ORTHOGRAPHIC LEARNING IN ARABIC-SPEAKING PRIMARY SCHOOL STUDENTS

by

Rima Balshe

A Thesis Submitted in Partial Fulfillment of the Requirements for the

Degree

Doctor of Philosophy

Department of Psychology and Human Development

UCL Institute of Education

London, England

Signed Declaration

I, Rima Balshe, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

The aim of this study was to examine how Arabic-speaking children construct orthographic representations and to identify cognitive/linguistic abilities that may facilitate novel word learning. The research involved first examining factors associated with single word reading and spelling accuracy in Arabic-speaking monolingual children and Arabic-English bilingual children, in order to separate universal from script-dependent predictors.

Because Arabic is diglossic (i.e., two varieties of the language, one spoken, and one for literary purposes), it was considered important to include print exposure as a measure in investigating factors associated with single word reading and spelling. Thus, Study 1 involved the development of Title Recognition Tests (TRT) in Arabic and in English. Participants were children from grades three to five; 86 students participated in the development of the lists in Study 1a, and 76 in the development of the revised lists in Study 1b. Both lists were reliable and were used in the subsequent studies.

Study 2 involved examining predictors of single word reading and spelling (receptive vocabulary, phonological processing, RAN, TRT, and orthographic matching) in 86 third- to fifth-grade bilingual children and 116 third-grade monolingual children. For the bilinguals, PA emerged as the strongest predictor of reading and spelling in Arabic. In English, verbal STM and orthographic matching were predictors for the younger bilinguals. PA was the strongest predictor of reading and spelling for the monolinguals.

In Study 3, novel word learning in Arabic was examined using a paired-associate learning task, orthography present or absent and varying visual complexity (ligature and diacritics). The 116 monolingual children from Study 2 participated. Child-related predictors of novel word learning were examined. Results revealed that presence of orthography facilitated learning. There was evidence that consonant diacritics are a source of difficulty, but diglossic phonemes may also be responsible for reading difficulties documented in Arabic.

Impact Statement

Approximately 300 million people worldwide are native Arabic speakers. Arab countries consistently rank at the very bottom of literacy achievement, regardless of wealth. Several variables have been put forward to account for the low levels of literacy attainment. The primary purpose of the present study was to examine written-form learning in Arabic, to shed light on variables that may be responsible. In order to address this aim, the research involved first examining factors associated with single word reading and spelling in Arabic-speaking monolingual and Arabic-English bilingual children. Comparison of the findings with the two groups of children allowed for the separation of variables involved in reading and spelling in each language. Once the Arabic factors were identified, a training study of how Arabic-speaking children acquire new vocabulary was conducted.

Based on the large body of evidence regarding reading habits and literacy achievement, it was considered that exposure to print could be one of the factors responsible for low literacy achievement among Arab countries. Data were collected and two age-appropriate reliable print-exposure assessments were developed for each language.

Phonological awareness emerged as the strongest and most consistent child-related factor influencing reading and spelling in Arabic, for both the bilinguals and monolinguals. Verbal short-term memory and phonological awareness were found to be associated with Arabic novel word learning. New words taught by seeing the word form were learned more efficiently than were words that were taught without printed word forms; however, the visual demands of Arabic also played a role in how well the words were recalled: words that were connected were recalled well, whereas words that had many diacritical marks (vowel markings) were the most difficult to recall.

The data demonstrate the importance of phonological awareness. This finding could have far-reaching implications for curriculum reform, which currently promotes a look-and-say method rather than a phonics approach. The data also demonstrate that instructional methods should take into account the visual complexities of Arabic that add to the challenges of learning to read and spell in Arabic.

The findings in this research are relevant for literacy acquisition in children learning to read Arabic. They underscore the importance of phonological awareness in reading instruction and how diglossia places constraints that need additional attention on the part of educators. For example, teachers could allocate more instructional time for phonemes that do not exist in children’s spoken vernacular. Teachers should also provide additional time for newly taught Arabic words that contain many diacritical markings to account for the additional effort required to learn them.

In summary, it was possible to identify factors that impede written-form learning in Arabic and that may contribute to reading and spelling difficulties often reported in Arabic. If it were possible to mitigate at least some of the difficulties in reading acquisition in Arabic highlighted in the research, then Arabic-speaking children may engage in reading for leisure, which would very likely lead to increased levels of literacy attainment.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1

Literature Review

The science of reading has dominated research in the cognitive sciences for the past four decades (Adelman, 2012a). Instant and efficient word reading (i.e., word recognition) has been referred to as the hallmark of reading. As such, many theories have been proposed to account for this seemingly effortless skill (Gough & Tunmer, 1986; Ehri, 1995, 2005a, 2005b; Frith, 1985; Share, 1995; Ziegler & Goswami, 2005; Perfetti & Hart, 2002; Pacton et al., 2001). While there is no shortage of theories attempting to explain how word recognition develops, the underlying mechanism involved is not fully understood. The theories that have been proposed over the years are not necessarily competing but rather complementary and demonstrate the complexities of reading development. Reading and spelling development are complex processes, involving the interaction of many cognitive abilities. These processes and accompanying theories are further complicated by the linguistic features of the language. Reading research has been almost exclusively based on findings within an English-speaking population; however, in recent years, a rich variety of cross-linguistic studies have confirmed, and sometimes challenged, the theories that dominate this field (Share, 2014, 2008; Ziegler & Goswami, 2005). Thus, it is necessary to disentangle the components that are language universal from those that are language specific in reading development. Reading research in other languages can be used to illuminate and advance our current understanding of reading development and the move toward a universal model of reading (Share, 2014; Nag & Snowling, 2013).

The primary purpose of this research was to examine orthographic learning in Arabic. However, in order to capture the interaction between the cognitive processes involved in reading and spelling and how these processes are influenced by the linguistic characteristics of Arabic in particular, it was deemed necessary to first examine predictors of reading and spelling. As a starting point, it was first necessary to separate language-specific predictors from those that are considered universal predictors of reading and spelling. Since reading research has, as noted, been dominated by studies carried out in English, the present research examined predictors in two samples: Arabic-speaking monolingual children, and a sample of Arabic (L1)-English (L2) bilinguals. The bilingual sample allowed for an examination of predictors using a within-participant, between-language design. The objective of the research was to gain insights into how the participating children construct word-specific representations in Arabic and the cognitive and linguistic variables that predict learning. The study involved investigation of orthographic learning in Arabic in relation to children's cognitive and linguistic skills that have been identified in past research as being important for reading.

The present context is discussed, demonstrating that the issue of the development of reading concurrently in both languages in this population warrants investigation. The discussion of the background begins with an overview of Gough and Tunmer's (1986) Simple View of Reading (SVoR) as a theoretical framework for the proposed research. The diglossic situation of Arabic is discussed within the SVoR framework, followed by theories of single word reading development that were formulated based on English orthography. Next, the orthographic depth hypothesis is discussed and Arabic's complex orthographic features are introduced. The self-teaching hypothesis is discussed as a framework for considering how orthographic learning occurs, irrespective of orthographic depth. While the self-teaching hypothesis offers an account of how children create orthographic representations, differences still exist in orthographic learning. Predictors of single word reading acquisition are discussed to account for differences in orthographic learning. Finally, research gaps are identified and research questions are proposed for the current study.

Simple View of Reading (SVoR)

The ultimate goal of reading is the comprehension of what is being read. SVoR involves two main components: word recognition and comprehension. According to the SVoR, Reading comprehension (R) = Decoding (word recognition) (D) x Language comprehension (L). According to this, reading comprehension is not increased if word recognition or linguistic comprehension is null (Gough & Tunmer, 1986).

Gough and Tunmer (1986) contended that decoding is defined as accurate and rapid word recognition and not as laborious grapheme (letter)-to-phoneme (sound) conversion. They acknowledged that the process of grapheme-phoneme conversion is fundamental in word recognition. Stuart et al. (2008) argued that decoding (as defined by Gough & Tunmer, 1986) involves two processes, one that involves instant sight word recognition, and the other that involves serial decoding. The use of one of these processes or a combination of the two will activate a word's pronunciation and meaning. This view is consistent with that of developmental reading theories, discussed below, that propose the use of multiple strategies to arrive at the correct pronunciation (Ehri, 2005a; Frith, 1985). Furthermore, Kirby and Savage (2008) also highlighted the role of fluency in word reading and argued that accuracy alone cannot lead to efficient reading comprehension. Clearly, SVoR is not so *simple*.

Hoover and Gough (1990) tested the merits of the SVoR model in a study conducted in the USA, using a sample of 557 students from grades one through four. The authors tested three predictions of the SVoR model. The first prediction was that a product of decoding and listening comprehension would predict unique variance in reading comprehension as opposed to a linear additive model. The second prediction was that both decoding and listening comprehension would be negative for less skilled readers. The final prediction was that the contribution of decoding and listening comprehension to reading comprehension would change in the course of reading development. Results confirmed all three predictions.

Chen and Vellutino (1997) attempted to replicate Hoover and Gough's 1990 findings by testing the three predictions. Using a sample of 460 first, second, sixth, and seventh graders, the authors found evidence supporting what they called a weaker version of the SVoR model. Chen and Vellutino’s sample comprised English monolingual speakers, whereas Hoover and Gough’s sample consisted of Spanish L1 and English L2 learners. This could explain why Chen and Vellutino were in favor of a weaker version of SVoR; they were unable to detect the extremes (i.e., children having little to no decoding and/or little to no oral comprehension) that Hoover and Gough detected. Chen and Vellutino argued that such extremes are rare; however, in recent years, some have argued that such L2 populations are becoming the norm rather than the exception (Durgunoglu, 2002; Melby-Lervag and

Lervag, 2011; Hammer et al., 2014). Indeed, the population under investigation in the current study would fall in these extremes in both their Arabic L1 due to diglossia (discussed in the following section) and L2, where linguistic comprehension and decoding are near null when they enter school.

Verhoeven and van Leeuwe (2012) selected a sample of 1,687 children from grades one through six from 72 schools in the Netherlands. Of this large sample, 394 children were Dutch L2 learners. The researchers employed a longitudinal design and assessed the children in listening comprehension and decoding in first, third, and fifth grades. They assessed reading comprehension in second, fourth, and sixth grades. Decoding in first-grade students predicted reading comprehension in second grade for both L1 and L2 learners. Decoding in third and fifth grades also predicted reading comprehension in fourth and sixth grades, however, to lesser degrees for both L1 and L2 learners.

As is demonstrated, SVoR is more complex than is implied by its name. Although SVoR provides a simple framework for identification and intervention, Hoover and Gough (1990) did not deny that many intricate components exist within word recognition and comprehension. SVoR offers a macro view of reading in general; however, for the purposes of this research the focus is on the underpinnings of word recognition (i.e., decoding according to SVoR). Adding to these complexities are environmental factors, such as learning to read in more than one language, and the characteristics of the language itself. The linguistic and orthographic features of Arabic are discussed next, and Psycholinguistic

Grain Size Theory (PGST) (Ziegler & Goswami, 2005) is proposed to account for the role of these features in single word reading.

Diglossia

The linguistic distance between Spoken Arabic (SA) and Modern Standard Arabic (MSA) has created a diglossic situation that researchers have described as akin to learning a foreign language when children enter school and begin formal reading acquisition in MSA (Abu-Rabia, 2000; Ibrahim, 2009; Ibrahim & Aharon-Peretz, 2005). The term “diglossia” is used to describe a situation where two distinct varieties of the same language are used for different purposes. One variety (MSA) is used for literary purposes and formal communication. The other variety (SA) is used for informal, day-to-day communication.

SA has many differing dialects depending on the country, or even regions within the same country, whereas MSA is the universal communication medium of all Arabic speakers. Prior to formal reading instruction, exposure to MSA is limited to religious sermons, news broadcasts, TV programs dubbed in MSA, and through print exposure by reading stories or having stories read to them (Albirini 2016; Saiegh-Haddad and Haj; 2018). Despite children’s exposure to MSA, this amount of exposure remains limited in comparison to that in other languages that are not diglossic.

This diglossic situation of Arabic creates two distinct varieties of the same language that differ yet overlap to a minimal degree in phonology, grammar, syntax, morphology, and vocabulary (Saiegh-Haddad, 2003; Saiegh-Haddad, 2018). Saiegh-Haddad (2018) argues that the linguistic distance between the lexical and lexico-phonoloigcal and MSA is substantial (only 20% of words are identical in both SA and MSA). Even though most of the phonemes are shared, most of the words sound different. Reading difficulties among native Arabic-speaking elementary school students have often been attributed to the diglossic situation of Arabic (Abu-Rabia, 2000; Asaad & Eviatar, 2013), which is consistent with the prediction from the SVoR model outlined in the last section. Based on SVoR, Arabic-speaking children learning to read are at a disadvantage because oral comprehension is severely compromised due to diglossia.

To investigate how diglossia interferes with PA, and ultimately reading, Saiegh-Haddad (2003) assessed Arabic-speaking children in kindergarten and first grade on a task that required them to manipulate phonemes that were either in their spoken vernacular or in literary Arabic. Isolating diglossic phonemes was significantly more difficult than was isolating spoken phonemes and nonwords that adhered to the spoken structure were read more accurately than were nonwords designed to mimic the literary syllabic structure. Saiegh-Haddad concluded that the linguistic distance between the spoken and literary varieties of Arabic does indeed impede the development of PA among Arabic-speaking children. The stark difference between spoken and diglossic syllables and phonemes illustrates how Arabic does not fit neatly in reading theories. According to SVoR, oral language is central to reading. Oral language facilitates phonological sensitivity, which is the hallmark of reading acquisition. Although there is some overlap in spoken and literary vocabulary, there is a great deal of difference in the phonological structure in Arabic.

To examine the effects of diglossia on the quality of phonological representations, Saiegh-Haddad and Haj (2018) manipulated MSA word conditions based on their phonological distance from SA. Participants were required to judge whether an orally presented word was pronounced correctly in MSA. Results showed that the greater the phonological distance between SA and MSA, the more challenging it was for children to determine the correct pronunciation. Indeed, there was an impact on the quality of phonological representations as a function of phonological distance for all grades (first, second, and sixth). These findings are consistent with the lexical quality hypothesis.

Diglossia determines orthographic depth in Arabic (Saiegh-Haddad & Henkin-Roitfarb, 2014). For example, the word may depict a transparent relationship between spelling and the phonological form it represents, but it may be psycholinguistically opaque because the phonological form it encodes is different from the form speakers harbour in their spoken lexicon (Saiegh-Haddad & Haj, 2018).

Although the primary focus of this study is on the decoding part of the SVoR model, it is impossible to separate oral comprehension, especially in a diglossic situation such as the case in Arabic. Of particular interest is the role of oral comprehension as it pertains to decoding (word recognition) and not reading comprehension. So, how exactly is oral comprehension accounted for in theories of word recognition? The following discussion will deal with decomposing the components of word recognition by examining theories of single word reading with an emphasis on the role of oral vocabulary.

Theories of Reading Development

Unlike SVoR, the reading theories discussed in this section provide a developmental account of single word reading. An understanding of how word reading develops and the assumptions made in such models is necessary for understanding reading difficulties and failure (Frith, 1985). Frith proposed three stages to word reading acquisition based upon observations of the errors children make when reading new words. A reading strategy dominates each stage. In the “logographic” stage, the strategy used is not governed by the phonological features of a word but rather by cues such as salient letter features. The second stage is termed “alphabetic”, whereby the letter-sound correspondences are learned and applied. Thus, the strategy used in this stage is governed by phonological recoding. The third and final stage of reading development, according to Frith (1985), is the “orthographic” phase, which is characterized by instantly recognizing words or morphemic units without the use of phonological recoding. Frith proposed that moving from one stage to another does not require the abandonment of the strategy used in the previous phase but rather a merging of the phases. Firth’s theory does not implicitly or explicitly address the role of oral language in reading development.

A dominant theory of reading development was proposed by Ehri (2005a, 2005b) which takes into account the role of oral language competence in the word recognition process. Ehri’s four phases of reading development are very similar to those of Frith (1985); however, Ehri’s fourth phase of word recognition is not a strategy used for reading but rather an involuntary way of reading once at this phase. What Ehri labels the “prealphabetic" phase is essentially the same as Frith's "logographic" phase. The next phase in Ehri's model is what she termed the "partial alphabetic" phase, in which children know a few letter–sound correspondences, usually involving the first and final sounds/letters of words. Children in this phase will have partial word representations. This partial alphabetic phase is where existing oral vocabulary aids in the successful decoding of words. It is proposed that children in this phase will partially decode and deduce the rest by relying on their oral lexicon. However, unlike readers of English, and most other languages, beginning readers of Arabic cannot rely on their oral vocabulary to help them partially decode words because of the linguistic distance between written language and spoken vernacular (Saiegh-Haddad, 2003; Asaad, & Eviatar, 2014).

Only when children are able to form complete letter–sound connections to store words in memory are they in the "full alphabetic" phase (i.e., of phonological recoding), which Frith (1985) termed "alphabetic." When children are able to segment and blend the individual phonemes, it leads to the final "consolidated" phase, which happens as a result of encountering a spelling pattern several times in the full alphabetic phase, thereby creating whole word representations in memory that are activated instantly when seen. Oral vocabulary, although implicit in Ehri's phases, plays a crucial role at the word-level starting at the partial alphabetic phase and the full alphabetic phase (Nation & Snowling, 2004; Ricketts et al., 2007), and continuing all the way through to the consolidated phase (Oullette, 2006). This is consistent with SVoR and underscores the role of role of oral vocabulary in reading development at the word level and not exclusively for reading comprehension. Indeed, Metsala and Walley (1998) propose that oral vocabulary growth is the mechanism by which phonemic awareness develops, which is the basis of word-level reading. Nonetheless, once children crack the alphabetic code, the bulk of vocabulary growth is a consequence of exposure to print. Thus, vocabulary in the auditory and visual modalities is important for facilitating word recognition.

Ouellette and Beers (2010) examined the role of oral vocabulary in regular and irregular word reading in a sample of English-speaking children from first grade and sixth grade. They found that oral vocabulary predicted decoding after controlling for PA only for the sixth graders, whereas PA predicted decoding for the first graders. The fact that oral vocabulary exerts an influence on word-level reading in later grades rather than during the early stages of reading development is intriguing and is counterintuitive to SVoR and to Ehri's phases.

There remain two critical issues. First, while Frith's (1985) and Ehri's (2005) phases offer a sequential account of reading development, neither explain the rapid move from the novice reader who decodes words with considerable effort to the efficient, effortless reader. Second, the bulk of reading research, as mentioned in the introduction, is based on English orthography, which in recent years has been labelled an "outlier" orthography (Share, 2008; 2014). Orthographic depth and orthographic learning via Share's self-teaching hypothesis and a psycholinguistic grain size theory will be discussed in the next section, to address both issues.

Orthographic Depth and Arabic Orthography

Ehri (2005a, 2005b) questioned whether phase theories, based on an opaque writing system such as English, apply to transparent writing systems. The consistency of spelling-to-sound relationships varies in different languages. In transparent/shallow orthographies (Katz & Frost, 1992; Ziegler & Goswami, 2005) such as pointed Hebrew, vowelized Arabic, and Greek, these relationships are very consistent, as phonemes have a nearly one-to-one mapping to graphemes. In contrast, in some orthographies, such as English, a phoneme might have several spellings, and a grapheme might have several pronunciations. To illustrate, consider the grapheme EA in English, which maps onto a different phoneme in different words, such as <steak>, <bread>, and <meat>. Writing systems such as English are known as deep or opaque.

Word-reading accuracy typically reaches a ceiling in the middle of the first grade in transparent orthographies, whereas for deep orthographies, such as English, it is achieved 2.5 years later (Seymour et al., 2003). Seymour et al. (2003) investigated literacy development across 13 European orthographies varying in depth. The authors assessed letter knowledge, high-frequency word identification, and decoding of simple nonwords. The sample consisted of 648 first and second graders. The findings of Seymour et al. were consistent with those of the orthographic depth hypothesis, demonstrating that foundation literacy skills were more easily learned in shallow orthographies as opposed to deep orthographies.

As mentioned, Ziegler and Goswami (2005) proposed PGST to account for reading development as modulated by orthographic depth. They argued that beginning readers have to contend with three issues: availability, consistency, and granularity. Availability refers to the fact that not all phonological units are explicitly available in oral language prior to reading instruction. Consistency refers to the level of consistency in the mapping between the phonological units and orthographic units. Granularity refers to the size of the orthographic units that must be learned in order to access the corresponding phonological units. According to PGST, the size of the lexical units determines the developmental strategy used for reading. In other words, children learning to read in a transparent orthography will rely on small grain size (at the phoneme level), whereas children learning to read in an opaque orthography will rely on more than one grain size (e.g., rhymes). This theory offers an explanation of the discrepancies in the development of word-level reading in different languages. Thus, the characteristics of Arabic’s orthography are discussed next in relation to PGST.

