Creative Education, 2022, 13, 3026-3052
https://www.scirp.org/journal/ce
ISSN Online: 2151-4771
ISSN Print: 2151-4755

The Reading Comprehension Process of Arabic Garden-Path Sentences: An Equivocal Effect

Abdullah M. Seraye Alseraye

Department of Curriculum & Instruction, School of Education, King Saud University, Riyadh, Saudi Arabia
Email: aalseraye@ksu.edu.sa

How to cite this paper: Alseraye, A. M. S. (2022). The Reading Comprehension Process of Arabic Garden-Path Sentences: An Equivocal Effect. *Creative Education, 13,* 3026-3052.
https://doi.org/10.4236/ce.2022.139191

Received: August 19, 2022
Accepted: September 26, 2022
Published: September 29, 2022

Abstract

The aims of this study were two-fold: first, to investigate the effect of a heterophonic homographic initial of a sentence (HP-HG) on the reading process of beginning Arabic readers as they read for comprehension; and second, to examine the role of the correct short vowels and diacritics in resolving garden-path (GP) ambiguity. Thirty-nine native monolingual Arab male fourth-grade students aged between 9 and 10, participated in reading 31 sentences representing three reading conditions, plain, vowelized-discretized, and wrongly vowelized, using a self-paced reading software program. The results showed that the GP structure of Arabic sentences did not affect the reading process of the beginning Arabic readers, nor was their process of reading such sentences significantly improved or shortened by the disambiguating short vowels and diacritics. However, examining the means descriptively shows that on average, the GP sentences took more time for the students to read but were comprehended better than their counterparts. Different proposed explanations were suggested, and a good-enough model is a good candidate to start with. One important factor that arose in all possible interpretations for the ineffectiveness of the GP structure on the reading comprehension process of Arab children (and adults) resides in the nature of Arabic morphology by supposedly maintaining the core root of the initial HP-HG words during GP sentence processing. Assessing the same phenomenon with different populations of low-skilled Arab children and learners of Arabic as a second language is recommended because they are both accustomed to reading a vowelized-diacritized script and their mastery of Arabic morphology is assumedly developing.

Keywords

Garden-Path Structure, Arabic Short Vowels, Good-Enough Approach, Heterophonic Homographic, Comprehension

1. Introduction

The modern Arabic language permits a flexible sentence order in which the sentence starts with a verb phrase (VP), noun phrase (NP), or prepositional phrase (PP), with no preference for one over the other (see Table 1). Classical Arabic and Modern Standard Arabic use various word orders for the verb (V), subject (S), and object (O), such as VSO, SVO, VOS, and OVS, based on stylistic variations (Mohammed, 2000). VSO is the dominant word order in Classical Arabic, whereas, for some linguists, SVO is the dominant order in Modern Standard Arabic (Fassi, 1993; Mohammed, 2000; Shormani, 2015; Watson, 2000).

For the basic verb-initial sentences (third person singular past-tense verb), it is expected that the initial word type would be a homographic (HG) word. This word, or to be exact according to Seraye (2004), the heterophonic homographic (HP-HG) word, may add some ambiguity to the sentence. In certain cases, this would garden-path readers, particularly when the initial HG word is relatively far from its disambiguating region (e.g., a five-to-seven-word distance, as in the adult studies by Seraye (2004), and Hermena et al. (2015); a three-word distance as, in Seraye's (2016) study of children). Arab readers will certainly encounter such sentences and are likely to have to go back to the initial HP-HG words, sometimes several times, to review their initial interpretations before they understand the sentences. However, the garden-path (GP) phenomenon is not omnipresent in Arabic. It occurs only when the initial word of the sentence is an HP-HG one when the sentence and the initial word are presented plainly without appropriate short vowels and diacritics.

The Arabic writing system is characterized as both a shallow and a deep orthography, based on the presence and absence of short vowels and diacritics. Short vowels and diacritics are represented voluntarily by signs that take the

Table 1. Word order variations.

Arabic sentence	Translation	Word order in the sentence
علي في المنزل	**Ali** [is] in the house المنزل في **علي** **[S[NP], [PP[NP]]**	S: NP, [VP], PP
في المنزل علي	**In the house** [is] Ali علي **المنزل في** [S[PP[NP], [NP]]	S: PP, [VP], NP
علي ترك المنزل	**Ali** left the house المنزل ترك **علي** **S[NP], [VP[NP]]**	S: NP, VP, NP
ترك المنزل علي	**Left** Ali the house المنزل علي **ترك** **S[VP, NP, NP]**	S: VP, NP, NP

Note. S = sentence; N = noun; P = preposition; V = verb.

shapes "ً," "ٌ," and "ٍ." Arabic script is distinguished by a large number of ligatures and the different shapes that its characters take. Depending on their positions in a text string and in relation to the surrounding characters, these letters take up to four different allographs: independent, word-initial, word-medial, and word-final. To illustrate, the letter "ك" can be written as /ك/, /ـكـ/, /ـسـكـ/, or /ـسـك/. Arabic is read and written from right to left in a cursive consonantal script based on 28 consonant symbols and three short vowels. These vowels, called *fatha*, "َ," *kasra*, "ِ," and *damma*, "ُ," are presented as marks above or below adjoining consonants (e.g., كَ, كِ, and كُ). The following diacritics represent other functions: *skun*, "ْ," is a subscript sign used to indicate that the letter is vowel-less; *shaddah*, "ّ," is a subscript sign that indicates a doubled consonant, or "germination," and *maddah* is represented by the symbol "~" and placed only over the consonant, "ا," to indicate the combination of two consonants, *alif*, "ا," and *hamza*, "ء." Finally, the case-ending markings are very small visual symbols that take the shapes "ْ," "ٌ," "ً" and "ٍ" to indicate syntactically different cases (in standard Arabic), including the nominative, genitive, and accusative cases. The short vowels might be doubled to indicate *nunation*, and take the following shapes: "ً-," "ٌ-," and "ٍ-" (Mahmoud, 1979).

When the Arabic text is vowelized (and also diacriticized), its orthography is considered to be transparent, with a consistent correspondence between the graphemes and phonemes; otherwise, it is considered to be a deep orthography. Unlike prepositional-initial sentences and non-HP-HG-initial sentences, this type of HP-HG-initial sentence needs the appropriate short vowels and diacritics to block the GP phenomenon when the sentence is read.

A potential GP sentence: {فـتح اللص خزانة البنك حدث في المساء}

The initial word, "فتح," is in its unvowelized form and is an HP-HG word that contains the roots of many words. It has four possible forms: it can be read as "فَتَحَ," a basic past-tense active-voice verb, "فَتْح," a basic gerund, "فُتِحَ," the basic passive-voice verb, and "فَتَّحَ," the geminate verb (causative) meaning, "I made it open." However, the passive-voice form is rarely applied unless the discourse requires it, and the last form, "فَتَّحَ," is reserved for special usage. Only the first two readings, especially the first one, can be expected to be used, as was observed in previous studies (Seraye, 2004, 2016) that included a running recording protocol while a connected text was read aloud (see **Figure 1**). Both Arab adults and children, including advanced and beginning readers, have been found to make

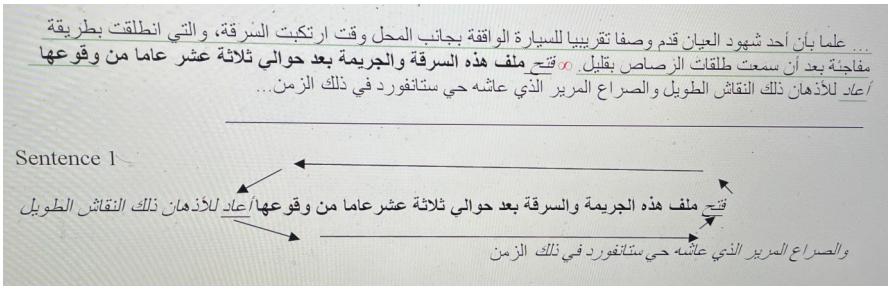

Figure 1. Diagram of the garden-path phenomenon of sentence 1.

use of the basic past-tense form for their initial readings. A syntactic descriptive analysis of a GP Arabic sentence follows:

{فتح اللص خزانة البنك حدث في المساء [فَتْحُ اللصِ ...؛ فَتَحَ اللصُ ...]}

- ***The thief {opening} the safe of the bank [happened] in the evening* (grammatically correct).**

Word order in Arabic: opening + the thief + the safe + of + the bank+ happened + in + the evening

Correct structure: S[NP[NP, NP, PP[P, NP]], VP[PP[NP]]

- ***{Opened} the thief the safe of the bank [happened] in the evening* (grammatically incorrect).**

1) Word order in Arabic: opened + the thief + the safe + of the bank + happened + in + the evening

Mistaken structure: S[VP[NP, NP, PP[P, NP]], VP[PP[P, NP]]

2) Word order in English: the thief opened the safe of the bank happened in the evening

In a previous study of skilled Arab adult readers, Seraye (2004) found that some participants would realize that their initial decisions were wrong and go back and reanalyze the sentence at its disambiguating region. As Ferreira et al. (2009) stated, "Subjects have a tendency to sacrifice reanalysis of the GP in order to keep up with the later material" (p. 416). However, very few participants would assign the correct reading form of the initial HP-HG word (as indicated by the curly brackets) in the first place and would not then be garden-pathed. However, the correct decisions they made would have been based on other factors, such as discourse, frequency, or experience, rather than on the text itself. To avoid such mistakes, all Arabic readers need to have appropriate short vowels and/or diacritics placed over some of the letters in the initial words.

