WRITING SYSTEMS RESEARCH, 2013
Vol. 5, No. 2, 156–168, http://dx.doi.org/10.1080/17586801.2013.862163

Routledge
Taylor & Francis Group

The effects of orthographic complexity and diglossia on letter naming in Arabic: A developmental study

Hanan Asaad and Zohar Eviatar

Psychology Department, University of Haifa, Haifa, Israel

The purpose of the study was to examine the influence of orthographic complexity and diglossia on letter naming and automaticity in Arabic. Two experiments were carried out by 31 first graders, 30 third graders, 34 fifth graders and 20 university students. In the first experiment we took advantage of the Arabic orthographic variation in letter shape, and compared the Stroop effect for correctly written and orthographically distorted words. All participants revealed a Stroop effect with both types of words, but only first graders showed the same degree of interference with distorted and correctly written words. We interpret these results to reflect the development of automaticity in reading. In the second experiment, six letter-naming tests were performed. The results showed that retrieval time of naming letters or the sounds that these letters represent decreased inversely with age. A different pattern was found between the school-age children and the university students. In children, the relationships between types of tests of retrieval speed remained constant: retrieval of letter names or sounds which do not have visual or phonological neighbours was the fastest, and of letter names representing sounds that do not exist in spoken Arabic was the slowest. There was no effect of changing letter shape. However, among the university students only changing letter shape affected the speed of responses. We interpret these results to reflect different representations of letter categories in adults and children. The findings have implications for models of reading development in Arabic.

Keywords: Arabic; Letter identification; Reading development; Diglossia.

Learning to read Arabic is especially challenging for native speakers of the language (e.g., Azzam, 1984; Ibrahim, Eviatar, & Aharon-Peretz, 2007; PISA 2009; The Ministry of Education and Culture, 1992). Two major reasons for this have been suggested: diglossia, and the visual characteristics of Arabic orthography. The purpose of the present research is to examine the influence of both these factors on the process of reading acquisition and on automaticity.

Diglossia

Arabic has two forms: the spoken form (*Ammia*—that has many local dialects) is used by speakers of the language in a specified geographical area for daily verbal communication, and is the native language of virtually all Arabic speakers. The literary form (*Fuṣḥa*), is the

Correspondence should be addressed to Hanan Asaad, University of Haifa, Haifa 31905, Israel. E-mail: hanan.assad@gmail.com

language in which all speakers of Arabic, from all over the world, read and write. This form of Arabic is universally used in the Arab world for formal communication and is known as 'Modern Standard Arabic' (MSA). The two forms have a different structure in all language domains—phonology, morphology, lexicon and syntax. No single spoken Arabic vernacular (SAV) has the same set of phonemes as MSA (Maamouri, 1998). At the level of the lexicon, MSA and SAV share many words, however, most of the words exhibit variable degrees of phonological disparity between their MSA and SAV forms. The inflectional system the MSA is richer and more complicated than the SAV. In addition, the MSA follows a VSO (verb-subject-object) word order while the SAV follows a SVO word order.

Several psycholinguistic studies have addressed the effects of diglossia directly. In children, Saiegh-Haddad (2008) examined the influence of phonological distance between the spoken and written language on basic reading processes. She found that the phonological distance between pronunciation in Ammia and Fuṣḥa directly affects phonological sensitivity to Fuṣḥa phonemes, auditory memory for Fuṣḥa words, reading fluency and auditory comprehension of Fusha.

In adults, Ibrahim (Ibrahim, 2009; Ibrahim & Aharon Peretz, 2005), examined the relationship between the two forms of Arabic by comparing auditory semantic priming and repetition effects on lexical decisions within the native language (Ammia) with the effects obtained when the primes were either in Fusha or in Hebrew (the participant's second language (L2)), and vice-versa. These studies showed that facilitation patterns were more similar between Fusha and Hebrew, than between either of these languages and Ammia. Ibrahim suggested that despite the intensive daily use and psychological proximity of Ammia and Fusha, they are represented in two different lexica in the cognitive system of the native Arabic speaker. Eviatar and Ibrahim (2000) examined this question in children, by exploiting the effects of the relationship between a bilinguals' languages and the emergence of metalinguistic skills in childhood. They asked whether preliterate and newly literate Arab children who are learning Fusha, behave like bilinguals. The study compared three groups of kindergarten and first-grade children: Hebrew-speaking monolinguals, Russian-Hebrew bilinguals and Arabic-speaking children. The Russian-Hebrew bilinguals showed the classic pattern resulting from exposure to two languages: higher performance levels in metalinguistic tests, and lower performance levels in the vocabulary measure as compared to monolinguals. The Arab children's performance levels mimicked those of the bilingual children for the most part, and suggested that exposure to Fusha in early childhood promotes metalinguistic skills to the same degree observed among bilingual children exposed to two different languages. In sum, previous research with both children and adults suggests that the diglossic situation in Arabic culture, where one version of the language is used for speech, while another is used mostly for reading and writing, has long lasting effects on both auditory and linguistic functions (e.g., reading).