Arabic orthography is referred to as an Abjad, a consonantal alphabetic language. It is written cursively from right to left. It has 28 letters; all are consonants except for three that are long vowels (Friedmann & Haddad-Hanna, 2014). Although short vowels exist as diacritical marks above and below letters, these diacritical marks are omitted in nearly all texts with the exception of those for the novice reader. Arabic orthography is considered to have a near one-to-one correspondence when diacritical marks are included; however, once these diacritical marks (vowels) are removed, Arabic orthography becomes ambiguous due to the under-representation of vowels, resulting in many homographs. Arabic orthography has been considered shallow/transparent when diacritics are present and opaque/deep orthography when diacritical marks are removed; thus, the source of its opaqueness is different from that of English orthography (Friedmann & Haddad-Hanna, 2014). Just as words are encountered enough times to become familiar enough to enjoy high-quality orthographic representation, diacritical marks are removed and this lexicalization process begins again.

Arabic orthography is believed to have originated from Nabatean script, which is a descendent of Aramaic (Saiegh-Haddad & Henkin-Roitfarb, 2014). Because Arabic has additional consonants that were not represented by the Aramaic alphabets, the letter shapes had to be modified. This resulted in many letters sharing the same basic shape, with dots (one, two, or three) appearing on, below, or in the letter. These are different from the diacritical marks that represent vowels. The placement and number of dots create an entirely different consonant. These diacritics are an integral part of the letter, just as the dot that appears on <j, i> in the Latin alphabet, but if the dot on either <j> or <i> is omitted this does not create another letter. In Arabic, however, this is the case. In addition, 22 of the 28 letters ligate (attach) and go through considerable change in their shape depending where they are positioned in the word (initial, medial, final) (Fragman, 2013). Only six letters are not permitted to ligate to letters that follow them but can ligate to letters preceding them (Saiegh-Haddad & Henkin-Roitfarb, 2014). The position and ligature of letters can create up to four distinct shapes that represent one phoneme.

The under-representation of vowels coupled with similarities in letter shapes and changing letter forms depending on the position in the word makes the Arabic script visually demanding (Khateb et al., 2014). Table 1 shows the Arabic letters in their different forms.

Table 1

The Position and Ligature of Arabic Letters

Final non ligated	Final ligated	Medial ligated	Initial (or medial non ligated)	IPA	Graphemic transcription
ا	ـا	ـا	ا	a	A
ب	ـب	ـبـ	بـ	b	B
ت	ـت	ـتـ	تـ	t	T
ث	ـث	ـثـ	ثـ	θ	θ
ج	ـج	ـجـ	جـ	ǯ	J
ح	ـح	ـحـ	حـ	ħ	Ħ
خ	ـخ	ـخـ	خـ	x	X
د	ـد	ـد	د	d	D
ذ	ـذ	ـذ	ذ	ð	ð
ر	ـر	ـر	ر	r	R
ز	ـز	ـز	ز	z	Z
س	ـس	ـسـ	سـ	s	S
ش	ـش	ـشـ	شـ	š	Š
ص	ـص	ـصـ	صـ	ṣ	ṣ
ض	ـض	ـضـ	ضـ	*ḍ*	*ḍ*
ط	ـط	ـطـ	طـ	*ṭ*	*ṭ*
ظ	ـظ	ـظـ	ظـ	ð	ð̣
ع	ـع	ـعـ	عـ	ʕ	ʕ
غ	ـغ	ـغـ	غـ	ɣ	Ɣ
ف	ـف	ـفـ	فـ	f	F
ق	ـق	ـقـ	قـ	q	Q
ك	ـك	ـكـ	كـ	k	K
ل	ـل	ـلـ	لـ	l	L
م	ـم	ـمـ	مـ	m	M
ن	ـن	ـنـ	نـ	n	N
ه	ـه	ـهـ	هـ	h	H
و	ـو	ـو	و	w/u:	W
ي	ـي	ـيـ	يـ	y/i:	Y
ة	ـة				Ḧ
ء (أ ؤ ئ)	ء (أ ؤ ئ)	ء (أ ؤ ئـ)	أ / إ	ʔ	ʔ

Note. Adapted from Friedmann and Haddad-Hanna, 2014

Researchers of Arabic have often suggested that the visually complex script is the source of necessary effort and impedes native Arabic speakers' fluency in reading (Dai et al., 2013; Ibrahim & Eviatar, 2012; Taha, 2013). For example, Khateb et al. (2014) examined the role of ligature (connectivity) on word recognition in Arabic, using a cross-sectional design of third-, sixth-, and ninth-grade students. They created three conditions: (a) non-connecting words (NCw), (b) partially connected (i.e., a mixture of connecting and non-connecting letters) (PCw), and (c) connected letter words (Cw). The researchers hypothesized that, because connecting letters have up to four shapes depending on their position in the word, such words would yield the longest reaction times (RTs). Indeed, this was the case for the less reading-experienced third-grade students. The opposite pattern was observed in the results for sixth-grade students. Connectivity did not seem to influence RTs for the more skilled readers in the ninth grade.

Dai et al. (2013) also examined the role of ligature in Arabic among third-grade students. They constructed nonwords that were either connected or non-connected and hypothesized that decoding the non-connected nonwords would be more accurate and fluent (as measured by RTs), but after exposure (i.e., once orthographic learning had occurred), connected items would be more accurate and faster to read. The results were inconsistent with their predictions and revealed that connected items were read over half a second longer than were unconnected items, showing that connectedness did not facilitate word recognition. In the second experiment, they manipulated the number of consonant diacritics (e.g., dots appearing in, above, or below consonants). Indeed, they found that many consonant diacritics were the source of errors.

Based on the definition of orthographic consistency in Ziegler and Goswami’s PGST, fully vowelized Arabic would be considered transparent, and so reading at the word level should develop with relative ease; however, this is not the case. As demonstrated from the studies above, the definition of consistency within contemporary theories of word recognition does not encompass all of the orthographic features of Arabic, and that appears to be the source of reading difficulties. Share and Daniels (2016) put forward this viewpoint. They argue that the current definition of consistency is based on European orthographies and does not account for other sources of orthographic depth. Thus, it may not be appropriate to define Arabic's orthographic depth by consistency alone.

In addition to Arabic's complex orthography, the language has a rich morphology that consists of roots and patterns. The root consists of trilateral (sometimes quadrilateral) consonants and carries the fundamental meaning but cannot be pronounced. The pattern is either derivational or inflectional (Saiegh-Haddad & Henkin-Roitfarb, 2014). The pattern consists of vowels affixed in a non-linear fashion to the root to signal grammatical information such as gender, number, and verb tense (Fragman, 2013). To demonstrate, the consonant root <KTB>, although unpronounceable, carries the basic meaning of *writing*; affixing a pattern to different slots of the root (e.g., <KATABA>) results in a pronounceable word that means *wrote* and is semantically related to the root <KTB> (Taha, 2013). However, in un-vowelized script, the pattern (the position of the vowels) can only be inferred through the context in which the word appears, because vowelization is via diacritical marks, which, as mentioned, are usually omitted from most texts for adults. This results in many homographs (Laiel et al., 2014; Fragman, 2013). The pattern does not always consist of vowels only. In addition to vowel information, the word could have additional consonants, such as <MAKTABA> (from the root <KTB>), which means *library*.

The rich morphology of Semitic languages has prompted researchers to suggest that the mental lexicon in Arabic, unlike in English and other Indo-European languages, is morphologically organized. This assumption stems from studies that demonstrate that Arabic and Hebrew (another

Semitic language) native speakers show priming effects when words overlap in morphology but not form (orthography). For English and Indo-European languages the opposite is found (Frost et al., 2005). Clearly, Arabic's complex orthography and morphological richness differ considerably from that of English and other European languages.

Although PGST offers a framework that explains reading development theories from a cross-linguistic perspective, Arabic's linguistic and orthographic characteristics may not be fully represented within this framework. Furthermore, individual differences exist among children learning to read, regardless of orthography. As discussed, SVoR leaves implicit many of the skills and abilities that determine word reading. Several cognitive, linguistic, and environmental factors have been identified to account for differences in reading acquisition and will be the focus of the following section.

Predictors of Reading and Spelling

Learning to read in any language is rooted in the mapping of spoken and written language. More specifically, in alphabetic orthographies, cracking the alphabetic code is contingent on leaning arbitrary symbols (orthography) and their corresponding sounds (phonology). It is therefore not surprising that letter knowledge and phonological processing abilities are considered universal in reading development. This constant interplay between phonology and orthography is what determines reading success. Phonological processing abilities have been identified by researchers as crucial for successful early reading success, namely phonological awareness (PA) (Snowling, 1998; Rose, 2009) and verbal short-term memory (STM). PA, which refers to the ability to manipulate units in spoken language, has been found to be strongly linked to the development of word reading and spelling (Hulme et al., 2002; Muter et al., 2004; Vellutino et al., 2007; Haigh et al., 2011) and to be a predictor of early word-reading development (Muter et al., 2004; Snowling, 2001). Commonly used tasks that assess PA involve isolating phonological units (e.g., what is the first sound in /cat/?), blending phonological units (e.g., what word is /s/, /k/, /u/, /l/?), and deletion (can you say /cat/ without /k/?). Verbal STM and working memory have been shown to be important in reading acquisition (Wagner et al., 1997). The conversion of grapheme-to-phoneme requires more than just identification of each unit; also, the temporary storage until subsequent units are decoded and finally assembled (Baddeley, 2012). Maintaining and recalling phonological units in a serial order is crucial for reading because the sequence of sounds that have been decoded need to be retained until the remaining sounds are decoded (Baddeley, 2012; Catts et al., 2006; Wagner et al., 1997). To assess phonological verbal STM, a commonly used task is nonword repetition, whereby participants are asked to repeat nonwords of increasing difficulty (Wagner et al., 1999). Rapid automatized naming (RAN), assessed through the rapid naming of letters, numbers, objects, or colors, has also been linked to reading development in various orthographies (Manis et al., 1999; de Jong & van der Leij, 1999; Kirby et al., 2003; Georgiou et al., 2016; Georgiou et al., 2021). These abilities are considered to constitute the universal predictors of reading and spelling. However, much debate exists on the role they play in reading development in various languages of varying orthographic depth.

Converging evidence from cross-linguistic studies demonstrates the universality of PA, verbal STM, and RAN as predictors of reading development (Georgiou et al., 2008; Ziegler et al., 2010; Caravolas et al., 2012; Moll et al., 2014). For example, Ehri et al. (2001) conducted a meta-analysis of the efficacy of phonemic awareness instruction on reading and spelling development. The authors reviewed 52 studies. Phonemic awareness instruction was found to contribute to reading and spelling acquisition in various alphabetic orthographies that included English, German, Dutch, Finnish, Hebrew, Norwegian, Spanish, Danish, and Swedish. Similarly, Caravolas et al. (2012) found that phonemic awareness was a strong predictor of reading and spelling across different European alphabetic writing systems.

Although the evidence for the universality of PA in reading is clear, Arabic's linguistic characteristics, such as diglossia, described earlier, may yield different findings from those reported in European languages and even Semitic Hebrew.

What is not yet clearly understood is the exact role and magnitude of these skills in the course of reading acquisition and development in varying orthographies. Thus far, cross-linguistic studies have demonstrated that the relationship between reading and PA is diminished in transparent orthographies beyond the early stages of reading acquisition (Georgiou et al., 2008). However, the contribution of PA has been shown to persist even in older grades in opaque orthographies (Deacon & Kirby, 2004; Wagner et al., 1997). In other words, PA is considered to be more important in opaque orthographies (Muter et al., 2004, Wagner et al., 1997) while RAN is considered a stronger predictor of reading development in transparent orthographies (e.g., Landerl & Wimmer, 2008). For example, Landerl and Wimmer (2008) found that the contribution of RAN was greater in transparent orthographies, and Caravolas et al. (2012) found that RAN longitudinally predicts reading in all four orthographies (English, Spanish, Czech, and Slovak). Caravolas and colleagues conclude that “RAN seems to tap a separable mechanism that is involved in forming associations between printed words and their pronunciations” (p. 684), and as such, is not influenced by orthographic depth.

Unlike PA that requires explicit manipulation of phonological units, verbal STM is considered an implicit phonological task that requires no reflection on the sound structure. In a meta-analysis, Melby-Lervag et al. (2012) examined the role and magnitude of both PA and verbal STM to reading development. They hypothesized that PA and verbal STM would be correlated to individual differences in reading skills; however, PA would emerge as the stronger predictor. Indeed, results of the meta-analysis revealed that PA was the strongest predictor of reading in studies of children with and without reading disabilities. Verbal STM was a unique predictor of reading, independent of the contribution of PA. Interestingly, these results were not modulated by orthographic depth. These results are consistent with Ziegler et al.’s (2010) findings that verbal STM is equally important in opaque orthographies. However, according to the PGST, verbal STM may be considered more important in transparent orthographies since units of smaller grain size must be decoded and assembled. Indeed, Caravolas et al., (2005) found verbal STM to predict reading fluency in transparent Czech but not in opaque English.

Saiegh-Haddad (2003) hypothesized that kindergarteners and first-grade native Arabic speakers would find difficulties in the acquisition of basic foundation skills for reading when they first encounter MSA in formal schooling due to the diglossic situation discussed earlier. Her results confirmed that children at both grade levels had difficulty isolating initial and final phonemes in MSA that did not exist in SA (phonemes that do not exist in their spoken variety), thereby confirming that the linguistic distance of both varieties of Arabic creates interference in basic reading skills. This study illustrates how Arabic's diglossic situation interferes with reading acquisition. As mentioned, diglossia impedes reading acquisition when children are at the partial alphabetic phase because they cannot depend on oral vocabulary to aid them to partially decode. This situation is cognitively taxing for the novice reader (Abu-Rabia, 2000; Asaad & Eviatar, 2013). Furthermore, Abu-Rabia et al. (2003) found verbal STM to be impaired in a sample of reading-disabled fifth-grade Arabic speakers. In contrast, Taibah and Haynes (2011) found no relationship between verbal STM and reading in Arabic in their sample of children from kindergarten to third grade. Due to the conflicting results from cross-linguistic studies and the literature on Arabic, PA and verbal STM will be examined in the current study to examine their potential to account for the variance in reading acquisition for the population in question.

In the Taibah and Haynes (2011) study, the researchers examined predictors of reading in Arabic among 237 children recruited from classes from kindergarten to third grade. They assessed PA, verbal STM, and RAN as predictors of Arabic reading accuracy and fluency. The authors reasoned that, due to the transparency of Arabic's fully vowelized script, RAN would be a stronger predictor than would PA. Contrary to their expectations, PA emerged as the strongest predictor of reading accuracy and fluency at every grade level. The predictive power of RAN showed a gradual increase in second and third grades.

Assad and Eviatar (2014) examined correlates of reading in Arabic in a sample of first, third, and fifth graders. They measured PA, RAN-letters, and visual abilities. PA was correlated with reading accuracy at all grade levels. The authors attribute the persistence of the influence of PA on reading in the older children to diglossia and the ambiguity of the Arabic script once vowels are removed. RAN was correlated to reading fluency for the first and fifth graders but not the third graders. The authors attribute this unusual finding to the transition from a vowelized script to a non-vowelized script, which creates an opaque script and thus a heavy reliance on PA.

Abu Ahmad et al. (2014) investigated linguistic and cognitive predictors of word recognition and reading comprehension in a sample of 194 kindergarten Arabic speakers who were followed longitudinally in second grade. Although the Arabic script is considered to be transparent when it is vowelized, the authors reasoned that due to diglossia and orthographic complexities, reading acquisition would be hampered. Indeed, unlike transparent European orthographies, where reading is acquired by the end of first grade, Arabic reading acquisition took longer. Their results revealed that in fully vowelized Arabic, sub-lexical factors such as phonemic awareness, phonological processing, and early literacy skills, were the best predictors of word recognition. Even after controlling for general ability and vocabulary, these sub-lexical abilities accounted for 25% of the variance in word recognition. In fact, phonemic awareness was found to be the strongest predictor in word recognition. This finding was unexpected for a transparent script such as vowelized Arabic since reliance on phonemic awareness diminished after first grade in other studies carried out with transparent writing systems. This inconsistent finding, coupled with the low reading accuracy scores of the second graders, reveal that transparency as defined by Indo-European scripts may not be adequate to fully account for considerable difficulties experienced by Arabic readers (e.g., similarities of letter shapes, diacritical marks, ligature).

In summary, the literature reviewed demonstrates the importance of PA, verbal STM, and RAN in word-reading acquisition, particularly in alphabetic languages. However, cross-linguistic literature on the roles of PA, verbal STM and RAN has yielded conflicting results. Furthermore, the overwhelming majority of cross-linguistic literature is based on English and European languages, which differ considerably from Arabic. Although Ziegler and Goswami's PGST offers a framework for reading development from a cross-linguistic perspective, it falls short of encompassing unique linguistic features like diglossia and other visual and orthographic features that are the source of orthographic depth in Arabic (Daniel & Share, 2014).

Orthographic Processing, Orthographic Sensitivity and Print Exposure

Just as there are individual differences in children's phonological processing and naming speed, there exist considerable differences in orthographic processing abilities, and these differences are also predictive of reading development (Cunningham & Stanovich, 1990; Apel, 2011). Existing literature contains ambiguous and inconsistent terms for describing orthographic knowledge. According to Apel (2011), orthographic knowledge refers to information stored in memory in written form to represent spoken words. It has been suggested that the two components of orthographic knowledge are mental orthographic representations—whole or partial word spellings stored in memory (Apel, 2011)—and orthographic pattern, which pertains to knowledge of rules that govern the English language. Finally, orthographic processing refers to the ability to learn, store, and retrieve orthographic representations and pattern knowledge (Apel, 2011).

Cunningham and Stanovich (1990) examined the unique contribution of orthographic processing in word recognition. They administered phonological processing tasks of phoneme deletion and phonological choice, whereby the participant had to judge which of the pair of pseudo words sounded like a real word (e.g., <kake>, <dake>) to 51 third-grade students and 47 fourth-grade students. The children were also given two orthographic choice tests. In the first, pairs of letter strings had the same phonology, but only one had the correct spelling (e.g. <lurn>, <learn>). In the second, a homophone choice test was administered. The experimenter asked “which one is a fruit?” and the child was presented with two printed words (e.g., <pair>, <pear>). Composite scores of the two phonological tasks and the two orthographic tasks were derived. They found that orthographic processing accounted for unique variance (10.2%) in word-recognition abilities, after accounting for the contribution of phonological processing.

Share (2004) proposed an orthographic sensitivity hypothesis that posits that individual differences in orthographic learning are attributable to differences in attention to orthographic detail. This sensitivity to orthographic detail is surely developed through experience with print. The amount of print exposure (time spent reading for leisure) a child experiences seems to play a crucial role in learning to read (Mol & Bus, 2011). Children with reading difficulties often exhibit a Matthew effect, demonstrating more depressed vocabularies due to less exposure to reading materials (Aguiar & Brady, 1991; Cunningham & Stanovich, 1997; Stanovich, 1986). Children who enjoy reading will often have better reading comprehension skills, whereas children who do not enjoy reading, possibly due to underlying reading difficulties, naturally shy away from reading, thereby widening the gap between good readers and poor readers (Stanovich, 1986). Anderson et al. (1988) used daily-activity diaries to estimate children's out-of-school reading habits and reported that print exposure (time spent reading for leisure) predicted variation in reading proficiency from the second to the fifth grades. They also determined that time spent reading explained 14.4% of the variance in the reading comprehension of the fifth-grade children. Cunningham and Stanovich (1990) examined the contribution of print exposure to word recognition in a sample of third- and fourth-grade children. The researchers found that print exposure accounted for 6.2% of the variance in word recognition after statistically controlling for age, general cognitive abilities, memory, and phonological processing skills. McBride-Chang et al. (1993) examined the effects of print exposure on word reading and reading comprehension in a sample of disabled and non-disabled readers of fifth- to ninth-grade students. Their findings suggest that print exposure can explain unique variance in word-reading abilities after controlling for phonological and orthographic skills for the non-disabled readers.

Print exposure might even be responsible in shaping the development of cognitive abilities, such as RAN. Powell et al. (2014) measured print exposure using an Author Recognition Test (ART) to determine whether the low RAN group scores in their study could be attributed to impoverished print exposure. No such evidence was found; the control and low RAN groups' performance did not differ on this measure. However, the authors adapted an ART that was developed by Ricketts et al. (2007) for a younger group of children (age 8;08 to 9;09), whereas Powell et al.'s sample consisted of two groups (the mean age of the younger group was 10.7 and the older group 11.2). Thus, the ART may not have been appropriate for the children in Powell et al.'s study. In addition, such checklists are ephemeral and need to be updated to maintain reliability (Stainthorp, 1997), meaning that the ART that was developed by Ricketts et al. (2007) may have been unreliable several years later.