Research into sentence processing, syntactic ambiguity, and syntactic complexity, particularly the GP structure, has been thoroughly discussed and different writing systems (English, Chinese, and Japanese) investigated to address the key questions. For example, how do "people cope with rampant ambiguity, especially syntactic ambiguity, as the linguistic signal unfolds over time? And how is sentence interpretation affected by variations in syntactic complexity?" (MacDonald & Hsiao, 2018: p. 173; for full coverage of the topic, see Ferreira & Çokal, 2016; Ferreira & Qiu, 2021; MacDonald & Hsiao, 2018; for the Japanese and Chinese writing systems, see Lin & Bever, 2011; Mazuka & Itoh, 1995; Mazuka et al., 1997). In the past two decades, extensive examinations have been done on the reading processes of individuals for words and texts in terms of their correlations to short vowels and context (for an overview, see Abu-Rabia, 2019; Ibrahim, 2013; Hermena & Reichle, 2020; Saiegh-Haddad, 2017; Seraye, 2004; Taha, 2016). However, very few studies published in the literature have addressed the processing of structural ambiguity in Arabic, particularly the GP structure, or examined how Arabic's GP structure affects the processes of Arabic readers in terms of accuracy and comprehension. The four studies in the English literature that provide relevant examples used different behavioral methodologies to look

at the processing of HP-HG-initial Arabic sentences in terms of accuracy (Abu-Rabia, 1995; Al-Fahid, 2000) and comprehension (Seraye, 2004 in particular; Hermena et al., 2015) in relation to the presence or absence of short vowels and diacritics[1].

Two of these studies used manual behavioral approaches for assessing their concerns and assumptions, including the masking/unmasking procedure (Abu-Rabia, 1995) and the Diacritic Placement Task (Al-Fahid, 2000). The other two studies involved the use of a self-passed moving window technique, with e-prime's method for clarification on the word level (Seraye, 2004), and the eye-movement technique (Hermena et al., 2015). In spite of the different tasks, techniques, and research paradigms adopted in those studies, their findings generally shared a common ground. The studies are examined extensively below, with illustrations.

The aim of Abu-Rabia's (1995) study was to assess whether Arabic readers whose skills are poor rely more heavily on context than skilled readers do. The study included 40 native Arabic speakers, aged 15 years, who first read aloud, for accuracy, the initial words of 20 HP-HG sentences (10 vowelized and 10 unvowelized), while the rest of the sentences were masked and corrected their readings after the remainders of the sentences had been unmasked. The reasoning was that, since Arabic is unvowelized, the initial words of sentences would be HP-HG words, and Arabic readers would be unable to read or comprehend them properly unless they went back. They would also need to rely on the sentence's linguistic context. The analysis revealed that readers with poor and good skills both needed the linguistic context to read these initial words correctly when they were presented as unvowelized.

Seraye's (2004) study did not support this claim but found that skilled adult Arabic readers initially used the simple form of the verb and then continued reading the sentence, and that most of them stumbled briefly over the disambiguating region in the sentence before continuing to read. However, their comprehension was not affected by the clear example of ambiguous structures in Arabic, which is referred to as "the GP phenomenon."

Because of the dual representation of Arabic orthography (i.e., letters and diacritics), Al-Fahid's study (2000) used qualitative modeling, the so-called Goodman's description of the reading process (1967-1997), to provide supportive evidence from Arabic. The study included 15 Saudi undergraduate students, all males aged from 19 to 25, who read five plain grammatical sentences that started with verbs and determined the possible meaning of only the initial word (the verb) by supplying possible short vowels and diacritics in writing. The aim was to find out "how Arabic readers assign phonology and inflectional features in reading unmarked (normal) texts," and "to test the hypothesis that they use their

[1]Indeed, there is an earlier eye-movement study by Roman, Pavard, and Asselah (1985), but it is in French, and I was not able to find it. Its primary concern (as extracted from the abstract and secondary sources, e.g., Hermena et al., 2015) is about the lexical access mechanisms of initial HP-HG verbs in the absence of short vowels and diacritics. However, as the subsequent study by Hermena et al. (2015) revisits this somewhat and incorporates the aspect of assessing reading comprehension, I believe that this study would suffice for the purpose of the current study.

linguistic knowledge to accomplish this task, and that skillful readers will be able to determine almost all the possible readings for unmarked sentences by placing the diacritics in various ways" (Al-Fahid, 2000: p. 95).

Al-Fahid's (2000) study showed that the participants assigned almost all of the possible grammatical forms to the initial verb, and that two of these, the active- and passive-voice forms, were the dominant choices. However, they also showed a preference for active reading over its passive counterpart. When asked to reflect on this, they attributed the choice of the active form to its common usage and the preference for it to ease of articulation when speaking. Clearly, the active voice sounds "more natural" because it is less marked.

Seraye's (2004) second experiment responded to the claims of previous studies, including the necessity to reread plain Arabic sentences to understand them. The study identified major factors involved in the process of reading Arabic sentences (reading time and comprehension) that need to be controlled to assess their effects in isolation and in terms of their correlation with other factors. These factors included HP-HG- vs. non-HP-HG initial words, absence vs. presence of short vowels and diacritics, economically vs. non-economically presentation of short vowels and diacritics, and lastly the GP vs. non-GP structures.

The study involved 35 native Arabic speakers, aged from 26 to 40, who were graduate or post-graduate students living temporarily in Pittsburgh and Indiana, PA. They read three sets of sentences (108 actual sentences and eight practice ones) in three sessions that took place in one setting.

The results showed that when Arab adults read an HP-HG- and non-HP-HG-initial plain sentence (i.e., ambiguous but with no-GP structure), their reading processes were not affected (p-values = .839 and .318, consecutively). However, their performance on the GP and non-GP sentences was affected in terms of the reading time variable (p = .016) but not for the reading comprehension variable (p = .053). Also, it took the participants longer on average to read the GP sentences than the non-GP ones (M = 6747.14 ms for the GP sentences with SD = 2071.86 ms; M = 6259.30 ms for the non-GP sentences with SD = 1413.28 ms). However, their performance with regard to reading comprehension was very good for both types of sentences (overall mean for the GP sentences was M = .89, SD = .17; overall mean for the non-GP sentences was M = .83, SD = .08). Therefore, a positive relationship exists between the time spent on reading the GP/non-GP plain sentences and the percentage of correct answers.

With regard to the word naming in the e-prime experiment, Seraye (2004) found that reading latency was "positively correlated with the gradual increase of the number of short vowels and diacritics" to the consonants (p. 214), and the GP sentences in Arabic can be resolved by adding some short vowels and/or diacritics to the consonants of the initial HP-HG words. The question became whether the economical representation of short vowels and diacritics that would help in blocking the GP phenomenon would also minimize the reading time by

speeding up the parser's checking process. Four reading conditions were compared: plain (rc1), short vowels-plus-shaddah (rc2), skun-only (rc3), and case-ending marking-only (rc4).

The analyses did not show any significant results for reading conditions with regard to either reading time (p = .283) or reading comprehension (p = .237). It took the participants the same amount of time on average to read the GP sentences, and the participants' comprehension was on average very good. The total means for the reading times were rc1, M = 6747.14, rc2, M = 7277.76, rc3, M = 6997.33, and rc4, M = 7230.64. The total means for the reading comprehension were rc1, M = .89, rc2, M = .80, rc3, M = .88, and rc4, M = .86. Note that it took the participants more time to read the GP sentences in rc2 and rc4, that might be attributed to the orthographically unfamiliar representation to which the participants are not accustomed in those two reading conditions.

In their study of eye movements, Hermena et al. (2015) examined the effect of Arabic orthographic representation on the reading processes of 25 adult native Arabic speakers as they silently read HP-HG verb-initial clauses (active vs. passive voice) that were embedded in complex structural Arabic sentences (40 target sentences) under five reading conditions. In the initial sentence in the active voice, the verb and the entire sentence were either plain or fully vowelized and diacritized. In the sentence using the passive voice, the verb and the entire sentence were plain or fully vowelized and diacritized or the verb alone was fully vowelized and diacritized. Only when the initial HP-HG word was passive and presented as plain would the reader be garden-pathed. Some of the sentences (25%) were followed by yes/no comprehension questions.

The example below of a sentence in the passive voice illustrates the four regions and the GP structure that were examined. Removing the PP (بيدي) changes it into the active voice:

سمع الجميع صرخة مدوية حينما دفعت1 **الطالبة التي كانت في طريقها إلى المعمل** بيدي2 زميلتها3 فهوت مغشيا4 عليها.

The analysis of the study by Hermena et al. (2015) revealed the following: Arab adults took longer to read the GP structure, but their comprehension was not affected (the score ranges = 70% - 100%), descriptively speaking, it took the participants more time on average to read the plain active-voice sentence (M = 8185 ms, and SD = 1451.15 ms) than its fully vowelized-diacritized counterpart (M = 8119 ms, and SD = 1190.58 ms), participants spent more time (disruption) on the disambiguating region (PP) when the passive verb was presented as plain than when it was vowelized-diacritized, there was more sensitivity to the GP structure in the PP region when the entire passive sentence was disambiguated by the presence of short vowels and diacritics than when only the passive verb was disambiguated, and the initial preference of the parser was to read the HP-HG past-tense simple verb as an active verb.

The results obtained from using the eye-movement technique are somewhat aligned with those for the offline tasks, particularly with the self-paced reading

procedure. The findings are as follows: there was an effect for the HP-HG initial structure (particularly the GP) in terms of the reading processes of the Arab adults, as reflected in their processing times for reading but not in their reading comprehension, the short vowels and diacritics, when presented indiscriminately, sometimes had a negative effect on the reading processes, as reflected in both the total reading times in the self-paced reading (and in the e-prime word naming in Seraye's 2004 third experiment in) and the fixation time/visual spans in the eye-movements, and the HP-HG verb in its base form was most frequently parsed as a verb in the active voice. Note that in the self-paced reading task (in Seraye's second experiment) the participant cannot reanalyze their first interpretation by going back to the HP-HG word; however, in eye-movement's (as in Hermena et al.'s, 2015, study), they can.

