was in first position.

In order to test the structural model for each letter position separately, we computed planned comparisons for each position. In the first letter position, the letter لwas missed more often when it was a function prefix (M=5.9%, sd=9.9) than when it was the first letter in a content word (M=3.03%, sd=7.2), $F(1,65)=7.18$,$\eta_p^2=.10$, $p<.01$. In the second letter position, the letter ل was missed significantly more often when it was in a function word (M=15.6%, sd=18.2) than when it was in a content word (M=4.11%, sd=5.5), $F(1,65)=31.41$,$\eta_p^2=.33$, $p<.0001$. Thus, syntactic role affected the miss rates in both positions, but the effect was much larger in second than in first position.

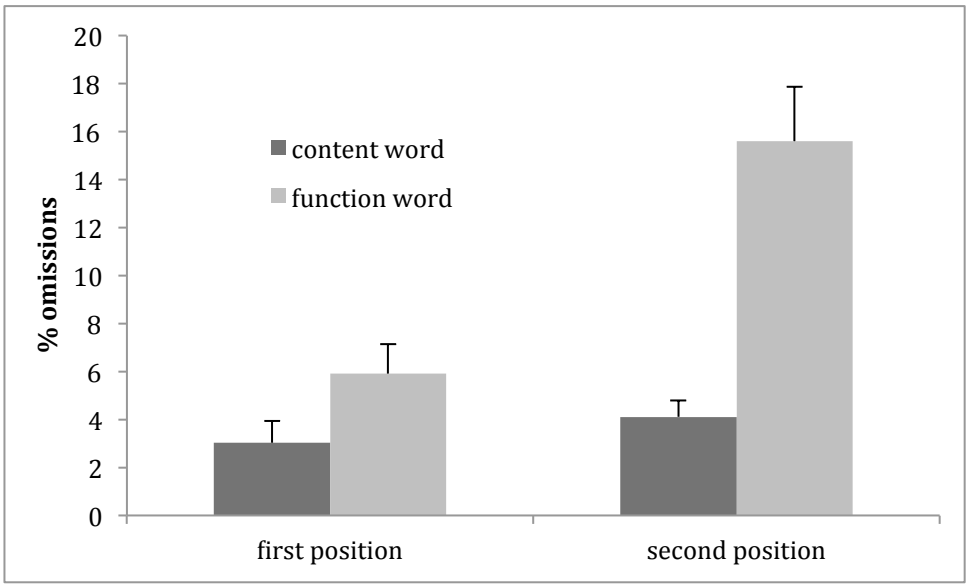

Figure 1: % Miss rates for the letter ل in the first position as a function prefix and in a content word and in the second position of the function word ألى or a content word. Error bars are standard errors.

Experiment 2: The Effects of Phonology

This experiment utilized one independent variable with three levels. We compared the rates of misses of the letter ل when it was part of the determiner أل (‘al’ - the) before a moon letter (where it would have been pronounced, had it been read out loud) versus when it appeared before a sun letter (where it would have been silent, had it been read out loud). These were compared to the rates of missing the letter لwhen it was part of a content word that began with the combination ‘al’ (a pseudo-prefix).These three conditions were compared using a 1-way GLM procedure, which revealed a significant main effect of condition, $F(2,130)=18.99$, $\eta_p^2=.23$, $p<.0001$. Planned comparisons revealed that overall, there was an effect of syntactic role: the letter ل was missed more often in the context of the function prefix ‘al’ (M=12.47%, sd=14.3) than when it was part of a content word (M=2.3%, sd=4.4), $F(1,130)=25.43$,$\eta_p^2=.16$,$p<.001$. This difference was also significant when ‘al’ in the context of a content word was compared to the condition in which the ل was pronounced (before moon letters, M=4.76%, sd=7.8), $F(1,130)=6.74$, $\eta_p^2= .05$, $p<.05$, and the effect

was larger when it was not pronounced (before sun letters, M=8.0%, sd=8.4), F(1,130)=37.67, η_p^2= .22, p<.0001. When the effect of phonology was tested directly, by comparing the rate of misses in the context of being pronounced vs. silent, this difference was also significant, F(1,130)=12.54, η_p^2= .09, p<.001. These patterns are seen in Figure 2. It can be seen that whether a letter is pronounced or not, has a strong effect on the probability that the letter will be missed.

