Does the visual attention span play a role in reading in Arabic?

Marie Lallier\textsuperscript{1,\*}, Reem Abu Mallouh\textsuperscript{2}, Ahmed M. Mohammed\textsuperscript{1}, Batoul Khalifa\textsuperscript{3}, Manuel Perea\textsuperscript{3}, and Manuel Carreiras\textsuperscript{1,4}

\textsuperscript{1} Basque Center on Cognition, Brain, and Language, Donostia, Spain,

\textsuperscript{2} Department of Methodology, Universitat de València, Valencia, Spain,

\textsuperscript{3} Psychological Sciences Department, Qatar University, Doha, Qatar,

\textsuperscript{4} Ikerbasque Basque Foundation for Science, Bilbao, Spain

\textit{*Correspondence:}

Marie LALLIER, PhD

Basque Center on Cognition Brain and Language

Paseo Mikeletegi 69, Planta, 2

20009 San Sebastián - Donostia, Spain

E-mail: \texttt{m.lallier@bcbl.eu}
Abstract

It is unclear whether the association between the visual attention span (VA span) and reading differs across languages. Here we studied this relationship in Arabic, where the use of specific reading strategies depends on the amount of diacritics on words: reading vowelized and non-vowelized Arabic scripts favor sublexical and lexical strategies, respectively. We hypothesized that the size of the VA span and its association to reading would differ depending on individual “script preferences”. We compared children who were more proficient in reading fully vowelized Arabic than non-vowelized Arabic (VOW), to children for whom the opposite was true (NOVOW). NOVOW children showed a crowding effect in the VA span task, whereas VOW children did not. Moreover, the crowding in the VA span task correlated with the reading performance in the NOVOW group only. These results are discussed in light of individual differences on the use of reading strategies in Arabic.

Keywords: Visual Attention Span, Arabic, Reading strategies, Morphology, Crowding
Introduction

Most cross-linguistic studies of reading have endeavored to pinpoint the contribution of orthographic depth (the complexity and predictability of grapheme-to-phoneme conversions) to literacy development (Lallier & Carreiras, 2017 for a review). These studies show that orthographic depth modulates the use of reading strategies and the corresponding underlying cognitive skills. For example, the regular letter-sound correspondences of shallow orthographies favor the use of sublexical strategies and the development of phonemic awareness (Ziegler & Goswami, 2005), whereas reading in deep orthographies boosts lexical strategies (Frost, Katz & Bentin, 1987) and favors a wider distribution of visual attention over letter strings (Lallier & Carreiras, 2017). The role of visual attentional processes for reading development (see Gori & Facoetti for a review) has only recently been approached from a cross-linguistic perspective (e.g., Lallier, Molinaro, Liazarzu, Bourguignon, & Carreiras, 2017). The present study focuses on the cross-linguistic modulations which affect the visual attention span (VA span) and its association with reading.

The VA span corresponds to the number of visual elements processed simultaneously in a multi-element array (Bosse, Tainturier, & Valdois, 2007) and plays an important role in reading acquisition through the build-up of orthographic knowledge (Bosse, Chaves, Largy & Valdois, 2015). Cross-linguistic studies suggest that both the size and the shape of VA span are affected by orthographic depth (Awadh, Phénix, Antzaka, Lallier, Carreiras, & Valdois; 2016; Lallier, Acha, & Carreiras, 2016). In addition, Awadh et al (2016)’s study in adults showed that there was a significant relationship between VA span and reading skills in a deep orthography (French), but not in two shallow orthographies (Spanish and Arabic). This result in shallow orthographies is at odds with similar studies carried out on children (e.g., Lallier, Valdois, Lassus-
This suggests that a wide distribution of attention might no longer be critical for expert reading in shallow languages. If age-related factors can explain differences between French and Spanish in Awadh et al’s study, additional variables have to be considered in the case of Arabic.

Arabic is unique since two oral forms - standard and colloquial Arabic - and two scripts – vowelized and non-vowelized - co-exist. The two scripts vary in the size and type of the orthographic chunks that have to be attended, and differ in the amount of vowels (diacritics) provided in words. Arabic scripts therefore vary from fully vowelized (each letter marked with a diacritic), to fully non-vowelized (no diacritics). Importantly, the pronunciation and meaning of words, which are derived from 3- or 4-consonant root morphemes, depend on their vowel structure. In the non-vowelized script, readers rely on fast root morpheme recognition, and word pronunciation and meaning are deduced from contextual information only (e.g., Abu-Rabia, 2007). In contrast, when reading vowelized Arabic, to decode and access the meaning of words, one can rely on the additional help from vowels using letter-by-letter decoding strategies (Weiss, Katzir, & Bitan, 2016).

