The Effect of Length on Word Recognition (in Reading): The Case of Arabic

Deia Ganayim and Shireen Ganayim

Abstract

In this study, two experiments were conducted to assess the role of word length in visual word recognition. In Experiment 1 two lists of Arabic three and five letter words were used in a print word-reading task which measured accuracy and reading time. In Experiment 2, three, four and five letter words were displayed in the center of fixation on a screen in a naming task measuring accuracy and naming time. In reading, two contrasting processes have been suggested: the holistic process and the analytical process. According to the holistic process, the recognition of a word is determined by its global features and configuration. Consequently, reading consists of the simultaneous processing of all the letters of a word in parallel. In contrast, according to the analytical process, reading is a sequential screening of all the letters within a particular word. Length effect—that is, short words are recognized more rapidly and accurately than long words—is the signature of analytical processing of the non-lexical route due to its seriality which is caused by assembled phonology. The results of both experiments revealed that the average reading time of Arabic words from paper and screen was affected by word length, reflecting certain analytical processes and the activation of a non-lexical route, in which letters are processed sequentially.

Keywords: Arabic, Word length, Reading, Visual Processing.

Disclaimer: The opinions and views expressed in the publications of Bethlehem University Journal are not those of the Journal’s editorial board and staff of the Dean of Research at Bethlehem University. The accuracy of the material and any errors in this publication are the sole responsibility of the respective authors and contributors.
الملخص

هدف هذا البحث إلى فحص العلاقة المتبادلة بين طول الكلمات والعوامل البصرية، مثل القالب العام للكلمة وأحرفها الخاصة وتأثيرها على القراءة في اللغة العربية. وتكونت عينة البحث من طلاب كلية لتأهيل المعلمين الذين يتمتعون بمستوى عالٍ من المهارة في القراءة إذ يتألفون الكلمات العربية والكلمات على اختلاف أطوالها. تعتمد القراءة على نوعين متعاكسين من المعالجة البصرية: المعالجة البصرية المتوازية لشكل الكلمة العام وأحرفها بالإضافة إلى المعالجة البصرية المتسلسلة المتراكمة لكل واحد من أحرفها. يعد تأثير طول الكلمة (الكلمات القصيرة تقرأ بشكل أدق وأسرع من الكلمات الطويلة) دليلاً على المعالجة التسلسلية المتراكمة لكل واحد من أحرف الكلمة. اعتمد هذا البحث على تجربتين تم في التجربة الأولى منها قياس دقة وزمن قراءة قائمنتي كلمات تضم كل واحدة منها 20 كلمة وحسب زمن قراءتها وفق طولها أي عدد أحرفها (3، 5 حرفًا). أما في التجربة الثانية فقد قيست دقة قراءة كلمات أطوال مختلفة (3 أو 4 أو 5 أحرف) تم عرضها على شاشة الحاسوب في مركز الرؤية. أشارت نتائج هاتين التجربتين إلى أن معدل زمن قراءة الكلمات عن ورقة وعن شاشة الحاسوب قد تأثر بطول الكلمة حتى عند المهررين، إذ أن معدل قراءة الكلمات القصيرة (3 أحرف) كان أقل من معدل قراءة الكلمات الطويلة (5 أحرف) مما يعكس معالجة غير قالبية للأحرف في الكلمة بل معالجتها بشكل تسلسلي.

كلمات مفتاحية: قراءة كلمات، العربية، معالجة بصريّة، طول الكلمة.

لا تعبير الأفكار الواردة في المخطوطة عن أفكار هيئة تحرير المجلة أو عمادة البحث العلمي في جامعة بيت لحم. يعتبر المؤلف المسؤول الوحيد عن مضمون المخطوطة أو أية أخطاء فيها.
Introduction

There is disagreement between psychologists and education professionals regarding the process of word recognition in reading. Two contrasting processes have been suggested: analytical and holistic (Balota, 1994; Lindell, Nicholls, Kwantes, & Castles, 2005). According to the supporters of the holistic process, the global features and the configuration of a word determine its recognition (Cattell, 1886). Consequently, reading consists of the simultaneous processing of all word letters in parallel (LaBerge & Samuel, 1974). In contrast, according to supporters of the analytical process, reading is a sequential screening of all the letters within a particular word (Bradshaw, Nettleton, & Taylor, 1981; Gough 1972; Whitney, 2001). If we think of the process of reading as a continuum, with analytical processing at one end and holistic processing at the other, the various extant word recognition models may be placed somewhere on this continuum (e.g., processing of supra-letter units) (Jordan, Patching, & Milner, 2000).