The descriptive analysis of the reading behavior of Arab adults as they read two paragraphs with embedded potential GP sentences for accuracy showed that almost all of them were garden-pathed, because their initial decisions were to assign the active-voice form to the initial HP-HG word. As **Figure 1** illustrates, the participants paused immediately after reading the final word in the previous sentence ("بقليل" followed by a period) and before reading the first word in the next sentence ("فتح"). They would always select the basic past tense until they arrived at the disambiguating region, where they would make an exclamation, sometimes sounding out the word "لا," meaning "oh no!" or "sorry!," indicating their awareness of their mistakes of assigning the initial HP-HG word to its past tense and not to its gerund form.

Some of the participants also tried other possible forms mentioned earlier before moving on to the second word. Indeed, participants reluctantly tried virtually all possible forms before finally selecting one, generally the incorrect one, that is, the active-voice basic verb. In addition, they would construct verb-initial sentences even when these began with nouns, gerunds, or prepositions; for example, reading the gerund "استئناف," which means appealing, as "استأنف," meaning to appeal, even though these words can be distinguished orthographically. However, they corrected their "miscues" immediately.

As mentioned above, the participants' understanding of the sentences was not affected, and this result does not agree with the claims of previous studies that Arabic readers must read a sentence twice to understand it (Abu-Rabia, 1995, for sentence level; Abu-Rabia, 1999, 2001, for text level). Therefore, the results of Seraye's (2004) study do not support the claim that Arabic readers must reread any plain Arabic sentence in general and GP sentences in particular to understand them. Indeed, this claim cannot be supported logically, because the sentence in Arabic could start with an HP-HG or a non-HP-HG word. However, within a special type of sentence that was presented as plain, the less skilled and even the highly skilled readers needed context to activate the correct form of the initial HP-HG word in the first place (i.e. in terms of *pronunciation/ sounding out* the word). Thus, skilled readers need the context for the plain initial HP-HG

words that turn a sentence structure into a GP structure. However, the readers' understanding of the meaning of the GP sentences was unaffected; their performances were expected to be the same regardless of the structure of the sentence. The question, then, is whether the same phenomenon would hold for beginning Arabic readers, as represented by fourth-grade students. Would they be able to use their morphological knowledge in understanding the GP sentence and activate all possible forms or interpretations initially and later on at the disambiguating region, to be able to arrive at the correct interpretation?

2. Method

2.1. Participants

A total of 39 fourth-grade native Arabic male students, aged 9 and 10, from three public elementary schools in Riyadh, Saudi Arabia, were chosen for the study and offered R20 as compensation for participating. Almost all of them had normal vision and reported no learning or reading difficulties The teachers' ratings and pre- and post-criteria procedures were taken into account, and only the skilled readers from the previous study (Seraye, 2016) who were at grade level in reading were included in the sentence-reading experiment. Also, only those who scored 40 out of the 50-word list were included. Furthermore, a post-criteria judgement (reading a short passage) was administered to each participant in the second session to ensure that only those who demonstrated reading fluency were included. Official approval and consent for participation were obtained beforehand.

2.2. Material

Two sets of sentences were constructed as the stimuli for the experiment: the first included 31 sentences, seven of which were for the practice session. Twelve of the 24 remaining sentences were plain without short vowels or diacritics (rc1), with only the consonants presented, and four of these 12 were potential GP sentences. The other 12 sentences were presented as fully vowelized and diacriticized (rc2), supplemented with short vowels, shaddah, and skun, and four of them were potential GP sentences. The short vowels and diacritics on the initial words in the sentences would convert into HP-HG words that would garden-path the readers only if they did not assemble them.

In the second set for the incorrectly vowelized condition (rc3), there were 13 sentences, one of which was used for the practice session. Four of the 12 remaining sentences were potential GP ones. Only the short vowels were manipulated incorrectly. Hence, by the end of both sessions, all 39 participants should have read 36 actual sentences and eight practice sentences representing three reading conditions (see the Appendix for more details). Some of these sentences were extracted from the participants' reading textbooks with modifications for control purposes.

Each sentence included a question followed by three responses: *true*, *false*, and

I don't know. The last option, *I don't know*, was given to prevent guessing and offer participants an option to respond when the sentence did not make sense to them. Only two possible responses were considered and collected: *true* and *false*, and only the participants' textually based comprehension was measured.

The sentences represent, to some extent, the form that Arabic sentences take and the types of sentences that Arabic readers encounter in connected texts. The target sentences were those with GP forms, and the remaining stimuli were used as filler items. The sentences were also equalized in terms of their lengths, numbers of words, and morphological units. Each sentence included seven words, and approximately three words separated the initial HP-HG word from the disambiguating area. All of the sentences included only words that the participants could see in their textbooks and texts designed for their age group.

In general, the sentences and questions were judged in terms of naturalness, authenticity, accuracy, age- and grade-level appropriateness, capability of capturing comprehension, and so on, by a team of Arabic fourth-grade teachers and graduate students in an Arabic teaching program (see the **Appendix** for the sentences and questions used in the experiment).

2.3. Measures

Two dependent variables were measured: reading time, measured to the nearest millisecond, and comprehension product, measured and coded as *true*, *false*, and *I don't know*. Each correct answer was assigned a 1; otherwise, the response (being false or I don't know) was given a 0.

2.4. Procedure

The study followed the procedure from the text reading study (Seraye, 2016) and identified participants based on their numbers of miscues and reading times, selecting only the skilled readers for the sentence processing study.

The task was conducted on three 14-inch display laptops, running a moving-window software program that I designed and tested (Seraye, 2004).

This was a word-by-word, self-paced reading procedure, in which participants used a button (space-bar key) that showed every word sequentially when clicked but hid the previous ones. Once the participants had finished the last word (which was followed by a stop sign) and pressed the space-bar key, a question would pop up with three options for responding, and the same process continued through to the final sentence, after which the participants were thanked and informed that their session had ended. There were no brackets between the words. The order of the reading conditions was counterbalanced, and the random presentation of the sentences was controlled by the software program.

Participants were asked to read the instructions and complete the practice part of the experiment before proceeding to the actual experiment. Further, they were informed orally before starting the practice and the actual experiment, about the nature of the reading task and how the self-paced reading software works. They

were all given the test individually in empty, secured rooms at their schools. The task was explained to them orally, and they were told to read the sentences at their own paces, word by word, and then to answer pop-up questions with three response options: *true*, *false*, and *I don't know*. They were told to choose "*I don't know*" only when they did not know the answer or wanted to guess. The Internal Research Board (IRB) protocols were also conducted accordingly

Because of the different materials and instructions, the task was given in two sequential sessions at the same meeting time, with the sessions including the same task and procedure. The only difference between the two sessions was in the reading materials: the students in the second session were informed that they would read sentences in which the words were presented with the wrong short vowels. They were also informed that assembling the wrong short vowels would lead to constructing words that had no meaning in Arabic; that is, the graphemic form (consonants) of the words was intact, but the phonological aspect was distorted. Assembling only the consonants and ignoring the short vowel signs would lead participants to read a real word in Arabic.

A small-scale pilot study was conducted on 10 participants who volunteered before the actual experiment. Some concerns about potential problems in the setting conditions were identified and resolved before the full study began.

2.5. Design and Analysis

Four analyses were conducted to evaluate the effects that the GP structure, by itself and in terms of the correlation with short vowels and diacritics, had on the reading processes of Arabic children, their reading times and comprehension. Two statistical procedures were employed separately, the dependent samples *t*-test, and the one-way repeated measures analysis of variance. The tests assumptions were checked before running the analyses.

3. Results

Four analyses were conducted to respond to the concerns that were raised, with a subset of the data, the plain reading condition, analyzed first using a dependent samples *t*-test. This involved a comparison between GP and non-GP sentences in terms of reading time and percentage of correct responses.

With regard to the data on reading times, the analysis (Table 2) did not reveal a significant difference for the GP structure, $t(38) = -.37$, $p = .710$. However, a difference in the mean values was observed (roughly 60 ms); it took the participants longer to read the GP sentences than the non-GP sentences ($M = 8172$ ms for the GP sentences; $M = 8113$ ms for the non-GP sentences).

Also, because the data were not normally distributed due to the outliers, a nonparametric test, the Wilcoxon matched-pairs signed-rank test, was used along with the dependent samples *t*-test analysis. The results did not reveal any significant differences between the means, *z* values (−.07), and *p* values (.944); therefore, only the results of the *t*-test are provided in Table 2.

Table 2. Results of the t-test on the reading times of GP and Non-GP sentences.

Non-GP sentences		GP sentences				
M	*SD*	*M*	*SD*	*t*	*df*	*p*
8112.7	2380.6	8172.3	2486.3	−.374	38	.710

Note. GP = garden-path.

These results, in terms of significance, are not in accordance with those of a study of adults (Seraye, 2004) that revealed (p = .016) that the GP structure affected the skilled Arabic readers as they read, comparing GP and non-GP sentences (M = 6747.1, SD = 2071.9; M = 6259.3, SD = 1413.3, respectively). The overall means showed real differences; that is, the participants took more time on average to read the GP sentences than the non-GP ones (487-ms difference). However, the sentences in our study had seven words, whereas they had about 11 words in the adult study. Furthermore, the distance between the initial HP-HG word of the sentence and the disambiguating region was five words on average in Seraye's (2004) adult study and three words in this study (except for one sentence, in which the distance was four words). Such distances are reported to have an effect on successful reanalysis of the GP sentences (Ferreira & Henderson, 1998; Ferreira et al., 2001: p. 4; for an illustration, see **Figure 1**).