Despite the evidence for the importance of print exposure in the course of reading development, it is seldom measured or controlled as a variable. This is due to the difficulty in measuring reading volume; however, this limitation has been addressed through the development of the Title Recognition Test (TRT) (Stanovich & West, 1989). Thus, an aim of this research was to develop a TRT in both Arabic and English to examine the role of print exposure as a predictor of reading and spelling in both languages.

Summary

Clearly, cognitive, linguistic, and environmental factors are intertwined in the course of reading development. The role and magnitude these factors play in reading development is influenced and shaped by the characteristics of the script. Most of the evidence, as mentioned, comes from English and European languages (Share, 2008; Share & Daniels, 2016), which do not reflect the linguistic and orthographic features of many other languages, and thus limits our understanding of the abilities that are universal versus those that are the result of the script. Therefore, predictors of Arabic and English reading and spelling will be addressed in this research in an attempt to separate the skills that are universal from those that are script dependent.

This chapter began with a macro, yet simple, view of reading to demonstrate that even the basic prerequisite, oral comprehension, is absent in Arabic due to diglossia. This creates a situation akin to learning a foreign language for the novice reader of Arabic and poses a challenge to developmental reading theories such as Ehri's four-phase theory of word recognition. Ehri's partial alphabetic phase is characterized by partial decoding that is facilitated by existing oral vocabulary to achieve the correct pronunciation, which then leads to the question of how Arabic speaking children learn to read in the absence of oral vocabulary. This account, though, is based on English, an outlier orthography (Share, 2008), whereby the dependence on existing vocabulary is necessary due to the opaqueness of its script. As described, vowelized Arabic has a near one-to-one correspondence between graphemes and phonemes; thus, Arabic's orthographic consistency is sufficient for correct pronunciation without the aid of existing oral vocabulary. Indeed, according to PGST, orthographic consistency facilitates reading development in transparent orthographies. However, PGST does not account for the visual and orthographic complexities of Arabic. In fact, the literature in vowelized Arabic indicates that reading development is similar to that found in opaque languages rather than in transparent ones. But even if we only consider cross-linguistic evidence on English and other European languages, there is variability in the role and magnitude of abilities found to underpin reading development. Thus far, the discussion has been dominated by the abilities that underpin reading development as modulated by orthographic depth; however, it is equally important to examine the process by which word representations are formed. The next section will present the literature on the most extensively studied cognitive ability of word recognition (Adelman, 2012a).

The Self-Teaching Hypothesis

The process of forming word-specific representations while reading is described by Share's self-teaching hypothesis (Share, 1995, 1999, 2004), which stipulates that, once the alphabetic rule is acquired, children will phonologically decode unfamiliar words (given that correct decoding occurs) and self-teach implicitly. This self-teaching mechanism provides the novice reader with the opportunity to decode numerous words encountered in print and acquire mappings between orthography and phonology, thereby building up the orthographic lexicon. Evidence in support of this hypothesis comes from several studies and varying orthographies, such as English (e.g., Cunningham, 2006; Nation et al. 2007; Ricketts et al., 2011), Dutch (e.g., de Jong & Share, 2007), and Hebrew (e.g., Shahar-Yames & Share, 2008; Share, 1999, 2004).

The self-teaching paradigm used in research, such as the studies listed above, mimics to a large degree how children naturally learn, and entails reading a story with a nonword (e.g., <yait>) embedded in text. The child reads the story, and then orthographic learning is assessed using an orthographic choice test, where the target stimulus is presented along with three distractors, and the child must choose the correct one. For example, if the target item is <yaite>, distractors will be a homophone (e.g. <yate>), a visually similar foil (e.g., <yaile>), and one involving transposition of two adjacent letters of the target (e.g., <yiate>). The task would involve timed reading aloud. Latencies to read targets are generally faster than to read homophonic foils. This indicates that word learning occurs rapidly even when presented with similar distractors. Finally, a spelling test is conducted in order to assess independent recall of the target item. The experimenter would ask the child, “Do you remember the story you read to me about the hottest town in the world?” (Share, 1999, p. 105). If the child is unable to recall the target, the experimenter would provide the first syllable, and if that prompt fails to elicit a response, then the word is provided orally by the experimenter and the child would then write down the target. These tasks are most commonly used to assess orthographic learning and have consistently demonstrated rapid orthographic learning in young children.

Using the methods mentioned, Share (1999) tested the self-teaching hypothesis in a sample of 40 second-grade Hebrew-speaking children who were shown to be good at decoding. The children read 10 short stories aloud with a target nonword embedded either four or six times in meaningful text. After the stories were read, children were asked five comprehension questions on each passage, to ensure they understood the stories. They were literal questions that could only be answered by understanding the text. Three days later the children’s orthographic learning was assessed using an orthographic choice test where the child saw the target and three foils and was asked to choose the one he or she had seen in the story. Next, the child was asked to read as quickly as possible the items that were presented on a computer screen, and naming latencies and accuracy were recorded. Finally, the children were asked to recall the targets, for example, they were asked “Do you remember the story you read to me about the hottest town in the world?” and then they were asked to write the target item down (Share, 1999, p. 105). On all three post-tests, the children showed orthographic learning; however, the effect of number of exposures (four times versus six) was not significant.

Utilizing this self-teaching paradigm, Cunningham et al. (2002) examined the role of self-teaching during reading in English. The results mirrored those obtained by Share (1999) in his sample of Hebrew-speaking children. Targets that were decoded correctly enjoyed better recall on post-tests, further confirming the central role of phonological decoding in acquiring word-specific representations. However, individual differences in orthographic learning were not solely attributed to successful decoding but also to existing orthographic knowledge as measured by an orthographic choice task. Share (1999) only assessed receptive vocabulary, pseudo-word accuracy, and fluency, and found no relationship between these background measures and orthographic learning.

Nation et al. (2007) examined the self-teaching hypothesis among eight- and nine-year-old native English speakers. They manipulated the number of exposures (once, twice and four presentation times of targets), context (targets embedded in a story versus presented in isolation), and durability (post-tests administered one day after presentation versus seven days after). Once again, orthographic learning was evident; however, unlike Share’s (1999) finding, orthographic learning increased with number of exposures. In a transparent orthography, like Hebrew and Arabic (when vowelized), it appears that number of exposures of the target unfamiliar letter string may not play a crucial role as it does in opaque English. Context had no effect on orthographic learning.

Ziegler et al. (2014) ran several simulations to test the merits of the self-teaching hypothesis using their computational connectionist dual process (CDP) model. The model was able to self-teach rapidly after letter–sound correspondences had been taught explicitly using a learning algorithm. After each successful pronunciation of a word, the orthographic lexicon was updated. The model was successful on 80% of the 32,735 words presented.

Ziegler et al. were equally interested in what happens when phonological decoding is impaired. They manipulated the model to include an incorrect learning condition where incorrect pronunciation was learned. The authors found that, even in the incorrect learning condition (incorrect decoding on the initial trial), the model was able to account for 45% of the words in the database, suggesting that if a word is not learned once, it might be during different exposures in different contexts. Such results support the strength of the self-teaching hypothesis. The authors stressed that continuous exposure to a given word in different contexts will be sufficient for establishing the correct orthographic representation.

Indeed, the findings of Ziegler et al. corresponded with Perfetti and Hart's (2002) lexical quality hypothesis. According to this, a high-quality representation of the word is established when all three constituents—phonology (sound), orthography (spelling), and semantics (meaning)—are activated. The more exposure an individual has to the word, the stronger the connections become among the constituents, establishing a unitary lexical entry. The quality of the word-specific representation defines how well it is identified and recalled (Verhoeven & Perfetti, 2011). Thus, the self-teaching hypothesis has been tested in various orthographies and more recently by reading modelers (Ziegler et al., 2014), showing consistently robust and rapid orthographic learning even when words are not initially decoded correctly. The finding that orthographic learning is possible even when words are decoded incorrectly is inconsistent with Share's (1999) suggestion that orthographic learning is parasitic on successful phonological decoding.

Support for the fact that successful phonological decoding may not be necessary for orthographic learning comes from studies demonstrating that incidental exposure to a word's form is sufficient for orthographic learning. For example, Ricketts et al. (2009) investigated the role of orthography in vocabulary acquisition in 58 eight- to nine-year-old children. Using a visual-verbal paired associated learning (PAL) paradigm, the children were trained to associate 12 nonwords with pictures of novel objects. Participants' learning of the pairings was assessed using spelling and nonword-and-picture-matching tasks. Ricketts et al. demonstrated that stimuli presented with orthography, but without drawing attention to it, resulted in better recall and more accurate spelling. In other words, implicit orthographic learning occurred in the absence of phonological decoding. Using the same paradigm, Rosenthal and Ehri (2008) reasoned that orthographic mapping to the phonological form (pronunciation) of a word and its semantic information should optimize learning and memory storage of new words. In an experiment involving 20 second graders (mean age = 7.7 years), similar to Ricketts et al.'s results, they found that the incidental presentation of orthography facilitated learning in both pronunciation recall and definition recall for low-frequency words. A second experiment conducted with 32 fifth-grade children mirrored results from the first experiment with the second-grade children, favouring the presence of orthography. The benefit of incidental presentation of orthography was also demonstrated in children with both high and low phonological decoding abilities learning English as a second language (ESL) by Hu (2008).

These studies demonstrate powerful and implicit learning from mere incidental exposure to orthographic representations and are consistent with statistical models of learning (Pacton et al., 2001; Qi et al., 2019). Perhaps the distinction in methodology between the self-teaching paradigm and the paired-associate learning paradigm in these studies best demonstrates the implicit mechanism that is not parasitic on phonological decoding but in fact connected to it through initial auditory statistical learning that begins in infancy at the phonemic level (Metsala & Walley, 1998), and leads to oral vocabulary acquisition (Smith & Yu, 2008), and visual statistical learning emerges with exposure to print.

The ability to implicitly learn linguistic statistical regularities such as detecting the boundaries of segments of speech is not a novel idea (Saffran et al., 1996). Similarly, links between statistical leaning and reading have been demonstrated (e.g., Cassar & Treiman, 1997; Pacton et al., 2001). Acquiring new vocabulary is rooted in the statistical co-occurrence between sound and referent (Smith & Yu, 2008) and thus, individual differences in vocabulary acquisition have been linked to differences in statistical learning. Despite the early links between statistical learning and reading and spelling, the topic has received little attention in the literature until it was revived recently (e.g., Spencer et al., 2015; Arciuli, 2018; Qi et al., 2019; Elleman et al., 2019). In fact, one could argue that connectionist models (discussed below) are models of implicit statistical learning. Because reading and spelling instruction at the word level could never be exhaustive in any given language, implicit statistical learning offers a compelling account of this rapid process (Apel et al., 2006). Further evidence for the implicit statistical nature of this leaning comes from studies that show that once words are consolidated into the lexicon, reading is no longer a strategy but rather an involuntary process (Ehri, 2005b) (e.g., the Stroop effect). This spontaneous and unconscious ability is consistent with the definition of statistical learning.

Interest in statistical learning as it pertains to orthographic learning is growing. Qi et al. (2019) examined statistical learning in both the auditory and visual modality in a sample of 36 children and 36 adults. They assessed sentence reading fluency as an outcome measure for children and adults, and nonword reading accuracy for the children only. They reasoned that auditory statistical learning would be more associated with reading and that the relationship of visual statistical learning to reading would be mediated by phonological processing abilities (e.g., PA, verbal STM). Auditory statistical leaning was associated with reading (nonword reading accuracy), and this relationship seemed to be mediated by phonological processing abilities. However, the nonword reading task is arguably assessing PA, and thus the relationship is not surprising. The authors take this as evidence reflecting the primary and persistent role of phonological input from one’s native language in infancy to reading acquisition. This is consistent with Metsala and Walley's (1998) claim that oral lexicon growth is contingent on phonemic and early reading ability.

Unlike all the reading models described thus far, statistical learning in both the auditory and visual modalities captures the interaction of both the cognitive processes (e.g., attention, perception, memory) and environmental influences (e.g., characteristics of the language, instructional strategies, exposure to print) involved in reading and spelling acquisition. Thus, tasks of incidental orthographic learning during reading (e.g., Ricketts et al., 2009, Rosenthal & Ehri, 2008) demonstrate the implicit acquisition of word-specific representations after just a few exposures. While the use of paradigms like PAL may be criticized for not being as naturalistic as the self-teaching paradigm, this is not the case in a diglossic situation, where novel word pairing is the norm and not the exception.

Gaps in the Literature

Several theories have been proposed to explain reading in Arabic. The literature reviewed in this chapter highlights the fact that theories of reading development are based on English and other European languages. First, SVoR was used as a starting point as a simple framework; however, the diglossic situation in Arabic complicates one of the core components of the model. How do Arabic-speaking children learn to read when oral comprehension is diminished? Ehri (2005b) provides insight in her developmental phase theory of word recognition. Ehri speculated that the partial alphabetic phase may not be necessary in transparent orthographies. To account for orthographic depth and its influence on reading development, PGST was used as a framework; however, as is clear from the literature reviewed on predictors of reading from a cross-linguistic perspective, Arabic deviates considerably from the standard definition of orthographic depth due to its underrepresentation of vowels and other visual complexities. Arabic's visual and orthographic features challenge our understanding and very definition of orthographic depth. While Ehri's phase theory coupled with Ziegler and Goswami's PGST provide a complimentary account of reading development as modulated by orthographic depth, they do not describe the process by which rapid orthographic representations are formed. The account of how children construct orthographic representations was the self-teaching hypothesis proposed by Share (1999). Share described this mechanism as parasitic on correct phonological decoding; however, computational models of reading (Ziegler et al., 2014), and behavioral data (e.g., Rosenthal & Ehri, 2008; Ricketts et al., 2009), demonstrate that this may not be a necessary prerequisite. Indeed, some computational reading modelers use an implicit statistical learning approach (e.g., Davis, 2010) that explains rapid lexicalization even when rules are not taught explicitly.

Despite evidence for the importance of print exposure in the course of reading development, it is seldom measured or controlled as a variable. As such, this is the first study to develop a Title Recognition Test (TRT) in Arabic. While the last two decades have seen a rise in research in Arabic, very few studies (e.g., Taibah & Haynes, 2011; Tibi & Kirby, 2019) have examined as many linguistic and cognitive factors as in the present study, and none have controlled for the role of print exposure as a predictor of reading. While this study adds to the existing correlational literature on reading in Arabic, Castles and Nation (2010) propose that the only way to address the underpinnings of the process of constructing orthographic representations is through training studies that use experimental designs to dissect the components of orthographic learning. This study addressed this by using a paired associate training paradigm to document the process of how Arabic-speaking children construct orthographic representations. Based on implicit statistical learning models, and the linguistic features of Arabic, an incidental word-learning experiment was deemed most appropriate in explaining how Arabic-speaking children construct orthographic representations in the absence of oral vocabulary and limited exposure to print prior to reading instruction. As such, this is the first study to examine incidental orthographic learning in Arabic.

Thus far, this literature review has demonstrated that cognitive, linguistic, and environmental factors are intertwined and difficult to separate in the course of reading development. Currently, there is no universal theory of reading development to account for all these factors. Thus, the primary aim of this study was to gain an understanding of how Arabic-speaking children construct orthographic representations and the underlying cognitive abilities that facilitate orthographic learning, and the interaction of linguistic and environmental factors that influence reading development.

Purpose of the Study

The primary purpose of the study was to examine orthographic learning in Arabic. Much of what we know about reading development comes from literature on English, an outlier script (Share, 2008; Share & Daniels, 2016). According to reading models, the diglossic situation of Arabic surely affects reading development (Saiegh-Haddad, 2018). It was speculated that environmental factors, such as exposure to print, might play an important role in reading development for children learning to read in diglossic Arabic (L1) and would be just as important for learning to read in English (L2). Based on statistical learning models, print exposure may facilitate reading through implicit learning, which would explain reading acquisition in the absence of oral language proficiency in both Arabic (L1) and English (L2). Thus, the development of comparable print exposure measures in Arabic and English was deemed necessary, and these were used in the subsequent research as a control measure.

The literature reviewed on reading development in various languages demonstrates that our current definition of orthographic depth may not be sufficient to explain the orthographic and linguistic complexities of Arabic. The second study involved the examination of predictors of reading in a sample of third-grade Arabic-speaking monolingual children, and Arabic-English bilinguals in the third, fourth, and fifth grades. Based on the literature, predictors of single word reading and spelling in both languages were examined in both samples, and the following questions were addressed:

1. Is print exposure (results from Study 1) related to children's single word reading and spelling in both samples and both languages?
2. What are the general abilities (nonverbal ability, visual concentration, vocabulary, verbal STM, working memory) related to children's single word reading and spelling in both samples and both languages?
3. What are the reading related abilities (PA, RAN, orthographic processing, and morphological awareness) that predict children's single word reading and spelling in both samples and both languages? What is

the role and magnitude of effect of these abilities as a function of development and orthographic depth?

In addition to diglossia being implicated as the source of literacy difficulties, the visual complexities of the Arabic writing system have been implicated. Yet, what is not well understood is whether these visual complexities stem from ligature or diacritics. The third study examined the facilitative effects of orthography on novel word learning using a paired associate learning task in a sample of third-grade Arabic-speaking monolinguals. This study isolated these two potential sources of visual complexity by using nonwords that manipulated conditions of ligature and consonant diacritics. Furthermore, this study examined participant-level predictors of novel word learning to determine which of children's existing cognitive and linguistic abilities might facilitate novel word learning. The third study addressed the following research questions:

1. What are the cognitive and linguistic abilities that may facilitate the learning of newly taught words?
2. Does incidental exposure of orthography facilitate recall (phonological, orthographic and semantic) identities of newly taught words?
3. Which item conditions (non-connected few consonant diacritics, connected few consonant diacritics, connected many consonant diacritics) are facilitated by exposure to orthography?

Organization of the Thesis

The general aim of this study was to examine how Arabic-speaking children construct orthographic representations and to identify the abilities that may facilitate novel word learning. Chapter 2 covers Study 1, the development of TRT in Arabic and in English in order to be able to control for print exposure in the following study. Chapter 3 (Study 2) presents the examination of the effect of predictors of single word reading and spelling in Arabic-English bilingual children and monolingual Arabic-speaking children. Chapter 4 (Study 3) presents the results of novel word learning in Arabic using a paired-associate learning task. Finally, Chapter 5 presents a discussion of the findings of the studies and their implications.

Chapter 2

Study 1: Development of a Title Recognition Test in Arabic and in English

Fostering good reading habits among young children has long been a goal of educators as a means of improving reading-related skills and reading comprehension (Share, 2014, 2008; Zeigler & Goswami, 2005). Stanovich (1986) was the first to coin the term “Matthew Effect” to describe the link between exposure to print and reading development. The Matthew Effect is a situation where children who acquire technical reading skills (e.g., letter-to-sound knowledge, PA, decoding) with relative ease tend to read more for pleasure, thus strengthening their phonological decoding skills, word recognition, vocabulary, and comprehension. The effect appears to be reciprocal. In contrast, children who struggle with reading acquisition tend to shy away from reading, resulting in inadequate word recognition, depressed vocabularies, and poor comprehension skills, hence the biblical reference from Matthew: “the rich get richer and the poor get poorer.”

Anderson et al. (1988) determined that an avid reader (98th percentile) may encounter more than 4 million words per year, while a less keen reader (2nd percentile) may encounter only about 8,000 words per year. Even pre-readers seem to benefit from book exposure. Davidse et al. (2011) asserted that book exposure could be used to predict vocabulary and letter knowledge among kindergarten children. In a longitudinal study of early literacy experience, Senechal and LeFvre (2002) found that early literacy skills predicted word reading at the end of first grade; print exposure predicted vocabulary growth and listening comprehension. Similarly, Zhang et al. (2020) followed children from kindergarten to second grade and found that a formal literacy environment (i.e., reading) at home predicted reading comprehension in Chinese. Thus, a closer look is warranted into children's print exposure and its relationship to literacy development.

Print exposure is rarely measured in studies examining reading development (Anderson et al., 1988), which is due to the complicated nature of quantifying such habits (Goldman & Manis, 2013). A search of the literature revealed that reading researchers have often implicated print exposure as a variable in explaining individual differences; however, most failed to examine the role reading experience plays in reading acquisition. Rather, they assumed that it shapes the development of reading and reading comprehension (Cain & Oakhill, 2011). The reason why researchers often do not control for reading exposure in their investigations is methodological limitations such as time to assess print exposure.