Reading acquisition techniques involve both of these identification processes (Berent & Shimron, 1997). Although the sequential-analytical process (letter-by-letter reading) is acquired first, the parallel-holistic process is acquired later as the basis for fast reading when readers are taught the grapheme-phoneme conversion rules (Gough, 1972; Leybaert & Content, 1995; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2002).

One of the most relevant reading models is the cascaded dual route model (Coltheart, 2005; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) that assumes the use of two processing routes during reading which may operate simultaneously and in parallel (Stuart, Masterson, Dixon, & Quinlan, 1999). The lexical route relies on activation of word form visual representation in the mental lexicon allowing the reading of familiar pre-stored words and raising the speed and accuracy of word recognition and spelling (Ehri, 2005; Frith, 1985; Vellutino, Fletcher, Snowling, & Scanlon, 2004). The non-lexical route relies on grapheme-phoneme conversion rules allowing the reading of unknown words (unfamiliar and novel words) and nonwords (or pseudowords) (Coltheart, 2005; Ehri & Snowling, 2005).
Problem Statement

The present study assesses the role of word length in visual word recognition. Assumes that instead of a sharp dichotomy between analytical processing (letter-based recognition, or the non-lexical route in the dual route model) and holistic processing (global word shape-based recognition, or the lexical route in the dual route model), a continuous processing spectrum exists with analytical attributes at one end, and holistic attributes at the other (Tanaka & Gauthier, 1997).

The effects of word length on visual word recognition have been examined through a variety of tasks (e.g., perceptual identification, lexical decision, naming, eye tracking), but the results have been inconsistent, ranging from inhibitory (longer words are harder) to null effects (New, Ferrand, Pallier, & Brysbaert, 2006). Previous developmental research, however, found no length effect in skilled readers, reflecting the effective connection of visual and lexical information (Aghababian & Nazir, 2000; Montant, Nazir, & Poncet, 1998; Nazir, 2000). Word length effect is considered evidence of sequential non-lexical processing (Eviatar & Zaidel, 1991; Iacoboni, & Zaidel, 1996; Reichle, Rayner, & Pollatsek, 2006), while the lack of word length effect by skilled readers suggests parallel lexical processing in word recognition (Aghababian & Nazir, 2000; LaBerge & Samuels, 1974).

Present study

The goal of the present study is to explore reading processes from different theoretical perspectives, and to assess the unique reading processes in Arabic, given its different word lengths. Most of the academic literature surrounding visual processing in reading has involved English speaking participants (Share, 2008). This has led to a restricted understanding of visual word recognition theory. Comparative research of different language families (e.g. Semitic and Germanic) is necessary in order to widen our knowledge of word recognition in the reading process. The present study is unique in exploring these effects from a theoretical perspective; the practical implications of such research, however, are essential not only to general reading programs, but also to interventions and treatment plans for individuals with reading disabilities (Lavidor, 2011).

In the present study, a word-reading task was used in two experiments to examine the reciprocal interaction between phonological encoding strategies and word length. Our participants comprised skilled Arabic readers familiar with different word lengths.
In order to study the effect of word length on word recognition in reading Arabic, three lists of Arabic words of varying word length were used. Accordingly, a phonological and/or orthographic encoding can be referred to in each word length.

**Hypotheses**

The present research is the first study of the effect of word length in the visual recognition of Arabic words. To do so, we employed different word lengths (3 vs. 5 letters) in a word reading task from paper (Experiment 1) and (3, 4, 5 letters) in a word reading task from central fixation display screen (Experiment 2). If word letters are processed in parallel in a lexical route, word length (letters numbers 3–5) will not affect its recognition. Otherwise, word length effect will be evidence of non-lexical processing.

The study question and hypothesis addressed the word length at word level in Arabic: Does the word length effect exist in Arabic even for skilled readers? We expected to find the length effect reflected in longer reading times for long words from paper (Experiment 1) and from central fixation display screen (Experiment 2), supporting a sequential processing theory of visual word recognition. We particularly expected to find this as evidence of non-lexical processing in reading even by skilled readers.

**Experiment 1**

Experiment 1 addressed the influence of the analytical processing reflected in the word length effect, on visual Arabic word recognition. The goal of this experiment was to explore the reading rate according to word length (3 vs. 5 letters). The word stimuli were divided into two lists according to their length. The participants were asked to read (recognize) all the words in each list.

