Perceptual load in the reading of Arabic: Effects of orthographic visual complexity on detection

Souad Abdelhadi
Learning Disabilities Department, University of Haifa

Raphiq Ibrahim
The Edmond J. Safra Brain Research Center for the Study of Learning Disabilities and Learning Disabilities Department, University of Haifa

Zohar Eviatar
Psychology Department and Institute for Information Processing and Decision Making, University of Haifa

Abstract

Previous research has suggested that reading Arabic is slower than reading Hebrew or English, even among native Arabic readers. We tested the hypothesis that at least part of the difficulty in reading Arabic is due to the visual complexity of Arabic orthography. Third- and sixth-grade native readers of Arabic who were studying Hebrew in school were asked to detect a vowel diacritic in the context of Hebrew words and nonwords, Arabic words and nonwords (including connected and unconnected Arabic letters), and nonletter stimuli that resembled Arabic or Hebrew letters. Participants were better at detecting target vowels in Hebrew than in any of the Arabic conditions. Moreover, target detection in Arabic was better for letter strings containing connected letters than for those containing unconnected letters. The findings extend previous results on Hebrew versus Arabic reading and support a perceptual load account of the source of processing difficulty in reading Arabic. Performance in the Arabic conditions did not reveal a word superiority effect, suggesting that even by sixth grade, reading is not automatized to the point where it can compensate for the the visual complexity of the orthography.

Correspondence:
Zohar Eviatar,
Psychology Department
and IIPDM,
University of Haifa,
Israel.
E-mail:
zohare@research.haifa.ac.il

Although the process of reading is known to be influenced by linguistic and by visuo-perceptual factors, the latter have not received as much attention until recently, thanks to findings from neuroimaging and other studies that suggest that a specific region of the visual cortex is responsive to written words (e.g. Cohen and Dehaene, 2004). Behavioral studies have also shown that the visual characteristics of words, such as the direction in which they are written or read (Lubow *et al.*, 1994; Eviatar, 1995; Vaid, 1988), word length (Aghbabia and Nazir, 2000), and their orthographic and morphological structure (Farid and Grainger 1996; Deutsch and Rayner, 1999), affect the way in which words are perceived.

Writing Systems Research, Vol. 3, No. 2, 2011, 117–127.

doi:10.1093/wsr/wsr014

On the basis of these findings, it is notable that quite a large number of experimental studies on reading in Arabic have reported that it is slower than reading in Hebrew, English, French, and Serbo-Croatian (Azzam, 1984; Frost *et al.*, 1987; Katz and Frost, 1992; Roman and Pavard, 1987; Bentin and Ibrahim, 1996; Eviatar and Ibrahim, 2004; Ibrahim *et al.*, 2007).

A number of studies have attempted to examine the possible sources of slowness in both reading acquisition and skilled reading of Arabic, in relation to other languages. The first source is related to the fact that the literate Arabic speaker is a bilingual *de facto* (e.g. Eviatar and Ibrahim, 2000). There is a diglossic situation in Arabic, referring to the existence of two forms/systems of the same language (Ferguson, 1959). Ammia—Spoken Arabic (SA) has no written form, while the written form, Fuṣḥa—Literary Arabic, more commonly referred to as 'Modern Standard Arabic' (MSA), is taught in schools in parallel to learning how to read and write. MSA is universally used in the Arab world for formal communication and writing and has become part of everyday life, as it is the language in which news is reported (both written and oral) and it is the language of prayer and of formal public occasions.

The second source focuses on the relationship between the specific characteristics of the Arabic orthographic system and cognitive processes that might be involved during word recognition and the acquisition of reading (e.g. Ibrahim *et al*., 2002; Eviatar *et al*., 2004; Ibrahim and Aharon-Peretz, 2005). There are two separate aspects of this relationship that may or may not be related. The first aspect is orthographic depth. This concept has to do with the relationship between letters and the sounds they represent (Katz and Frost, 1992). Orthographies in which this relationship is straighforward (such as Spanish) are considered 'shallow', whereas orthographies in which it is not (such as English), are considered 'deep'. The second aspect of the orthographic system that can affect reading processes is the visual complexity of the letters themselves. Recently, a study by Rao *et al*. (2011) examined the effects of both orthographic depth and visual complexity in Urdu and Hindi. They measured speed and accuracy of reading single words in Urdu (in which the deep orthography is based upon a modification of Perso-Arabic script), and in Hindi, which uses a shallower, and less visually complex orthography (see below for definitions of orthographic complexity), in Urdu–Hindi adult bilinguals. They report that despite the fact that Urdu was the participants' native language and the language in which most of their schooling took place, responses to Urdu were consistently slower and more error prone than for Hindi. These authors suggested that this is due not only to the differences in orthographic depth in the two languages, but also because Urdu is visually more complex than Hindi. Arabic and Hebrew allow us to disentangle orthographic depth and visual complexity. In terms of orthographic depth, they are similar: Arabic and Hebrew have two versions of orthography (see below for more details). Vowelled Arabic and Hebrew texts are shallow, in the sense that all of the phonological information necessary for reading is represented in the text. Unvowelled Arabic and Hebrew texts are deep, because information about vowels must be inferred from contextual and lexical representations. In terms of visual complexity, Arabic and Hebrew are different, as letter shapes and word forms are visually more complex in Arabic than in Hebrew. In the research presented below, we explored the effects of this visual complexity on a very basic task: the detection of a vowel diacritic in the shallow form of the two orthographies.