Orthography

Arabic is an abjad (Daniels, 1990), in which letters denote consonants. There are three letters that denote long vowels as well, but it is not always clear to the reader whether these letters denote a vowel or a consonant. When short vowels do appear (in children's books and poetry), they are denoted by diacritics above or below the letters, and the orthography is considered shallow because all of the phonological information, necessary for reading aloud, is represented in the text. Unvowelled Arabic texts are considered deep, because information about the vowels must be inferred from contextual and lexical representations.

The second factor affecting the complexity of Arabic orthography is the visual complexity of the letters themselves. Dots comprise an integral part of many letters, and there are many sets of letters that have a similar or even identical structure, and are distinguished only on the basis of the existence, location and number of dots. Panel A of Table 1 shows examples of these. These factors may have far-reaching effects on skilled reading in Arabic. Previous research with adult skilled readers has suggested that this characteristic of Arabic orthography disallows the involvement of the right hemisphere in letter identification (Eviatar & Ibrahim, 2004). Analyses of a cross-language lexical-decision task have suggested that while the right hemisphere is involved in this task in English and in Hebrew, it is not involved in lexical decision in Arabic (Ibrahim & Eviatar, 2012). In addition, some letters have phonological neighbours, in the existence of allophonic variants, which are represented by different letters. Examples of these are shown in panel B of Table 1. This characteristic of Arabic orthography and its implications for spelling are discussed in more detail in Saiegh-Haddad (this volume).

A third aspect of complexity is reflected in the variability of letter shapes which is dependent on their placement in the word. Twenty-three of the 29 letters in the alphabet have four shapes each (word initial, medial when they follow a connecting letter, final when they follow a non-connecting letter and final when they follow a connecting letter, for example, the phoneme /h/ is represented by the graphemes: (هـ ـهـ ه ـه), and 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. An additional characteristic of the Arabic orthography is that the majority of letters must be connected to their neighbours from both sides (right and left), except for six letters (ا ـد ـذ ـر ـز ـو). These are connected only from their right side (the side from which reading progresses).

Recently, a study by Rao, Vaid, Srinivasan, and Chen (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), in Hindi (which uses a shallower, and less visually complex orthography) and 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.

The hypothesis tested here is that these characteristics of Arabic affect both the fluency and speed with which letters are identified and the achievement of automaticity in reading

TABLE 1
Some examples of visual and phonological neighbours in Arabic

A. Visual neighbours	/b /= بـ /θ/ = ثـ /t/ = تـ /y/= يـ /n/= نـ
	/k/= قـ /f/= فـ
	/j/= جـ /ḥ/= حـ /x/= خـ
	/z/= ز /r/= ر
	/s/ =ص /ḍ/ =ض
B. Phonological neighbours	/t/̈/̈/ = ط ت
	/d//ḍ/= ضـ د
	/s//ṣ/= صـ سـ
	/k//q/= ق ك
	/z//ð·/= ظ ز

acquisition. In order to test this, we used well-known psycholinguistic tests. In order to test automaticity, we used the word version of the Stroop task (Stroop, 1935), in which the fast and automatic process of reading in skilled readers interferes with colour naming. In this Stroop task, participants name the colour in which words are written, and the time to name the ink colour of words denoting other colours (e.g., the word 'red' written in blue ink) is compared to the time to name the ink colour of neutral words (e.g., the word 'rat' written in blue ink). This paradigm has been used extensively to examine automatic processes (see MacLeod, 1991 for a review). The names of the colours are the same in Ammia and Fusha, allowing us to examine reading fluency independently of the effects of diglossia. In order to test speed of letter identification, we used the Rapid Automatic Naming (RAN) paradigm designed by Denckla and Rudel (1974). Here, participants name a series of 50 letters as quickly as they can. Much research with this task has suggested that it reflects reading fluency (e.g., Wolf & Segal, 1992).

The Stroop task allowed us to test the effects of the variability of letter shapes on automatic reading. We created two Stroop tasks—a regular one, in which the naming of the ink colour of words denoting colours was compared to the time to name ink colours of neutral words, and a distorted one, in which the names of the colours were spelt correctly, but the shapes of the letters were wrong: for example, the word-initial shape was used in the middle of the word. Differences in the amount of interference between these versions indicate global strategies in reading and automaticity.

Letter knowledge is the best predictor of reading acquisition in many languages (e.g., Levin, Shatil-Carmon, & Asif-Rave, 2006; Share, Jorm, Maclean, & Matthews, 1984; Treiman & Kessler, 2004). The thesis is that letter knowledge involves learning that a family of shapes that share prototypical features belong to a distinct category that is labelled by the name of the letter (Ehri, 1986). It has been suggested that in addition to promoting phonological awareness, knowledge of letter names helps children understand the relationship between spoken and written words, and the idea that in alphabetic scripts, letters stand for sounds. To our knowledge, with the exception of the studies undertaken by Levin and her colleagues (e.g., Levin et al., 2006; Levin, Patel, Margalit, & Barad, 2002), who studied preliterate Israeli children learning Hebrew letters, all of the studies examining the relationship between reading acquisition and skill have been run on children learning to read in Roman alphabets. Interestingly, Levin and her colleagues show patterns of learning the names and the sounds of letters that are somewhat different from those shown for children learning Latin letters, and interpret these in the context of the differences between writing systems.