With regard to the data on reading comprehension, the analysis did not reveal any significant differences between the means of the percentages of correct responses to both types of sentences, $t(38) = -1.66$, p = .105 (**Table 3**). The participants' correct responses did not differ significantly on average between the GP and non-GP sentences (the overall mean for the non-GP sentences was M = .73; the overall mean for the GP sentences was M = .80). Since the data were not normally distributed but extremely skewed because comprehension was good on the whole and because of the outliers, a nonparametric test, the Wilcoxon matched-pair signed-rank test, was used along with the dependent samples t-test analysis. The results did not reveal any significant differences between the means, z values (−1.17), and p values (.242); therefore, only the result of the t-test is provided in **Table 3**. However, based on the overall means, the participants scored better on average on the GP sentences than on the non-GP sentences (M = .80 (80%) and M = .73 (73%), respectively).

These results, showing no significant differences between the two, are in accordance with those of Seraye's (2004) study on adults, which used both parametric and nonparametric tests (p = .053) and revealed that on average the GP structure did not affect the comprehension of skilled Arabic readers; their performances on both structures were on average very good (M = .89, SD = .17 for the GP and M = .83, SD = .08 for the non-GP). Furthermore, the results of this study were in line with those of the adult study in terms of the differences between the overall means in favor of the GP structure. Both studies showed that Arabic readers took more time to read GP sentences than non-GP sentences and then answered the comprehension questions correctly.

Table 3. Results of the t-test on the reading comprehension of GP and Non-GP sentences.

Non-GP sentences		GP sentences				
M	*SD*	*M*	*SD*	*t*	*df*	*p*
.73	.22	.80	.32	−1.66	38	.105

Note. GP = garden path.

Because the GP sentences can be resolved and understood by adding the correct short vowels and diacritics to the initial HP-HG words, a one-way repeated measures analysis of variance was conducted on a subset of the data to respond to the second concern of whether such a supplement of short vowels and diacritics would make any difference. The GP sentences were compared using the *a priori* or potential GP sentences supplemented with the correct short vowels and diacritics that would, if being processed, convert the sentences into non-GP ones or lead only to phonological distortions when supplemented with the wrong short vowels, while the graphemic representation (consonantal structure) remained intact. The GP and the potential GP sentences in reading conditions 1, 2, and 3 were compared in terms of reading time and comprehension. The vowelized condition that was incorrect was included for control purposes; that is, if the participants were not affected by adding the wrong short vowels to the consonants, this indicated that beginning Arabic readers might not process such sub- and superscripts, and that their reading habitually depended on processing only the consonants in the words (Seraye, 2004). Therefore, the main question would be whether supplementing with short vowels and diacritics would have any value for the reading process of GP sentences.

For the data on reading times, the analysis showed that it did not matter whether the GP sentences were supplemented with the correct short vowels and diacritics or presented as plain ones. The participants' reading times were on average the same ($F(2, 76) = .58$, $p = .565$). Also, it took the participants on average 8172.33 ms to read the GP sentences that were presented as plain ones, 8007.64 ms to read the GP sentences that were supplemented with the correct short vowels and diacritics, and 7882.74 ms to read the GP sentences that were supplemented with the wrong short vowels (**Table 4**).

The overall means showed that it took the participants more time on average to read the plain sentences than their vowelized-diacritized counterparts. This can be attributed to the short vowels and diacritics that helped the participants to process the sentences more rapidly. However, the fact that it took less time on average to process the potential GP sentences with the incorrectly vowelized sentences than the other two types of sentences presents a problem. Different scenarios have been suggested to resolve this. The first involves a description of the contextual abnormality in the reading conditions, in which rc1 and rc2 were given together in one setting, whereas rc3 was given by itself in one setting due to the different instructions. Furthermore, the participants in rc3 were provided with a prompt about the reading condition and informed that assembling those

Table 4. Overall means on reading time for GP sentences.

	Reading condition (Sentence stimuli)	GP sentences	
		M	*SD*
Group	Plain (no short vowels or diacritics)	8172.3	2486.4
	Fully vowelized and diacritized	8007.6	2274.3
	Wrong short vowels	7882.7	2273.0

incorrect short vowel signs would result in the construction of words with no meaning in Arabic. The results showed that it took considerably less time to compare this reading condition to the others. Therefore, could this be an extra attention-based process that resulted in the participants giving the sentences a deeper reading? However, in addition to accepting this last assumption, it would be reasonable to anticipate that the reading comprehension would, at the very least, not be affected, and this extra attention-based process would turn into accurate comprehension of the GP sentences, which was not the case since the participants scored the lowest with this reading condition (Table 5). Furthermore, the reading time values in rc3 were not even close to their counterparts in rc1 (see Figure A1 in the **Appendix**).

Any assumption that Arabic readers benefit from the supplemented vowels and diacritics in rc2 when processing the GP sentences, as reflected in the shorter times, cannot be substantiated, knowing that they scored lower on the reading comprehension compared to their performance in rc1 (Table 5).

Finally, the carryover (and order effects) could become confounded with the effect of the independent variable in a within-subjects design with a repeated measures analysis of variance. However, although the current study involved a counterbalancing procedure that was conducted between reading conditions, and random presentations of sentences were controlled by the software program, a follow-up experiment with new materials and designs (between-subjects design with the same target sentences) is needed to determine whether the assumptions can still hold.

The analysis of the data on reading comprehension revealed a significant difference between the three reading conditions ($F(2, 76) = 3.85$, $p = .026$). Furthermore, pairwise comparisons showed a significant difference between reading condition 1, the plain one, on one side, and reading condition 2, the vowelized and diacriticized one ($p = .045$), and reading condition 3, the wrongly vowelized one ($p = .012$), on the other side. However, there was no significant difference between reading condition 2 and reading condition 3 ($p = .618$). To some extent the participants scored on average higher on reading condition 1 ($M = .80$), where the GP sentences were presented as plain ones, than on reading condition 2 and reading condition 3 (Table 5). Indeed, descriptively speaking, the participants scored the worst on average on the incorrectly vowelized condition (see Figure A2 in the **Appendix**). In his study of adults, Seraye (2004) also found scores to be higher on the plain reading condition (the mean values in Seraye's

Table 5. Overall means on reading comprehension for GP Sentences.

	Reading condition (sentence stimuli)	GP sentences	
		M	*SD*
Group	Plain (no short vowels or diacritics)	.80	.32
	Short vowels-plus-diacritics	.69	.29
	Wrong short vowels	.66	.25

2004 study were .89 for the plain reading condition and .80 for the vowelized-diacriticized condition). Adding the correct short vowels and diacritics to turn the GP sentences into non-GP sentences did not seem to have a major effect on the reading comprehension of experienced or beginning Arabic readers. However, the analysis did show that the participants comprehended the actual GP sentences better when they were presented as plain sentences than when they were supplemented with the correct short vowels and diacritics. This is unexpected result, because there is no extra benefit from the presence of short vowels and diacritics that should limit the potential readings to one correct reading and subsequently make the structure easier to understand. Since the fourth grade children are accustomed to reading texts provided with short vowels and diacritics, the "lack of familiarity" interpretation might not be endorsed.

4. Discussion

The results of our study showed that the GP structure of Arabic sentences did not affect the reading processes of beginning Arabic readers as indicated by their reading times and responses to the comprehension questions. Nor were their processes for reading such sentences significantly improved or shortened by the supplementation of the correct and appropriate short vowels and diacritics that should have eliminated the ambiguity that the GP structure caused. Although the design of the self-paced reading software program prevented the participants from returning to other parts of the sentences to clarify or verify their understanding of them, the GP structure did not affect their comprehension. As the mean values indicate, the participants did well on both GP and non-GP structures, and this is consistent with Seraye's (2004) study of highly skilled Arabic readers.

To provide an overview of the topic of sentence processing that responds to the central questions mentioned previously regarding how the syntactic ambiguity is handled as the linguistic signal unfolds over time and how it affects sentence interpretation, MacDonald and Hsiao (2018) arranged the responses into two broad theoretical approaches. These included an approach that emphasized innate processing mechanisms and one that focused on the role of prior linguistic experience (p. 173). They then put the theoretical models into a very concise diagram with two intersecting axes: one representing "the model's general emphasis in explanation for comprehension data"; the second representing "the type of materials most typically studied by proponents of the approach" (Mac-

Donald & Hsiao, 2018: p. 174).

Based on experimental and observational studies, researchers have proposed different models to explain the process of converting a string of words into a structural representation (Mitchell, 1994; for a review, see MacDonald & Hsiao, 2018). The experimental studies cite the effect of garden-pathing, and different models have been suggested for explaining this (Mitchell, 1994; Ferreira & Cokal, 2016; Wonnacott et al., 2016; Ferreira & Qiu, 2021).

Observational studies that involve reading aloud, such as Seraye's (2004) study and the current one, explicitly demonstrate the effects of GP for both highly skilled Arabic adults and beginning-level Arabic children. Both groups of readers hesitated over the initial HP-HG words of the GP sentences and activated one of the possible forms of the initial HP-HG words. When they arrived at the disambiguating regions, they showed stunned expressions and then went back to the beginnings of the sentences to reanalyze their first choices. It was expected that this reanalysis process would result in delays and subsequently lead to some extra reading time to process the sentence. This extra reading time can be explained on the basis of two assumptions: "on the basis of the implicit checking process that operates with a delay cost, or on the basis of the processing load in the ambiguous region." The effects were demonstrated by several studies that employed different techniques, eye-tracking studies (Ferreira & Henderson, 1990, Experiment 1), first fixation data (Frazier & Rayner, 1982), self-paced reading tasks (Mitchell et al., 1994, Experiment 1, as cited in Mitchell, 1994: p. 381; Seraye's 2004, Experiment 1), and brain imaging (Mason et al., 2003; Seraye, 2004: p. 178).