At the beginning of reading acquisition for children in Qatar, diacritics are used to foster decoding development. Around Grade 3, consonantal roots must be memorized as children are mostly presented with non-vowelized texts. Lastly, a non-frequent fully vowelized script, used for poetry and religious texts, must be mastered from Grade 4 onwards. For expert readers of the non-vowelized script, fully vowelized Arabic is composed of non-familiar orthographic forms (Weiss, Katzir, & Bitan, 2015) for which letter-by-letter recoding strategies have to be applied (see Weiss et al., 2016) to prevent the automatic access to root morphemes from interfering with reading.
Therefore, fully vowelized and non-vowelized scripts are read with distinct strategies (Frost & Bentin, 1992), which might also be associated with distinct visual attention distribution modes. Whereas fully-vowelized Arabic should require narrow visual attentional captures on each letter to access all the vowel information (Weiss et al., 2016), efficient non-vowelized script reading may require the homogenous distribution of attentional resources across the words to promote automatic access to root morphemes as wholes (Frost & Bentin, 1992; Frost, et al, 1987; Katz & Frost, 1992). Although this arguably suggests that a large VA span should be critical for non-vowelized reading, Awadh et al (2016) could not report any relationship between the VA span and non-vowelized text reading skills. We suggest that a feature specific to Arabic contributed to this result.

Given that Arabic readers master both fully-vowelized and non-vowelized reading from Grade 4 onwards, the VA span-reading relationship might be different between readers with distinct preferred reading strategy: one may be better at reading in one script compared to the other, and favor this “preferred” strategy when reading in either script. Thus, the VA span-reading relationship should be visible in a subgroup of readers who are more expert in a script where lexical strategies and a wide distribution of attention over letter strings contribute to fluent reading (i.e., non-vowelized script). On the other hand, the reading performance of Arabic speakers whose preferred script requires letter-by-letter decoding (i.e., fully vowelized script) should not depend as much on the number of visual elements processed under one attentional capture.

Here, we examined the VA span and reading skills of Qatari children who are in Grade 4. If script and reading strategy preferences are associated with specific visual attention distribution modes, children who have better reading skills to deal with the non-vowelized script than the fully vowelized one should exhibit a more homogeneous
distribution of attention over letter strings compared to the other group of children. In addition, we predicted that there would potentially be a stronger VA span-reading relationship for those readers who are better at reading non-vowelized Arabic than fully vowelized Arabic.

Material and Methods

Participants
Fifty-nine Grade 4 native Arabic speaker children from Al-Bayan Independent School in Doha, Qatar, participated in this experiment (right-handed females; mean age = 127 months ± 4). We focused on this age group because children are supposed to master both non-vowelized and vowelized reading at this age. All of the participants had normal or corrected-to-normal vision and none were reported to have reading difficulties or developmental disorders. The Qatar Foundation ethical committee approved the experiment. The written consent of each child’s legal tutor was obtained.

Task battery.

Non-verbal IQ. Non-verbal IQ skills were measured with the Egyptian version of WISC non-verbal IQ subtests (Ismael, & Maleka, 1993) including the picture completion, picture arrangement, block design, object assembly, digit-coding symbol, and mazes subtests. The standard score was calculated.

Fully vowelized and non-vowelized text reading. Two Arabic texts were created. For both texts, a fully vowelized version and a non-vowelized version were created, and presented to participants on a different sheet of paper. The four texts (text 1 vowelized, text 1 non-vowelized, text 2 vowelized, text 2 non-vowelized) were
administered in two sessions spread over two weeks. In the first session, children read the non-vowelized version of text 1 and the vowelized version of text 2. One week later, children read the non-vowelized version of text 2 and the vowelized version of text 1. They were instructed to read the text out loud as well as possible and they knew that they were being timed. Participants were given a maximum of six minutes to read each text. Both the time taken to read the text and the number of words correctly read were recorded. The number of correct words read per minute was computed in order to obtain comparable outcome measures for the four texts. For each participant, the mean z-score obtained from the vowelized texts was subtracted from the mean z-score obtained from the non-vowelized texts. Participants scoring above the median were assigned to the NOVOW group (i.e., children with better non-vowelized than vowelized reading skills) and those scoring below the median were assigned to the VOW group (i.e., children with better vowelized reading skills than non-vowelized reading skills; see Figure 1). Table 1 presents further information about the groups’ characteristics and reading performance.
Figure 1. Caterpillar plot depicting the distribution of the difference scores obtained from the reading performance on the non-vowelized and the vowelized texts. White dots represent children in the VOW group (n=30) and black dots represent children in the NOVOW group (n=29).