We believe that the situation of diglossia—together with the visual complexity, and the complexity of the relationship between graphemes and phonemes—in Arabic, may complicate the acquisition of letter prototypes. Thus, we also hypothesised that the complex relationship between graphemes and phonemes, and phonemic diglossia, would impede retrieval of letter names. We manipulated the characteristics of the letters in the RAN tests, and expected these characteristics to affect the time to name the letters or the sound they represent. Three tests examined the effects of visual and phonological neighbours, one test examined the effects of diglossia, and two tests examined the effects of shape variation.

We examined readers at four levels of skill. In the Arabic-language elementary schools in Israel, in first and second grade, school time is used to develop letter recognition, decoding and vocabulary in MSA. In the middle of third grade the major focus of study moves on from basic reading skills, and children are required to use reading for learning other topics. In addition, the transition from fully vowelled text (including both long and short vowels) to partially vowelled text (including only long vowels) begins in third grade.

Reading unvowelled Arabic text is expected to be fluent by the end of fourth grade. In the present study, we examined beginning readers (first graders), readers in the process of the transition mentioned above (third grade), more fluent readers (fifth grade) and skilled readers (university undergraduates).

METHOD

Participants

The participants were 31 first graders (17 girls, with a mean age of 7.02 years), 30 third graders, (16 girls mean age of 8.94 years) and 35 fifth graders, (17 girls mean age of 10.88 years), from an elementary school in Nazareth, and 20 undergraduate students from the University of Haifa. All participants were native Arabic speakers. The official language of the elementary school is Arabic, and the population is of middle to high socioeconomic status. Based on teacher reviews, none of the children suffered from developmental or acquired neurological, learning, emotional or attention disorders. In addition, a hearing test was performed on all the children in a quiet room, using a portable audiometer Siemens SD-25 and TDH-50P headphones. Those diagnosed as having hearing within normal range, were chosen to take part in the research. This was done because it has been shown that small deficiencies in hearing can have consequences for phonological representations, and on performance on phonological tasks (e.g., Briscoe, Bishop, & Frazier Norbury, 2001). The research was approved by the ethics committee of the Ministry of Education in Israel. The University students were screened for learning disabilities (in Hebrew) and grew up in or around Nazareth. This is to ensure that they spoke the same dialect as the children.

The language of teaching at the university is Hebrew. All of our participants were learning Hebrew and English. In this elementary school, the children begin to learn English in first grade and Hebrew in third grade. All literate Arabic speakers in Israel are minimally quadrilingual: Ammia is their spoken L1, and Fusha, Hebrew and English are languages in which they become literate (Eviatar & Ibrahim, 2012).

Stimuli and materials

Stroop task

Three cards were constructed. On the ‘control’ card, 16 4–5 letter words naming concrete objects were written in coloured ink in two columns. On the ‘classic Stroop’ four names of colours (red, yellow, green and blue) were written in a different colour (e.g., the word ‘blue’ was written in red ink) and repeated four times in the same format as the control card. On the ‘distorted Stroop’, the same colour names were written in the same sequence and ink colour as on the ‘classic Stroop’ card, but the shape of the words was distorted. For example, it is possible to write the word احمر (red) using the correct letters but in their incorrect forms (اح م ر, اح مر, احم ر).

Letter-naming tasks (RAN)

Six tests were created, each comprised of a card with letters arranged in five rows, 10 letters in a row. The tests differed in the identity of the letters.

1. *‘different letters’*—The test consisted of five letters, each repeated 10 times randomly, that have no visual or phonological neighbours (و, م, ل, ي, هـ).
2. *‘visual neighbours’* – The test consisted of five letters, each repeated 10 times randomly, that have visual neighbours but without phonological neighbours (غ, فـ, حـ, شـ, ب).

None of the neighbours were included in the test. Visual neighbours are defined as letters that represent different phonemes, with identical overall shapes, that differ only in the placement and number of dots around these shapes.

3. *'visual and phonological neighbours'*—The test consisted of five letters, each repeated 10 times randomly, that have visual and phonological neighbours (ك, ز, ضـ, ت, س). None of the neighbours were included in the test. Phonological neighbours are letters that represent an identical or very similar phoneme, such as regular and emphatic versions of sounds, but are written differently.
4. *'Fusha only'*—The test consisted of four letters, each repeated 12 or 13 times randomly, that have visual and phonological neighbours but represent sounds that are not used in the spoken language (Ammia) in Nazareth dialect (ق, ظ, ة, ث).