Traditionally, researchers have employed teacher and parent questionnaires, interviews, and reading diaries/logs to determine the individual differences in print exposure (Anderson et al., 1988; Senechal & LeFevre, 2002). Yet, such methods are time-consuming and are subject to contamination due to desire to provide socially acceptable responses (Echols et al., 1996). Based on these observations, the following section contains the rationale for the research.

The purpose of Study 1 was to develop measures of print exposure in Arabic and English appropriate for the sample of children under study, using methods from previous research (Cipielewski & Stanovich, 1992; Cunningham & Stanovich, 1990; Stainthorp, 1997; Masterson & Hayes, 2007; Stanovich & Cunningham, 1992). Print exposure may be of relevance, particularly in Arabic, where oral comprehension may not aid reading comprehension due to the diglossic situation (two varieties of the language, one spoken, and one for literary purposes). Furthermore, constructing such a measure in English for use with English-language learners may additionally shed light on how children, in the present context, gain vocabulary knowledge when they are known to lag behind in their vocabularies in comparison to their English native speaker counterparts (Cipielewski & Stanovich, 1992; Cunningham & Stanovich, 1990; Stainthorp, 1997; Masterson & Hayes, 2007; Stanovich & Cunningham, 1992). Based on the literature reviewed below, this study will demonstrate the importance of measuring print exposure when investigating predictors of reading and reading-related abilities.

Literature Review

Adult Print Exposure and Relationship to Reading-related Skills

Motivated by the fact that some children and adults lag behind in their word recognition skills despite adequate phonological processing skills, Stanovich and West (1989) decided to investigate this discrepancy. They addressed the following two questions: Can orthographic processing skills account for individual differences in word recognition after controlling for phonological processing skills? Can print exposure account for individual differences in orthographic processing? However, traditional methods of print exposure tend to be both time-consuming and susceptible to contamination by participants over-estimating their reading volume. To remedy issues of providing socially desirable answers on questionnaires, Stanovich and West (1989) developed an experimental test, ART, wherein a checklist of real authors and foils (names of people who are not authors) were mixed, and the participant was asked to check the authors' names. The method of using targets and foils was used as way of checking for participant guessing. Scores were calculated on the basis of the number of correct responses (hit rate) and false responses (false alarm rate) (Snodgrass & Corwin, 1988). Stanovich and West (1989) used this kind of checklist in conjunction with a traditional reading habits questionnaire to establish construct validity. The list was constructed to exclude authors who were regularly studied in educational settings. This was done to tap into the participants' leisure reading habits. Similarly, they designed a Magazine Recognition Test (MRT) using the same methods and logic.

In their first study, they recruited 61 undergraduate students and administered two spelling tasks, a reading and media habits questionnaire, and two newly developed print exposure checklists, ART and MRT. Print exposure as measured by ART was correlated with the composite score of the two spelling tests. Interestingly, scores from ART were particularly highly correlated with exception word items (words that deviate considerably from sound-to-print rules) (Stanovich & West, 1989). In their second study, the authors addressed whether print exposure effects are independent of phonological processing skills in a sample of 180 undergraduate students. They administered standardized measures of word recognition and reading comprehension, print exposure (ART, MRT), spelling tasks (standardized spelling test and experimental spelling task), phonological processing (phonological choice task, pseudo-word naming), and orthographic processing (orthographic choice, naming regular, exception, and strange words). The sample was split into good and poor readers according to their scores on the standardized word recognition test prior to analysis.

Results of the orthographic measures did not explain any variance in word recognition once the phonological tasks were entered into the equation (Stanovich & West, 1989). Orthographic processing accounted for an additional 4.1% variance in irregular word spelling, 5% in morphophone spelling (e.g., doubling the *r* in *conferring* but not in *conference*), and 6.9% in spelling of exception words, after statistically controlling for phonological processing. ART predicted 9.5% of variance in word recognition after phonological processing and orthographic processing were entered into the equation. After partialling out phonological processing, ART was a unique predictor of both orthographic processing tasks (4.2% for orthographic choice and 5.4% for homophone choice task). For spelling, ART predicted 4.1% of variance only for exception words, after statistically controlling for phonological processing.

In summary, this was the first study to develop such print exposure checklists. Stanovich and West (1989) posited that such checklists are linked to reading and reading-related skills, unlike traditional methods of questionnaires that are susceptible to contamination by providing socially desirable answers. ART was a sensitive measure of print exposure and its relationship to phonological processing, orthographic processing, and spelling and word recognition. However, the researchers did not statistically control for differences in reading comprehension scores, which could have inflated the contribution of ART to word recognition. Chateau and Jared (2000) addressed this limitation by replicating and extending the work of Stanovich and West.

Chateau and Jared (2000) recruited 64 undergraduate students and divided them into high and low print exposure as measured by ART. The results confirmed Stanovich and West's results by demonstrating that groups of high and low print exposure were significantly different on phonological and orthographic tasks despite the fact that participants were matched on standardized reading comprehension scores. This method ensured that differences observed were, in fact, due to print exposure and not general verbal ability. The high print exposure group was faster than the low print exposure group in pseudo-word naming (phonological task/decoding), suggesting that individuals who read more have better spelling-to-sound representations. Similarly, the high print exposure group was significantly faster and more accurate on the homophone choice task (orthographic task) than was the low print exposure group. Results confirmed findings obtained by Stanovich and West on the effects of print exposure by using more stringent measures of orthographic and phonological processing and by statistically controlling for verbal ability as measured by reading comprehension.

Goldman and Manis (2013) examined the relationships between reading skills, print exposure, and cortical thickness (more neural resources) in 28 adults. The researchers found that print exposure, as measured by ART, TRT, and MRT, consistently correlated with cortical thickness in the reading network of the brain. Print exposure accounted for a unique difference in cortical thickness even after controlling for reading skills. Although the findings are intriguing, their study, like many other studies examining print exposure, was correlational, and it cannot be assumed that reading experience influences cortical thickness. Yet, such findings confirm the notion that Stanovich and colleagues have long had, that avid readers differ from those who seldom read.

Children's Print Exposure and Relationship to Reading-Related Skills

Anderson et al. (1988) selected 155 fifth-grade children and used daily-activity diaries to estimate their out-of-school reading habits. To conceal the purpose of the study, the authors included time estimates of other out-of-school activities such as time spent on doing chores or watching television. They collected these daily diaries for a median duration of 57 days (different schools were recruited for different durations). In addition, the children were administered a reading comprehension test, a vocabulary test, and a reading fluency test. Finally, second grade standardized reading tests were obtained from school records for the participating fifth- grade students. The authors confirmed that out-of-school reading predicted differences in reading proficiency from the second to the fifth grades. They also found that time spent reading explained 14.4% of the variance in the reading comprehension of the fifth- grade students. As suspected, reading experience did explain individual differences in reading and reading-related skills; however, the time to collect the diaries was very time-consuming and required a substantial commitment from participants.

Following the development of ART and TRT with adults using checklists, Allen et al. (1992) attempted to establish construct validity and reliability of print exposure measures (ART and TRT) with children. To establish validity, the authors recruited 63 fifth- grade students. Several print exposure measures were administered, daily-activity diary, activity preference questionnaire, a literacy and media habits questionnaire, and recognition checklists (ART, TRT, and comics recognition tests). Several vocabulary tasks were administered, but due to the overlap, scores were combined to make a composite score. Similarly, measures of general knowledge were combined to make a composite score.

Results showed that the mean reading time of fifth-grade students calculated from their activity diaries was 10.2 minutes. High correlations were found between reading time as measure by diary and TRT and ART, suggesting that both of these methods tap into print exposure. Although the reading time derived from the activity diary could be used to predict variance in vocabulary knowledge, it did not explain any additional variance once TRT and ART were entered into the equation, demonstrating an overlap between print exposure checklists and absolute reading time from the diary. In summary, Allen et al. (1992) were able to establish validity for TRT and ART as quick and easy to administer measures of print exposure. Allen et al. (1992) also used a daily-activity log, like the one used by Anderson et al. (1988), in conjunction with ART and TRT checklists and found that the tests were correlated, suggesting that the use of checklists is a quick and reliable measure of non-school print exposure. Although checklists such as ART and TRT do not provide an estimate of time spent reading, they do provide an effective, reliable, and valid way of measuring print exposure for correlational purposes.

In a longitudinal study, Cunningham and Stanovich (1997) assessed first-grade children in a variety of literacy skills and followed up 10 years later with subsequent tests. Of the original sample of 56 first-grade students, only 27 were available for the follow-up study. The data collected when they were in first grade were two measures of cognitive ability

(Raven’s Colored Matrices and PPVT) as well as reading ability measures that included Wide Range Achievement Test and reading comprehension. Also, there was data available from school records for their third, fifth, and tenth grade MAT scores (reading comprehension). In the eleventh grade, the 27 participants underwent reading comprehension, written vocabulary (participant reads an incomplete sentence and must select the appropriate word from five words), PPVT (a receptive vocabulary measure), Raven’s Matrices (a nonverbal intelligence measure), print exposure measures (ART, MRT, Activity Preference Questionnaire), and general knowledge measures. Print exposure, measured at grade eleven, explained 23% of the variance in comprehension measure in grade eleven.

Cunningham and Stanovich (1997) then examined whether first-grade variables and print exposure could explain variance in grade eleven outcomes. Print exposure did explain 10.2% of the variance in grade-eleven comprehension ability. Print exposure explained 44.8% of the variance in vocabulary after partialling out early comprehension ability. Using data on reading comprehension from the third grade, print exposure (measured in grade eleven) explained 9.7% of the variance in third-grade reading comprehension after controlling for first-grade reading comprehension scores. Similarly, using fifth-grade reading comprehension scores, print exposure explained 17.8% of the variance in comprehension ability from first to fifth grades and between fifth and tenth grades (12.8%).

The next analysis was to determine whether reading fluency in first grade was responsible for reading volume in later years. All reading measures in first grade predicted print exposure in eleventh grade (10%) even when eleventh grade reading comprehension ability was partialed out.

In summary, print exposure measured in eleventh grade was a reliable predictor of reading ability in first, third, and fifth grades, even after controlling for comprehension ability in the eleventh grade. These results are consistent with Anderson et al.'s (1988) results demonstrating that print exposure can explain unique variance in comprehension ability and vocabulary growth, even when the measures of print exposure used were vastly different.

In a longitudinal study, Cain and Oakhill (2011) examined the effects of reading comprehension and reading experience (print exposure) on vocabulary development. They recruited a sample of 102 students in year three (ages 7 to 8) and assessed them again in years six (n=83), nine (n=52) and eleven (n=40). In years three and six, the participants were assessed in reading ability, vocabulary knowledge, cognitive ability, and reading habits. In years nine and eleven, participants were assessed in reading ability, vocabulary knowledge, and print exposure. Receptive vocabulary measured in year three and year six was strongly correlated with reading ability. The researchers also identified poor and good comprehenders at age eight. The poor comprehenders did not differ from the good comprehenders in reading measures or comprehension ability but did significantly differ in vocabulary growth and exposure to print. Print exposure, as measured by ART and reading comprehension, consistently predicted vocabulary growth in years six, nine, and eleven, for the sample as whole (Cain & Oakhill, 2011).

Cunningham and Stanovich (1990) examined whether orthographic processing could account for a unique contribution to word recognition after statistically controlling for other literacy related variables. Fifty-one third-grade and 47 fourth-grade children were assessed in single word reading (Word Identification of Woodcock-Johnson Reading Mastery Tests), Raven's Matrices (nonverbal intelligence), phonological processing measures, orthographic processing measures, paired associate memory task, and print exposure (TRT). Orthographic processing accounted for 10.2% of the variance in word identification even when age, nonverbal intelligence, phonological processing, paired associate learning, and print exposure were entered into the equation. TRT explained 9.5% of the variance in orthographic processing after statistically controlling for age and phonological processing. Finally, TRT accounted for 6.2% of the variance in word recognition.

In summary, orthographic processing predicted unique variance in word recognition after statistically controlling for all other variables. Print exposure predicted orthographic processing abilities after controlling for all other variables. Results demonstrated that environmental factors, such as print exposure, do predict unique variance in orthographic processing, and ultimately, word recognition.

Similarly, Cunningham et al. (2002) examined the role of orthographic processing and print exposure in word recognition. The authors found that differences in orthographic processing skills were not entirely dependent on phonological processing skills. When they entered the print exposure measure (TRT), the phonological processing composite score, and the orthographic processing composite score, they found that only orthographic processing contributed a unique variance (11.2%) to word recognition in third-grade students. In addition, TRT made a unique contribution to word recognition (11.7%) even after accounting for variability of phonological and orthographic processing abilities. In fact, print exposure explained differences in orthographic processing skills. Cunningham et al. (2002) assessed children in first, second, and third grades, but the TRT was only administered in the third grade. They argued that print exposure measure is a retrospective one, meaning that it is not a measure of print exposure at the time of testing but is rather an accumulative assessment of print exposure during earlier years.

McBride-Chang et al. (1993) investigated the relationships between word recognition, orthographic processing, vocabulary, reading comprehension, and scores in a revised version of the TRT of Cunningham and Stanovich (1990), to examine the effects of print exposure on word reading and reading comprehension in a sample of fifth- to ninth-grade reading-disabled and typically developing students. Orthographic processing measures were correlated to TRT scores only for the typically developing students.

McBride-Chang et al. (1993) then examined whether TRT could predict variance in word recognition before and after phonological and orthographic abilities were entered into the equation. For the reading-disabled (RD) group, TRT did not account for additional variance in word recognition once orthographic processing was entered into the equation, suggesting an overlap between orthographic processing and TRT. When TRT was entered after age and phonological processing, it accounted for significant variance in word recognition. TRT also accounted for variance in orthographic processing for the RD group. For the nondisabled readers, TRT contributed unique variance in word recognition after phonological processing and orthographic processing were statistically controlled. These findings suggest that print exposure could explain unique variance in word reading abilities after controlling for phonological and orthographic skills for non-disabled readers. However, when orthographic skills were entered into the regression equation, the TRT could not predict additional variance, suggesting that TRT overlaps with orthographic skills. Furthermore, TRT also explained unique variance in reading comprehension abilities.

Unlike the Cunningham and Stanovich (1990) findings, TRT did not predict orthographic processing for the non-disabled readers. This may be due to the age differences of the two samples or to ceiling effects in McBride-Chang et al.'s (1993) sample on the orthographic measures. TRT accounted for unique variance in reading comprehension for the disabled readers but not for the nondisabled readers. In the McBride-Chang et al. (1993) final analysis, they examined whether TRT could predict vocabulary knowledge for the non-disabled readers. Once word recognition was entered into the equation, TRT did not predict vocabulary knowledge. These results are inconsistent with those obtained by Cain and Oakhill (2011) despite the similar age of both samples.

Cipielewski and Stanovich (1992) examined the contribution of print exposure to reading and reading comprehension growth in a sample of third- and fifth-grade students using a longitudinal design. Third-grade standardized scores in reading comprehension were available for the participants and were used in the analyses. In fifth grade, the school also administered assessments of reading comprehension, reading rate (fluency), and phonological decoding. In addition, the researchers administered print exposure measures (TRT and ART). When scores on reading comprehension in third grade were entered into the equation followed by TRT and ART scores, print exposure explained unique variance in reading comprehension scores in fifth-grade students (11% for TRT and 8% for ART). Both TRT and ART accounted for a unique variance in reading rate at the fifth-grade level even when reading comprehension at grade three was partialed out. Finally, even when phonological decoding was partialed out, TRT only was a unique predictor of reading rate. In summary, print exposure was a reliable measure in predicting individual differences in reading comprehension growth from third to fifth grades.

In a two-year longitudinal study, Echols et al. (1996) followed fourth-, fifth- and sixth-grade students to determine the effects of print exposure on verbal cognitive skills using the ART and TRT checklists. This longitudinal design allowed the researchers to examine whether differences in print exposure measured at Time 1 could explain variance at Time 2 and Time 3 of the testing. Assessments were administered during two years and included receptive vocabulary, general knowledge, spelling, reading comprehension, and print exposure (TRT, ART). They found that print exposure measured at Time 1 explained significant and unique variance (7.8%) in vocabulary growth at Time 3 even when it was entered last in the regression equation. When print exposure measured at Time 2 was entered into the equation, the variance increased to (10.3%) in vocabulary growth at Time 3. Finally, Echols et al. found that print exposure measured at Time 3 explained unique variance in vocabulary measured at Time 3 (10.6%). Thus, the TRT (not ART) consistently and reliably predicted variance in receptive vocabulary over the span of two years. TRT was a unique predictor (explaining 7.6% of variance) of reading comprehension scores. All of the above correlations were statistically significant but only for TRT.

Similar patterns of results were obtained for TRT and general knowledge growth and spelling. Thus, print exposure, as measured by TRT, explained unique and statistically significant variance in general knowledge and spelling abilities over the two years of the study.

By using a stringent and longitudinal design, Echols et al. (1996) were able to demonstrate that print exposure, as measured by TRT, can uniquely predict individual differences in vocabulary and comprehension abilities. ART was not as sensitive to individual differences, because lower scores were obtained on this task as opposed to TRT scores. These results are consistent with the findings of Cipielewski and Stanovich (1992) that TRT is a better predictor of children’s reading and reading-related abilities. It seems that TRT is more suitable for use with children, as they have not yet developed a preference for authors (Echols et al., 1996).

Twenty-six first-grade students were recruited by Cunningham and Stanovich (1993) to investigate the effects of print exposure on phonological and orthographic abilities. The measures included word reading, spelling, phonological processing, orthographic tasks, and print exposure (TRT). There were several significant correlations between TRT and word reading (.64), spelling (.69), letter-string (orthographic task) (.39), and experimental spelling (orthographic task) (.50). None of the phonological processing tasks correlated with TRT. The two orthographic measures were correlated with word reading even when phonological processing was statistically controlled. This implies that orthographic processing skills are separable from phonological processing.

These results led Cunningham and Stanovich to investigate variance in orthographic processing skills. TRT accounted for unique variance in the experimental spelling task (21.2%) after all phonological measures were partialed out. Similarly, scores in TRT were a significant predictor of individual differences in spelling (43.2%) and word reading (34.9%) after statistically controlling for phonological measures. In summary, print exposure, as measured by TRT was highly to modestly correlated with differences in orthographic processing (letter string choice .39 and experimental spelling .50), reading (.64) and spelling (.69) but not with phonological processing.

Since Stanovich and West's (1989) development of ART, and Cunningham and Stanovich's (1990) development of the TRT, there have been numerous researchers who developed their own versions of the ART and other similar checklists (e.g., Stainthorp, 1997; Masterson and Hayes, 2007). These checklists were subsequently found to be valid and reliable in assessing individual differences in print exposure. Such checklists demonstrate an ability to explain variance in word recognition, spelling, orthographic processing, reading fluency, vocabulary knowledge, and reading comprehension. They also have the advantage of being quick to administer and require little commitment on the part of the participant to complete, unlike activity diaries. Finally, these recognition checklists are less susceptible to socially desirable answers than are reading questionnaires that tend to inflate the time spent reading. There are, however, two drawbacks to these checklists: first, they need to be updated to reflect changes in reading trends, and second, it is necessary to take into consideration the context in which they will be used (they need to be culturally sensitive).

Cultural Sensitivity of the Checklists

Several print exposure checklists have been used for different populations and in different languages. For example, Ecalle et al. (2009) developed several checklists for use with French native-speaking first- to fifth-grade students to examine the effects of print exposure on literacy development. Similarly, Masterson and Hayes (2007) developed UK versions of TRT and ART for native English-speaking adults. Stainthorp (1997) developed a UK version of ART for children and found that the checklist was more reliable than was the list developed by Cipielewski and Stanovich (1992) for North American children. Chen and Fang (2013) developed a Chinese version of ART and found it to have the strongest predictive power of reading achievement among college students (26 were freshmen, 86 sophomores, 79 juniors, and 57 senior students). Stainthorp concluded, “it would therefore seem more valid to use a test which is culture-specific, but designed as an equivalent form” (p. 154). In addition, Stainthorp (1997) found that such checklists need to be updated because they tend to be ephemeral, a point alluded to by Cipielewski and Stanovich (1992).

The Present Study

It has long been recognized that exposure to print can enhance literacy-related skills. A search of the literature confirmed that observation. Yet, controlling for print exposure is a time-consuming task. The development of print-exposure checklists has alleviated the problems commonly associated with traditional methods of measuring print exposure. The literature reviewed has demonstrated that such checklists provide a valid and reliable measure that tap into print exposure. Furthermore, print-exposure checklists can explain unique individual differences in reading even when phonological and orthographic abilities have been statistically controlled. As a result, the present study aimed to develop a TRT in both Arabic and English for Arabic native speakers (third, fourth, and fifth graders) learning ESL. From the literature reviewed (e.g., Cipielewski & Stanovich, 1992; Echols et al., 1996), it is apparent that TRT is a more suitable measure of print exposure for use with children, maybe because children have not yet developed a preference for particular authors (Masterson & Hayes, 2007). To establish construct validity for the checklists, a reading questionnaire was adapted from Cunningham and Stanovich (1997). The present study implemented the same (where possible) protocols as did Cunningham and Stanovich (1997) for the development of a valid and reliable TRT in both languages. Two studies (1a and 1b) were carried out in the course of developing the Arabic and English versions of the TRT and were used as measures of print exposure in Study 2.