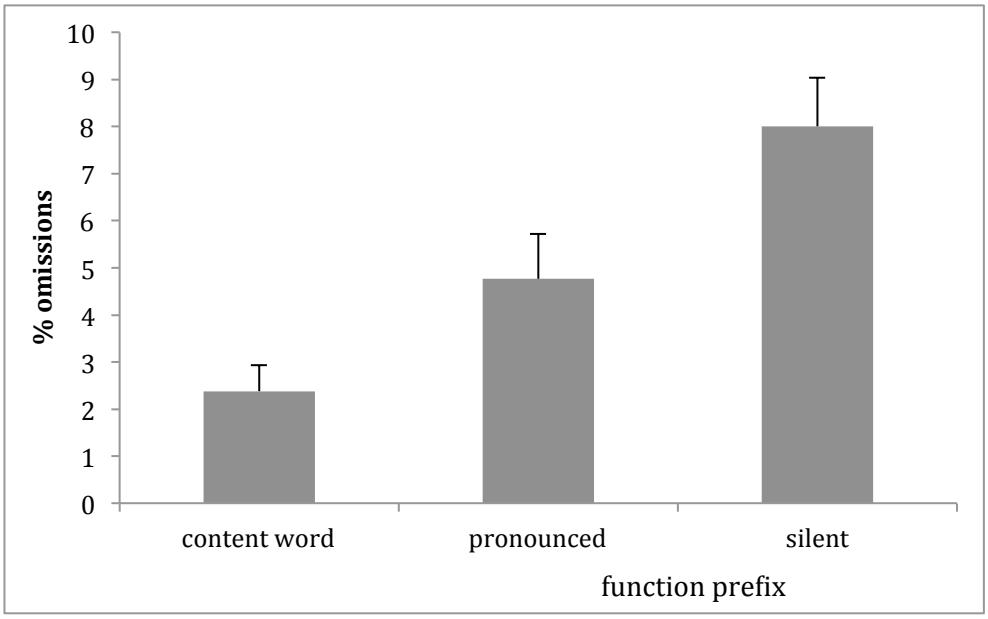

Figure 2: Effects of phonology: % Miss rates for the letter ل when it appears in the second position in a content word or in the determiner ألـ before a moon letter (where it would be pronounced) or before a sun letter (where it would be silent). Error bars are standard errors.

Experiment 3: The Effects of Letter Identity, Position and Context

Perusal of Figure 1 reveals that the rate of misses of the letter ل when it was in the first position in the word was smaller than when it was in the second position. In order to explore the generalization of this, we used

these data, and the responses from Booklet3, in which the target letter was أ, the first letter in the function word ألى. We also examined the effect of position in content words, and we compared the rates of misses of أ in content words in which the second letter is ل, thus creating a pseudo-prefix 'al', versus when the second letter did not create a pseudo-prefix.

a. Letter Identity and Syntactic Role

The first analysis used a 2x2 ANOVA with letter identity (ل vs. أ) and syntactic role (content vs. function) as within-subject independent factors. In all of the stimuli, the target letters were in initial position in the words. The analysis revealed that overall, there was a significant main effect of syntactic function, where letters in function words (M=3.49%, sd=7.7) were missed more often than letters in content words (M=1.52%, sd=5.3), $F(1,65)=13.00$, $\eta_p^2=.20$, $p<.001$. In addition, and as can be seen in the top panel of Figure 3, the letter ل (M=4.47%, sd=8.8) was missed more often than the letter أ(M=0.54%, sd=2.3),$F(1,65)=18.58$,$\eta_p^2=.22$, $p<.001$. However, the interaction between these factors is not significant, $p>.13$, suggesting that the same process is occurring for both letters.