A comparison of the two overall means of the times for the beginning readers in this study to process the GP sentences and their non-GP counterparts showed that on average it took them somewhat longer (60 ms difference) to process the GP sentences (**Table 2**). However, as was pointed out earlier, this difference was not statistically significant. According to Seraye (2004), there are two models of explanations that can be proposed to account for such results: the resource-free parallel model and the resource-limited parallel model. "By saying that there is no difference between the reading time of the GP sentences and the non-GP sentences, then the former model can be suggested to account for the finding of no difference. However, by accepting that, there is a discrepancy in the reading time between GP and non-GP sentences, only a resource-limited parallel model can account for such discrepancy in reading time between the two types of structures (Mitchell, 1994). That is, the existence of unexplored options might somehow be tagged or marked at the choice point, perhaps providing the basis for relatively efficient re-analysis procedures (Frazier & Rayner, 1982, cited in Mitchell, 1994: p. 378). It is suggested that tagging or marking, for Arabic, is based on the core element, the root, that the alternative forms of [the heterophonic]-homographic initial share (p. 178).

Arabic morphology is based on the trilateral/quadrilateral root, and when

Arab readers are challenged with a consonant-based writing system they are expected to exploit their knowledge of word formation in Arabic when accessing the mental lexicon of the word representations (Abu-Rabia, 1995, 2001, 2012). Accordingly to Seraye (2004), "very often, at the core of all activated potential forms of the HP-HG, there will be a trilateral/quadrilateral-root which indicates the core semantic element that is very often shared by all activated forms. In addition, the form/pattern of the word (its skeletal tier/word pattern/binyan, McCarthy, 1979; 1981) narrows the possible readings of the HP-HG word (p. 240). In fact, as Seraye (2004) put it, "the predictability/productivity of word forms/patterns; affixation, etc., compensate for the lack of short vowels and diacritics in print" (p. 259).

The role of Arabic morphology roots in the reading process has recently been revisited, consolidated, and thoroughly documented (see, for example, Boudelaa & Marslen-Wilson, 2005, 2015; Mahfoudhi, 2007; Mahfoudhi et al., 2010; Abu-Rabia & Abu-Rahmoun, 2012; Taha & Saiegh-Haddad, 2016, 2017; and for a recent overview of the claim, see Saiegh-Haddad, 2017; Abu-Rabia, 2019; Hermena & Reichle, 2020; Wattad & Abu-Rabia, 2020; Abu-Rabia, 2021).

Two explanations, one stem-based and the other morpheme-based, have been proposed to account for the linguistic unit of analysis in Arabic word processing, lexical access, and lexical representation. With regard to the first, the verb stem was considered as the linguistic unit of analysis (Benmamoun, 1999); for the morpheme-based one, the root and word pattern were viewed to be the linguistic unit of analysis (Boudelaa, 2014).

After merging three data types of evidence drawn from three techniques (priming, neuropsychological, and neuroimaging), Boudelaa (2014) concluded that "lexical processing in Arabic evolves around roots and word patterns, and that the extraction of these units during spoken language comprehension and reading is subserved by an obligatory decomposition mechanism" (p. 47). The existence of sublexical accessibility in word recognition (Cole et al., 1997; Taft, 1981a, 1981b) has also been documented. Taft's (1981b) experiments demonstrated that "prefix stripping occurs in word recognition and this, in turn, implies that prefixed words are accessed through a representation of their stem" (p. 296).

The study conducted by Mahfoudhi et al. (2010) revealed that the possession of good morphology skills among elementary students is related to better reading comprehension and could explain the independent variability in their comprehension levels. This exploitation of the morphological knowledge, word roots, and patterns in word processing and spelling were also found in second graders (Taha & Saiegh-Haddad, 2016, 2017). Even among highly skilled Arab adults, word roots were found to be helpful in reading morphologically complex words, whether they were presented as plain or vowelized (Abu-Rabia, 2012). However, in a study of the role of morphological knowledge (root and pattern "units") and awareness in the reading acquisition of primary-grade children, El Akiki and Content (2020) found word patterns to be a playful factor in Arabic word recognition. Nevertheless, they offer the following caution: "the absence of

a root frequency effect should not be taken to conclude that children are insensitive to root properties. In fact, it would seem hard to explain how [the children] could isolate patterns without at the same time extracting the roots" (p. 13).

This "awareness" was also documented by Badry (1982), whose study revealed that Moroccan children, aged 3 to 6 years, were aware of the underlying morphological roots in their spoken language, and this awareness was reflected in the production stage of their acquisition. The role of correct and appropriate short vowels and diacritics in disambiguating the initial HP-HG word in a GP sentence is clear. Adding these to the consonants in the initial HP-HG words will turn them into one legal reading, and thus, the sentence must not garden-path readers on the condition that the readers assembled such visible sub/super signs with the consonants. However, the results of our study revealed that reading GP and *a priori*/potential GP sentences was essentially the same. Although the participants took more time on average to read the GP sentences (M = 8172.33 ms), they took less time on the GP sentences that were provided with the correct short vowels and diacritics (M = 8007.64) and much less time on the GP sentences that were provided with the incorrect short vowels and diacritics (M = 7882.74). However, although these observations could indicate that Arab readers are accustomed to reading consonant-based orthography because if they assemble the wrong diacritical signs their reading times increase, this is not the case. Furthermore, there is evidence that people's previous experiences with linguistic and nonlinguistic input play a central role and "strongly shape" their online interpretations of ambiguity in sentences (MacDonald & Hsiao, 2018: p. 176).

This lack of a role for the short vowels and diacritics in connected texts can be explained by the fact that subjects, as Ferreira et al. (2009) state, "have a tendency to sacrifice reanalysis of the garden-path in order to keep up with the later material. This pattern of results is consistent with the assumptions of the good enough theory of language processing, which assumes that processing resources are limited, and; therefore, predicts that garden-path reanalysis processes will be curtailed if upcoming material must also be processed ." (p. 416). The special GP structure in Arabic that was assessed in this study and in Seraye's (2004) provides the experimenter/researcher (in a reading-aloud task) with a window into the behavior of the Parser (with uppercase "p"),—the tendency to sacrifice the reanalysis processes.

Examining the nature of the GP phenomenon in Arabic seems to be a unique feature of the literature. This involves the view that the incorrect first reading of the initial word of a GP sentence would always lead readers astray, because they would start building structures that are not compatible with the entire sentence once they arrive at the disambiguating region. Then, when readers discover that their initial assignments are wrong, it seems that all they can do to compensate is to replace the noun with a verb or vice versa (the "Adjust" principle in Fodor's model; Fodor & Inoue, 2000) and supposedly restructure the sentence in a way that fits the correction. Although this restructuring would not affect the meaning extracted from the GP sentence, it would affect the type of structure that is built,

whether the sentence starts with a noun phrase or a verb phrase. This assumption is supported by the fact that Arab readers, as mentioned before, exploit their morphological knowledge by supposedly maintaining the core root of the initial HP-HG word during sentence processing. Furthermore, it can be justified by assuming that the sentence processing system [may go] into "reanalysis mode," attempting to adjust the syntactic structure that has been built so as to create a grammatical analysis (Ferreira & Henderson, 1991; Fodor & Ferreira, 1998; Fodor & Inoue, 1994). Ease of analysis depends on the extent to which the sentence processing system can find lexical and grammatical information that motivates an alternative structure (Ferreira & Cokal, 2016: p. 13).

The parser may also handle ambiguity by leaving the interpretation open and "making a specific attachment decision only once it is necessary to do so" (Swets et al., 2008). Evidence regarding underspecified representations has also led to suggestions of a tendency within the processing system to delay interpretations of HP-HG words that have multiple forms by initially activating an underspecified structure and then filling out the semantics once contextually disambiguating information becomes available.

Therefore, an account that presupposes an activation of various types of information that are being brought to bear on the interpretation of Arabic GP sentences is still legitimate. This is supported by the fact that the Arabic parser (with lower case) would activate one root for two pattern forms (verbal vs. nominal).

Reiterating the previous claim of the uniqueness of the Arabic GP structure further emphasizes the uniqueness of its orthography, since its ability to be segregated allows the researcher to examine the GP structure easily. Essentially, when reading such structures aloud in their plain representation, the reader will assign reading to the initial word of the GP sentence on the spot, giving it a syntactic/thematic role that would immediately determine whether the reader was garden-pathed or not. As in the examples above (**Figure 1**), if the initial word "فتـح," was written as "فَتَحَ" (the simple past tense), and this was the default reading, readers would be garden-pathed until reaching the disambiguating region. They would then go back, giving only an exclamation or continuing to read the sentence as if there was nothing wrong.

Therefore, using a behavioral approach for assessing assumptions about sentence processing, the task of reading a GP sentence aloud from the beginning would be a short-cut tool for investigating some of the hypotheses on sentence processing. In terms of silent reading of the GP sentence, the question the researcher proposes should target the type of structure readers build and then determine whether the reader was garden-pathed or not (particularly in a self-paced reading).

In summary, it is recommended that the same study be applied to Arabic readers with poor reading skills and non-Arab learners of Arabic, since Arab children who do not have strong skills and Arabic learners are accustomed to

reading vowelized and diacriticized reading materials, and their mastery of Arabic morphology is developing. The justifications for the recent findings of this study can be assessed in accordance with this.

To conclude, I propose that the parser behavior in processing ambiguous Arabic structures (initial HP-HG words) is driven by the two mechanisms (Macdonald and Hsiao, 2018): the modular/innate one (i.e., the morphological/syntactic knowledge), in which the root and/or pattern forms are activated, and the experiential one (i.e., the prior linguistic experience), in which the simple past-tense form is activated as the default word form while building the structure of an HP-HG-initial sentence.