Two additional tests were constructed in order to examine the effect of letter shape variation:

5. *'different different shape'*—The test consisted of the same letters that are in the 'different letters' test, but they appear in different positional variants (ـمـ, مـ, ي, ـي, ـيـ, يـ, ل, ـل, ـلـ, لـ و, ـو, هـ, ه, ـه, ـهـ, م, ـم).
6. *'visual neighbours different shape'*—The test consisted of the same letters that are in the 'visual neighbours' test, but they appear in different positional variants (فـ, ـفـ, ـف, ح, ـح, حـ, ـحـ, ش, ـش, شـ, ـشـ ب, بـ, ـب, ـبـ غ, ـغ, غـ, ـغـف).

Procedure

Participants were tested individually in a quiet room at their school during regular school hours. The instructions were given in Ammia. During the meeting all children were given the Stroop tests, followed by RAN tests. University students were tested individually in a quiet room at the university. Verbal responses were recorded on audiotape.

The Stroop effect was measured as the difference in naming time between the 'classic' condition and the 'control' condition, and between the 'distorted' Stroop and the 'control' condition.

In the RAN tests, participants were required to retrieve names or sounds of the letters as quickly as they could. The time it took them to name all the letters, in each list, was used as an index of automatised naming speed.

RESULTS

Stroop effects

We computed two Stroop effects: the regular Stroop is the difference between the time to name the ink colour of the neutral words, and the time to name the ink colour of the correctly spelt (but incongruous) colour names. The 'distorted' Stroop is the difference between the time to name the ink colour of the neutral words, and the colour of the words naming ink colours in which the words were graphemically distorted.

These data were analysed with a two-way mixed ANOVA with age as a between-subjects factor (first grade, third grade, fifth grade and adults) and test type (regular vs. distorted) as a within-subject factor. The results revealed a significant interaction between the factors, $F(3, 112) = 3.58$, $p < .05$, $\eta_p^2 = .09$. There is also a significant main effect of type of test, $F(1, 112) = 39.46$, $p < .0001$, $\eta_p^2 = .26$, and a main effect of age, $F(3, 112) = 4.27$, $p < .01$, $\eta_p^2 = .10$. The cell means can be seen in Figure 1.

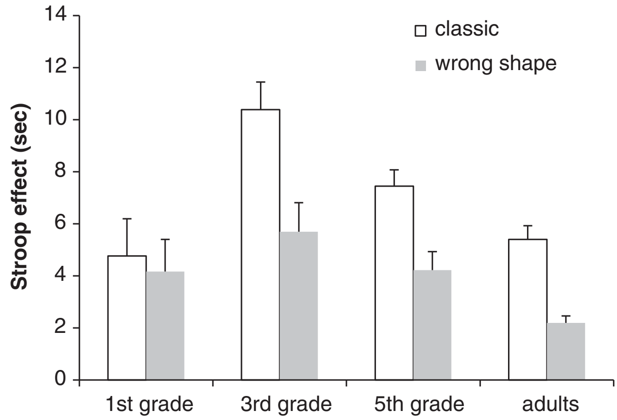

Figure 1. Stroop effects for regular and distorted words. Error bars indicate standard errors.

Post-hoc tests were computed to test three effects, as shown in panel A of Table 2. We tested whether the Stroop effect in each test was significantly different from 0, and also whether the effects were significantly different from each other. As can be seen in the table, all of the groups evinced significant Stroop effects in both conditions. However, all of the groups except first grade, revealed significantly larger Stroop effects when the words were written correctly than when one of the letters was in the wrong shape. In the first grade tasks, both tests resulted in an equivalent Stroop effect. Interestingly, for the regular Stroop effect, the simple main effect of age was significant, $F(3, 112) = 6.21$, $p < .001$, $\eta_p^2 = .14$,

TABLE 2
Panel A: Mean Stroop effects (in seconds) and results of significance tests. One-sample *t*-tests show whether the Stroop effect is significantly different from 0. The two-sample *t*-test compared the Stroop effects in the regular and distorted conditions. Panel B: Pair-wise comparisons among the four groups of participants in each of the Stroop conditions

	Regular Stroop		*Distorted Stroop*		*Regular vs. Distorted*	
A.	*Mean (SEM)*	*Significance 1 sample* t-*test*	*Mean (SEM)*	*Significance 1 sample* t-*test*	*Mean (SEM)*	*Significance 1 sample* t-*test*
1st grade	4.77 (1.4)	$t(30) = 3.4$ $p < .002$	4.16 (1.2)	$t(30) = 3.43$ $p < .002$	0.6 (0.94)	$t(30) = 0.64$ $p > .5$
3rd grade	10.38 (1.05)	$t(29) = 9.91$ $p < .0001$	5.7 (1.09)	$t(29) = 5.22$ $p < .0001$	4.68 (1.12)	$t(29) = 4.18$ $p < .0005$
5th grade	7.45 (.62)	$t(34) = 12.16$ $p < .0001$	4.22 (.703)	$t(34) = 6.0$ $p < .0001$	3.23 (0.79)	$t(34) = 4.08$ $p < .0003$
Adults	5.4 (.51)	$t(19) = 10.48$ $p < .0001$	2.2 (.25)	$t(19) = 8.54$ $p < .0001$	3.2 (0.51)	$t(19) = 6.25$ $p < .00001$