b. Letter Identity and Position

The second analysis examined the effect of letter position in the word for the two target letters. Thus, we used a 2x2 ANOVA with letter identity (ل vs. أ) and letter position (initial or second) as independent factors. All of the letters appeared in content words, so that syntactic role was not manipulated. Analysis of the rates of misses in content words, in which the target letter was in first or second position revealed that there is a significant effect of letter position, $F(1,65)=12.5$,$\eta_p^2= .16$, $p<.001$, where letters were missed more often in second position (M=3.32%, sd=5.2) than in first position (M=1.52%, sd=5.3). In addition, here too, the letter ل (M=3.57%, sd=6.4) was missed more often than the letter أ (M=1.26%, sd=3.6), $F(1,65)=13.69$,$\eta_p^2= .17$, $p<.0005$. However, the interaction between these factors is not significant, $p>.24$, suggesting that the same process is occurring for both letters. This pattern is seen in the lower panel of Figure 3.

c. Effects of orthographic context

Finally, we examined the effect of context on rates of misses in content words, when the letter أ was in the context of a pseudo-prefix 'al' versus when it was not in such a context. A within-subjects t-test revealed that the difference between the rate that participants missed the letter أ when it was in the context of a pseudo-prefix (M=0.30%, sd=.014) was marginally more than when it was not (M=0), $t(65)=1.76$,$p=.083$.