However, it is worth mentioning that the findings of the current study (and Seraye, 2004) must not be interpreted as a shift from the mainstream instruction method of teaching reading in Arabic (i.e. phonics) to a whole language approach. That is, the "experience" explanation that is suggested here should not be overinterpreted in a way that would entail a subsequent recommendation of using the whole language approach in teaching reading in Arabic instead of the mainstream one (i.e. phonics). The current findings should be restrictedly interpreted within the issue proposed here, the GP phenomenon in Arabic. The ineffectiveness of the supplemental short vowels and diacritics to the consonants in Arabic GP sentences on the reading comprehension process, must not underestimate the role that short vowels and diacritics play in the accurate representation of speech in writing Arabic, and subsequently reduces the reading processing time on the reader. Further, the current results need not convey a message to teachers in presenting reading materials that should move the child from transparent to deep orthography. Indeed, owing to the "experience" explanation, pupils should be encouraged and trained to pay more attention to the supplementation of short vowels and diacritics in the reading materials. Moreover, the writers and textbooks designers need to identify the ambiguous regions in the reading materials, and provide them with the necessary short vowels and diacritics in order to remove the ambiguity from texts and thus helping in reducing the time load of reading by speeding up the parser's checking process.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

Abu-Rabia, S. (1995). Learning to Read in Arabic: Reading, Syntactic, Orthographic & Working Memory Skills in Normally Achieving and Poor Arabic Readers. *Reading Psychology: An International Quarterly, 16,* 351-394. https://doi.org/10.1080/0270271950160401

Abu-Rabia, S. (1999). The Effect of Arabic Vowels on Reading Comprehension. *Journal of Psycholinguistic Research, 28,* 93-101. https://doi.org/10.1023/A:1023291620997

Abu-Rabia, S. (2001). The Role of Vowels in Reading Semitic Scripts: Data from Arabic &

Hebrew. *Reading & Writing: An Interdisciplinary Journal, 14,* 39-59. https://doi.org/10.1023/A:1008147606320

Abu-Rabia, S. (2012). The Role of Morphology and Short Vowelization in Reading Morphological Complex Words: Evidence for the Domination of the Morpheme/Root-Based Theory in Reading Arabic. *Creative Education, 3,* 486-494. https://doi.org/10.4236/ce.2012.34074

Abu-Rabia, S. (2019). The Role of Short Vowels in Reading Arabic: A Critical Literature Review. *Journal of Psycholinguistic Research, 48,* 785-795. https://doi.org/10.1007/s10936-019-09631-4

Abu-Rabia, S. (2021). The Impact of Diacritic Marking on Listening Comprehension Skills in Arabic Orthography. *The Journal of Educational Research. 114,* 64-73. https://doi.org/10.1080/00220671.2021.1872476

Abu-Rabia, S., & Abu-Rahmoun, N. (2012). The Role of Phonology and Morphology in the Development of Basic Reading Skills of Dyslexic and Normal Native Arabic Readers. *Creative Education, 3,* 1259-1268. https://doi.org/10.4236/ce.2012.37185

Al-Fahid, J. M. (2000). *The Goodman Psycholinguistic Model of English Reading and Its Applicability to Semitic languages*. Unpublished Doctoral Dissertation, University of Arizona.

Badry, F. (1982). *The Centrality of Root in Semitic Lexical Derivation: Evidence from Children's Acquisition of Moroccan Arabic* (Papers and Reports on Child Language Development, No. 21. 9-15). Stanford University. (ERIC Document Reproduction Service No. ED 222046).

Benmamoun, E. (1999). Arabic Morphology: The Central Role of the Imperfective. *Lingua, 108,* 175-201. https://doi.org/10.1016/S0024-3841(98)00045-X

Boudelaa, S. (2014). Is the Arabic Mental Lexicon Morpheme-Based or Stem Based? Implications for Spoken and Written Word Recognition. In E. Saiegh-Haddad & M. Joshi (Eds.), *Handbook of Arabic Literacy: Insights and Perspectives* (pp. 31-54). Springer. https://doi.org/10.1007/978-94-017-8545-7_2

Boudelaa, S., & Marslen-Wilson, W. (2005). Discontinuous Morphology in Time: Incremental Masked Priming in Arabic. *Language and Cognitive Processes, 20,* 207-260. https://doi.org/10.1080/01690960444000106

Boudelaa, S., & Marslen-Wilson, W. (2015). Structure, Form, and Meaning in the Mental Lexicon: Evidence from Arabic. *Language, Cognition and Neuroscience, 30,* 955-992. https://doi.org/10.1080/23273798.2015.1048258

Cole, P., Segui, J., & Taft, M. (1997). Words and Morphemes as Units for Lexical Access. *Journal of Memory and Language, 37,* 312-330. https://doi.org/10.1006/jmla.1997.2523

El Akiki, C., & Content, A. (2020). Early Sensitivity to Morphology in Beginning Readers of Arabic. *Frontiers in Psychology, 11,* Article ID: 552315. https://doi.org/10.3389/fpsyg.2020.552315

Fassi, A. (1993). *Issues in the Structure of Arabic Clauses and Words* (Vol. 29). Kluwer Academic Publishers

Ferreira, F., & Çokal, D. (2016). Sentence Processing. In G. Hickok, & S. L. Small (Eds.), *Neurobiology of Language* (pp. 265-274). Academic Press. https://doi.org/10.1016/B978-0-12-407794-2.00022-5

Ferreira, F., & Henderson, J. (1998). Syntactic Reanalysis, Thematic Processing, and Sentence Comprehension. In J. D. Fodor, & F. Ferreira (Eds.), *Reanalysis in Sentence Processing* (pp. 73-100). Kluwer Academic Publishers. https://doi.org/10.1007/978-94-015-9070-9_3

Ferreira, F., & Henderson, J. M. (1990). Use of Verb Information in Syntactic Parsing: Evidence from Eye Movements and Word-By-Word Self-Paced Reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 555-568. https://doi.org/10.1037/0278-7393.16.4.555

Ferreira, F., & Henderson, J. M. (1991). Recovery from Misanalyses of Garden-Path Sentences. *Journal of Memory and Language, 30,* 725-745. https://doi.org/10.1016/0749-596X(91)90034-H

Ferreira, F., & Qiu, Z. (2021). Predicting Syntactic Structure. *Brain Research, 1770,* Article ID: 147632. https://doi.org/10.1016/j.brainres.2021.147632

Ferreira, F., Christianson, K., & Hollingworth, A. (2001). Misinterpretations of Garden-Path Sentences: Implications for Models of Sentence Processing and Reanalysis. *Journal of Psycholinguistic Research, 30,* 3-20. https://doi.org/10.1023/A:1005290706460

Ferreira, F., Engelhardt, P. E., & Jones, M. W. (2009). Good Enough Language Processing: A Satisficing Approach. In N. Taatgen, H. Rijn, J. Nerbonne, & L. Schomaker (Eds.), *Proceedings of the 31st Annual conference of the Cognitive Science Society* (pp. 413-418). Cognitive Science Society.

Fodor, J. D., & Ferreira, F. (1998). *Reanalysis in Sentence Processing.* Kluwer. https://doi.org/10.1007/978-94-015-9070-9

Fodor, J. D., & Inoue, A. (1994). The Diagnosis and Cure of Garden Paths. *Journal of Psycholinguistic Research, 23,* 407-434. https://doi.org/10.1007/BF02143947

Fodor, J. D., & Inoue, A. (2000). Garden Path Reanalysis: Attach (Anyway) and Revision as Last Resort? In M. de Vincenzi, & V. Lombardo (Eds.), *Cross-Linguistic Perspective on Language Processing* (pp. 21-61). Kluwer. https://doi.org/10.1007/978-94-011-3949-6_2

Frazier, L., & Rayner, K. (1982). Making and Correcting Errors in Sentence Comprehension: Eye Movements in the Analysis of Structurally Ambiguous Sentences. *Cognitive Psychology, 14,* 178-210. https://doi.org/10.1016/0010-0285(82)90008-1

Hermena, E. W., Drieghe, D., Hellmuth, S., & Liversedge, S. P. (2015). Processing of Arabic Diacritical Marks: Phonological-Syntactic Disambiguation of Homographic Verbs and Visual Crowding Effects. *Journal of Experimental Psychology: Human Perception and Performance, 41,* 494-507. https://doi.org/10.1037/xhp0000032

Hermena, E., & Reichle, E. (2020). Insights from the Study of Arabic Reading. *Language and Linguistics Compass, 14,* 1-26. https://doi.org/10.1111/lnc3.12400

Ibrahim, R. (2013). Reading in Arabic: New Evidence for the Role of Vowels Diacritics. *Creative Education, 4,* 248-253. https://doi.org/10.4236/ce.2013.44036

Lin, C.-J., & Bever, T. (2011). Garden Path and the Comprehension of Head-Final Relative Clauses. In H. Yamashita et al. (Eds.), *Processing and Producing Head-Final Structures, Studies in Theoretical Psycholinguistics* (Vol. 38, pp. 277-297). Springer. https://doi.org/10.1007/978-90-481-9213-7_13

MacDonald, M., & Hsiao, Y. (2018) Sentence Comprehension (Chp 8). In S.-A. Rueschemeyer, & M. G. Gaskell (Eds.), *The Handbook of Psycholinguistics* (pp. 171-196). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198786825.013.8

Mahfoudhi, A. (2007). Roots and Patterns in Arabic Lexical Processing. In E. Benmamoun (Ed.), *Perspectives on Arabic linguistics XIX* (pp. 123-150). John Benjamins. https://doi.org/10.1075/cilt.289.11mah

Mahfoudhi, A., Elbeheri, G., Al-Rashidi, M., & Everatt, J. (2010). The Role of Morphological Awareness in Reading Comprehension among Typical and Learning Disabled

Native Arabic Speakers. *Journal of Learning Disabilities, 43,* 500-514. https://doi.org/10.1177/0022219409355478

Mahmoud, Y. (1979). *The Arabic Writing System and the Sociolinguistics of Orthographic Reform*. Unpublished Doctoral Dissertation, Georgetown University.

Mason, R., Just, M., Keller, T., & Carpenter, P. (2003). Ambiguity in the Brain: What Brain Imaging Reveals About the Processing of Syntactically Ambiguous Sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 1319-1338. https://doi.org/10.1037/0278-7393.29.6.1319

Mazuka, R., & Itoh, K. (1995). Can Japanese Speakers Be Led down the Garden Path? In R. Mazuka, & N. Nagai (Eds.), *Japanese Sentence Processing* (pp. 295-329). Lawrence Erlbaum Associates.