**Method**

**Variables.** The independent variable was word length (3 vs. 5 letters). The dependent variables were the accuracy rate and the response times for each word list.

**Participants.** A total of 105 students from Sakhnin Teachers’ College participated in the study (average age: 25.3, SD = 6; 10 males and 95 females). All were right-handed, native Arabic speakers of a middle socio-economic status, and had normal or corrected-to-normal vision in both eyes. None had a history of neurological or emotional disorders. As college students, all participants were assumed to read at satisfactory skilled levels, and none had been formally diagnosed as having any reading impairments.
The stimuli consisted of two lists of 20 words, each of which was composed of 3 and 5 letters. The words were nouns selected from school textbooks (grades 1-12) and the dictionary, with a medium frequency of occurrence (2.5-3.5), as judged by 52 students from Sakhnin Teachers College on a 1-5 frequency scale (1 = very rare, 2 = rare, 3 = average, 4 = frequent, 5 = very frequent). The words were displayed in black 20-point Simplified Arabic Fix font on white A4 size paper.

Procedure. Each participant was tested individually, in a random presentation sequence of the two word lists. The participants were instructed to read the words at a fast rate, and the reading time for each list was measured.

Results
Because the accuracy percentage exceeded 99% for all conditions, an analysis of the accuracy percentage was not conducted. The differences between the two word lists according to word length (3 vs. 5 letters) were tested through the RM-ANOVA. We found a significant word length effect ($F (1,104) = 83.677, p < 0.005$), with an average reading time for three letter words (Mean=407, SD=103) being significantly shorter than that for five letter words (Mean=499, SD=151). (Figure 1).

![Figure 1 Average reading time per word as a function of word length (3, 5 letters). Note: Error bars represent standard error.](image-url)
Discussion
The findings of Experiment 1 revealed that the average reading times for Arabic words were affected by word length, since the average reading time was shorter for three letter words than for five letter words reflecting analytical processes and the activation of the non-lexical route, which processes letters sequentially in letter strings.

Experiment 2
Experiment 2 addressed the influence of word length in Arabic on visual word recognition. The goal of this experiment was to explore the effect of three-letter, four-letter and five-letter words on reading. A central fixation display on a PC screen paradigm was used to present words. The participants were given a naming task and were asked to read all the words in a naming task.

Method
Variables. The independent variable was word length (3, 4, 5 letters). The dependent variables were the accuracy rate and response time.

Participants. A total of 41 university students participated in the study (average age: 22.7, SD = 2; 20 males and 21 females). All were right-handed native Arabic speakers of middle socio-economic status who displayed normal or corrected-to-normal vision in both eyes. No participants had had a history of neurological or emotional disorders. Being university students, the participants were assumed to read at a skilled satisfactory level, and none had been formally diagnosed as having any reading impairments.

Stimuli. The stimuli were three lists of 20 three-letter, four-letter and five-letter words each. All words were presented in the center of fixation on a PC screen. The words were nouns selected from primary and secondary school textbooks, as well as those occurring in the dictionary with a medium frequency (2.5–3.5), as judged by 52 university on a 1–5 frequency scale (1 = very rare, 2 = rare, 3 = average, 4 = frequent, 5 = very frequent). The words were displayed in white 24-point Simplified Arabic Fix font on the black background of a PC screen. The words were introduced randomly.

Procedure. A CRT display (19 inch) was placed at a viewing distance of 60 cm from the participant. There were 60 trials. Each trial contained the following steps using super lab software:
• Two vertical fixation lines were presented in the middle of the screen for 300 MS.
• The word stimulus was presented for 150 MS between the lines with the central letter that was to be fixated on placed between the lines.
• The fixation lines remained on the screen until the voice key registered a response, or until a time-out of 1500 Ms was reached (Figure 2). A break was provided whenever the participant indicated that (s)he needed a break.

Figure 2. Time course of one trial in the optimal viewing position task.

The participants were informed that there would be an Arabic word between two vertical lines in the middle of the screen. It was stressed, explicitly and repeatedly, that it was important to fixate between the two lines, when these lines were presented. The participants were asked to name the words as quickly and as accurately as possible. They were told that they could ask for a break whenever they wanted.

Each participant was tested individually and with a different random presentation sequence of the words for each participant. The response time was a measurement of the time between the presentation of a word and the onset of a spoken response, where \textit{onset} is defined acoustically. The word noted /said/articulated by the participant was written by the experimenter.
Results

Because the accuracy percentage exceeded 99% for all conditions, an analysis of the accuracy percentage was not conducted. A RM-ANOVA analysis was conducted of the reading times according to word length (3, 4, 5 letters) as a within-subjects factor. A word length effect was found ($F(2,95) = 6.4, p < 0.05$), with reaction time increasing with increase in word length. (Figure 3).