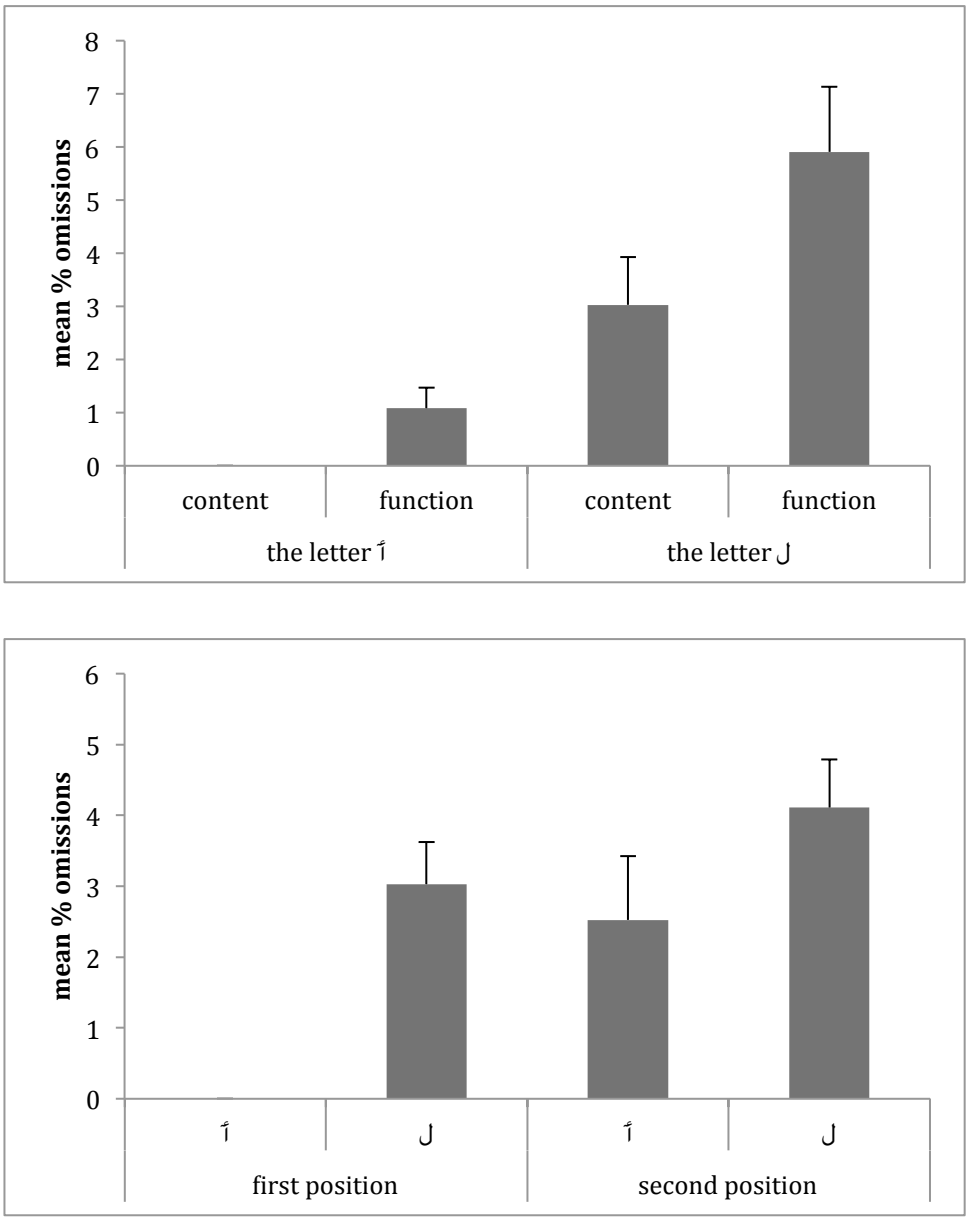

Figure 3: ***Top panel*****: % Miss rates for the letters أ and ل when they are in first position in content words versus when أ is the first letter in a function word (ألى – to), or ل is in a position of a function prefix. Error bars are standard errors.**

Bottom panel**: % Miss rates for the letter أ in content words that begin with a pseudo-prefix (the first two letters are ألـ) versus when the second letter is not ل and the word does not contain a pseudo-prefix. Error bars are standard errors.**

DISCUSSION

The results of the present study replicate the finding that letters are missed more often in function words and prefixes than in content words, in a third, non-Latin orthography. As the first experiment clearly shows, the letter ل is missed significantly more often when it is in a function word or is a function prefix, than when it is in the same positions in content words. These findings support the assumption of the structural and GO models: that the syntactic role of the word in which the target letter occurs affects the degree to which it is detected in a letter search task. The second and third experiments reveal that letter position, phonological manifestation, and shape, also affect the rate of letter misses. These findings support the position of the unitization model, showing effects of the physical and phonological characteristics of the words on letter detection. Thus, our results show both pre-lexical and post-lexical effects on letter detection. Our results do not speak to the recent suggestion of Roy-Charland, Saint-Aubin, Laurence, & Klein (2009) who suggested that the assumption of all the models, that letter processing is truncated when higher levels have identified the word, is wrong. We remain agnostic on this aspect of the models.

Another important goal of the experiment was to test the hypothesis that phonology is a mandatory stage in silent reading of Arabic, and to see if it affected the rates of letter misses. Although participants were reading the texts silently, they missed the letter ل significantly more often when it was in an orthographic context that excluded an acoustic manifestation of the sound (before sun letters) than when it would have been pronounced (before moon letters), with all other aspects of the sentences held constant. Our findings support the hypothesis that phonological recoding is a mandatory stage in skilled reading in Arabic.

Another specific goal of our study was to test the effects of orthographic (letter position) and visual characteristics (gestalt 'goodness') of the letters on miss rates in Arabic. We found that both of these factors affected miss rates: participants missed letters in the first position less often than in the second position, in both function and content words. This replicates the results reported by Ibrahim (2011) for Arabic, by Goldman and Healy (1985) for English, and by Tao and Healy (2002) for Chinese. These are three very different writing systems that differ in many dimensions. The first is the manner in which speech is represented in the orthography, where English is alphabetic, such that all sounds are graphically (albeit opaquely) represented, Arabic is an *abjad*, where the letters represent consonants and vowels are mostly missing, and Chinese is

a logography, where the phonology is not directly represented by the characters at all. A second dimension, which may be even more critical, is that these three orthographies differ greatly in degree and type of visual complexity. Both English and Arabic texts present words as separate units: in English separate letters are grouped into words, with larger spaces between words; in Arabic, most words are constructed of ligatured (connected) letters, and again, there are larger spaces between words. In Chinese, on the other hand, characters represent morphemes, and most words are represented by two characters. However, in text, characters are presented one by one, with equal amount of space between all. Thus, words are not visually distinguished on the page at all. The fact that position effects are found in these three very different writing systems, suggests a linear aspect to the process of recognizing words that is independent of the writing system. This suggests that the processing of graphemes as visual objects is similar in very different orthographies. Thus, although languages have very different ways of representing speech, the manner in which the orthography is processed may have universal aspects.