B.	Regular Stroop	Distorted Stroop
Simple main effect of age	$p < .001$	ns. $p > .14$
1st grade vs. 3rd grade	$F(1,112) = 16.17$, $p < .001$	ns, $p > .25$
1st grade vs. 5th	$F(1,112) = 3.99$, $p < .05$	ns, $p > .96$
1st grade vs adults	ns, $p < .6$	ns, $p < .19$
3rd grade vs. 5th grade	$F(1,112) = 4.67$, $p < .05$	ns, $p > .25$
3rd grade vs. adults	$F(1, 112) = 10.02$, $p < .001$	$F(1, 112) = 5.44$, $p < .05$
5th grade vs. adults	ns, $p > .18$	ns, $p > .16$

with planned comparisons revealing the effects shown in panel B of Table 2. It can be seen that the Stroop effect was significantly larger in third graders than in all of the other groups and smaller in first graders than in fifth graders. However, in the distorted Stroop condition, the main effect of age is not significant, $p > .14$, although as shown in Table 2, the difference in the size of the effect is significant between the third graders (who had the highest amount of interference) and adults (who had the lowest amount of interference), $F(1, 112) = 5.44$, $p < .05$, $\eta_p^2 = .05$.

Letter naming (RAN)

We computed two separate analyses on the letter naming times from the Rapid Automatic Naming Test. In the first analysis we examined the effects of visual and phonological neighbours on the access to letter names, and also the effect of the existence of the sound represented by the letter in the participant's spoken dialect. In the second analysis we examined the effects of the variability of the shapes of the letters.

Neighbours and familiarity

We analysed the time taken to name the letters in four conditions: letters that do not have visual or phonological neighbours ('different letters'), letters that have visual neighbours, letters that have visual and phonological neighbours and letters that represent sounds that do not exist in the spoken dialect of the participants. The analysis revealed a significant interaction between age and test, $F(9, 333) = 9.63$, $p < .0001$, $\eta_p^2 = .21$; a main effect of test, $F(3, 333) = 120.58$, $p < .0001$, $\eta_p^2 = .52$; and a main effect of age, $F(3, 111) = 92.01$, $p < .0001$, $\eta_p^2 = .71$. These effects can be seen in the top panel of Figure 2.

Planned comparisons were computed to examine the differences between the tests in each age group. We compared pairs of tests in rising order of difficulty; first, letters with no neighbours ('different letters') to letters with visual neighbours, letters with visual neighbours to letters with visual and phonological neighbours, and the latter group, to letters naming sounds that appear only in Fusha. These comparisons for each age group are shown in Table 3. It can be seen that among the children, all of the differences are marginally significant or highly significant, especially the difference between letters naming sounds that occur in their spoken dialect versus those that do not. Thus, although there are effects of visual and phonological complexity, the largest effect is of the familiarity of the sound named by the letter. Among the adults, only the existence of neighbours affected the speed of responses.

Effects of variations in letter shape

For two of the tests, the one including unique letters ('different letters') and the one including letters that have visual neighbours, we ran two additional conditions. These were identical to the original test, but the letters appeared in different shapes. The 'different shapes' test included exactly the same five letters that were included in the original test, but here they appeared in different forms—once in their final form when they follow a connecting letter (ـه), once in their initial form or mid-word form when they follow a non-connecting letter (هـ), once in their mid-word form when they follow a connecting letter (ـهـ), once in their final form when they follow a non-connecting letter (ه). This tested the effects of the variation in form on the ability to access the letter name. To examine this, we compared the naming times for the single-form test to the multiple-form test.

For letters that have no neighbours, we analysed response times to the two tests (where the letter appears multiple times in the same shape, versus the version in which the letter appears in different shapes), for the four groups. The analysis revealed an interaction between test type and age, $F(3, 111) = 3.92$, $p < .05$, $\eta_p^2 = .10$, no main effect of test