Mazuka, R., Itoh, K., & Kondo, T. (1997). Processing down the Garden Path in Japanese: Processing of Sentences with Lexical Homonyms. *Journal of Psycholinguistic Research, 26,* 207-228. https://doi.org/10.1023/A:1025013716381

McCarthy, J. (1979). *Formal Problems in Semitic Phonology and Morphology*. Doctoral Dissertation, Massachusetts Institute of Technology.

McCarthy, J. (1981). A Prosodic Theory of Nonconcatenative Morphology. *Linguistic Inquiry, 12,* 373-418.

Mitchell, D. C. (1994). Sentence parsing. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 375-409). Academic Press.

Mohammed, M. (2000). *Word Order, Agreement in Arabic*. John Benjamins B.V.

Roman, G., Pavard, B., & Asselah, B. (1985). Traitement perceptif des phrases ambigües en Arabe [Perceptive analysis of ambiguous Arabic sentences]. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition, 5,* 5-22.

Saiegh-Haddad, E. (2017). Learning to Read Arabic. In L. Verhoeven, & C. Perfetti (Eds.), *Learning to Read across Languages and Writing Systems* (pp. 104-126). Cambridge University Press. https://doi.org/10.1017/9781316155752.006

Seraye, A. M. (2004). *The Role of Short Vowels and Context in the Reading of Arabic, Comprehension, and Word Recognition of Highly Skilled Reader*. Unpublished Doctoral Dissertation, University of Pittsburgh. http://d-scholarship.pitt.edu/9821/

Seraye, A. M. (2016). Short Vowels versus Word Familiarity in the Reading Comprehension of Arab Readers: A Revisited Issue. *International Electronic Journal of Elementary Education, 8,* 481-506.

Shormani, M. M. (2015). Is Standard Arabic a VSO Language? Evidence from Syntax and Semantics. *Al-Qalam Journal, 3,* 1-49. https://doi.org/10.35695/1946-000-003-012

Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Underspecification of Syntactic Ambiguities: Evidence from Self-Paced Reading. *Memory & Cognition, 36,* 201-216. https://doi.org/10.3758/MC.36.1.201

Taft, M. (1981a). Lexical Access via an Orthography Code: The Basic Orthographic Syllabic Structure (BOSS). *Journal of Verbal Learning and Verbal Behavior, 18,* 21-39. https://doi.org/10.1016/S0022-5371(79)90544-9

Taft, M. (1981b). Prefix Stripping Revisited. *Journal of Verbal Learning and Verbal Behavior, 20,* 289-297. https://doi.org/10.1016/S0022-5371(81)90439-4

Taha, H. (2016). Deep and Shallow in Arabic Orthography: New Evidence from Reading Performance of Elementary School Native Arab Readers. *Writing Systems Research, 8,* 133-142. https://doi.org/10.1080/17586801.2015.1114910

Taha, H., & Saiegh-Haddad, E. (2016). The Role of Phonological versus Morphological

Skills in the Development of Arabic Spelling: An Intervention Study. *Journal of Psycholinguistic Research, 45,* 507-535. https://doi.org/10.1007/s10936-015-9362-6

Taha, H., & Saiegh-Haddad, E. (2017). Morphology and Spelling in Arabic: Development and Interface. *Journal of Psycholinguistic Research, 46,* 27-38. https://doi.org/10.1007/s10936-016-9425-3

Watson, J. (2000). Word Order in Arabic [Book Review]. *Journal of Semitic Studies, 45,* 389-391.

Wattad, H., & Abu-Rabia, S. (2020). The Advantage of Morphological Awareness among Normal and Dyslexic Native Arabic Readers: A Literature Review. *Reading Psychology, 41,* 130-156. https://doi.org/10.1080/02702711.2020.1768973

Wonnacott, E., Joseph, H., Adelman, J., & Nation, K. (2016). Is Children's Reading 'Good Enough'? Links between Online Processing and Comprehension as Children Read Syntactically Ambiguous Sentences. *Quarterly Journal of Experimental Psychology, 69,* 855-879. https://doi.org/10.1080/17470218.2015.1011176

Appendix

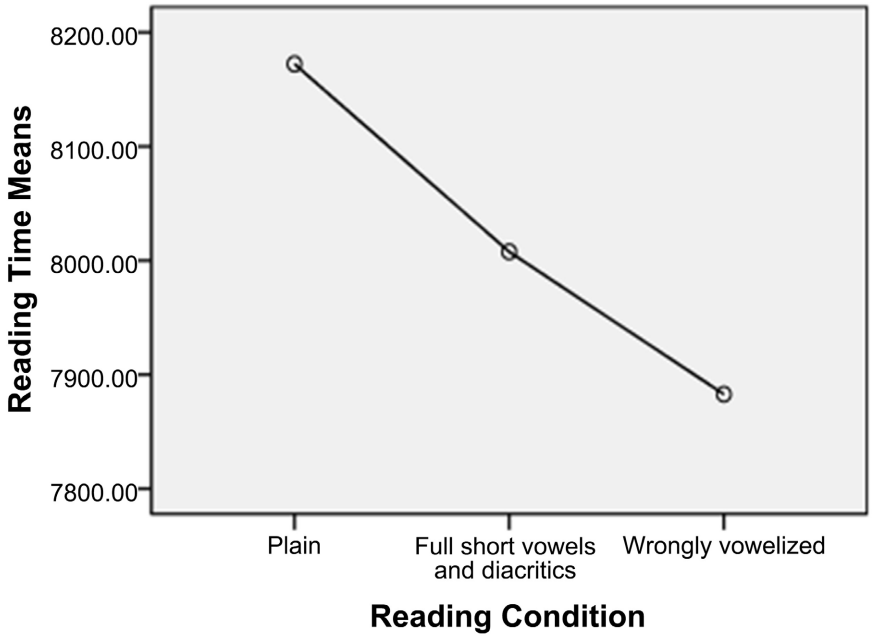

Figure A1. Overall means on reading time for GP sentences.

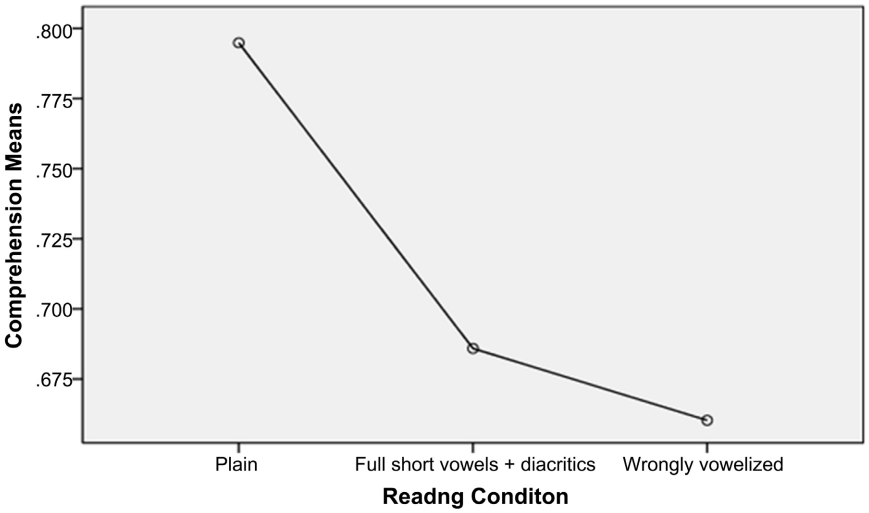

Figure A2. Overall means on reading comprehension for GP sentences.

The experimental items: Sentences and questions

Set I (RC1: Plain; and RC2: Vowels + diacritics, sentences)