1 Diglossia

Although sharing a limited subgroup of words, the two forms of Arabic are semantically, phonologically, and syntactically different. For example: the word balcony in English is برندة /bˆra:ndˆ/ in SA, while it is شرفة /ʃorfˆ/ in MSA. Ibrahim (e.g. Bentin and Ibrahim, 1996; Ibrahim and Aharon-Peretz, 2005) has addressed this issue directly, and has shown that the two forms of Arabic function as two separate language systems, such that a literate Arabic speaker is essentially bilingual. Eviatar and Ibrahim (2000) showed that young Arab children who have been exposed to Literary Arabic function as bilinguals on tests of meta-linguistic awareness.

Saiegh-Haddad (2003, 2004, 2005) has shown that phonological differences between MSA and SA affect reading acquisition. In addition, when Ibrahim *et al.* (2007) compared reading measures in Arab, Hebrew monolingual, and Hebrew–Russian bilingual first-grade children, they found that Hebrew–Russian bilinguals and the children learning to read Arabic revealed better performance on tests of phonological awareness than monolingual Hebrew speakers. However, text reading revealed a significant effect of the language on reading performance, with children learning to read Hebrew (monolingual and Russian–Hebrew bilinguals) outperforming the children learning to read Arabic. Phonological ability predicted reading performance differently among the language groups: over 60% of the variance in text reading (speed and accuracy) was predicted by phonological tests for children learning to read Hebrew (both monolinguals and bilinguals), and only 30% for Arabic readers. Thus, even though the Arab children had higher scores than monolinguals on tests of phonological awareness, those abilities did not facilitate text reading performance. The authors suggested that this could be due to both the diglossic situation and aspects related to the orthography complexity of Arabic.

2 **Orthographic Complexity in Hebrew and Arabic**

In Hebrew and Arabic, all verbs and most nouns are written primarily as consonantal roots that are differently affixed and vowelled to form the words of the lexicon (Berman, 1978). Most written materials in both languages do not include vowels. When vowels do appear (in poetry, children's books and liturgical texts), they are signified by diacritical marks above, below, or within the body of the word. Inclusion of these marks completely specifies the phonological form of the orthographic string, making it transparent (shallow) in terms of orthography/phonology relations. The effects of the omission of vowels on skilled reading in Hebrew has been shown in an interesting study by Shimron and Sivan (1994), who found that adult Hebrew–English bilinguals read text more quickly in English than in unvowelled Hebrew, but not more quickly than Hebrew text that includes the vowel diacritics. Thus, in Hebrew, even though addition of vowels results in a somewhat more complex visual form of the text, it facilitates both the speed and the comprehension of reading.

In Arabic, there are three diacritics signifying short vowels: two are positioned above the letter: *fatHa* <َ>=a, *damme* <ُ>=u, and one is positioned below: *kasra* <ِ>=i. Although these vowels are not letters, their combinations with consonants form CV syllables. In addition, there are *double fatHa* <ً>='an,'; *double damme* <ٌ>='on,'; and *double kasra* <ٍ>='in.' These vowel signs also have a syntactic role as they are used to mark indefinite subjects (e.g. subjects that in English would not be preceded by 'the'). For example: *fatHa* or *double fatHa* on the last letter of a word signifies it as the object of the sentence, while *damme* or *double damme* signifies it as the subject of the sentence. In addition to the diacritics for short vowels, there are four other reading signs: the *skoon* <ْ> which signals absence of a vowel, *shada* <ّ> which signals doubling of a consonant, *maddah* <~> which signals doubling of the letter alif and *hamzeh* <أ> which signals the glottal-stop sound.

In Hebrew, the diactrical marks are dots and strokes that are usually positioned below the letters, and sometimes above or in the middle of the letters. Historically there are different diacritics in Hebrew for long and short vowels. However, because modern Hebrew has not retained the long vowels, the phoneme <a> is represented by the *kamatz* <ָ> and *patax* <ַ>, the phoneme <e> is represented by *tsere* <ֵ> and *segol* <ֶ>, and the phoneme <i> represented by *xeereek* <ִ>.

There are three letters in Arabic (ي ,و ,ا) and four in Hebrew (י, ו, ה, א) which, in addition to signifying specific consonants, also specify long vowels. Thus it can be difficult for the reader to determine whether these dual-function letters represent a vowel or a consonant.