Table 1. Characteristics of the two groups.

<table>
<thead>
<tr>
<th></th>
<th>NOVOW (n=29)</th>
<th>VOW(n=30)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M(SD)</td>
<td>Range</td>
</tr>
<tr>
<td>Age (months)</td>
<td>126.6(3.5)</td>
<td>121-131</td>
</tr>
<tr>
<td>IQ (standard score)</td>
<td>103.1(11.3)</td>
<td>80-118</td>
</tr>
<tr>
<td>Arabic script reading (cwpm)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-vowelized</td>
<td>76.2(19.0)</td>
<td>47-126</td>
</tr>
<tr>
<td>Fully vowelized</td>
<td>48.8(12.5)</td>
<td>27-78</td>
</tr>
<tr>
<td>Difference score</td>
<td>27.4(9.9)</td>
<td>14-48</td>
</tr>
</tbody>
</table>

Cwpm: correct words per minute; Difference score= (Non-vowelized)-(Fully vowelized)

VA span. A visual 1-back paradigm was used. Stimuli were created using 13 Arabic consonants in their isolated form (ص/ر/ي/ل/ن/م/ط/ك/ع/ف/ب). The consonant strings did not include clusters corresponding to root morphemes and consonants could only be included once in the same string. The task included 104 5-consonant strings presented on a white screen in black upper-case Arial font. Children were seated 70cm away from the screen so that stimulus width was between 5.3° and 5.55° of visual angle and the center-to-center distance between each adjacent letter was 1.2°. At the start of each trial, a central fixation point was displayed for 1000ms followed by the centered consonant string during 200ms. Then, a white screen lasting 100ms was presented followed by a consonant (target) appearing below the median
horizontal line. Target consonants were presented in red with a bold-italic font. Children were instructed to respond as fast as possible by pressing the ‘M’ key when the target was present in the previous consonant string and the ‘Z’ key when it was absent. The target disappeared after the child’s response, and a screen with a question mark was presented until the experimenter decided to initiate the next trial. The 104 trials included 65 trials in which the target was present in the consonant string (each consonant presented five times as target, once at each position) and 39 trials in which the target was absent (each consonant presented three times as target). The task was preceded by 5 practice trials. Trial order was randomized. For each position, target detection accuracy (%) and sensitivity (d’) were computed.

Data Analyses
All analyses included non-verbal IQ and age as co-variates. In order to quantify group differences on reading skills, an ANCOVA on text reading z-scores was conducted with Group (NOVOW, VOW) as the between-subject factor and Script (vowelized, non-vowelized) as the within-subject factor. Then, ANCOVAs on the mean d prime (d’) and percentage scores from the VA span task were conducted with Group (NOVOW, VOW) as the between-subject factor and Target Position (1, 2, 3, 4, 5) as the within-subject factor. Post hoc tests were performed using Bonferroni corrections. Lastly, partial correlations within each group were performed between VA span and reading z-scores for the vowelized and the non-vowelized texts separately.

Results
Vowelized and non-vowelized text reading
Reading scores are presented in Table 1. There was a Group by Script interaction ($F(1,55) = 81, p < 0.001, \eta^2 = 0.60$). Post hoc exploration showed that the VOW group was better at reading the vowelized than the non-vowelized texts, and the NOWOV
group was better when reading the non-vowelized than the vowelized texts ($p < 0.001$). The two groups had similar reading performance on the fully vowelized texts ($p > 0.05$), whereas the NOVOW groups was better at reading the non-vowelized texts than the VOW group ($p = 0.01$).

**Figure 2.** Reading performance on the two scripts in the two groups of participants. Vertical bars denote 95% confidence intervals corrected with Cousineau (2005)’s method.