We also found that أ was missed less often than ل in both first and second position, and suggest that this is a result of the visual characteristics of the letters, with أ having better 'Gestalt goodness', being recognized faster, and thus having a longer processing time. أ is one of the 6 letters in Arabic that does not connect to other letters that follow it. When it is in first position, it stands out easily by not being connected to any other letters, as in أليافاً بنسبة, whereas when the letter ل is in first position, it is connected to the letter that follows it, thus merging with the main body of the word: لميس. When the two letters appear in second position, أ also stands out because it causes a gap in the ligatured word: as in سائلاً. Whereas the target letter ل is connected to both preceding and following letters (ملحاً), and may thus take longer to identify. Given that differential miss rates for different letters within a writing system have been found in all of the writing systems tested, it may be reasonable to assume that this, also, is a general characteristic of reading.

Roy-Charland and colleagues (Roy-Charland, Saint-Aubin, Klein, & Lawrence, 2007) have tested the predictions of the GO model with eye-tracking measures and latencies. They report that overall, the GO model is supported by the omissions data from these different paradigms, and also by gaze durations, such that words in which letters that were missed had shorter gaze durations than words in which the target letters were detected. These authors have suggested (Roy-Charland, Saint-Aubin,Lawrence, &Klein 2009) that comparisons of the gaze durations on words when the letter was missed, versus when the word was read, without the detection

task, result in a contradiction of the hypothesis that omissions result from the truncation of processing in highly frequent and function words. They suggest that the longer durations on words in which the target letter was detected than on words in which it was missed, are due to processes involved in the production of the response, not to truncation of processing. Our findings do not speak to this issue. In fact, our findings support elements of all of the models proposed to account for the MLE. Thus, consistent with the unitization model, our findings show effects the visual and phonological characteristics of letters; and consistent with the structural model, our findings show that the syntactic role of the words in which the letters occur are critical to the rates of missing them.

In Chinese, it has been shown that frequency on the level of the components of the characters, of the characters themselves, and of compounds that represent words, all affect character detection (Tao & Healy, 2002). In addition, character position within the words also affects detection, with skilled readers of Chinese having lower rates of omissions when the target character was in first position rather than in final position in a compound character. We have not found studies directly testing the effects of phonology on character detection in Chinese, although Seidenberg (1985) reported that phonology affects naming latencies of low frequency words in Chinese, but not of high frequency words. Our findings with Arabic replicate previous findings with other languages in terms of the effects of syntactic function, phonology, letter position, and the visual characteristics of letters, on letter detection. Thus, the MLE has now been shown in readers of orthographies based on the Latin alphabet (English, French, German), the Chinese logography, and Hebrew and Arabic, which are *abjads*. This finding is important, as it suggests that the underlying source of the MLE is manifested in very different orthographic systems that represent the sounds of speech in very different ways. Thus, although languages are different from each other, and writing systems are even more different from each other, similar processes seem to be occurring in human minds during reading.

REFERENCES

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.

Asaad, H. & Eviatar, Z. (2013).The effects of orthographic complexity and diglossia on letter identification (naming or retrieval) in Arabic: Developmental changes. *Writing Systems Research, 5 (2)*, 156-168.

Asaad, H. & Eviatar, Z. (2014).Learning to read in Arabic: The long and winding road.. *Reading and Writing*. In Press.

Azzam, R. (1984). Orthography and reading of the Arabic languge. In J. Aaron & R. M. Joshi (Eds.), *Reading and writing disorders in different orthographic systems* (pp. 1–29). Kluwer Academic.

Bentin, S., & Ibrahim, R. (1996). New evidence for phonological processing during visual word recognition: The case of Arabic. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 309-323.

Chomsky, N.(2002). *Syntactic Structures*. Mouton de Gruyter: Berlin

Corcoran, D.W.J. An acoustic factor in letter cancellation. *Nature, 214,* 685.

Daniels, P.T. (1990). Fundamentals of grammatology. *Journal of the American Oriental Society, 110* (4), 727-731

Evans, N.& Levinson S.C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Science, 32*, 429-492.