Two additional factors add to the complexity, particularly of Arabic. The first has to do with the role of dots. In Hebrew, dots are a diacritic used as a stress-marking device (*dagesh*). This stress-marking device (which does not appear in the unvowelized script) changes the phonemic representation of the letters from fricatives (v, x, f) to stops (b, k, p for the letters ב ק פ, respectively). However, בּ and ב represent the same letter, which has two phonemic representations. In the unvowelized form of Hebrew script, these different phonemes can be disambiguated by their place in the word, as only word and syllable initial placement of the letters indicate the stop consonant, otherwise they are pronounced as fricatives.

In Arabic, dots in themselves do not have phonetic value but comprise an integral part of many letters. For example, letters having a similar or even identical structure are distinguished on the basis of the existence, location, and number of dots (e.g. the Arabic letters for /n/ /θ/ /t/ and /b/ are represented by the graphemes: ب ث ت ن, respectively).

An additional characteristic that contributes to the complexity of the two orthographies is that many letters are represented by different shapes, depending on their placement in the word. In Hebrew, there are five letters that change shape when they are word final: (ם-מ, ן-נ, ף-פ, ך-כ, ץ-צ). In Arabic, 23 of the 29 letters in the alphabet have four shapes each (word initial, medial, final, and when they follow a nonconnecting letter. For example, the phoneme /h/ is represented by <ﻫ> when it is word-initial, by <ﻬ> when it is in the middle of the word, by <ﻪ>, when it is word final, and <ﻩ> when the letter is preceded by a nonconnecting letter. The remaining six letters have two shapes each, final and separate. Thus, the grapheme–phoneme relations are quite complex in Arabic, with similar graphemes representing quite different phonemes, and different graphemes representing the same phoneme.

3 The Present Research

Thus, although it is clear that phonology plays an important role in the reading of Arabic, as it does in English, a number of studies (e.g. Abu-Rabia, 1996, 1998) have suggested that morphology, syntax, and orthographic abilities play a large role in reading in Arabic. In fact, even though the largest deficit found to date in Arabic reading disabled children is in phonological measures, Abu-Rabia *et al.* (2003) suggest that "… the visual-orthographic reading route is still probably the most effective reading strategy due to the orthographic nature of the Arabic language". (Abu-Rabia *et al.*, 2003, p. 437). Therefore, the goal of the present research is to explore visual and orthographic processing in good readers of Arabic, at different levels of skill.

The hypothesis tested in this study is that one of the reasons for slowness in reading acquisition in Arabic in normal readers is due in part to the large perceptual load in vowelled text. We used a detection task in Hebrew and Arabic in which the target was a vowel diacritic (*fatHa* in Arabic stimuli and *patax* in Hebrew stimuli). The target appeared in three conditions: in a word, in a nonword, and among a series of shapes patterned to look like letters. The particpants were native Arabic speakers in third and sixth grades. The language of schooling is Arabic, with Hebrew and English learned as foreign languages. Arab children begin to learn to read Hebrew in third grade.

We tested whether detection of an identical target (a small horizontal line) would be easier in the context of Hebrew than in Arabic. In addition, we manipulated the visual complexity of the Arabic stimuli by asking participants to detect the line when it was presented in the context of words, pseudowords, and nonletter stimuli, to examine how this can affect perception for children at different levels of reading skill.

4 Method

4.1 Participants

Forty 6th graders (19 boys and 21 girls, mean age 11.5) and forty-two 3rd graders (22 boys and 20 girls, mean age 8.6) who are native Arabic speakers participated in this study. All children were recruited from an elementary school in which Arabic is the official language. The school is in a village located in the lower Galilee, with middle socioeconomic status.

The participants began learning to read Arabic in first grade and Hebrew in third grade. All began to learn to speak Hebrew in second grade. The third graders typically read both Arabic and Hebrew words with vowel diacritics. Diacritics in Arabic are gradually phased out during fifth grade, but sixth graders are still reading Hebrew with vowels. Since the children study most of the lessons in their native language (Arabic), the exposure to written Arabic is higher than the exposure to written Hebrew for all of the participants.

All the third and sixth graders in the school were tested, but only the results of 'good readers' in both languages were considered. These were chosen by their teachers: we asked the teachers of the classes (both Arabic and Hebrew) to identify students who read well, not too slowly, and do not make many errors. None suffered from developmental or acquired neurological, learning, emotional, or attention disorders.

4.2 Stimuli and materials

The stimuli were created with three levels of lexicality: real words, orthographically legal pseudowords, and nonlinguistic stimuli– letter-like patterns of matched lines and curves. These were scanned for presentation on a laptop screen. The stimuli were presented in the center of the black screen in gray (680 pixels×512 pixels for each stimulus). Times New Roman font 150 was used for Arabic stimuli. Tahoma font 150 was used for Hebrew stimuli.