Plain; and Vowels + diacritics, sentences	Questions/answers
Practice sentences (1 - 7)	
	هَلْ يُحِبُّ مُحَمَّدٌ أَنْ يَلْعَبَ بِالْكُرَةِ كُلَّ يَوْمٍ ؟ 1
	خَالِدٌ يُحِبُّ أَكْلَ التُّفَّاحَ؟ 0
	رَاشِدٌ اِبْتَعَدَ عَنْ بَلَدِهِ لِمُدَّةٍ قَصِيْرَةٍ؟ 0
	الإِكْتِئَارُ مِنْ عَصِيْرِ اللَّيْمُونِ يَضُرُّ بِالأَسْنَانِ؟ 0
1. مُحَمَّدٌ يُحِبُّ أَنْ يَلْعَبَ بِالْكُرَةِ كُلَّ يَوْمٍ.	شَاهَدَ عَبْدُاللَّهِ بَرْنَامَجاً عَنْ مَخَاطِرِ الحُرُوْبِ ؟ 0
2. خَالِدٌ لاَ يُحِبُّ أَكْلَ التُّفَّاحَ أَبَدَاً.	هَلْ حَصَلَ مُنْتَخَبُ كُرَةِ الطَّائِرَة عَلَى المَرْكَزِ الثَّانِي؟ 0
3. جَاءَ رَاشِدٌ مِنْ سَفَرِهِ بَعْدَ غِيابٍ طَوِيْلٍ.	هَلْ احترقت كتب خالد بسبب الحريق؟ 0
4. الإِكْتِئَارُ مِنْ عَصِيْرِ اللَّيْمُونِ لاَ يَضِرُّ بِالأَسْنَانِ.	الَّذِي رَسَمَ الصُّورَةَ الْجَمِيْلَةَ كَانَ تِلْمِيْذَاً صَغِيْرَاً؟ 0
5. شَاهَدَ عَبْدُاللَّهِ بَرْنَامَجاً تِلْفَازِيَّاً عَنْ أَضْرارِ التَّدْخِينِ.	الْبَيْعُ وَالشِّرَاءُ بِجَانِبِ مَبْنَى الْوَزَارَةِ مَمْنُوعٌ؟ 0
6. حَصَلَ مُنْتَخَبُ كُرَةِ الطَّائِرَة عَلَى المَرْكَزِ الثَّالِثِ.	اِسْتِقْدَامُ الْعَامِلاَتِ مِنْ أي بَلَدٍ أَمْرٌ يَسِيرٌ؟ 1
7. كتب خالد الموجودة في المكتبة لم تحترق.	أَتَى خَالِدٌ مِن السُّوقِ رَاكِبَاً سَيَّارَةً؟ 0
Plain sentences (8 - 19)	السَّهَرُ كَانَ مِنْ أَسْبَابِ مَرَضِ مُحَمَّدٍ؟ 1
	هَلْ ذَاكَرَ الإِبْنُ دُرُوسَهُ ؟ 0
8. الذي رسم الصورة الجميلة كان رجلا عجوزا.	خَوْفُ السُّكَّانِ كَانَ بِسَبَبِ صَوَاعِقَ شَدِيْدَةٍ؟ 0
9. إن البيع والشراء بجانب مبنى الوزراة مسموح.	سَعْدٌ هُوَ الْمُعَلِّمُ لِخَالِد؟ 0
10. استقدام العاملات من كل البلدان أصبح يسيرا.	البَائِعُ فَتَحَ مَحِلَّهُ فِيْ وَقْتِ الصَّبَاحِ وَلَكِن مُتَأَخِّراً؟ 0
11. جاء خالد من السوق ماشيا على قدميه.	أَحْمَدُ رَجُلٌ سَلِيْمٌ لاَ يَشْتَكِي مِنْ أَيِّ أَمْرَاضٍ؟ 0
12. من أسباب مرض محمد عدم نومه ليلا.	أَزْعَجَ الأَطْفَالُ السُّكَّانَ بِلَعِبِهِم الْكُرَة صَبَاحاً؟ 0
13. غضب الأب من ابنه لأنه لم يذاكر.	الطَّبِيْبُ كَشَفَ عَلَى خَالِدٍ مِنْ غَيْرِ مُسَاعَدَةِ أَحَدٍ؟ 0
14. حدث خوف للسكان بعد حصول الزلزال القوي.	هَلْ الَّذِي كَتَبَ الْقِصَّةَ كَانَ طِفْلا صَغِيْرَاً؟ 1
15. علم خالد أخاه سعدا كتابة حروف العربية.	اللَّعْبُ وَالسِّبَاقُ بِجَانِبِ مَبْنَى الْوَزَارَةِ مَمْنُوعٌ؟ 0
Garden-path sentences (16 - 19)	الإِخْتِبَارُ فِي مَادَّةِ الْقِرَاءَةِ كَانَ سَهْلاً؟ 0
	قَادَ صَالِحٌ السَّيَّارَةَ بِسُرْعَةٍ؟ 1
16. فتح البائع محله في الصباح كان مبكرا GP.	خَالِدٌ لاَ يُحِبُّ أَكْلَ الْخُضْرَوَاتِ؟ 1
17. مرض أحمد بمرض السكري كان لكثرة جلوسه GP.	اِبْتَسَمَ الأَبُ لابْنِهِ لأَنَّهُ كَانَ يَسْتَمْتِعَ بِلُعْبَتِهِ؟ 0
18. لعب الأطفال بالكرة صباحا لم يزعج السكان GP.	هُرُوبُ رِجَالِ الإِطْفَاءِ كَانَ قَبْلَ سُقُوطِ الجِدَارُ؟ 0
19. كشف محمد مرض خالد كان بمساعدة الممرضة GP.	فَهْدٌ يَعْرِفُ الإِنْجِلِيْزِيَّةَ أَكْثَرَ مِنْ أَحْمَدَ؟ 0
Vowelized-diacritized sentences (20 - 31)	خَالِدٌ يَعْرِفُ كَيْفَ يَرْبُطُ أَسْلاَكَ الْكَهْرَبَاءِ؟ 0
	لَمْ يَجِدْ أَحْمَدُ مَفَاتِيْحَ الْمَحَلِ, فَقَامَ بِكَسْرِ الْقُفْلِ؟ 1
20. الَّذِي كَتَبَ الْقِصَّة الْجَمِيْلَة كَان طِفْلاً صَغِيْرا.	لاَ يَسْتَطِيْعُ أَحْمَدُ تَنَاوُلَ عَصِيْرِ الْبُرْتُقَالِ الطَّازِجِ؟ 1
21. إِنَّ اللَّعِب وَالسِّبَاق بِجَانِب مَبْنَى الْوَزَارَة مَسْمُوح.	حُصُولِ خَالِدٍ عَلَى الْمَالِ الْكَثِيْرِ كَانَ مِنْ عَمَلِهِ؟ 0
22. اِخْتِبَار الطُّلاَّب فِي مَادَّة الْقِرَاءَة كَان صَعْبا.	
23. جَاء صَالِح مِن الْعَمَل مُسْرِعا فِي سَيَّارَته.	
24. مِن أَسْبَاب مَرَض خَالِد عَدَم أَكْله الْخُضْرَوَات.	
25. صَرَخ الأَب عَلَى اِبْنه لأَنَّه لَم يَسْتَمِع.	
26. فَرّ رِجَال الإِطْفَاء بَعْد سُقُوط الْجِدَار الطَّوِيْل.	
27. دَرَّس فَهْد صَدِيْقه أَحْمَد كِتَابَة اِسْمِه بِالإِنْجِلِيْزِيَّة.	
Potentially Garden-path sentences (28 - 31)	
28. رَبْط خَالِد أَسْلاَك الْكَهْرَبَاء كَان خَطَأ كَبِيْرا Potentially GP.	
29. كَسْر أَحْمَد قُفْل الْمَحَلّ كَان لِضَيَاع الْمِفْتَاح Potentially GP.	
30. شُرْب عَصِيْر الْبُرْتُقَال الطَّازِج مَمْنُوع عَلَى أَحْمَد Potentially GP.	
31. جَمْع خَالِد الْمَال الْكَثِيْر كَان مِن وَالِدَيه Potentially GP.	

Set II (RC3: Wrongly vowelized sentences)

Wrongly vowelized sentences	Questions/answers
Practice sentence (1)	
	مَنْزِلُ سَعِيْدٌ مَنْزِلٌ صَغِيْرٌ وَبَاعَهُ بِسِعْرٍ رَخِيْصٍ؟
	0
	خَالِدٌ اِشْتَرَى لُعْبَةً غَيْرَ مَكْسُورَةٍ؟
	0
	الشُّرْبُ وَالأَكْلُ فِيْ الْحَدِيْقَةِ مَمْنُوعٌ؟
	0
	أَدَاءُ خَالِدٍ فِيْ مَادَّةِ التَّارِيْخِ كَانَ مُمْتَازَاً؟
	0
	جَاءَ فَهْدٌ مِن الْحَدِيْقَةِ مَاشِيَاً عَلَى قَدَمَيْهِ؟
	0
	هَلْ أَحْمَدُ يَكْرَهُ أَكْلَ اللَّحْمِ؟
	1
	عَدَمُ مَجِيْئِ الْمُوَظَّفِ إِلَى الْعَمَلِ أَدَّى إِلَى فَصْلِهِ؟
	1
	الْغُبَارُ كَانَ بِسَبَبِ سُقُوطِ الْمَبْنَى الْكَبِيْرِ؟
	1
	اِسْتَلَمَ مُحَمَّدٌ الرِّسَالَةَ مِن الْمُوَظَّفِ؟
	0
	سَعْدٌ لاَ يَرْكُضُ فِي صَالَةِ الْبَيْتِ؟
	0
	الْمُدِيْرُ كَانَ يُعَامِلُ العَامِلَ بِلُطْفٍ وَإِحْسَانٍ؟
	0
	سَعْدٌ لاَ يَسْتَطِيْعُ أَنْ يَأْكُلَ الأَسْمَاكَ إِذَا أَرَادَ؟
	1
	سَعْدٌ مَعْرُوفٌ بِأَمَانَتِهِ؟
	1
Wrongly vowelized sentences (2 - 9)	
1. بُاع سْعِيد مُنزُله الكُبَيَر بَسَعَر رُخُيص جُدًا.	
2. الِذِي بُاع خُالُدا اللَعَبُة الِمِكِسَورُة كُان كُاذُبا.	
3. إنُ الأكَل والشَّرُب فُي الُحَدُيقُة غُيُر مُمُنَوع.	
4. أُهُمُال خُالُد فُي مُادُة الُتارُيُخ كُان واضُحا.	
5. جُاء فهُد مُن الُحَدُيقُة راكِبًا عُلُى دِرُاجُتِه.	
6. مُن أسِبِاب مرُض أُحُمُد عِدَم أكَلِه الُلُحِم.	
7. فِصِل مَدُير الشِرُكُة الَمَوظُف لأنه لُم يِحِضِر.	
8. حِصِل غِبُار لَلنِاس بُعُد سِقُوط الُمُبَنى الُكُبير.	
9. سِلِم مَحُمُد رُسِالة الُمِدُير الِمِهِمة إلِى الَمَوظف.	
Potentially garden-path (10 - 13)	
10. رِكِض سُعَد فُي صُالُة الَبِيُت كُان مَزُعِجا Potentially GP.	
11. ظِلِم العُامُل فُي الشِرِكة كُان بَسِبُب الُمِدُير Potentially GP.	
12. أِكل سُعَد الأسِمِاك البُحُرَية غُيَر مُسِمَوح لِه Potentially GP.	
13. كِسِب سُعَد السَمِعُة الَحِسِنُة كُان بَسُبُب أِمانُته Potentially GP.	

Note: GP= garden-path sentence; Potential GP = if only the consonants of the initial HP-HG word of the sentence are assembled during reading, the sentence should often garden path the reader.