![Figure 3 Average reading time per word as a function of word length (3, 4, 5 Letters). Note: Error bars represent standard error.](image)

Discussion

Similar to Experiment 1, the findings of Experiment 2 revealed that the average reading times for Arabic words were affected by word length, since the average reading time increased as the word length increased reflecting analytical processes and the activation of a non-lexical route, which processes letters in letter strings sequentially.
General discussion

The present study explored how word length affects the processing of print listed and on screen centrally presented Arabic words. Variations in recognition and processing time were measured as a function of word length using words of 3, 5 letters) in Experiment 1 and of 3, 4, 5 letters in Experiment 2 when participants were asked to perform a recognition task. The word length effect is consistent with the results of previous studies on English (Joseph, Liversedge, Blythe, White, & Rayner, 2009; Plummer & Rayner, 2012; Rayner, Slattery, Drieghe, & Liversedge, 2011). The word length effect has been considered by several researchers as evidence for a sequential process (Eviatar & Zaidel, 1991; Iacaboni, & Zaidel, 1996; Reichle, Rayner, & Pollatsek, 2006), while its absence in skilled readers has been taken to indicate parallel letter processing in word recognition (Aghababian & Nazir, 2000; LaBerge & Samuels, 1974). Its presence in recognition of Arabic words supports the analytic-sequential processing and the non-lexical route in reading even by skilled readers. Our findings provide strong evidence suggesting an inhibitory effect of word length in Arabic where longer words are harder to recognize.

The findings of the present experiments present evidence to support the hypothesis of a word recognition spectrum in Arabic visual word recognition, with sequential-analytical processing at one end (lexical route) and a parallel-holistic (non-lexical route) at the other. The inhibitory word length effect found in Arabic word recognition, however, must be taken into account because longer words are harder to recognize supporting non-lexical processing. A regression analysis revealed that Arabic word length predicted visual word recognition reaction time. Data on word length showed a negative linear relationship (b = -0.278, p < 0.05), and indicated that the longer reaction time was associated with longer words (r = -0.394). Thus, 15% of the variance in reaction time was shared with word length ($R^2 = 0.15$).

From a theoretical standpoint, the present study suggests that instead of a dichotomous distinction between analytical-sequential processing (Balota 1994; Lindell et al., 2005), or letter-based recognition, and holistic-parallel processing (Cattell 1886; LaBerge, & Samuels, 1974), or global word shape based recognition, a continuum exists of visual word recognition processes (Tanaka & Gauthier, 1997). This supports the assumption of lexical and non-lexical routes of the dual route model in reading. The three representational forms of words—phonological, orthographic, and semantic—capture different proportions of processing, and the balance between them may shift according to word length. Thus, in normal reading
development, the effect of word length is expected to have a notable effect on reading performance even by skilled readers in Arabic. This stems from utilizing efficient mappings between orthographic and phonological representations in the task. When word length is manipulated (e.g., when different word lengths are used, we can expect anomalies in orthographic and phonological representation to affect reading recognition (Lavidor, 2011). Although skilled readers depend on the parallel-holistic process as the basis for fast reading when they are taught the grapheme-phoneme conversion rules (Gough, 1972; Leybaert & Content, 1995; Rayner, Foorman, Perfetti, Pesetsky, & Seidenberg, 2002), it appears they cannot dispense with sequential-analytical process (i.e., letter-by-letter reading).

**Conclusion**

The findings of the present study have practical implications for both general reading instruction and reading disability-targeted interventions and treatment plans (Lavidor, 2011). In particular, the present results suggest that reading acquisition techniques should involve both reading processes (Berent & Shimron, 1997) even for skilled readers. However, the choice of short words by publishers of reading materials and software companies may be optimal for beginning readers. Additionally, diagnostic separate word reading tasks should comprise various word lengths in order to reflect reading performance.

**Study Limitations**

In the present study we used only three, four, five letter words. Despite the fact that most Arabic words are of the above word lengths (Belaid & Choisy, 2006), we note that the studied word lengths do not comprise the full range of Arabic word lengths. This may not allow the generalizability of our findings to other word lengths (such as two, six, seven, etc.) that would benefit from future research. Furthermore, to our knowledge, no systematic data exist on word form frequency in Arabic, so we cannot address any possible effects of word form frequency with our findings.
References