All stimuli were strings of three letters or semi-letters (patterns that look similar to letters) and diacritical marks. Most short vowels from the two languages were included. The diacritics were presented beneath and/or above the letters according to the acceptable rules in Arabic and beneath the letters according to the acceptable rules in Hebrew orthography.

Arabic stimuli differed in their complexity as follows: (1) 'Simple'—Stimuli constructed from letters that do not connect to other letters, and do not have dots as an integral part (e.g. دُرَرٌ /dur^run/); (2) 'Connected'—Stimuli constructed from letters that do connect to others, but do not have dots as an integral part (e.g. مُعْدٍ /muʕdIn/); (3) 'Complex'—Stimuli constructed from connecting letters that do include dots as an integral part (e.g. نَتَجَ /nətəʒ^/).

In Hebrew, there are no connecting letters and all dots are diacritics. Thus all the Hebrew stimuli are perceptually equivalent to the Arabic 'simple' stimuli.

Table 1 Examples of stimuli in the Arabic conditions with transliterations for words and pseudowords

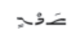
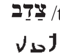
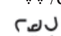
Arabic stimuli: lexicality levels	Orthography groups	With (´) diacritic	Without (´) diacritic
Real words	(1) simple	دُرَرٌ /dur^run/	إرْمِ /ˁrmI/
	(2) connected	مَطَرٌ /m^tˁ^run/	مُعْدٍ /muʕdIn/
	(3) complex	نَتَجَ /nətəʒ^/	خُبْزٌ /xubzun/
Pseudowords	(4) simple	وَرَحٌ /wərəħun/	وِدْمٌ /wIdmun/
	(5) connected	عَسْمٌ /ʕəsmun/	لُكدٌ /lukdun/
	(6) complex	بَشْيَ /bəʃjə/	فُخِض /fuxIdˁ/
Nonletters	(7) simple		
	(8) connected		
	(9) complex		

Table 2 Examples of stimuli from the Hebrew conditions with transliterations for words and pseudowords

Stimuli in Hebrew	With (ַ)	Without (ַ)
Real words	שִׂיחַ /sej^ħ/	עֶרֶב /ʕˁrIv/
pseudowords	צַדַב /ts^dˁv/	דֶרֶר /dˁrIr/
Nonletters		

The target diacritic, the *fatha* (´) existed in half of the stimuli in Arabic. The *patax* (ַ) existed in half of the stimuli in Hebrew stimuli.

The Arabic conditions used a 3×3 within subjects factorial design: three levels of lexicality (real words, pseudowords, nonletter stimuli)×three levels of complexity (simple, connected, and complex). Each condition consisted of ten stimuli with *fatha* and ten without *fatha.* Hence, there were 180 stimuli in Arabic.

The experiment in Hebrew had a single factor design with three levels of lexicality (real words, pseudowords, nonletters stimuli). Each condition consisted of ten stimuli with *patax* and ten without *patax.* Hence, there were sixty stimuli in Hebrew.

The real words in Arabic and Hebrew were chosen based upon two pretests based on frequencies. The first pretest was determined by fifty tenth-grade readers on a scale from 1 (lowest frequency) to 7 (highest frequency). Words rated as having a frequency of 4.5–5.5 were chosen.

The words and pseudo words were presented in a lexical decision task to nineteen native Arabic-speaking university students. Real words and pseudowords were chosen in such a way that the mean reaction times to recognize them did not differ.

4.3 Procedure

Participants were tested individually in a quiet room at their school during regular school hours. The instructions were given in spoken Arabic. The participant sat on a chair ~50-cm away from the computer screen. In the first set, all the Arabic stimuli (180 stimuli) appeared one by one on a computer screen randomly from the nine groups of Arabic stimuli, with a break after sixty stimuli (total of two breaks). The length of the breaks was not controlled.

The task of the participants was to detect the *fatha* (َ) by pressing the right square parentheses ']' if they see *fatha* in the stimulus in the center of the screen, and by pressing the left square parentheses '[' if the *fatha* does not appear in the presented stimulus. This session was about 15 min long.

In the second set, all the Hebrew stimuli (sixty stimuli) appeared one by one on a computer screen randomly from the three groups. Children were asked to detect the *patax* (ַ). This session was conducted without a break and lasted about 5 min.

In both sets (Arabic and Hebrew), participants were requested to use either their dominant hand or both hands for pressing the buttons. We measured the time between the appearance of the stimulus and the response on the keyboard. All participants completed the two sessions in the same order.

A training session preceded each set to verify that the child understood the task. Twenty trials were given before the set of Arabic stimuli and fourteen trials prior to the set of the stimuli in Hebrew. The stimuli for the practice trials were selected randomly from the different categories. During the practice trials, feedback was provided following each response. During the training session, the participant could ask any question and when necessary, the task was explained again. However, no feedback was given during the experimental trials and the participant was not allowed to talk.