Study 1a Method

Study 1a was conducted to generate potential titles for Arabic TRT and English TRT.

Participants

A total of 86 third- to fifth-grade students participated in Study 1a. Since measures of print exposure as indexed by a title recognition test are a measure of children's reading for leisure, it was reasoned that children would need to be older (third grade and above) in order to have developed reading preferences. Studies assessing print exposure in children using the title recognition test have used samples of children in third, fourth, and fifth grades (Cipielewski & Stanovich, 1992). Thus, the children taking part in Study 1a were third, fourth, and fifth graders. There were 23 third graders (mean age 8.6, SD=0.5), of whom 11 were female; 31 fourth graders (mean age 10.3, SD=0.4), of whom 15 were female; and 32 fifth graders (mean age 11.1, SD=0.5), of whom 17 were female.

All participants were randomly selected from the same international[1] school situated in an upper-middle-class neighborhood in Cairo, Egypt. The school offers a British curriculum delivered in English. International schools charge high tuition fees but have the advantage of providing another country's curriculum (most commonly British and American). Only Arabic, religion and social studies are taught in Arabic following the national Egyptian curriculum. Arabic is taught daily for 40 minutes, and religion and social studies are taught once a week for 40 minutes each. All study participants were native Arabic speakers. Participants with sensory-motor and/or cognitive impairments were excluded from the study.

Ethical committee approval was granted by the UCL Institute of Education to conduct this research. Permission to conduct the study at the school was obtained in response to a request sent to the school principal (Appendix A). Upon receipt of approval from both the university and the school principal, letters of parental and student consent were sent to potential participants and their parents (Appendix B). All students were reminded of their right to withdraw from participation without consequence.

Materials

Assembling Arabic and English Book Titles for Test Construction

In generating potential titles of Arabic and English children's books, a number of sources were used. Egyptian publishing houses, bookstores, public libraries, schoolteachers, and school librarians were contacted to provide a list of popular book titles for children ages eight to twelve. This was done to generate a larger than needed list of book titles to weed out the items with low reliability. Everyone contacted was asked to exclude book titles that were part of the school curriculum (in the case of teachers and school librarians) and to further exclude books that were fairy tales or those that had been made into movies or television shows. This was done to ensure the validity of the construct.

Ten Egyptian publishing houses were contacted by email to provide book titles of their bestselling children's publications. With the exception of one, none replied to the researcher's request. The publishing house that did respond indicated that their most popular publications are those promoted on their website. Those listed on their website and the websites of other publishing houses were used in the initial book-title-gathering process.

Similarly, four major bookstore chains were contacted via email. These bookstore chains were located across major cities in Egypt and all sell both English and Arabic books. Unlike the request made to the publishing houses, bookstores were asked to provide a list of their bestselling children's books for both Arabic and English books. Two bookstores replied to the researcher's request and provided lists of their most frequently sold children's books. One of the bookstores provided two lists of 20 titles each: one for Arabic books, and one for English book titles.

The other bookstore provided a list of 43 bestselling Arabic children's books. The remaining bookstores did not respond to the researcher's request.

The Integrated Care Society, which is a non-governmental organization (NGO), was consulted pertinent to establishing lists. The NGO runs 12 public libraries across Egypt and is a chapter of the Anna Lindh Foundation that is funded by the Swedish International Development Cooperation Agency. One of the aims of the Anna Lindh Foundation is to promote children's literature and reading in the Arab region. The Foundation has 7,323 member children in Egypt, 4 to 15 years old. Their database identified 833 member children, 8 to 12 years old. Two lists were generated, one for most frequently signed out children's Arabic books, and one for English book titles most frequently borrowed by children 8 to 12 years old. The database generated 37 English book titles and 85 Arabic book titles.

Arabic Teachers

Arabic teachers were consulted to provide a list of book titles they believed were popular with their students but were not part of the Arabic curriculum or classroom reading activities. The Arabic teachers were told to exclude book titles that had been adapted into television shows/movies and vice versa. They were also advised not to include fairy tales, as many have been made into movies and/or are frequently narrated by parents. Finally, it was also decided to exclude biblical book titles because many of the participating children are exposed to these stories in mosques and churches, religion class, and frequently narrated by parents.

Classroom Teachers and School Librarians

Classroom teachers and school librarians were asked to provide book titles popular with children. They were explicitly instructed to exclude book titles used for classroom reading activities. It was ensured that the titles selected were not part of the curriculum since the construct is a measure of out-of-school reading. The same reasoning was used by Cunningham and Stanovich (1990) to ensure their construct’s validity. If titles chosen were part of curriculum or classroom activities, it would weaken the test’s ability to predict individual differences in print exposure (reading experience for leisure).

Development of the TRT Measures

The purpose of Study 1a was to eliminate items that were either recognized by the majority of participants or targets that were seldom selected by the majority of participants. Second, it was important to match both Arabic and English lists in number of items and difficulty. The titles provided from the various sources were compared and analyzed. It was reasoned that analyzing the book titles from all the various sources would reveal the most frequently read books in Arabic and English. It was assumed that using this method of generating book titles would yield the most-read books by children of the indicated ages; however, the various sources provided very different lists with little or no overlap, whether in English or Arabic. This proved to be a challenge.

Because the book titles varied tremendously depending on the source, it was necessary to pilot them to eliminate those that had very low reliability (titles that were unfamiliar to most children). Lists were constructed from the various sources listed above and resulted in 40 English titles and 60 Arabic titles. To make the lists manageable for children during piloting, the English titles were randomly divided into List A and List B, resulting in 20 items per list. Similarly, the Arabic titles were randomly divided into three lists, A, B, and C, 20 items on each list. Ten fictitious English book titles and 10 fictitious Arabic book titles were created and checked on internet search engines to ensure that they were not real titles of books, movies, etc. The foils were embedded randomly but were in the same position for the English List A and List B. Similarly, the Arabic foils were first inserted randomly into List A and then subsequently inserted into List B and List C in the same positions.

Procedure

Students were seen in groups of four at a time in a quiet room on school premises. The participants were instructed as follows:

You will see two lists of book titles, one in English and one in Arabic. Some of the titles are the names of real books, and some are not. Please read the names and put a check mark next to the names of those that you know are books. Do not guess, but only check mark or tick those that you know are actual books. Remember, some of the titles are not those of real books, so guessing can easily be detected (Cipielewski & Stanovich, 1992, p. 80).

If a child did not understand the instructions, the researcher gave an example for clarification. The lists were then distributed and children were told that the task had no time limit for completion.

Study 1a Results

Scores for correct selection of targets and selection of foils can be found in Appendix C by percentages for the Arabic and English TRT lists.

A summary of the TRT corrected scores (targets minus foils) for each grade level (third to fifth) is given in Table 2 for the three Arabic TRT lists and the two English TRT lists.

Table 2

Arabic and English TRT Corrected Mean Scores and Standard Deviations

(SD; max score = 20)

		List A		List B		List C		Overall Mean[a]	
Grade	*N*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*	*M*	*(SD)*
Arabic									
Third	23	1.00	(1.41)	.57	(2.07)	2.50	(2.56)	1.39	(2.15)
Fourth	31	.67	(1.87)	2.18	(4.47)	2.00	(3.19)	1.68	(3.38)
Fifth	32	.75	(1.42)	1.20	(1.69)	1.60	(1.58)	1.16	(1.55)
All	86	.79	(1.52)	1.43	(3.12)	2.00	(2.49)	1.41	(2.48)
English									
Third	23	1.64	(3.30)	1.89	(1.45)	--	--	1.74	(2.68)
Fourth	31	1.31	(1.18)	3.00	(1.82)	--	--	2.29	(1.77)
Fifth	32	3.00	(1.63)	2.56	(1.03)	--	--	2.78	(1.36)
All	86	2.05	(2.05)	2.60	(1.51)	--	--	2.33	(1.96)

Note: Mean[a] = These scores were used in the 2 x 3 mixed-ANOVA.

To determine whether the overall mean-corrected TRT scores (the average of the corrected Lists A and B scores for the Arabic titles; the average of the corrected Lists A and B scores for the English titles) differed across language and grade and whether there might be an interaction between language and grade, a 2 x 3 mixed-ANOVA was conducted. The within-subjects variable was language (Arabic vs. English), and the between-subjects variable was grade (third, fourth, vs. fifth).

The findings revealed that, firstly, overall mean corrected TRT scores did not differ across grade levels, $F(2,83) = .55$, $p = .581$, partial $\eta^2 = .013$. Secondly, overall mean corrected TRT scores differed significantly across language, $F(1,83) = 6.55$, $p = .012$, partial $\eta^2 = .073$. Students had higher overall mean-corrected English TRT scores ($M = 2.33$, $SD = 2.48$) than they had overall mean-corrected Arabic scores ($M = 1.41$, $SD = 1.96$). Lastly, there was no significant interaction between grade level and language, $F(2,83) = 1.38$, $p = .257$, partial $\eta^2 = .032$.

Study 1a Discussion

The purpose of Study 1a was to generate potential titles for Arabic TRT and English TRT. Because overall mean-corrected English scores were significantly higher than were overall mean-corrected Arabic scores, the frequencies of correct targets and foils for the Arabic and English lists (reported in Appendix C) were inspected to see whether certain titles should be removed. The most frequently selected target on the Arabic lists had a hit rate of 51.7% (صلاح الدين الأيوبي بطل حطين). Three titles were removed from the English lists because they had very high hit rates (over 79%) (e.g., *Matilda*). Some foils in both Arabic and English lists were removed because they had a high false positive rate. One of the Arabic foils (قصص من زمن

العصر الحجري) was falsely selected by 32.6% of the participants, whereas one of the English foils (*The Dark Horse*) was selected by 31.4% of the participants. Since it was necessary to develop parallel measures of print exposure in both languages, the frequencies of targets in the Arabic lists and in the English lists were compared to ensure the items on both tests were equivalent in difficulty. For example, all the English target titles that had a hit rate of 55% and higher were deleted because the most frequently selected target on the Arabic lists had a hit rate of approximately 52%. This procedure was used to match item by item (both targets and foils) on both the English and Arabic lists. By using this method, the lists were amended, and produced the lists used in Study 1b. The lists are reported in Appendix C.

Study 1b Method

Study 1b was conducted to establish reliability and validity of the final versions of Arabic TRT and English TRT. The revised lists comprising an Arabic list of 20 targets and 10 foils and an English list of 20 targets and 10 foils were administered to a separate group of children in Study 1b, with the aim of establishing construct validity and reliability of the items. Targets were randomised and foils were distributed randomly in both lists.

Participants

A second sample of 76 third- to fifth-grade students participated in Study 1b. Like the sample in Study 1a, all participants were randomly selected from the same international[1] school that Study 1a participants were selected from.

The children taking part in Study 1b were 76 third, fourth, and fifth graders. There were 27 third graders (mean age 8.5, SD=0.3), of whom 12 were female; 27 fourth graders (mean age 9.7, SD=0.4), of whom 12 were female; and 22 fifth graders (mean age 10.8, SD=0.5), of whom 15 were female.

Procedure

The same administration and procedures were used as in Study 1a. In addition, the participating children were asked to complete a reading habits questionnaire adapted from Cunningham and Stanovich (1997). The version used is given in Appendix D. The questions were presented in a forced choice format and instructions were taken from Cunningham and Stanovich (1997) as follows:

Below you will be given a choice of two activities. Please put a check mark next to the one that you prefer. Please mark only one. That is, even if you like both activities, please mark only the one you like better. Similarly, even if you dislike both activities, mark the one that you would prefer to do. For each item, please mark only one choice Cunningham and Stanovich, 1997, p. 938).

This questionnaire served to establish convergent validity for TRT; in other words, children were scored from 0 (they never chose reading over another activity) to 6 (they always preferred reading over other activities). The reading habits questionnaire was administered in English since all the participants were fluent English speakers. It took approximately five minutes to complete. Participants were seen in groups of four and completed the reading habits questionnaire, followed by the TRT lists. The order of the TRT list presentation was rotated across students (Arabic list followed by the English list and vice versa).

Study 1b Results

The scores (in percentages) for correct selection of targets and selection of Arabic and English foils are shown in Appendix C. The corrected TRT scores (targets minus foils) for each grade level are presented in Table 3. The corrected mean scores (targets selected minus foils selected) for Arabic and English TRT as well as the descriptive statistics for the Reading Habits questionnaire are shown in Table 3.

Table 3

Arabic and English TRT Corrected Mean Scores (max score=20) *and Cronbach's Alpha and Reading Habits Scores* (max score=6)

		Arabic TRT[a]			English TRT[a]			Reading Habits	
Grade	*N*	*M*	*(SD)*	α	*M*	*(SD)*	α	*M*	*(SD)*
Third	27	4.85	-3.26	0.82	6.33	-3.77	0.83	2.33	-1.62
Fourth	27	3.59	-2.87	0.79	4.22	-3.14	0.81	2.33	-1.57
Fifth	22	3.95	-2.77	0.73	4.59	-2.26	0.64	2.68	-1.62
All	76	4.15	-3	0.79	5.08	-3.27	0.81	2.43	-1.59

Note. TRT[a] = These scores were used in the 2 x 3 mixed-ANOVA.

Reliability was assessed by means of Cronbach's alpha. According to Hair et al. (2010), a measure is reliable if its alpha is above .60. As shown in Table 2, the Arabic and English TRT measures can be considered to be reliable since α in all cases was above .60.

Construct validity was evaluated using the scores from the Reading Habits Questionnaire adapted from Cunningham and Stanovich (1997). There were no significant correlations between the Reading Habits Questionnaire scores and corrected mean scores of Arabic TRT and English TRT.

To determine whether the overall mean corrected TRT scores (the average of the corrected scores for the Arabic titles and the average of the corrected scores for the English titles) differed across language and grade and whether there might be an interaction between language and grade, a 2 x 3 mixed-ANOVA was conducted. The within-subjects variable was language (Arabic vs. English) and the between-subjects variable was grade (third, fourth, vs. fifth).

The findings revealed that, firstly, overall mean corrected TRT scores differed across grade levels, $F(2,72) = 5.28$, $p = .007$, partial $\eta^2 = .089$. Post-hoc Tukey comparisons indicated that third-grade students had significantly higher scores ($M = 11.19$, $SD = 5.57$) than did fourth-grade students ($M = 7.81$, $SD = 4.84$; $p = .038$). Secondly, overall mean corrected TRT scores differed significantly across language, $F(1,72) = 4.10$, $p = .047$, partial $\eta^2 = .054$. Students had higher English corrected scores ($M = 5.04$, $SD = 3.27$) than Arabic corrected scores ($M = 4.15$, $SD = 3.00$). Lastly, there was no significant interaction between grade level and language, $F(2,72) = .55$, $p = .577$, partial $\eta^2 = .015$.

Study 1b Discussion

The purpose of Study 1b was to establish the reliability and validity of the final versions of Arabic TRT and English TRT derived from Study 1a. Reliability for both English and Arabic lists ranged from .64 to .83. Construct validity was evaluated using the Reading Habits Questionnaire. Contrary to Cunningham and Stanovich's (1997) results, there were no significant correlations between the Reading Habits Questionnaire scores and scores of Arabic TRT and English TRT.

Although every measure was taken to ensure that the English and Arabic items were similar in difficulty, students had higher scores on the English TRT. The Arabic TRT proved more difficult.

Discussion

The objective of Study 1 was to develop a print exposure measure (TRT) in both Arabic and English. While there have been many such tests developed in English and other languages (e.g., Cunningham & Stanovich, 1997; Echols et al., 1996; Ecalle et al. 2009; Chen & Fang, 2016), no such test has been developed in Arabic. This section contains a summary of the development of the TRT and results, followed by a discussion of the challenges and limitations of developing such measures in an Egyptian context.

One aim of Study 1b was to establish reliability of the Arabic and English TRT measures. The reliability of Arabic TRT items ranged from .79 to .82, whereas the reliability of English TRT items ranged from .64 to .83. Thus, both lists can be considered reliable since α was above .60. Pearson correlations were also conducted to establish validity. The scores for the Reading Habits Questionnaire did not correlate significantly with corrected mean scores of Arabic TRT and English TRT. This lack of correlation is perhaps due to the nature of the questionnaire, which measures "attitudes about reading rather than indicators of actual reading behaviors"

(Cipielewski & Stanovich, 1992, p. 497). In other words, the TRT seems to reflect an index of actual reading, whereas the Reading Habits Questionnaire is a subjective measure of attitudes about reading. Finally, the mixed-ANOVA results revealed that corrected mean scores differed across language, students having higher English corrected scores than Arabic corrected scores. The corrected mean scores also differed across grades, third-grade students having significantly higher scores than fourth-grade students.

The corrected mean scores from Study 1b for both Arabic and English lists were very similar to those found in other studies that have developed such checklists for children. For example, in the study of Cunningham and Stanovich (1990), the TRT corrected mean score for a sample of third- and fourth-grade students was 5.8. Similarly, in the study of Ecalle et al. (2009), the corrected mean score for the TRT was 5.23 for the third-grade students, 6.2 for the fourth-grade students, and 4.87 for the fifth-grade students.

Finally, the inter-item reliabilities of Arabic TRT and English TRT are nearly identical to those obtained by other researchers when developing TRT. For example, the TRT developed by Cunningham and Stanovich (1990) had a reported reliability of .81. Allan et al.'s (1992) version of the TRT had a reported reliability of .80.

Limitations

Despite the effort deployed in developing the checklists, the Arabic titles provided by bookstores, librarians, and publishers varied considerably. Through informal interviews with bookshop owners, librarians, and Arabic teachers, it was evident that reading in Arabic was not promoted or encouraged by parents and educators. Publishers seldom publish children's books in Arabic because they do not generate revenue. Reading Arabic literature only develops in adulthood if it develops at all. This is evident from the Arabic book titles that were generated from the above-mentioned sources; most of the books bought in Arabic are biographies of famous characters or historical accounts rather than fiction. These books are often purchased for completion of school projects. For example, the most commonly identified target on the Arabic TRT in both the first and second pilot studies was Salah El Din (identified 52%). This historic figure is taught in Arabic, social studies, and religion classes.

English book titles that proved most reliable were those generated by school librarians based on the most frequently borrowed books. The English titles that were provided by bookstores were often activity books (e.g., colouring books) with minimal text. Such titles mostly consisted of television characters (e.g. Dora, SpongeBob Square Pants). Book titles generated from school libraries included many titles of Roald Dahl; however, some were removed for having been made into movies (e.g., *Charlie and the Chocolate Factory*) and had a hit rate of 79% in Study 1a.

The difficulty in generating Arabic and English titles, coupled with comments from educators, librarians, and bookshop mangers/owners seemed to be consistent with the notion that Arabic-speaking children seldom read for leisure. Indeed, studies that have investigated reading habits in Arabic-speaking children have encountered similar challenges. For example, Feitelson et al. (1993) designed an intervention study to examine the role of MSA story reading for Arabic-speaking five- and six-year-old children in 12 kindergartens across northern Israel. Ironically, the researchers could not find age-appropriate storybooks in MSA at the time of the study and had to translate popular Hebrew storybooks for their intervention, which illustrated the lack of exposure to MSA prior to reading acquisition.

Iraqi (1990, cited in Feitelson et al., 1993) found that only 1.8% (five out of 290 families) of Arabic-speaking parents read books to their preschool children. These findings are consistent with the observations in the present study when generating Arabic book titles and confirm that reading habits do not seem to be not fostered culturally as a whole. However, despite the challenges, reliable checklists were developed for use in the present context.

In summary, it is important to investigate how reading development occurs with limited exposure to print, particularly in a diglossic situation where oral language does not aid reading acquisition.

Chapter 3

Study 2: Predictors of Single Word Reading and Spelling

Reading, although seemingly simple, requires intricate and complex cognitive abilities that work together in a highly coordinated manner (Kirby et al., 2010). As such, reading is the most researched cognitive ability (Adelman, 2012a). Literature on reading has been dominated by research in English, an outlier writing system (Share, 2008). By contrast, research in Arabic, although growing, remains scant in comparison to research in English and other European languages. As more research emerges in Arabic and other languages, our understanding of universal predictors and those that are specific to the characteristics of a given language is enriched. Thus, this study examined the linguistic skills and cognitive abilities that underpin single word reading and spelling in several ways.