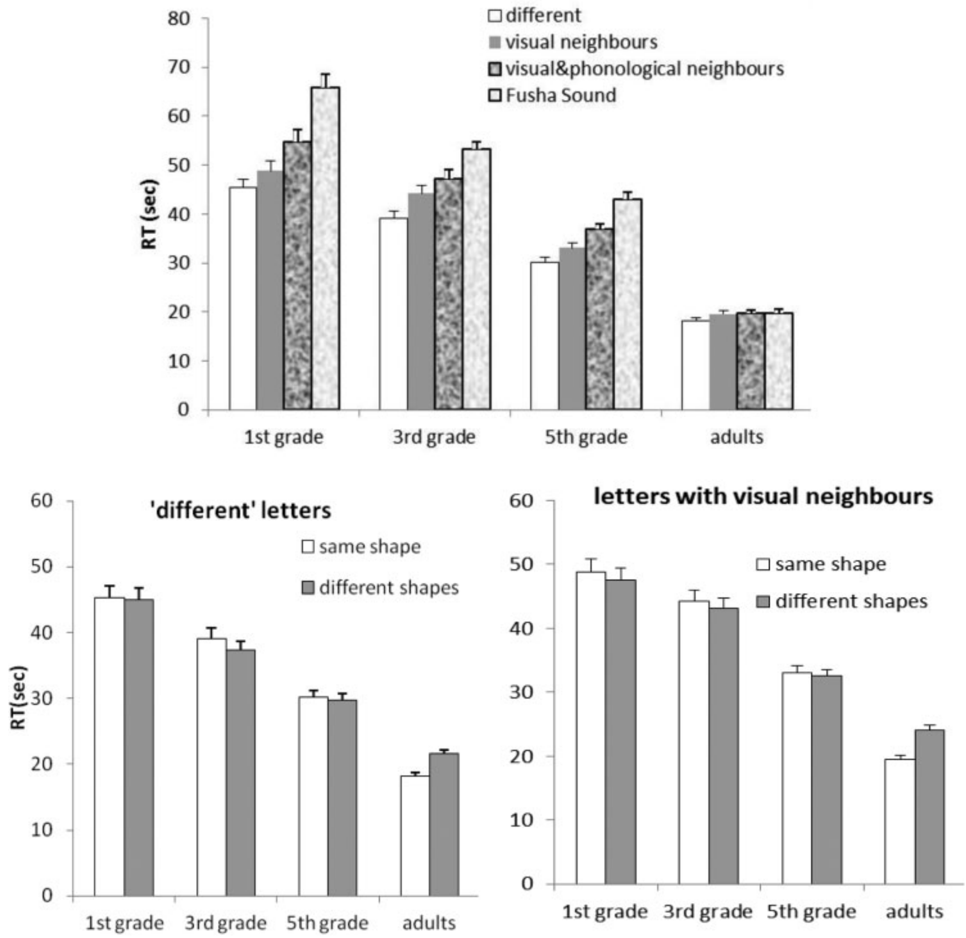

Figure 2. Top panel: Time taken by each group to name the letters in each test. Bottom left panel: Comparisons between versions of the 'different' test with uniform or varied letter shapes. Bottom right: Comparison between version of the 'visual neighbors' test with uniform or varied letter shapes. Error bars indicate standard errors.

($p > .6$), and a main effect of age, $F(3, 111) = 62.34$, $p < .0001$, $\eta_p^2 = .63$. These cell means are shown in the left lower panel of Figure 2. Planned comparisons revealed that among the children, none of the groups showed a significant effect of type of test, whereas the adults show a significant effect, $F(1, 19) = 33.58$, $p < .0001$, $\eta_p^2 = .64$, with the variation in shape resulting in slower responses.

TABLE 3
Planned comparisons for response times on the RAN tests

	1st grade (N = 31)	*3rd grade* (N = 30)	*5th grade* (N = 34)	*Adults* (N = 20)
Different vs. visual neighbours	$F = 2.87$, $p = .09$	$F = 15.7$, $p < .0001$	$F = 7.8$, $p < .01$	$F = 5.54$, $p < .05$
Visual neighbours vs. visual and phonological neighbours	$F = 8.19$, $p < .01$	$F = 5.34$, $p < .05$	$F = 13.32$, $p < .0001$	ns $p > .8$
Visual and phonological neighbours vs. Fusha only	$F = 27.95$, $p < .0001$	$F = 21.55$, $p < .0001$	$F = 34.43$, $p < .0001$	ns $p > .7$

For letters that have visual neighbours the same analysis revealed a significant interaction between age and test, $F(3, 111) = 5.57$, $p < .01$, $\eta_p^2 = .13$, no main effect of test ($p < .9$) and a main effect of age, $F(3, 111) = 61.62$, $\eta_p^2 = .62$. These cell means are shown in the lower right panel of Figure 2. Planned comparisons again revealed that among the children, none of the groups showed a significant effect of test type, whereas in the adult group, this effect was highly significant $F(1, 19) = 76.05$, $p < .0001$, $\eta_p^2 = .80$.

DISCUSSION

The results of the two experiments reveal interesting changes in the development of global versus local reading strategies, the development of letter retrieval, and the relationship of these changes with automaticity in reading. The Stroop experiment revealed expected differences between first graders and more skilled readers, but also interesting reflections of specific processes in third grade. These are detailed below. The RAN experiments showed the expected effects of visual complexity and familiarity on the ability of children to retrieve letter identity, but also the unexpected effects of letter shape on adult letter retrieval. As detailed below, we believe these findings speak to the creation of letter prototypes in Arabic.

Stroop effects

The results of the Stroop tests revealed that from at least third grade on, children and adults read in an automatic and global manner, such that distorting the shapes of words affects the degree of interference with colour naming. Although these participants still revealed a significant effect of the meaning of the distorted words on the time taken to name the ink colour, this effect was significantly smaller than with correctly written words. Our first graders show an equivalent Stroop effect with both types of words, those written correctly and distorted words. This suggests that although reading is occurring automatically, as the meaning of the colour words interferes with naming the ink colour, reading is based on local letter-by-letter identification of the words, such that changes in the global aspects, for example, letter shapes, do not affect it.