5 Results

Accuracy was high in both grades (third grade errors = 8.9%; sixth-grade errors = 6.13%). The correlations between median RT and percent error revealed no speed–accuracy tradeoffs. Sensitivity was calculated as the signal detection measure d′. This is the difference between the *z*-scores for the probability of hits (the vowel was there and the child pressed 'yes') and for false alarms (FA, the vowel was not there and the child pressed 'yes'). Higher values reflect better performance. We used correction computations for probability values of 1, which was changed to $1–1/(2N)$, and of 0, which was changed to $1/(2N)$, based on the suggestions of Macmillan and Creelman (1991).

5.1 Vowel perception in Arabic

To analyze the results in Arabic, we used a mixed ANOVA with grade (third versus sixth) as a between groups factor, and lexicality and complexity as within groups factors. Latency (means of median response time for correct responses), percent error, and sensitivity (d′) were dependent variables.

The RT and error data show similar patterns: a significant two-way interaction between complexity and lexicality [RT, $F(4,396) = 6.95$, $p < .001$; percent errors, $F(4,396) = 4.93$, $p < .001$]. This interaction is shown in the top row of Fig. 1. It can be seen that in RT, this is due to a similar pattern for words and nonwords, and a different pattern for nonletter stimuli. For errors, it can be seen that there is a different pattern for each stimulus type: for words, the least number of errors were made for complex stimuli, whereas for nonletter stimuli, the least number of errors were made for connected stimuli. In addition, both measures reveal a main effect of grade that is significant in RT and marginal in errors [RT, $F(1,79) = 32.13$, $p < .0001$; percent errors; $F(1,79) = 3.03$, $p = .080$], with sixth graders responding faster and more accurately than third graders, and a significant main effect of complexity [RT, $F(2,158) = 5.10$,