The literature review will begin with an overview of PGST as a backdrop (Ziegler & Goswami, 2005). First, a review of the literature on predictors of reading and spelling based on findings from cross-linguistic studies is presented, followed by a review of the literature on predictors of reading and spelling in Arabic and similar languages. The literature review will conclude by summarizing findings from the literature and presenting the aims of the study. The following section presents the design, measures, and methodology employed. The Results section of this chapter presents predictors of reading and spelling in Arabic L1 and English L2 in a sample of third-, fourth- and fifth-grade bilinguals, in order to investigate the universality of predictors using a within-participant design. A subsample of the bilinguals was compared to a sample of third-grade monolinguals on reading and spelling measures and examined whether their reading and spelling are predicted by the same or different abilities. This chapter concludes with a discussion of the findings and the limitations inherent in this type of research.

Literature Review

There is a consensus among reading researchers that the development of the relationship between letters and sounds is a compulsory prerequisite when learning to read in an alphabetic writing system. However, cracking the alphabetic code is contingent on the characteristics of the writing system. There is considerable variability in the depth with which phonemes map onto graphemes in alphabetic writing systems, and this orthographic depth interacts with reading development (Katz & Frost, 1992).

Reading and spelling development are one of the most important components of learning in any language (Pollor et al., 2008). Reading, however, has received far more attention than has spelling, and oftentimes spelling development is explained using theories about reading development (Treiman & Cassar, 1997). Much like studies in reading development, studies that examine spelling development from a cross-linguistic perspective offer a great deal of insight into the universal underpinnings of spelling acquisition and development (Caravolas, 2004)

Predictors of Reading and Spelling

As mentioned in Chapter 1, Ziegler and Goswami (2005) proposed PGST to account for the influence of orthography on reading development. Despite the influence of the writing system on reading acquisition, it appears that a set of processes underpinning reading development is similar across languages of varying orthographic depth. Thus, the next section will begin with an overview of the predictors most often implicated as crucial to reading development, followed by a review of the literature on reading in various languages, from a cross-linguistic perspective. The review of the literature concludes with studies examining Arabic and Hebrew (both Semitic languages) to determine the cognitive and linguistic predictors that are universal versus those that are determined and influenced by the linguistic and orthographic features of the language, using the PGST as a backdrop.

Various reading models have focused more or less on a finite set of processes, some of which are distal (e.g., general ability, auditory and visual memory), and some proximal that are directly involved in reading (e.g., PA, orthographic processing, vocabulary) (Castles et al., 2018; Kirby et al., 2008). Other factors that are often implicit in these models involve environmental factors, such as exposure to print and home literacy (Castles & Nation, 2010). Studies examining two or more languages varying in orthographic depth have generally examined PA, verbal and working memory, naming speed, orthographic processing, morphological awareness, and vocabulary as predictors of reading accuracy and fluency, and in some cases, spelling (e.g., Ziegler et al., 2010; Caravolas et al., 2012). Such studies offer evidence of the universality of such predictors and their developmental contributions as mediated by orthographic depth. Despite their invaluable contribution to our understanding of reading development in different languages, cross-linguistic studies have inherent problems, such as the issue of equating measures across the different languages, controlling for instructional differences, and controlling for between-participant differences. These methodological issues are of great importance when interpreting the findings from cross-linguistic research, discussed later in this chapter.

Phonological Awareness

It has been well established that the ability to identify and manipulate sounds in spoken words is essential for reading (Wagner et al.,1997; Manis et al., 2000; Mutter et al., 2004; Ehri, 2005a, 2005b). Although this ability appears to be important to all orthographic systems (Ziegler et al., 2010; Landerl et al., 2019), some researchers have suggested that the influence of PA may have been inflated, since most of these findings are based on English, which has been described as an outlier orthography (Share, 2008). It has been suggested that PA may be more important in opaque languages, since grapheme-phoneme conversion in transparent orthographies is acquired early in reading acquisition (Seymour et al., 2003). Cross-linguistic research has demonstrated that the contribution of PA to reading diminishes in the early years of reading acquisition in transparent orthographies (e.g., Landrl & Wimmer, 2000; Papadopoulos et al., 2009; Furnes & Samuelsson, 2011), whereas the contribution of PA to reading persists in opaque orthographies. The contribution of phonological influence on reading is mediated by the depth of the orthography (Wimmer & Goswami, 1994; Melby-Lervag et al., 2012). Nonetheless, there is a consensus that reading acquisition in an alphabetic system is reliant on PA at some point, regardless of orthographic depth.

Verbal Short-term Memory

Another phonological processing ability often implicated in reading and vocabulary acquisition is verbal STM. Perez, Majerus, and Poncelet

(2012) propose that verbal STM capacity probably aids in the decoding process during the early stages of letter-sound conversion and blending, which may foster the creation of new orthographic representations. It would seem that, although verbal STM is essential in the early years, once the alphabetic code is cracked and frequently encountered words have lexical representations, the cognitive demand of maintaining a phonological code in memory while decoding is diminished. As with the role of PA described above, conflicting views regarding the role of verbal STM in reading acquisition and vocabulary knowledge may be influenced by both development and orthographic depth. Regardless of the relationship, verbal STM appears to be a robust predictor of reading acquisition at the beginning stages of reading (Baddeley, 2012, 2015; Wagner & Torgesen, 1987; Torgesen et al., 1997; Melby-Lervag et al., 2012; de Jong & Van der Lei, 1999).

Working Memory

Unlike the passive storage of verbal STM, working memory requires the storage and explicit manipulation of information (Baddeley, 2012). Melby-Lervag et al. (2012) argue that verbal STM is an implicit phonological task that does not require explicit reflection on phonological units. Based on this distinction between verbal STM and working memory, the temporary storage of phonological units until subsequent units are decoded and blended involved in reading is consistent with the definition of working memory rather than verbal STM. Thus, in theory, working memory is considered to play an important role in reading since it is responsible for the integration of phonological and visual information (Peng et al., 2018). Based on this definition, beginning readers may rely heavily on working memory because of the limited lexical representations available to them in long-term memory. Despite the apparent theoretical link between working memory and reading, it is often overlooked in the reading literature, and even when examined, the research has produced inconsistent results (Peng et al., 2018).

Visual Attention

Spatial STM is another component of memory within the visuo-spatial sketchpad. This ability is usually tested using tasks that require sequential visual recall (Baddeley, 2012; 2015). Reading is clearly a visual task that requires attention and the processing of letter strings. The novice reader must attend successively to each letter in the word. As reading develops, larger units become familiar and automatic and are processed at a glance (Ehri, 2005b), thus the need for lesser demands on visual attention. However, Franceschini et al., (2012) found that the role of visual attention increased from first grade to third grade for nonword reading, whereas the contribution of visual attention remained stable across grades for irregular word reading. The authors take this to suggest that visual attention processes are involved in the reading task from the beginning of learning to read to a persistent long-term influence of orthographic learning.

RAN

Although auditory and visual span (verbal STM, working memory, and visual attention) are components related to reading, it is the efficient synchronization of these abilities that is assumed to be required for reading. The speed at which one can retrieve and pair the visual representation with its phonological code is predictive of reading and reading difficulties (Wolf & Bowers, 2000; Manis et al., 2000; Georgiou et al., 2022). The ability to access verbal codes from visual information rapidly (RAN) is a robust predictor of reading across different orthographies and is considered a universal predictor (Nag & Snowling, 2012; Araujo et al., 2015).

While there exists unequivocal evidence of the relationship between RAN and reading (Inoue et al., 2020; Georgiou et al., 2022), there is much debate regarding the underlying cognitive underpinnings of this relationship (de Jong, 2011). Some have argued that RAN is just another construct of phonological processing (e.g., Wagner et al., 1994; Lervåg & Hulme, 2009; Ziegler et al., 2010; Martinez et al., 2021). Others have argued that it taps into orthographic processes (e.g., Bowers & Wolf, 1993; Wolf & Bowers, 1999). It has been proposed that it measures the ability to form visual-verbal connections such as associations between orthographic units and their pronunciations (Kirby et al., 2010). Finally, it has been proposed that RAN taps into general cognitive processes; however, even when studies control for verbal and nonverbal abilities, RAN continues to predict reading (Araujo et al., 2015). These ambiguities concerning the underpinnings of the RAN-reading relationship are further complicated by the course of reading development and orthographic depth.

Orthographic Processing

Another skill that is broadly termed “orthographic processing” has been found to contribute to reading ability. Orthographic processing refers to the ability to learn, store, and retrieve orthographic representations and pattern knowledge (Apel, 2011). As the definition suggests, there are two types of knowledge, one for mental orthographic representations—whole or partial word spellings stored in memory (Apel, 2011)—and one for orthographic pattern, which pertains to knowledge of rules that govern the writing system of a given language. In fact, orthographic processing has been found to predict word recognition above and beyond other variables (Cunningham & Stanovich, 1990). Just as infants are sensitive to the phonological regularities of language, so are children, once introduced to print. Thus, once children are exposed to print, orthographic statistical regularities are learned, and variability in this ability may account for differences in reading ability.

Print Exposure

It has long been recognized that variability in reading ability is not solely dependent on cognitive and linguistic factors but also on environmental ones. Consistently, results have demonstrated that environmental factors, such as print exposure, do predict unique variance in orthographic processing and, ultimately, word recognition (Cunningham & Stanovich, 1990; Cunningham et al., 2002). Exposure to print has also been linked to variance in vocabulary acquisition and growth (Cain & Oakhill, 2011).

Vocabulary

According to SVoR, oral language competencies, such as vocabulary (Kirby et al., 2008), are essential to reading. This is also consistent with the developmental reading models (Ehri, 2005a; 2005b). It is proposed that existing oral vocabulary can be used as an aid for the beginning reader by partially decoding and approximating the proper pronunciation based on existing representations in the oral lexicon. Although the role of oral vocabulary is well established in reading research, the exact mechanism underlying this relationship is not well understood. The role of vocabulary in reading development may even be more relevant in Arabic due to diglossia and thus warrants investigation.

Morphology

Like vocabulary, morphological awareness has been linked to reading outcomes in English (Bowers, Kirby, & Deacon, 2010) and even more so in Semitic languages such as Arabic and Hebrew (Abu-Rabia, 2003; Saiegh-Haddad & Geva, 2008; Ravid & Schiff, 2006; Fumero & Tibi, 2020; Tibi et al., 2020). In fact, Boudelaa (2014) argues that the Arabic lexicon is organized and accessed based on morphemes. The relationship between vocabulary and morphology is intertwined, so growth in one promotes the other reciprocally (McBride-Chang et al., 2005; Kirby et al., 2008).

Summary

It is abundantly clear that the processes involved in reading are intertwined in a complex manner that is further complicated by development and environmental factors. A finite set of both distal (e.g., nonverbal ability, vocabulary, verbal and visual STM, working memory) and proximal predictors (PA, orthographic processing, RAN, morphological awareness) are relatively undisputed. However, there are several issues to consider based on PGST (Ziegler & Goswami, 2005). If reading in a transparent orthography, then PA should only predict reading at the early stages of reading acquisition. Thus, PA should be more important in opaque orthographies. Similarly, if reading in a consistent orthography promotes the use of grapheme-phoneme conversion strategy, then one would expect a greater role of verbal STM and working memory because children need to decode fine-grained units and these units must be available until the individual units are blended. Thus, verbal STM and working memory should be more important in consistent orthographies. Similarly, if RAN is considered to tap into the speed of phonological retrieval from long-term memory, then it should be more important in consistent orthographies, since phonological representation of each grapheme is unambiguous and should be retrieved very quickly. If, however, RAN is considered to tap into orthographic processing, then RAN should be more important in inconsistent orthographies, since larger phonological units have to be sufficiently specified to the corresponding orthographic representations. The cross-linguistic evidence below will shed light on these issues. Thus, the next section will review cross-linguistic literature.

Cross-linguistic Evidence

In one of the first cross-linguistic studies to examine the influence of orthographic depth on reading development, Seymour et al. (2003) measured emerging literacy skills in English-speaking children and children from 12 European countries. This study did not investigate predictors of reading but rather the literacy skills involved in reading acquisition in various orthographies. Letter identification speed was slower in the English-speaking children. While reading accuracy for all European languages ranged between 74% and 98%, English accuracy scores were at 34% and were not included in the analysis. The analysis of reading speed showed a similar pattern. The English-speaking children had a substantially slower speed than did children reading in the other European orthographies, and this was mediated by orthographic depth. Nonword reading revealed the same pattern: the English-speaking children were the slowest and least accurate. However, these results should be interpreted with caution, since the children in English sample were a year younger than were the other children in the study. Nonetheless, this cross-linguistic study did demonstrate that the time required to establish foundation literacy is hampered in languages with inconsistent orthographies, such as French, Portuguese, and Danish. The findings in this study are consistent with PGST, which states that reading development is contingent on the orthographic consistency of the language.

Georgiou et al. (2008) examined predictors of decoding and reading fluency in English and Greek. The authors reasoned that, despite mounting evidence for the importance of phonological processing skills to reading acquisition, these relationships vary as a function of orthographic depth. They longitudinally assessed children in first grade on parallel measures of PA (elision), RAN (color, digit), verbal STM (digit span), orthographic processing (orthographic choice), nonword reading, and reading fluency. The children were assessed only on the reading measures in second grade. PA measured in the first grade predicted reading accuracy and fluency in first and second grades in both English and Greek. However, PA was a stronger predictor in English than in Greek. RAN-digit predicted reading fluency in both English and Greek. Verbal STM predicted first-grade nonword reading only in Greek. These findings are consistent with the PGST, since in Greek finer-grain phonological information must be retained in STM for blending. The same granularity argument could be made for the lack of contribution of STM to nonword decoding in English. The strongest predictors of English nonword reading were PA and orthographic processing, whereas RAN was the strongest predictor of nonword reading in Greek. However, the orthographic task used in this study has been criticized for being a measure of word recognition and not a pure measure of orthographic processing (Castles & Nation, 2010).

Ziegler et al. (2010) examined the universal predictors of early reading in five European orthographies varying in orthographic consistency. They assessed 1,265 Finnish, Hungarian, Dutch, Portuguese, and French second graders on measures of PA, RAN (object), reading, and nonword reading accuracy and fluency. They also assessed nonverbal ability, vocabulary, and verbal STM. The authors were particularly interested in addressing whether PA is important for all languages regardless of orthographic depth and if its importance is modulated by orthographic consistency. Furthermore, they were interested in addressing the claim that RAN is a stronger predictor of reading in transparent orthographies than in opaque ones. PA was the strongest predictor of reading accuracy and speed, and this was modulated by orthographic depth; PA had a stronger influence for inconsistent orthographies. The opposite pattern emerged for vocabulary. Predictive power was stronger for consistent orthographies than for inconsistent ones; however, these effects were no longer significant when Finnish was removed from the analysis. Furthermore, verbal STM predicted reading and decoding accuracy for consistent orthographies. The authors concluded that PA is important to all languages; however, its magnitude is modulated by orthographic depth. RAN was found to weakly predict reading and nonword reading speed, and this effect was not modulated by orthographic consistency. The authors argue that literature demonstrating a weaker impact of PA and a stronger impact of RAN in transparent orthographies is due to ceiling effects on measures of phonological measures, which in turn leaves the “lion’s share” of shared variance to RAN. In other words, RAN's contribution to reading in transparent orthographies is due to insufficient measures of PA. However, just as the authors highlight that the role of PA may have been obscured by the choice of measure, the RAN (object) measure may have obscured the magnitude of the RAN-reading relationship (Kirby et al., 2010). Furthermore, the Dutch, French, and Portuguese samples were more than a year younger than the sample of Finnish and Hungarian children, and thus the results should be interpreted with caution. The results of PA and RAN-reading relationships may be developmental rather than a reflection of orthographic depth.

Caravolas et al. (2012) longitudinally examined the role of predictors of reading and spelling in four alphabetic languages (English, Spanish, Slovak, and Czech). Seven hundred thirty-five kindergarten children were assessed at Time 1 on measures of vocabulary, nonverbal ability, PA (phoneme isolation and blending), RAN (object and color), and verbal STM (repeating a list of monosyllabic words in the same order). They were also assessed on measures of reading and spelling: two reading measures (picture-word matching reading test, and one-minute reading test of high-frequency words), a spelling-to-dictation-test, and letter writing. Ten months later, at Time 2, children were assessed only on reading, spelling, and letter-writing measures. Results revealed that letter-sound knowledge, PA, and RAN measured at Time 1 predicted reading and spelling in all four languages at Time 2. Verbal STM did not predict reading or spelling; however, this task had a low reliability according to the authors. The low reliability of the task is not surprising, since they used real monosyllabic words to assess verbal STM, which may have existed in the children's oral vocabulary. The reason the authors chose to assess verbal STM in that manner is puzzling, since the monosyllabic items were probably familiar to the participants and they were too short to be able to assess capacity. Another limitation to this study is that the English-speaking children in the sample were nearly a year younger than the children in the other three samples, as was the case in Seymour et al. (2003). Although Ziegler et al. (2010) did not include English-speaking children in their sample, the French (most opaque) children were the youngest. Caravolas et al. argue that the English children had more formal reading instruction (5 to 6 months more) than did the other samples and thus still provides equitable comparisons. In conclusion, Caravolas et al.'s study indicates that, at least in alphabetic orthographies, letter-sound knowledge, PA, and RAN are universal predictors irrespective of orthographic consistency, at least for the early stages of reading acquisition. Although Caravolas et al., like Ziegler et al. (2010), used a non-alphanumeric RAN task, this may have been more appropriate since the participants in Caravolas et al. were kindergartners, whereas Ziegler et al.'s were second graders.

Moll et al. (2014) examined predictors of reading accuracy, reading fluency, and spelling in five varying orthographies (English, French, German, Hungarian, and Finnish). They assessed 1,062 participants from the second to seventh grades on measures of PA, verbal STM, RAN, and reading accuracy, reading fluency, and spelling. PA had a low impact on reading fluency; however, this was not clear-cut across the five languages. RAN was more influential in English reading accuracy than in other languages; however, it is important to note that reading accuracy in this study was a timed measure. RAN also exerted more contribution to spelling in English than to other transparent orthographies. The results indicated that PA and verbal STM predicted reading accuracy and spelling, whereas RAN was the strongest predictor of reading fluency. The results indicate that English behaves much like transparent languages, with two exceptions: 1) PA, verbal STM, and RAN accounted for greater variance in English than in transparent orthographies, suggesting that more cognitive demands are placed on the learner in an inconsistent orthography; and 2) RAN predicted reading accuracy and spelling in English, accounting for 14% and 16% respectively.

Landerl et al. (2019) examined the predictive role of PA and RAN to reading in English, French, German, Dutch, and Greek longitudinally in a sample of 1,120, first and second graders. The children were assessed at three time points: at the beginning of first grade, at the end of first grade, and at the end of second grade. The children were assessed in PA measured by a deletion task of real and nonwords. RAN was assessed using color and digit naming and reading fluency (nonword decoding efficiency and sight-word reading efficiency). Findings revealed that RAN consistently predicted reading in all orthographies. This indicates that the RAN relationship to reading fluency is universal, irrespective of orthographic depth. Furthermore, there was an interaction between RAN and PA. In both English and German, RAN measured at T1 predicted PA at T2, and in Greek, PA at T1 predicted RAN at T2, suggesting that these two constructs are not entirely separable from one another. Unlike RAN, PA failed to show a consistent predictive pattern. PA did not predict reading in transparent Greek and Dutch, whereas PA measured at T2 predicted reading at T3; thus, the relationship is unidirectional. In English, German, and French, the relationship between PA and reading was interactive (predictive pattern going in both directions) (Castles & Coltheart, 2004). The authors lend support to the view that early phonological skills increase as a function of the orthographic complexity being acquired. However, these findings should be interpreted with caution because reading was assessed with reading fluency only.

Summary of Cross-linguistic Evidence

Cross-linguistic research has yielded a finite set of universal predictors of reading and spelling, yet there are conflicting results on the importance of these predictors to reading acquisition as a function of orthographic depth. Although PA is a universal predictor of reading and spelling, it seems to be a stronger predictor in less consistent orthographies (Georgiou et al., 2008; Ziegler et al., 2010). RAN is also a universal predictor of reading and spelling (Caravolas et al. 2012); however, it is not clear whether this relationship is modulated by orthographic depth (Ziegler et al., 2010; Moll et al., 2014). Verbal STM appears to be a universal predictor although to a lesser degree than the influence of PA and RAN. The discrepancies between studies are due to methodological issues discussed at the beginning of this chapter, such as variability in measures used and differences in outcome measures (e.g., accuracy versus fluency). For example, in some consistent alphabetic orthographies, PA has been shown not to be an important predictor of reading; however, this may be due to the choice of tasks that are not sensitive enough (ceiling effects) in consistent orthographies (Ziegler et al., 2010). RAN has been shown to have a stronger influence in consistent orthographies and to be a consistent predictor of reading in transparent orthographies above and beyond PA.