**VA span**

The ANCOVA on d prime values did not reveal any main or interaction effect. However, a significant interaction between Target Position and Group was found on percentage scores ($F(4,220) = 3.07, p = 0.017, \eta^2 = 0.05$; Figure 2). Post hoc tests did not show any group difference at any position (all $ps > 0.50$). In addition, the VOW group had similar performance across positions ($ps > 0.50$), whereas the NOVOW group
exhibited better performance at position 5 than positions 2 and 4 ($p < 0.001, p = 0.04$, respectively) reflecting a significant serial position function (Grainger, Dufau & Ziegler, 2016) henceforth referred to as “crowding” effect although we are aware that it also reflects acuity effects. An ANCOVA controlling for overall sensitivity (mean d’ across positions) resulted in similar results.

**Figure 3.** Mean percentage scores obtained on the VA span task for each group of participants. The horizontal dashed line indicates the 0.50 chance level. Vertical bars denote 95% confidence intervals corrected with Cousineau (2005)’s method.

**Correlations between VA span and reading in Arabic**

Crowding effects in reading refer to perceptual difficulties at identifying letters within words, that stem from the interference produced by lateral masking between adjacent letters. In a 5-letter word, letters in position 2 and 4, which are surrounded by letters in both sides, will be the most affected by crowding.
There was no significant link between overall sensitivity (d’) on the VA span task and text reading scores within the whole group and the VOW group (ps > 0.30). In the NOVOW group, the lower target detection sensitivity, the better the reading of both non-vowelized (r(29) = -0.43, p = 0.025) and fully vowelized texts (r(29) = -0.41, p = 0.03).

Follow-up analyses were also conducted to explore the “crowding” effect that differentiated group performance on the VA span task. Individual crowding scores were computed by subtracting the mean d prime over positions 2 and 4 from the mean d prime over positions 1, 3, and 5 (see the serial position function in VA span tasks, Ziegler, Pech-Georgel, Dufau, & Grainger, 2010), so that positive scores corresponded to large crowding effects. Small crowding effects were associated with low target detection sensitivity within the whole group (r(59) = 0.91, p < 0.001), the VOW group (r(30) = 0.95, p < 0.001) and the NOVOW group (r(29) = 0.87, p < 0.001). Lastly, smaller crowding effects were also linked to better reading skills in the NOVOW group only (non-vowelized text - NOVOW: r(29) = -0.37, p = 0.05; VOW: r(30) = 0.04, p > 0.50; whole group: r(59) = -0.06, p > 0.50; fully vowelized text – NOVOW: r(29) = -0.35, p = 0.07; VOW: r(30) = -0.07, p > 0.50; whole group: r(59) = -0.10, p > 0.50; see Figure 3). The correlation coefficients tended to differ between the groups for the vowelized text (p = 0.06; non-vowelized text: p = 0.14).
Figure 4. Scatter plots of the correlations between crowding effects on the VA span task and reading performance obtained on the fully vowelized (A.) and non-vowelized (B.) scripts, within each group. Reading z-scores are plotted on the x axis and crowding effects on the y axis.

Discussion

The present study aimed to explore the relationship between the VA span and Arabic reading skills, focusing on a novel factor: individual script preferences. To do so, the performance of children who were better at reading the non-vowelized texts than the fully vowelized texts (NOVOW group) was compared to the performance of children for whom the opposite was true (VOW group).
Whereas both groups performed similarly on the reading task in the fully-vowelized script, the NOVOW group was better at reading the non-vowelized text (see Figure 1). Since the quality of root morpheme representations may contribute to the ease at which non-vowelized script is read (Bar-on & Ravid, 2011; Saiegh-Haddad & Geva, 2008; Saiegh-Haddad & Henkin-Roitfarb, 2014), the poorer non-vowelized reading skills of the VOW group might stem from weaker internalization of these morphemes. In contrast, the two groups performed similarly on the fully vowelized text, suggesting that reading with diacritics may require skills other than morphological abilities alone, such as bottom-up grapheme-to-phoneme mappings (Weiss et al., 2015, 2016).

**Reading preferences and VA span distribution mode**

Since different strategies are associated with non-vowelized and fully vowelized Arabic scripts (lexical and sub-lexical, respectively), we first predicted that higher reading expertise in one script compared to the other would be associated with different VA span behaviors\(^2\). No group difference was found on the overall target detection sensitivity (d’) on the VA span task at any of the five positions within the consonant string. This is in line with data showing no quantitative VA span differences between children across languages (see Lallier et al., 2016). However, qualitative group differences linked to the distribution mode of VA span skills (spatial bias affecting target detection) emerged: whereas the NOVOW group showed a significant response bias to uncrowded consonant targets (first, third and fifth positions), the position did not affect the target detection in the VOW group\(^3\). Contrary to our predictions, children who preferred reading in a script requiring letter-by-letter strategies (VOW group)

\(^2\) Performance on our VA span task may not only engage visual attention, but also visual working memory or fine grain visual perception skills. Therefore, our results may be interpreted in the context of visual theories of reading other than the VA span hypothesis alone.