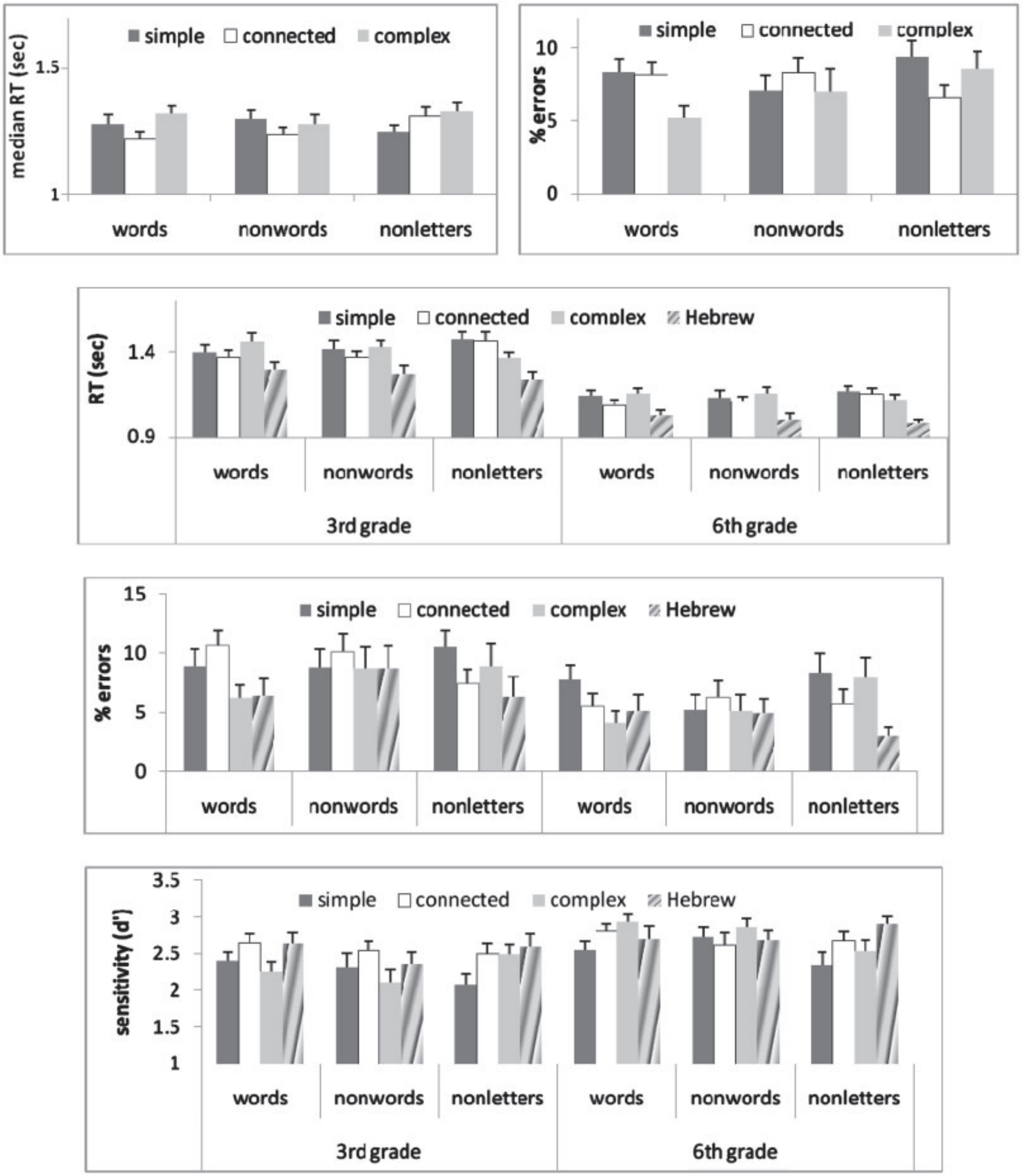


Fig. 1 Detection measures of the target in different levels of orthographic complexity in Arabic and in Hebrew. The top panels depict the lexicality by complexity interaction in the Arabic stimuli in median RTs and error percentages. The lower panels depict cell means in the Arabic and Hebrew conditions. Responses to Hebrew are striped. Error bars are standard errors.

$p<.010$; % errors; $F(2,158)=3.92$, $p<.050$]. In RT, connected stimuli were responded to most quickly (1.257 s versus 1.288 s for the complex and 1.294 s for simple stimuli). In errors, participants made the least number of errors on complex stimuli (6.87%) and more on simple (8.28%) and connected stimuli (7.67%). No other effects or interactions were significant. These effects can be seen in the middle panels of Fig. 1 (responses to Arabic stimuli are represented in the solid fill bars). It can be seen that there was no word superiority effect in either RT or error scores, hence, lexical information did not facilitate perception of the *fatha* in real words as compared with meaningless stimuli.

The pattern of results for sensitivity scores was different. First, we tested to see whether sensitivity was better than chance by computing one-sample *t*-tests on the d′ scores. All of the participants showed a pattern indicating sensitivity to vowels and better-than-chance performance in all of the conditions. ANOVA of the d′ scores revealed a significant three-way interaction between grade, lexicality, and complexity [$F(4,316)=2.79$, $p<.05$]. This pattern is shown in the bottom panel of Fig. 1. In addition, the d′ scores revealed a significant complexity by grade interaction, $F(2,158)=5.65$, $p<.05$, a main effect of complexity, $F(2,158)=10.37$, $p<.001$, and a main effect of grade, $F(1,79)=4.72$, $p<.05$.

It can be seen that the three-way interaction in the d′ scores is due to differences between the grades in the effects of complexity for the different types of stimuli. For sixth graders, there is no interaction between complexity and lexicality ($p>.100$), with complexity showing a significant main effect, $F(2,76)=7.39$, $p<.005$ [where complex and connected stimuli were responded to with greater efficiency than simple stimuli over all levels of lexicality (words, nonwords, and nonletters)], and lexicality showing a marginal main effect, $F(2,76)=3.01$, $p=.055$ (where words and nonwords were responded to with greater sensitivity than nonletter stimuli).

Third graders reveal a significant interaction between complexity and lexicality, $F(2, 164)=2.79$, $p<.050$, a significant main effect of complexity, $F(2,82)=8.46$, $p<.0001$, and no effect of lexicality ($p>.400$). Planned comparisons revealed that the lexicality of the stimulus behaved differently at the different levels of complexity: when the stimuli were complex, third graders were best at detecting the target when the stimli were not letters; when the stimuli were connected, there was no difference between words, nonwords, and nonletter stimuli; and for simple stimuli, the target was detected better in words than in nonletter stimuli ($p<.05$), but not better in words than in nonwords In both grades, sensitivity scores do not show a word superiority effect at any level of complexity.

5.2 Vowel perception in Hebrew

The median reaction times percent errors, and sensitivity scores from the Hebrew conditions were analyzed with a mixed ANOVA, using grade as a between groups factor and lexicality as a within groups factor. The analyses revealed the patterns illustrated in the striped bars in Fig. 1. No measure revealed an interaction between grade and lexicality. In addition, although the main effect of lexicality was significant in errors [$F(2,158)=4.42$, $p>.05$] and in d′ [$F(2,158)=3.52$, $p<.05$], and marginal in RT ($p=.080$), it can be seen that there is no word superiority effect. The main effect of grade is significant only in RT [$F(1,82)=28.99$, $p<.0001$ (with sixth graders (1.007 s) responding faster than third graders (1.267 s)].