Reading accuracy can be at ceiling by first grade in transparent orthographies. Moll and colleagues (2009) posit that reading fluency in transparent orthographies provides a purer assessment of reading that is not contaminated by accuracy. However, the variance picked up by RAN could have simply been the result of the weak PA tasks. PA tasks could be too easy, so ceiling effects may obscure results in consistent orthographies. But according to Araujo et al. (2015), the influence of RAN was greater in opaque orthographies although it decreased with age. The vast majority of cross-linguistic studies are based on European languages (Share, 2008; Share & Daniels, 2016). Our understanding of orthographic depth comes from evidence that is based on English and other European languages. Thus, it is important to examine non-European languages such as Arabic in order to shed light on how these predictors and reading development interact with the specific features of the language. The following section will review the literature on reading and spelling in Arabic and some literature in Hebrew.

Predictors of Reading and Spelling in Arabic

The Arabic language has linguistic, orthographic, and visual features that may impede reading and spelling acquisition. As discussed in Chapter 1, Arabic has several visual features that have been described as the source of difficulty (Taha, 2013). At first glance, the vowelized Arabic script would seem to be transparent in orthographic depth (Katz & Frost, 1992) and definitions of consistency (Ziegler & Goswami, 2005). When diacritical marks are removed, Arabic is considered opaque (Friedmann & Haddad-Hanna, 2014). Arabic orthography is characterized by extensive diacritics for both consonants and representation of short vowels. It has many letters that have the same basic shape, and letters change shape depending on their position in a word (allography). The visual complexities of Arabic, such as letters having similar shapes, consonant diacritics, ligature, and under-representation of vowels in the unvowelized script, are assumed to be the source of opaqueness in Arabic, which is different from that of English and other European orthographies (Friedmann & Haddad-Hanna, 2014).

Indeed, Khateb et al. (2014) demonstrated the effects of ligature (connectivity) on word recognition in Arabic in a sample of third, sixth, and ninth graders. The researchers hypothesized that, because connecting letters have up to four shapes depending on their position in the word, such words would yield the longest reaction times (RTs) in a lexical decision task. They manipulated ligature using three conditions: (a) non-connecting words, (b) partially connected words (a mixture of connecting and non-connecting letters), and (c) connected letter words. The third graders had slower reaction times for the connected words, but there was no difference in reaction times between partially connected and non-connected words. For the sixth graders, there was a facilitative effect for partially connected and connected words, whereas the non-connected words yielded the longest reaction times. Connectivity did not seem to influence RTs for the more skilled readers in the ninth grade.

Dai et al. (2013) also examined the role of ligature in Arabic among third-grade students. They constructed nonwords that were either connected or non-connected and hypothesized that decoding the non-connected nonwords would be easier (accuracy and fluency, as measured by RTs), but after exposure, once orthographic learning had occurred, connected items would be faster and more accurate to read. The results were inconsistent with their predictions and revealed that connected items took over half a second longer to read than did unconnected items, showing that connectedness did not facilitate word recognition. In the second experiment, the researchers manipulated the number of consonant diacritics (dots appearing in, above, or below consonants). They found that many consonant diacritics were the source of errors.

Based on the definition of orthographic consistency in Ziegler and Goswami's PGST, fully vowelized Arabic would be considered transparent, and as such reading at the word level should develop with relative ease; however, this is not the case. The definition of consistency does not encompass the majority of Arabic's orthographic and visual features that appear to be at the source of reading difficulties. Indeed, Share and Daniels (2016) have argued that the current definition of consistency is based on European orthographies and does not account for other sources of orthographic depth. Thus, it may not be appropriate to define Arabic's orthographic depth by consistency alone. As noted, reading and spelling difficulties among native Arabic speakers have often been attributed to diglossia (Abu-Rabia, 2000; Saiegh-Haddad, 2003, 2005; Asaad & Eviatar, 2014). Thus, diglossia constitutes an additional burden that readers of Arabic must contend with during reading acquisition.

As briefly covered on page 22, Saiegh-Haddad (2003) investigated the effects of diglossia on PA and reading by assessing 23 Arabic-speaking children in kindergarten and 42 in first grade on isolation of initial and final phonemes from an auditorily presented word or syllable. The task was manipulated so that the phonemes were either spoken or diglossic (phonemes that only exist in literary Arabic). Similarly, the syllables were constructed in such a way that the children were asked to isolate a phoneme from a syllable that was either in their spoken vernacular or syllables in literary Arabic. In both conditions, isolating diglossic phonemes was significantly more difficult than was isolating spoken phonemes, and isolation of phonemes from a spoken syllable was significantly easier than it was from a literary syllabic structure. Finally, a nonword reading test was administered for the first graders and revealed the same pattern of results: nonwords that adhered to the spoken structure were read more accurately than were nonwords designed to mimic the literary syllabic structure. Saiegh-Haddad concluded that the linguistic distance between the spoken and literary varieties of Arabic does indeed impede the development of PA among Arabic-speaking children. Thus, the effects of diglossia may contribute to orthographic depth that is not accounted for in the current definition (Share & Daniels, 2016). Indeed, diglossia has been shown to impede spelling in Arabic (e.g., Abu Rabia & Taha, 2006). Spelling in Arabic requires segmenting spoken words into their respective phonemes. Since the mappings of the orthographic units and phonological units are near one to one, the standard expectation would be that spelling for Arabic children should develop with relative ease (Caravolas, 2004). This is not the case. According to Arabic researchers, spelling development is affected by diglossic phonemes (e.g., Abu Rabia & Taha, 2006). This has been demonstrated by languages that have sounds in the dialect that differ from the written form (e.g., *akshara*). This inconsistency in sound-to-spelling linguistic features makes it difficult to spell, even when the script is transparent (Nag, 2011). Indeed, this is the case in diglossic Arabic.

There is also the presence of several emphatic phonemes that are similar to other “soft” phonemes but are pronounced and articulated in the same way (e.g., ت, ط , /t/), which creates phonetic spelling errors despite the transparency (Saiegh-Haddad, 2013).

Taibah and Haynes (2011) investigated predictors of Arabic reading accuracy and fluency in a cross-sectional sample of 237 children from kindergarten to third grade. They measured verbal STM (nonword repetition and digit span), PA (elision and blending), and RAN (object, color, letter, and digit). They also assessed reading accuracy (word recognition) and fluency (nonword reading fluency, text reading fluency) as reading outcome measures, and a comprehension fluency test that required the children to retell the story read in the text-reading fluency as fast as possible in one minute. Verbal STM did not predict any of the reading measures. PA emerged as the strongest predictor of all outcome measures, and that was true irrespective of grade level. RAN also predicted reading, but to a far lesser degree than did PA, and showed the most influence in the third grade. Even when RAN was entered before PA, the influence of PA was superior. There was a trend of PA making more of a contribution to reading in the early grades, whereas the influence of RAN seemed to increase with grade level. The authors argue that these findings are consistent with findings that transparent orthographies demonstrate a greater role of PA during the early years of reading instruction and an increase in the role of RAN in later grades (e.g., Landerl & Wimmer, 2008).

Taha (2013) examined the phonological and visual processing abilities in poor and typically developing readers and their relationship to reading in Arabic. An initial sample of 67 children was divided into the two groups based on their scores on a real-word reading test. Children scoring 70% or below were considered poor readers. This resulted in 32 poor readers (mean age 12.04) and 35 typically developing readers (mean age 12.12). They were all assessed on measures of RAN (letter, digit, and object), phonological processing (elision, blending, verbal STM, working memory), morphological awareness task (morphological decision of selecting the word that is not related to the other four words), visual processing tasks (visual perception, visual search, and visual search of a series of digits), and reading of real words and nonwords. Phonological processing (deletion, blending, working memory) were consistent predictors of real-word reading and nonword reading for both the poor and typically developing readers. Of the visual processing tasks, only the visual search task, requiring the participant to mark the series *529* on a sheet with many distractors, was a consistent predictor, and more so for real words than for nonwords. This is somewhat surprising considering that the author argued that, due to the visual complexity of the Arabic orthography, unfamiliar words (nonwords) would require a great deal of visual attention until the word is parsed and read; however, visual attention, or at least as measured by this task, was more related to real-word reading. Nonetheless, the study does demonstrate the dominance of phonological processing skills in reading in Arabic, and this phonological influence is seen as the consequence of diglossia. RAN predicted real-word reading and nonword reading even after the contribution of phonological processing was accounted for in the typically developing readers. Unlike the findings of Taibah and Haynes (2011), Taha found verbal STM predicted word reading and nonword reading in both groups of readers.

Asaad and Eviatar (2014) examined several cognitive and linguistic predictors of reading accuracy and fluency in Arabic. They assessed children in the first, third, and fifth grades in visual-processing tasks, RAN letter, PA (phoneme segmentation, blending phonemes, syllable deletion, sound deletion), and text reading (accuracy, and fluency). PA was the only predictor of reading accuracy for all grades. PA and RAN were the only two predictors of reading fluency; however, the contribution of RAN to reading fluency was limited to first and fifth graders only. PA appears to be a consistent predictor of reading in Arabic.

As argued by Castles and Coltheart (2004), training studies are needed in order to understand the relationship between PA and reading. Ibrahim (2013) conducted a training study to determine whether phonological training in kindergarten would improve PA in first grade and whether this improvement would be reflected in reading skills. Two groups were selected: 30 children in the intervention group and 27 children in the control group (received no alternate intervention). The training lasted eight weeks and was delivered three times a week for 30 to 45 minutes each time. All measures of PA, nonverbal ability, verbal STM, and receptive vocabulary were measured in kindergarten and after intervention in first grade. Reading was assessed using a word recognition test (matching a picture referent to the corresponding word), a syllable reading test, a nonword reading test, and word reading test, administered at the end of the intervention. The intervention group improved on all measures of PA tasks between pretest and post-test; however, there was no significant difference in reading test scores in first grade between the two groups. The author attributed the non-significant results in reading performance to the linguistic characteristics of Arabic, such as diglossia, the opaque orthography, and visual complexities. These explanations seem unlikely. First, exposure to literary Arabic has been shown to improve children's reading acquisition (e.g., Abu-Rabia, 2000). Second, vowelized Arabic is very transparent, and the reading tests given to first graders in this study would have been fully vowelized. In fact, it is assumed that Arabic's near one-to-one mapping between graphemes and phonemes is the means by which children develop PA in diglossic Arabic (Taha, 2013). The training program in this study included rhymes, blending, and segmenting that were delivered through language games; however, grapheme-phoneme correspondence was never taught. Therefore, it is not surprising to detect gains in phonological tasks but not in reading.

In a review of the literature in Arabic, Al Ghanem and Kearns (2015) examined the relationship between phonological, orthographic, and morphological skills in Arabic word reading. The authors hypothesized that orthographic skills would be more relevant in reading in Arabic than would phonological or morphological skills, due to the visual and orthographic complexities of Arabic. Based on their selection criteria, 12 studies were selected. All 12 examined phonological skills and reading, 5 studies examined orthographic skills and reading, and 3 studies examined morphological skills and reading. Phonological skills were found to contribute to reading in both young and older children. Furthermore, phonological skills were important for both vowelized and unvowelized reading. Since vowelized Arabic is considered to be transparent based on the conventional definition of orthographic depth (Share & Daniels, 2016), this finding is inconsistent with findings in the existing literature on the importance of PA in opaque orthographies rather than transparent ones. Orthographic skills were also found to be related to reading, in particular in studies that used measures of orthographic pattern, and this relationship was found to hold for both vowelized and unvowelized printed words. Limited inferences can be made about the role of morphological skills since only three studies that examined this skill were included in the review; however, the data indicated that morphological skills are important to reading, especially in older grades.

Abu Ahmad et al. (2014) longitudinally examined predictors of word recognition in a sample of children that were first assessed in kindergarten and again in second grade. They hypothesized that phonological processing (RAN, and verbal STM), PA, and morphological awareness would predict word recognition in Arabic. They also hypothesized that, due to Arabic's visual complexities, orthographic-visual processing would play a significant role in predicting word recognition. In kindergarten, the children were administered a visual perception task, short-term symbol memory (where the assessment had very low reliability, .44), PA, (isolation, identification), RAN (objects and colors), and nonword repetition. They also assessed letter naming (12 letters in their non-ligated form), concepts of print (questions related to print conventions), morphological awareness (using pairs of pseudo-words, one with a suffix and one without; the participant had to choose the one that is plural). General ability was also measured (nonverbal ability, receptive vocabulary, working memory, and syntactic awareness). In the second grade, the children were assessed on single word reading (50 vowelized words in increasing difficulty), nonword reading (50 fully vowelized nonwords), and semantic categorization (reading a list of words silently and circling the ones that belonged to a given category). PA was the strongest predator of word recognition. Once again, the pattern of results is what would be expected from an opaque orthography rather than a transparent one. Phonological processing (RAN, nonword repetition), preschool literacy measures (letter knowledge, concepts of print), and morphological awareness also made significant contributions to word recognition. Ahmad et al. argue that the weaker role of RAN in word recognition in comparison to that of PA is also more consistent with findings from opaque orthographies. Another striking finding was the low accuracy rates in word recognition and nonword reading (67% and 63% respectively). These results are also more consistent with what is usually found in an opaque orthography (e.g., Seymour et al., 2003). All these findings suggest that we need to reconsider the criteria of orthographic depth to encompass languages such as Arabic. The authors suggest that Arabic could be considered a case of “semi-modularity” (p. 189), meaning that it is semi-transparent even when it is fully vowelized.

Batnini and Uno (2015) investigated the basic cognitive predictors of reading and spelling in Arabic. One hundred sixteen third graders participated in the study. The authors administered a nonverbal task, phonological processing (nonword repetition and reverse-order repetition of real words), visual cognitive processing (copy drawing, immediate recall, and delayed recall), RAN (objects), receptive vocabulary, single word reading accuracy, nonword decoding accuracy, text reading (timed fluency measure), word spelling to dictation test, and nonword spelling to dictation. RAN was the strongest predictor for all reading measures (single word reading, nonword reading, and text reading) and word spelling.

Working memory (reverse-order repetition) was a predictor of word reading, text reading (timed measure), and word spelling. None of the other measures contributed to reading and spelling measures.

The authors dichotomized the sample into 100 good readers versus 16 bad readers, and 105 good spellers versus 11 poor spellers, to determine the cognitive predictors of reading and spelling difficulties. The poor readers scored significantly lower than did the good readers on the reverse-order repetition task (working memory) and significantly slower on the RAN task. Similarly, poor spellers scored significantly lower than did the good spellers on reverse-order repetition. The authors take these results to indicate that reading and spelling difficulties in Arabic can be predicted by RAN and working memory tasks; however, these group comparisons should be interpreted with caution, as the sample sizes are extremely unequal between good and poor readers and spellers. A serious drawback to this research is the researchers' failure to measure PA, and thus the results of this study should be interpreted with great caution. PA has been found to be the strongest predictor of reading in Arabic (e.g., Assad & Eviatar, 2014; Taha, 2013) and has been found to be the strongest predictor of spelling in various orthographies (Furnes & Samuelsson, 2011); however, this skill was not measured. Thus, the finding that RAN was the strongest predictor of reading and spelling could be the result of not having a PA measure that left the lion's share of variance to RAN (Ziegler et al., 2010).

Bar-Kochva and Brentiz (2013) longitudinally examined predictors of reading as a function of varying orthographic depth in Hebrew. Hebrew, like Arabic, is a Semitic language and has both a vowelized transparent version and an unvowelized opaque version. This allowed the researchers to examine orthographic depth using a within-language design. Eighty-five children were assessed in third grade (T1) and again in fourth grade (T2). Measurement of cognitive skills included receptive vocabulary, PA (elision and phoneme segmentation), morphological awareness (Analogies Test), verbal STM (digit span forward and backwards), RAN (letters and digits), and visual processing speed. Reading tasks were assessed using single word reading (accuracy and speed) and text reading (accuracy and speed) of vowelized, partially vowelized, and unvowelized. Nonword decoding was assessed using only fully vowelized words, as nonwords can be read correctly in multiple ways when unvowelized. PA was the strongest predictor of reading accuracy in general and more specifically in the fully vowelized script. PA in the fourth grade was also the strongest predictor of reading accuracy and showed an increase in predictive power for the partially vowelized and unvowelized scripts from the third grade, revealing equal predictive power in all three versions of the script. Morphological awareness also predicted reading accuracy equally across the three scripts, but only in the third grade. RAN predicted only a small variance in reading accuracy only in the fully vowelized script. Fluency was predicted by RAN in the third graders for the vowelized and partially vowelized scripts only, whereas RAN explained a large amount of variance for the fourth graders in all three versions of the script. PA also predicted reading fluency but only in fourth graders for all three forms of the script. Verbal STM and visual processing did not explain any variance in reading accuracy or fluency. The authors hypothesized that PA and verbal STM would be more related to reading accuracy, whereas vocabulary, morphological awareness, and visual processing would be more related to unvowelized (opaque) script; however, the predictors were very similar, irrespective of the orthographic depth of the script. The importance of PA to fully vowelized reading in Hebrew is similar to findings of its importance in Arabic, which is inconsistent with findings from transparent languages. Such comparisons of Arabic and Hebrew are important because despite being very similar orthographically, diglossia only applies to Arabic. Taken together, it appears that diglossia is not the root cause of this inconsistency but rather the definitional parameters of orthographic depth (Share & Daniels, 2016).

Because Hebrew and Arabic have similar scripts, the comparison of findings from both allows for the separation of factors related to diglossia from factors that are related to orthographic depth. Making use of the same orthographic depth feature as in Hebrew (Bar-Kochva & Brentiz, 2013), Asadi and Khateb (2017) examined the predictive role of PA and RAN in both vowelized and unvowelized reading in Arabic. The Arabic script is both transparent when vowelized and opaque when unvowelized, a feature that allowed the authors to compare the effects of orthographic depth in the same language and using the same participants and tools. A total of 458 children in the first and second grades were assessed in word reading (vowelized and unvowelized), PA (phoneme deletion and segmentation), RAN (objects and letters), and vocabulary (expressive and receptive). Like Bar-Kochva and Brentiz's (2013) results, Asadi and Khateb's results indicated that PA was the strongest predictor for both vowelized and unvowelized reading, and RAN's contribution was weak in both vowelized and unvowelized reading for both grades. Unlike the findings from Bar-Kochva and Brentiz (2013), vocabulary predicted both versions of the script, but the contribution of vocabulary was most pronounced in the second grade and for unvowelized reading. It is possible that vocabulary plays a more significant role in reading in Arabic than Hebrew due to diglossia.

Again, the finding that PA contributed similarly to both vowelized and unvowelized reading and is similar to findings in Hebrew is inconsistent with findings in the existing literature that PA is a stronger predictor of reading in opaque orthographies (Ziegler et al., 2010). The authors explain this odd finding by the characteristics of vowelized Arabic which contains all the phonological information, including short vowels, which forces the reader to pay close attention to every phonological detail. When vowels are removed, readers must rely again on phonological information due to the under-representation of vowels. Indeed, if this is the case, then our conceptualization of orthographic depth is not inclusive of non-European languages. Based on these results, the authors suggest that Arabic's orthographic depth status should be reconsidered and that it should not be solely based on vowelization.

Based on existing literature on the contribution of RAN, the authors predicted that RAN would be a strong predictor of reading vowelized words; however, this was not the case. RAN predicted both vowelized and unvowelized reading and only weakly. While this finding is consistent with Bar-Kochva and Brentiz's (2013) findings, it is inconsistent with findings in the literature on transparent orthographies. The authors attribute this finding to the fact that both reading measures were untimed, which supports claims that RAN is a measure of reading speed (retrieval rate). However, their finding of RAN being equally important regardless of orthographic depth is in line with Landerl et al.'s (2019) cross-linguistic study, described earlier, which concluded that RAN should be considered a universal cognitive ability that is involved in reading, independent of orthographic depth.

Although vocabulary predicted reading of vowelized and unvowelized words at both grade levels, its strongest contribution was to reading of unvowelized words and for the second graders. This finding is consistent with the findings of Farran, Bingham, and Matthews (2016) that vocabulary only predicted unvowelized word reading accuracy. This is consistent with the authors' prediction and with literature that suggests the strong role of vocabulary in opaque orthographies; however, vocabulary contributed to reading even in transparent vowelized Arabic, which the authors attribute to diglossia. Furthermore, the contribution of vocabulary was stronger in the second grade in both the vowelized and unvowelized reading tasks, which suggests that there is a great developmental shift in vocabulary acquisition after two years of formal instruction in MSA.