As shown in Table 2, the degree of interference in both versions of the Stroop do not differ between first graders and adults. We believe that this happened for two different reasons: in the regular Stroop test, first graders show lower levels of interference than third and fifth graders, because their reading has fewer automatic features, and thus the colour names interfere less (this is pursued in more detail below). Adults show less interference than the third graders on the regular Stroop because they are better at inhibiting the colour names than these younger participants. In the distorted version of the test, we believe that first graders are using the same reading strategy as in the regular version, and thus show the same degree of interference. In the adults, on the other hand, the distortion disrupts the normal automatic flow of reading, resulting in significantly less interference than in the regular Stroop test.

Our results suggest that the processes underlying reading in third grade differ from those in the other age groups. Recall that in the regular Stroop condition, third graders show significantly larger effects than all the other groups, and in the distorted Stroop condition, they also show the largest Stroop effect (see Figure 1), which is significantly different from the one shown by the adults. We believe that the difference between third and first grade reflects greater skill in reading. In first grade, children are still at the stage of letter learning and their reading is based mainly on decoding a series of phonemes and their inter-connections, in order to arrive at the meaning of a word. They also use this decoding process in order to create phonological representations of words they have read (it must be remembered that due to the diglossic features of Arabic, in general, first graders cannot depend on phonological representations to accelerate their reading pace). Therefore, they show smaller Stroop effects than third graders in both Stroop conditions. However, why do third graders differ from fifth graders and adults? We suggest two (nonexclusive) explanations. The first has to do with the development of inhibition. Adults and fifth graders are better at inhibiting the meaning of the colour words than third graders, and therefore, show smaller Stroop effects in both conditions. The second possible explanation is more relevant to the distorted condition, and has to do with differences in reading strategies among readers with differing levels of skill. Third-grade students are in the middle of the reading acquisition process, transitioning from analytic (as shown by the first graders) to global reading. It is likely that due to the visual complexity of written Arabic, the students will try to read more globally, rendering them 'pseudo-global' readers. They thus depend on the phonological representations that have been learned in the previous three years of schooling which provide them with an exact phonemic pattern in order to read in a global-like manner. This allows for bridging between less than perfect global reading and the need to accelerate reading. Supporting this hypothesis, in other research, we have shown that among third graders, phonological awareness significantly predicts text reading speed, whereas it does not predict this in first and fifth grade, where RAN scores were the significant predictor (Asaad & Eviatar, 2013). Adults and fifth graders, on the other hand, are reading more globally, and thus also show smaller Stroop effects than third graders in the distorted condition.

Letter name retrieval

Not surprisingly, the speed of letter name retrieval grew monotonically with age—older and more skilled readers were faster than younger and less skilled readers. There are three major findings that we would like to emphasise. The first is that among all our groups, the names of letters that do not have visual or phonological neighbours were retrieved significantly faster than in the other tests. This finding supports the hypothesis that visual complexity is a significant factor in the reading of Arabic at all levels of skill. The second finding is that among the children, there is a monotonic decrease in letter naming speed: retrieval of letter names or sounds which do not have visual or phonological neighbours was the fastest and of letter names representing sounds that are not represented in spoken Arabic was the slowest. Retrieval of letters names or sounds having visual and phonological neighbours was slower than letters which have only visual neighbours. Thus, in children, phonological neighbours add complexity over and above visual neighbours, and phonological unfamiliarity adds additional difficulty, whereas this pattern is not found in adults, where, as shown above, only visual uniqueness has effects.

The third surprising finding is that in children, variability of letter shapes did not affect retrieval speed, whereas it did so for adults. One possible explanation for this is that for children, a sound is represented by a 'family' of shapes, which are considered equivalent representations. Thus, the name of the letter is equally retrievable from all its shapes. Whereas for adults, the various shapes of letters may not be equivalent representations. This may be because non-connected final shapes are the prototypical shape representing letters: these are the letter shapes on keyboards and mobile phones, and in all alphabeticised lists (such as in the phone book). We interpret these results to reflect differences in the structure of letter categories in adults and children.

The underlying assumption is that reading is essentially a process of pattern recognition, with perceptual expertise developing as individuals become skilled readers. In fact, Burgund, Schlaggar, and Petersen (2006) explicitly show that perceptual expertise for letters is directly related to reading skill. This perceptual expertise is thought to include a perceptual prototype (either feature or template based) to which individuals compare the stimulus. We show that for children, all versions of letters are responded to with equal speed. However, for adults, variability of letter shape in the RAN tests resulted in slower responses than when the shape was held constant. Recall also that children showed effects of visual and phonological neighbourhoods, and of diglossia, while adults did not. These differing patterns suggest that different types of visual complexity have different effects on access to letter names during development.

The findings have implications for our models of reading development in Arabic. It is necessary to discover when and how letter categories are developed. We can see that at least until fifth grade, all of the letter shapes are equivalent, but in university students they are not. It is clear that more research is needed focusing on the manner in which Arabic readers represent letter shapes, letter names and letter sounds.