5.3 Comparing Hebrew with Arabic

We compared the children's performance in Hebrew[1] with each of the levels of complexity in the Arabic conditions. Perusal of Fig. 1 reveals that for RT, responses to Hebrew stimuli were always faster than to Arabic stimuli, in both grades: for simple stimuli, $F(1,79)=77.37$, $p<.0001$; for connected stimuli, $F(1,79)=45.31$, $p<.0001$; for complex stimuli, $F(1,79)=54.32$, $p<.0001$. It can also be seen that students made less or equal numbers of errors in Hebrew compared to the best category in Arabic, with the same pattern shown for d′. For simple stimuli, in errors, $F(1,79)=7.12$, $p<.01$, in d′, $F(1,79)=7.46$, $p<.01$. For connected stimuli, in errors, $F(1,79)=5.47$, $p<.01$, d′, ns; for complex stimuli, the differences between Hebrew and Arabic are not significant in errors and d′.

6 Discussion

The goal of our experiment was to examine the effects of the visual complexity of Arabic orthography. Our hypothesis was that this complexity results in a high perceptual load, contributing to the difficulty and slowness of processing in reading in Arabic. We used a simple detection task, while manipulating the lexicality of the stimulus (letter strings that are meaningful versus those that are not, and strings of nonletters). The results revealed no word superiority effect and an interesting effect of complexity.

We found a main effect of grade in all of the measures in Arabic, with sixth graders responding significantly faster and more accurately than third graders. Interestingly, in Hebrew, the advantage of the sixth graders is seen only in RT. In general, the advantage of the sixth graders may be accounted for by natural developmental differences, but we think that this difference is less distinct in Hebrew because of differences in the extent of exposure to Hebrew and Arabic between third and sixth graders. The gap in exposure to the Hebrew writing system is smaller than for Arabic.

In Arabic, we found a significant effect of complexity in all dependent variables. The particular pattern found suggests that some very low-level reading mechanisms were engaged to the extent that visual complexity itself was not the sole or even determining factor in the pattern of responses. Recall that we had defined stimuli of the simple type, such as دُرَرٌ /durˆrun/, as less complex than stimuli of the connected type, such as مَطَرٌ /mˆtˤˆrun/, because the former are comprised of three unconnected letters, whereas the latter are comprised of connected letters. We had assumed that unconnected stimuli are visually less complex than connected stimuli. However, we found that children responded fastest to the connected stimuli when they were letters, and most accurately to the connected nonletter stimuli (see the top panels of Fig. 1). The explanation for this finding might be the nature of Arabic orthography. In Arabic, the majority of letters must be connected to their neighbors from both sides (right and left), except for six letters (ـذ ـر ـز ـو ا ـد). The unique aspect of these six letters is the fact that they can only be connected from their right side. There are no letters in Arabic that cannot be connected from both sides; therefore, in order to compose words with disconnected letters, all letters in the words have to be from the set of six letters which cannot be connected from the left. Thus, most words in the language contain at least some connected letters, with letter strings composed of separate letters being very infrequent. It may be that what we are seeing is an effect of orthographic frequency. The participants were detecting the target more efficiently when the global aspect of the stimulus was most similar to the majority of words in Arabic. However, this effect must be at a very low level, because real words did not facilitate detection of the target (connected letter-like nonsense shapes were also more facilitating than nonconnected shapes), nor was there an effect of lexical frequency—the target was not detected more efficiently in the context of real words.

In addition, the fact that responses were equally or more efficient in Hebrew than in Arabic also supports the hypothesis that reading mechanisms were engaged at a very low level or not at all. These results complement those reported by Rao *et al.* (2011), where native Urdu readers read Urdu more slowly than Hindi. These authors suggested that this is due to two factors: orthographic depth, where the relations between graphemes and phonemes are more regular in Hindi than in Urdu (which has a consonantal script, like Arabic and Hebrew); and the greater visual complexity of Urdu orthography than Hindi orthography. Our results suggest that in elementary school children, visual complexity has an effect over and above orthographic depth. This is because in terms of orthographic depth, Arabic and Hebrew are similar. Both orthographies have an unvowellized form in which words are written as consonants, and the vowels and the identity of the word must be inferred from the context. Both Hebrew and Arabic contain a very large number of heterophonic homographs (words in which the visible consonants are identical, but the invisible vowels are different). It is clear that this type of orthographic depth is different from that evinced by English, which is considered a deep orthography because there is variability in letter–sound relations. The findings of Shimron and Sivan (1994) suggest that unvowellized Hebrew is harder to read than English (even for native Hebrew readers), whereas vowelized Hebrew (in which the orthography is shallow) is not.

The present results, together with our previous findings that native Arabic speakers process the Hebrew orthography more easily than Arabic orthography, converge to suggest that visual complexity is a major determinant of reading in Arabic. Here, we have shown that vowels are more difficult to detect in Arabic than in Hebrew, and that even by sixth grade, reading is not automatized to the point that lexical knowledge can facilitate detection.

References

Abu-Rabia, S. (1996). The role of vowels and context in the reading of highly skilled native Arabic readers. *Journal of Psycholinguistic Research*, 25: 629–41.