\(^3\) A visual inspection of Figure 2 shows that a similar but much reduced serial position function effect can be observed in the VOW group.
distributed their visual attention more homogeneously across stimuli than children who preferred reading in a script where multi-letter units have to be extracted (NOVOW group).

A closer look at the crowding (and serial position function) literature gives some hints to interpret this unpredicted finding. We know that the identification of crowded elements is facilitated by top-down orthographic knowledge (Grainger, Tygdat & Issele, 2010; Montani, Facoetti & Zorzi, 2015). We know from another line of work that top-down contextual mechanisms help us to resolve perceptual noise and ambiguity through predictive mechanisms (Ahissar & Hochstein, 2004; Guediche, Blumstein, Fiez, & Holt, 2014; Mattys, Davis, Bradlow & Scott, 2012; Panichello, Cheung, & Bar, 2012;). In our task, participants could barely rely on lexical or contextual feed-back to reduce perceptual masking between the five consonants which were presented simultaneously. Interestingly, non-vowelized Arabic reading demands a constant feed-back between root morpheme identification and top-down processes in order to access the meaning of words (Abu-Rabia, 1997, 2001). Therefore, the children in the NOVOW group, who rely on contextual information when reading, might have experienced the greatest difficulties at inhibiting perceptual noise in our task because of the absence of contextual help. Accordingly, the NOVOW group showed significant crowding effects, whereas the VOW group did not. It is noteworthy that two interpretations of these group differences are possible, implying opposite causal directions: (i) Reading preferences resulted in a specific VA span behavior: strongly relying on contextual top-down information whilst reading increased the impact of crowding on orthographic processing; (ii) Vulnerability to crowding resulted in specific reading preferences: visual noise exclusion deficits encouraged the alternative use of context for lexical and semantic access. Future studies are needed to differentiate between the two alternatives.
We hypothesized that the contribution of VA span skills to reading would be different between groups of readers with distinct reading preferences. Although it will be important to replicate these results in larger samples, they suggest that smaller crowding effects were associated with better reading in NOVOW readers only. The absence of a significant VA span-reading relationship in the VOW group may be an indicator of weaker non-vowelized reading skills (see Figure1) and weaker reliance on top-down processes whilst reading. Contrary to the NOVOW group, children in the VOW group might have resorted primarily to vocalization to mediate the access to morphemes and word meaning (Schiff, 2012), thus tending to weaken the association between crowding and reading in Arabic. Interestingly, the participants tested in Awadh et al (2016) were highly educated PhD students. Arabic-speaking individuals with this educational level are likely to be frequently exposed to religious and poetry texts and be fluent in fully-vowelized reading. We cannot rule out that the absence of VA span-reading relationship in Arabic in Awadh et al (2016) might partly result from their participants’ expertise in reading fully-vowelized Arabic. Altogether, these findings suggest that inter-individual variability regarding how much top-down knowledge is used during reading might affect the strength of the contribution of VA span skills to reading within and across languages (Awadh et al., 2016). Taking into account individual reading profiles as well as crowding effects in addition to overall sensitivity in VA span tasks should shed light on these “cross-linguistic” questions.

Conclusion

By studying the unique properties of Arabic, we showed that the dominant reading strategy of individuals (lexical/top-down over sublexical/bottom-up, and vice-versa) is
associated with different distribution modes of VA span skills and different VA span-reading relationship strengths. One interpretation of these results is that the relatively stronger reliance on top-down contextual information compared to letter-by-letter bottom-up strategies whilst reading increases the vulnerability to crowding and the strength of its association with reading. Future studies should thrive on developing experimental designs testing such causal assumption. Lastly, our findings suggest that some strengths and weaknesses of Arabic readers could be identified on the basis of their script preferences, which, in the long run, should contribute to designing tailored reading interventions for struggling readers.

**Acknowledgements**

This report was made possible by a NPRP award [Grant NPRP No. 6-378-5–035] from the Qatar National Research Fund (a member of The Qatar Foundation). The statements made herein are solely the responsibility of the authors.
References