REFERENCES

Asaad, H., & Eviatar, Z. (2013). *Learning to read in Arabic: The long and winding road*. Submitted.

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

Briscoe, J., Bishop, D. V. M., & Frazier Norbury, C. (2001). Phonological processing, language, and literacy: A comparison of children with mild-to-moderate sensorineural hearing loss and those with specific language impairment. *Journal of Child Psychology and Psychiatry*, 42, 329–340. doi:10.1111/1469-7610.00726

Burgund, E. D., Schlaggar, B. L., & Petersen, S. E. (2006). Development of letter-specific processing: The effect of reading ability. *Acta Psychologica*, *122*, 99–108. doi:10.1016/j.actpsy.2005.11.005

Daniels, P. T. (1990). Fundamentals of grammatology. *Journal of the American Oriental Society*, *110*, 727–731. doi:10.2307/602899

Denckla, M. B., & Rudel, R. G. (1974). Rapid automatized naming of pictured objects, colors, letters, and numbers by normal children. *Cortex*, *10*, 186–202. doi:10.1016/S0010-9452(74)80009-2

Ehri, L. C. (1986). Sources of difficulty in learning to spell and read. *Advances in developmental and behavioral pediatrics 7*, 121–195.

Eviatar, Z., & Ibrahim, R. (2000). Bilingual is as bilingual does: Metalinguistic abilities of Arabic-speaking children. *Applied Psycholinguistics*, *21*, 451–471. doi:10.1017/S0142716400004021

Eviatar, Z., & Ibrahim, R. (2004). Morphological and orthographic effects on hemispheric processing of nonwords: A cross-linguistic comparison. *Reading and Writing: An Interdisciplinary Journal*, *17*, 691–705. doi:10.1007/s11145-004-2659-8

Eviatar, Z., & Ibrahim, R. (2012). Multilingualism among Israeli Arabs and the neuropsychology of reading in different languages. In M. Leikin, Y. Tobin, & M. Schwartz (Eds.), *Current issues in bilingualism: Cognitive and socio-linguistic perspectives* (pp. 57–74). Heidelberg: Springer-Dordrech.

Ibrahim, R. (2009). The cognitive basis of diglossia in Arabic: Evidence from a repetition priming study within and between languages. *Psychology Research and Behavior Management*, *12*, 95–105.

Ibrahim, R., & 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), 51–71. doi:10.1007/s10936-005-3631-8

Ibrahim, R., & Eviatar, Z. (2012). The contribution of the two hemispheres to lexical decision in different languages. *Behavioral and Brain Functions*, *8*(3).

Ibrahim, R., Eviatar, Z., & Aharon- Peretz, J. A. (2007). Metalinguistic awareness and reading performance: A cross language comparison. *Journal of Psycholinguistic Research*, *36*, 297–317. doi:10.1007/s10936-006-9046-3

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning, advances in psychology*, *94* (pp. 67–84). Oxford: North-Holland.

Levin, I., Patel, S., Margalit, T., & Barad, N. (2002). Letter names: Effect on letter saying, spelling, and word recognition in Hebrew. *Applied Psycholinguistics*, *23*, 269–300. doi:10.1017/S0142716402002060

Levin, I., Shatil-Carmon, S., & Asif-Rave, O. (2006). Learning of letter names and sounds and their contribution to word recognition. *Journal of Experimental Child Psychology*, *93*, 139–165. doi:10.1016/j.jecp.2005.08.002

Maamouri, M. (1998). *Language education and human development: Arabic diglossia and its impact on the quality of education in the Arabic region*. Discussion paper prepared for The World Bank. The Mediterranean Devlopment Forum.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163–203. doi:10.1037/0033-2909.109.2.163

PISA. (2009). *Results: Executive summary*. http://www.oecd.org/pisa/pisaproducts/46619703.pdf

Rao, C., Vaid, J., Srinivasan, N., & 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*, 679–695. doi:10.1007/s11145-010-9256-9

Saiegh-Haddad, E. (2008). On the challenges posed by diglossia to the acquisition of basic reading processes in Arabic. *Language and Literacy*, 101–126. (Hebrew)

Share, D. L., Jorm, A. F., Maclean, R., & Matthews, R. (1984). Sources of individual diVerences in reading acquisition. *Journal of Educational Psychology*, *76*, 1309–1324. doi:10.1037/0022-0663.76.6.1309

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662. doi:10.1037/h0054651

The Ministry of Education and Culture. (1992). *The report on the results of the national survey of reading comprehension for third grades*. Jerusalem: The national department for learning programs at the Ministry of education.

Treiman, R., & Kessler, B. (2004). The case of case: Children's knowledge and use of upper-and lowercase letters. *Applied Psycholinguistics*, *25*, 413–428. doi:10.1017/S0142716404001195

Wolf, M., & Segal, D. (1992). Word finding and reading in the developmental dyslexia. *Topics in Language Disorder*, *31*(1), 51–65. doi:10.1097/00011363-199211000-00007