Abu-Rabia, S. (1998). Reading Arabic texts: Effects of text type, reader type, and vowelization. *Reading and Writing: An Interdisciplinary Journal*, 10: 106–19.

Abu-Rabia, S., Share, D. L., and Mansour, M. S. (2003). Word recognition and basic cognitive abilities among reading-disabled and normal readers in Arabic. *Reading and Writing: An Interdisciplinary Journal*, 16: 423–42.

Azzam, R. (1984). Orthography and reading of the Arabic language. In Aaron, J. and Joshi, R. M. (eds), *Reading and Writing Disorders in Different Orthographic Systems*. London: Kluwer Academic, pp. 1–29.

Bentin, S. and Ibrahim, R. (1996). New evidence for phonological processing during visual word recognition: the case of Arabic. *Journal of Experimental Psychology*, 22(2): 309–23.

Cohen, L. and Dehaene, S. (2004). Specialization within the ventral stream: The case for the visual word form area. Neuroimage, 22, 466–76.

Eviatar, Z. and Ibrahim, R. (2000). Bilingual is as bilingual does: Metalinguistic abilities of Arabic-speaking children. *Applied Psycholinguistics*, 21(4): 451–71.

Eviatar, Z., Ibrahim, R., and Ganayim, D. (2004). Orthography and the hemispheres: Visual and linguistic aspects of letter processing. *Neuropsychology*, 16(3): 322–6.

Farid, M. and Grainger, J (1996). How initial fixation position influences visual word recognition: A comparison of French and Arabic. *Brain and Language*, 53: 351–68.

Ferguson C. A. (1959). Diglossia. *Word*, 15: 325–40.

Frost, R., Katz, L., and Bentin, S. (1987). Strategies for visual word recognition and orthographical depth: A multilingual comparison. *Journal of Experimental Psychology: Human Perception and Performance*, 13: 104–14.

Ibrahim, R., Eviatar, Z., and Aharon-Peretz, J. (2002). Do the characteristics of Arabic orthography slow its cognitive processing? *Neuropsychology*, 16(3): 322–6.

Ibrahim, R. and Aharon-Peretz, J (2005). Is literary Arabic a second language for native Arab speakers? Evidence from a semantic priming study. *The Journal of Psycholinguistic Research*, 34(1): 51–70.

Ibrahim, R., Eviatar, Z., and Aharon-Peretz, J. (2007). Metalinguistic awareness and reading performance: A cross-language comparison. *Journal of Psycholinguistic Research*, 36(4): 297–317.

Katz, L. and Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. Orthography, phonology, morphology, and meaning. In Frost, R. and Katz, L. (eds) *Orthography, Phonology, Morphology, and Meaning, Advances in Psychology*, vol. 94. Oxford, England: North-Holland, pp. 67–84.

Lubow, R. E., Tsal, Y., Mirkin, A., and Mazliah, G. (1994). English and Hebrew letter report by English- and Hebrew-reading subjects: Evidence for stimulus control, not hemispheric asymmetry. *Brain and Cognition*, 25: 34–51.

Macmillan, N. A. and Creelman, C. D. (1991). *Detection Theory: A User's Guide*. Cambridge: Cambridge University Press.

Rao, C., Vaid, J., Srinivasan, N., and Chen, H.-C. (2011). Orthographic characteristics speed Hindi word naming but slow Urdu naming: Evidence from Hindi/Urdu biliterates. *Reading and Writing: An Interdisciplinary Journal*, 24(6): 679–95.

Roman, G. and Pavard, B. (1987). A comparative study: How we read Arabic and French. In O'Regan, J. K. and Levy-Schoen, A. (eds), *Eye Movements: From Physiology to Cognition*. Amsterdam, The Netherlands: North Holland Elsevier, pp. 431–40.

Saiegh-Haddad, E. (2003). Linguistic distance and initial reading acquisition: The case of Arabic diglossia. *Applied Psycholinguistics*, 24: 431–51.

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

Saiegh-Haddad, E. (2005). Correlates of reading fluency in Arabic: Diglossic and orthographic factors. *Reading and Writing: An Interdisciplinary Journal*, 18: 559–82.

Shimron, J. and Sivan, T. (1994). Reading proficiency and orthography: Evidence from Hebrew and English. *Language Learning*, 44(1): 5–27.

Vaid, J. (1988). Asymmetries in tachistoscopic word recognition: Scanning effects reexamined. *International Journal of Neuroscience*, 42: 253–8.

Note

1 In order to test the hypothesis that better performance on Hebrew stimuli is not due to a practice effect that goes beyond the practice trials, we reanalyzed the responses to Arabic stimuli with block as an independent variable. If practice improves performance in this task, we should see improvement in performance between blocks 1 and 3. The analyses showed no effects of block on either RT or error scores. We therefore suggest that performance levels with Hebrew stimuli are due to the differences in orthographies between Arabic and Hebrew, and not to a practice effect.