Eviatar, Z., Ibrahim, R., & Ganayim, D. (2004). Orthography and the hemispheres: Visual and linguistic aspects of letter processing. *Neuropsychology, 18(1),* 174-184.

Eviatar, Z., & Ibrahim, R. (2000). Bilingual is as bilingual does: Metalinguistic abilities of Arabic-speaking children. *Applied Psycholinguistics, 21(4),* 451-471.

Eviatar, Z., & Ibrahim, R. (2004). Morphological and orthographic effects on hemispheric processing of nonwords: A cross-linguistic comparison. *Reading and Writing*, *17,* 691-705.

Greenberg, S.N.,Healy,A.F., Koriat, A., &Kriener , H. (2004). The GO model: A reconsideration of therole of structural units in guiding andorganizing text on line. *Psychonomic Bulletin & Review, 11 (3),* 428–433.

Healy, A. F.(1976). Detection errors on the word THE: Evidence for reading units larger than letters. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 235-242.

Healy, A. F., Drewnowski, A. (1983). Investigating the boundaries of reading units: Letter detection in misspelled words. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 413-426.

Healy, A. F., Oliver, W. L., & McNamara, T. P. (1987).Detecting letters in continuoustext: Effects of display size.*Journal of Experimental Psychology: Human Perception and Performance*, *13*, 279-290.

Ibrahim, R. (2011). How does dissociation between written and oral forms affect reading: An investigation of auxiliary verbs in Arabic. *Journal of Research in Reading*, *34(2),* 247–262.

Ibrahim, R., Eviatar, Z., & Aharon-Peretz, J. (2002). The characteristics of Arabic orthography slows its processing. *Neuropsychology*, *16(3)*, 322-326.

Koriat, A., Greenberg, S. N., &Goldshmid, M. (1991). The missing-letter effect in Hebrew: Word frequency or word function? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 66-80.

Koriat, A., & Greenberg, S. N. (1994). The extraction of phrase structure during reading: Evidence from letter detection errors. *Psychonomic Bulletin & Review, 1(3),* 345-356.

Müsseler, J., Nißlein, M., & Koriat, A. (2000).Letter-detection patterns in German: A window to the early extractionof sentential structure during reading. *Memory & Cognition, 28(6),* 993-1003.

Müsseler, J., Nißlein, M., & Koriat, A. (2005).German capitalization of nouns and the detection of letters in continuous text. *Canadian Journal of Experimental Psychology*, *59(3),* 143-158.

Prunet, J.F., Be´land, R., &Idrissi, A. (2000) The mental representation of Semitic words. *Linguistic Inquiry*, *31,* 609–648.

Roy-Charland, A., Saint-Aubin, J., Lawrence, M. A. et Klein, R. M. (2009). Solving the chicken and egg problem of letter detection and fixation duration in reading. *Attention, Perception, & Psychophysics*, *71,* 1553-1562.

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 words and pseudowords in a diglossic context. *Applied Psycholinguistics, 25*, 495-512.

Saint-Aubin, J., & Poirier, M. (1997).the influences of word function in the missing letter effect: Further evidence from French. *Memory & Cognition, 25(5)*, 666-676.

Seidenberg, M. (1985).The time course of phonological code activationin two writing systems. *Cognition*, *19*, 1-30.

Schneider, V. I., & Healy, A. F. (1993).Detecting phonemes and letters in text: Interactions between different types and levels of processes. *Memory & Cognition*, *21*, 739-751.

Simon, G., Bernard, C., Lalonde, R., &Rebai, M. (2006). Orthographic transparency and grapheme- phoneme conversion: An ERP study in Arabic and French readers. *Brain Research, 1104,* 141-152.

Tao, L., & Healy, A.F. (2002).The unitization effect in reading Chinese and English text. *Scientific Studies of Reading*, *6*, 167–197.

Taouk, M. &Coltheart, M. (2004). The cognitive process involved in learning to read in Arabic. *Reading and Writing: An Interdisciplinary Journal, 17,* 27-57.

Tao, L.& Healy, A. F.(2002). The unitization effect in reading Chinese and English text. *Scientific Studies of Reading*, *6(2),* 167–197.

(Manuscript received: 27 March 2014; accepted: 21 November 2014)