Taha (2016) examined whether spelling and word recognition processes develop in parallel or whether different skills underlie spelling in a transparent orthography such as vowelized Arabic. A sample of 143 participants from the second, fourth and sixth grades was assessed on fully vowelized word recognition, a spelling to dictation task, and an orthographic decision task, where the participant was asked to choose the correct spelling from pseudo-homophones. Results indicated that, for the second and fourth graders, there was a significant difference between performance on the orthographic task and spelling and word recognition tasks, whereas there was no significant difference between performance on the reading and spelling tasks. For the sixth graders, results indicated that there was a significant difference between performance on the orthographic task and word recognition only. The author concludes that, in earlier grades (second and fourth), spelling and reading development occur in parallel when there is a high reliance on phoneme–grapheme mappings, whereas in the sixth grade, spelling development exceeds reading, and suggests that the reason for this is the “orthographic density” of vowelized words for the experienced reader. However, the conclusions should be interpreted with caution, since scores on the orthographic task (98.1%) and spelling task (99.2%) were at ceiling for the sixth graders.

Asadi, Ibrahim and Khateb (2017) sought to examine cognitive and linguistic predictors of spelling in Arabic in a large cohort of first- to sixth-grade (N= 1,278) native Arabic speakers. For nearly all grades, PA, morphological, and orthographic knowledge consistently predicted spelling. Orthographic knowledge was the strongest predictor of spelling ability, followed by morphological knowledge and PA.

Saiegh-Haddad and Taha (2017) sought to investigate the predictive role of morphological and PA in word spelling and reading in Arabic in a sample of RD children and age-matched typically developing children (TR) in the first to fourth grades (N = 160). Several morphological tasks were employed in both oral and written modalities that targeted the root and word pattern derivational system of Arabic. The PA measures included tasks of deletion, segmentation, and blending. Spelling was assessed using both real- and pseudo-word dictation tasks. As expected, the results demonstrated that PA was the strongest predictor of reading, followed by morphological awareness. However, morphological awareness was found to be a stronger predictor of spelling than was PA. More specifically, PA explained a larger variance in pseudo-word spelling than in morphological awareness, whereas morphological awareness explained a larger variance in real-word spelling. Furthermore, the predictive power of PA in real-word spelling was completely diminished when morphological awareness was accounted for.

In a representative sample of 1,305 participants ranging from the first to sixth grades, Asadi et al. (2017) examined cognitive and linguistic predictors of reading in Arabic-speaking children. They assessed single word reading accuracy and fluency for fully vowelized words as their outcome measures. They assessed nonverbal ability, visual perception, verbal STM, phonological working memory, RAN (object and letter), PA (deletion, segmenting), orthographic processing (detecting word boundaries task, orthographic choice), morphological awareness (inflecting verbs and nouns, derivation of words in context, root awareness, and pattern awareness), vocabulary (receptive and expressive), syntactic knowledge (sentence judgment, personal pronoun affinity). Across all grades, predictor measures contributed similarly to both reading accuracy (48% to 66%) and fluency (46% to 65%) of explained variance. PA was a consistent predictor of reading accuracy, whereas RAN was a consistent predictor of reading fluency. Orthographic knowledge predicted both reading accuracy and fluency; however, its predictive role was dependent on outcome measure and age; the predictive role of orthographic knowledge reached a maximum by the second grade for reading accuracy, whereas for reading fluency its predictive power continued to increase in higher grades. Morphological awareness was only a significant predictor of reading accuracy and fluency in the first and fourth grades. Unlike the findings of Asadi and Khateb

(2017), vocabulary and syntax did not predict either reading fluency or accuracy at any grade. Taken together, their results indicate that, whereas verbal STM and PA contributed to reading accuracy, RAN and orthographic knowledge contributed to reading fluency.

Despite Arabic's visual complexities, the visual perception tasks did not contribute to reading fluency or accuracy past the first grade, and even then it was very marginal. The authors highlight that visual and orthographic processes should be considered separately rather than together, as if one ability. Verbal STM and working memory were consistent predictors in all grades for reading accuracy but not for fluency. The authors attributed that to the fact that Arabic-speaking children's heavy reliance on phonological decoding (fine-grained strategy) puts a strain on the number of units that must be processed through working memory. RAN was the strongest predictor of reading fluency and made no contribution to reading accuracy through all grades. This result supports the argument that RAN is not a phonological processing skill but rather a measure of fluency. In addition, their finding is at odds with findings in the literature that claim that RAN is more important than PA is in reading accuracy in transparent orthographies.

In contrast, PA made more of a contribution to reading accuracy than to reading fluency. Furthermore, this contribution was not limited to younger grades, as suggested in cross-linguistic literature on transparent languages. The authors suggest that this may be due to the fact that all words were vowelized, which forces the reader to pay closer attention to the phonological information. Thus, the contribution of PA to reading accuracy and the lack of contribution to reading fluency highlight the proposition that accuracy development precedes reading fluency. Orthographic knowledge contributed to both reading fluency and accuracy. Orthographic knowledge made more of a contribution in grades five and six, and the authors take that finding to indicate that a reciprocal relationship exists between reading and orthographic knowledge. Morphological awareness only predicted reading accuracy and fluency for the first and fourth graders and made very small contributions to each. The contribution of morphological awareness may have been obscured by the reading outcome measures and may be more important for reading comprehension rather than reading accuracy or fluency. Indeed, Tibi and Kirby (2019) found that morphological awareness had the strongest contribution to reading comprehension, whereas Saiegh-Haddad and Geva (2008) found that morphological awareness does not make any additional contribution to reading once PA is accounted for in word reading. Vocabulary and syntactic knowledge did not predict reading accuracy or fluency in all grades.

In order to investigate the role of RAN and PA and their relationship to reading, Tibi and Kirby (2018) assessed native Arabic third graders to address several key questions. Based on previous research, they hypothesized that RAN and PA would make independent contributions to reading and that PA would be stronger for accuracy, whereas RAN would be more strongly associated with fluency measures. Indeed, after controlling for nonverbal ability, gender, and receptive vocabulary, RAN and PA made unique contributions to each of the five reading measures (word reading accuracy, nonword reading accuracy, word reading fluency, text reading fluency, and reading comprehension); however, the contribution of RAN was less than that of PA on measures of word reading and nonword reading accuracy (untimed reading measures). The strong role of PA observed in this study calls into question the very definition of orthographic transparency (Share & Daniels, 2016), since all reading measures were fully vowelized, making it transparent. Like Share and Daniels (2016), Tibi and Kirby (2018) argue for a multidimensional conceptualization of orthographic depth to encompass dimensions such as missing vowel information in unvowelized script, ligature, diglossia, allography, and visual density. Indeed, receptive vocabulary, measured as a control variable, made a unique contribution to all five reading measures, underscoring the importance of oral language in reading development, particularly in diglossic Arabic.

Tibi and Kirby (2019) examined predictors of reading and reading comprehension in a sample of 201 Arabic-speaking third grade students. They included measures of vocabulary, PA, RAN (objects and digits), which had very low reliability, orthographic processing, and morphological awareness. They also assessed some distal processes such as nonverbal ability, verbal STM, and working memory. They included five outcome measures of reading: word reading (isolated fully vowelized words), nonword reading (fully vowelized), word-reading fluency (number of fully vowelized words read in 60 seconds), text-reading fluency (number of fully vowelized words read in a passage in 60 seconds), and a reading comprehension task. PA predicted every reading measure and emerged as the strongest predictor of word reading and nonword reading, whereas RAN had the strongest influence on the fluency tasks; morphological awareness and vocabulary had the greatest influence on the reading comprehension tasks. Receptive vocabulary predicted every outcome measure, underscoring the importance of oral competence in the SVoR model. PA also predicted every outcome measure, demonstrating that fully vowelized Arabic behaves much like what would be expected in an opaque orthography. For the entire sample, RAN made the strongest contribution to the two timed reading measures (fluency). However, when the sample was divided into poor and good decoders, RAN predicted every reading outcome in the last step of the model (except maze comprehension) for the poor decoders. The authors argue that this is consistent with findings in previous studies showing that RAN is predictive of reading in poor readers. Orthographic processing also consistently predicted all outcome measures, with the exception of reading comprehension. Morphological awareness predicted all outcome measures, having the greatest impact on reading comprehension. These results demonstrate the importance of morphology in Arabic.

Summary of Predictors of Reading and Spelling in Arabic

As mentioned, there are inherent problems with cross-linguistic studies, such as equating measures across the different languages, controlling for environmental differences, and controlling for between-participant differences. The various experimental measures used in the Arabic studies also complicate the interpretation of the results. Thus, these issues, in both cross-linguistic research and research in Arabic, make it difficult to ascertain the skills and abilities that contribute to reading development. Nonetheless, a finite set of consistent predictors can be identified. Table 4 presents a summary of the literature reviewed.

Table 4

Summary of Predictors of Reading and Spelling

Study	Languages	Grade	Predictors of reading accuracy	Predictors of reading fluency	Predictors of Spelling
Georgiou et al. (2008)	English and Greek	1st–2nd	English: PA, orthographic processing Greek: PA, VSTM	English: RAN Greek: RAN	N/A
Ziegler et al. (2010)	Finnish, Hungarian, Dutch, Portuguese, and French	2nd	Finnish: PA, VSTM, vocabulary Hungarian: PA, VSTM Dutch: PA Portuguese: PA French: PA	Finnish: PA, vocabulary Hungarian: PA, RAN Dutch: PA, RAN Portuguese: PA, RAN French: PA, RAN, vocabulary	N/A
Caravolas et al. (2012)	English, Spanish, Slovak, and Czech	Age: 60.27 months–71.86 months	English: letter-knowledge, PA, RAN Spanish: letter-knowledge, PA, RAN Slovak: letter-knowledge, PA, RAN Czech: letter-knowledge, PA, RAN		English: letter-knowledge, PA, RAN Spanish: letter-knowledge, PA, RAN Slovak: letter-knowledge, PA, RAN Czech: letter-knowledge, PA, RAN
Moll et al. (2014)	English, French, German, Hungarian, and Finnish	2nd–7th	English: PA, RAN, VSTM French: PA, RAN, VSTM German: PA, RAN, VSTM	English: PA, RAN, VSTM French: PA, RAN, VSTM German: PA, RAN, VSTM Hungarian: PA, RAN, VSTM	English: PA, RAN, VSTM French: PA, RAN, VSTM German: PA, RAN, VSTM

			Hungarian: PA, RAN, VSTM Finnish: PA, RAN, VSTM	Finnish: PA, RAN, VSTM	Hungarian: PA, RAN, VSTM Finnish: PA, RAN, VSTM
Landerl et al. (2019)	English, French, German, Dutch, and Greek	1st– 2nd	N/A	English: RAN, PA French: RAN, PA German: RAN, PA Dutch: RAN Greek: RAN	N/A
Taibah & Haynes (2011)	Arabic	Kindergarten–3rd	PA, RAN	PA, RAN	N/A
Taha (2013)	Arabic	Poor readers (mean age 12.04) and typically developing readers (mean age 12.12)	PA, RAN, VSTM	N/A	N/A
Asaad & Eviatar (2014)	Arabic	1st, 3rd, and 5th	PA	PA, RAN	
Abu Ahmad et al. (2014)	Arabic	Kindergarten–2nd (longitudinal)	PA, RAN, VSTM, letter knowledge, concepts of print, and morphological awareness	N/A	N/A
Batnini & Uno (2015)	Arabic	3rd	RAN, working memory	RAN, working memory	RAN, working memory
Bar-Kochva & Brentiz (2013)	Hebrew	3rd–4th	PA, RAN morphological awareness	RAN, PA	N/A
Asadi & Khateb (2017)	Arabic	1st–2nd	PA, RAN, Vocabulary	N/A	N/A

Asadi, Ibrahim, & Khateb (2017)	Arabic	1st–6th	N/A	N/A	PA, morphological, and orthographic knowledge
Saiegh-Haddad & Taha (2017)	Arabic	1st–4th (good/poor readers)	PA, morphological awareness	N/A	PA, morphological awareness
Asadi et al. (2017)	Arabic	1st–6th	PA, verbal STM, orthographic knowledge, morphological awareness	RAN, orthographic knowledge, morphological awareness	N/A
Tibi & Kirby (2018)	Arabic	3rd	PA, RAN, vocabulary	PA, RAN, vocabulary	N/A
Tibi & Kirby (2019)	Arabic	3rd	PA, RAN, verbal STM, orthographic processing, morphological awareness, vocabulary	PA, RAN, verbal STM, orthographic processing, morphological awareness, vocabulary	N/A

Based on PGST (Ziegler & Goswami, 2005), and in light of the literature on predictors of reading in Arabic reviewed above, it appears that Arabic's orthographic and linguistic characteristics challenge our current conceptualization of orthographic depth (Katz & Frost, 1992). Consistently, PA was the strongest predictor of reading (e.g., Taibah & Haynes, 2011; Taha, 2013; Assad & Eviatar, 2014; Asadi & Khateb, 2017; Asadi et al., 2017; Tibi & Kirby, 2018, 2019). RAN also made unique contributions to reading but to a lesser degree than did PA (e.g., Taibah & Haynes, 2011; Asadi & Khateb, 2017; Tibi & Kirby, 2018, 2019) and made the greatest contributions in older grades and for reading fluency (e.g., Taibah &

Haynes, 2011; Assad & Eviatar, 2014). These results mirror the findings for English, an opaque orthography (e.g., Kirby et al., 2003; Powell & Atkinson, 2021). PA, RAN, and their relationship to reading have received considerable attention in the literature, and thus the role of PA and RAN was an aim of the current research. Furthermore, the role of verbal STM and working memory in reading in Arabic seems to be inconclusive, some studies showing a robust relationship (e.g., Tibi & Kirby, 2018, 2019), and some showing no relationship (e.g., Taibah & Haynes, 2011). It is important to note that, although reading research in Arabic is growing, none of the measures used in any of the studies reviewed are standardized. Due to these inconsistencies, the aim of this research was to address the role of verbal STM and working memory in reading.

It is not surprising that vocabulary in Arabic has been found to predict reading (e.g., Asadi & Khateb, 2017; Tibi & Kirby, 2018, 2019). Diglossia may be the source of the relationship between vocabulary and reading. However, another possibility for this relationship could be through morphology, which overlaps with vocabulary (Kirby et al., 2012). Morphological awareness has been found to predict reading even after controlling for PA, RAN, and other cognitive abilities (Tibi & Kirby, 2019). Furthermore, reading disabilities in Arabic have been characterized by poor PA, morphological awareness (e.g., Abu-Rabia et al., 2003; Tibi & Kirby, 2019), and RAN (e.g., Saiegh-Haddad, 2005; Tibi & Kirby, 2019). In addition, poor performance on visual processing tasks and verbal STM tasks has been reported in poor readers of Arabic (e.g., Abu-Rabia et al., 2003; Taha, 2013). Based on our current understanding of orthographic depth (Katz & Frost, 1992), vowelized Arabic behaves much like opaque languages rather than transparent ones. The distal and proximal predictors (verbal STM, working memory, vocabulary, PA, RAN, orthographic processing, and morphological awareness) in English and various European languages apply to the vastly different Arabic script and are considered to be universal.

It appears that spelling is also affected by the characteristics of the Arabic language and script. Consistent with research in English (Caravolas et al., 2001), and other languages (Landerl & Wimmer, 2008), it appears that PA is the strongest predictor of spelling in Arabic. There is also evidence that morphological awareness may be as important for spelling in Arabic (e.g., Saiegh-Haddad & Taha, 2017). Thus, an aim of this study was to examine predictors of spelling in Arabic.

The Present Study

Ziegler and Goswami (2005) proposed that reading development is influenced by *availability* (which refers to the fact that that not all phonological units are explicitly available in oral language prior to reading instruction), *consistency* of mappings between orthographic and phonological units, and *granularity* (unit size). Against this backdrop, both proximal and distal predictors of single word reading and spelling have been identified, and several issues have been raised regarding Arabic. Thus, predictors of reading and spelling were reviewed in relation to the three issues outlined in PGST (Zeigler & Goswami, 2005), and Arabic's linguistic characteristics and orthographic features.

According to PGST, the diglossic situation of Arabic creates insufficient availability. There are many sounds that exist in the spoken dialect that do not have a written form, and many sounds in the written form that are not used in the spoken form (Saigh-Haddad, 2005). Zeigler and Goswami (2005) argue that "phonological characteristics of the spoken language have significant effect on phonological development" (p. 7), which is compromised in the diglossic context. Reading acquisition theories are challenged in Arabic where the phonological form of the written word is not consistent with its spoken form (Saiegh-Haddad, 2005). The development of PA may be affected because of the availability problem. There are phonemes that only exist in literary Arabic that are not in the spoken vernacular "as a result of the diglossic context, the orthography often maps onto novel phonological structures that are not familiar to beginning readers from their oral language vernacular" (Saiegh-Haddad, 2005, p. 571). Such phonemes are referred to as diglossic phonemes (Taha, 2013). Thus, availability may be more pronounced in Arabic due to diglossic phonemes and as a result would affect PA.

Second, the role of *consistency* may not be straightforward in Arabic, since its orthographic inconsistency stems from the under-representation of vowels and other visual and linguistic features, rather than the inconsistency of mappings between orthographic units and their corresponding phonological representations. According to our current classification of orthographic depth, Arabic is a transparent orthography when fully vowelized and is considered opaque when diacritics (vowels) are removed. However, based on the persistent role of PA in fully vowelized Arabic, researchers have challenged the conceptualization of orthographic depth (Share & Daniels, 2016; Abu Ahmad et al., 2014; Tibi & Kirby, 2019). Arabic's complex visual features of the orthography may provide a challenge to the assumptions underlying the role of RAN. As mentioned, some researchers have suggested that RAN is a better predictor of reading than is PA in transparent orthographies (e.g., Moll et al., 2009). Thus, Arabic's consistency problem may shape the relationship between PA, orthographic processing, and RAN with reading differently from the way the results obtained from European orthographies indicate. Indeed, Wolf and Bowers (2000) raised this issue by highlighting that there are unanswered questions regarding the role of orthography and its relationship to naming speed and reading.

Finally, due to the challenges posed by *availability* and *consistency* problems in Arabic, the *granularity* utilized in reading Arabic may also be very different from the granularity in other languages. Verbal STM should be more important in transparent writing systems that rely on fine-grained decoding which would require the reader to maintain and recall smaller phonological units, whereas in opaque orthographies, a lesser effect of verbal memory would be observed because the reader is only required to maintain larger-grained phonological units. Vowelized Arabic is considered transparent, and thus it could be assumed that verbal STM and working memory would be important for reading. Based on this assumption, verbal STM should be significant for reading in vowelized Arabic, since the reader relies on decoding smaller phonological units. Arabic's letter shape similarities may force young readers of Arabic to read in a fine-grained fashion because they need to pay extra attention to the consonant diacritics. The visual complexities may hamper the development of a full consolidated phase of word recognition, since one consonant diacritic can change a word into an entirely different word.

Ziegler and Goswami (2005) proposed that, when using fine-grained units for reading in an inconsistent orthography, the need for PA and vocabulary become crucial, as the dependence on PA alone will be insufficient. This is similar to what Ehri (2005a) described as the partial alphabetic phase, in which children can partially decode and draw on their oral vocabulary to arrive at the correct pronunciation. However, this is challenged in Arabic because the children's oral vocabulary is limited due to diglossia.

The role of verbal STM may be more persistent in Arabic due to diglossia. Since many of the words encountered when learning to read are novel, the beginning reader in Arabic cannot rely on a partial decoding technique and existing oral vocabulary. Thus, phonological sensitivity (PA and working and verbal STM) and vocabulary were examined. Indeed, Taha (2013) argues that the transparency of the grapheme-phoneme mappings in vowelized Arabic is a feature that alleviates the constraints of the availability problem. In other words, according to PGST, the problem of availability is remedied by Arabic's orthographic consistency, paving the way for phonological representations to develop through exposure to reading instruction. However, the transition of fully vowelized-transparent Arabic to a non-vowelized script is expected to disrupt the fully consolidated words that were once of high lexical quality into novel words again.

This study examined the predictors of single word reading and spelling in Arabic and English in a sample of third, fourth, and fifth graders who are Arabic native speakers learning ESL and in a sample of monolingual third graders. Based on the literature reviewed above, the research aimed to answer the following broad questions:

1. What are the predictors of reading and spelling in both Arabic and English in the sample of bilinguals? This allowed for a cross-linguistic investigation of universal predictors and language-specific predictors.
2. What are the predictors of Arabic reading and spelling in the sample of third-grade monolinguals?
3. What are the underlying cognitive abilities and reading-related skills that predict reading accuracy, fluency, and spelling? Do these differ between the bilingual and monolingual samples?

It was hypothesized that Arabic and English predictors (PA, RAN, orthographic processing) would be more or less the same in the sample of bilinguals. However, the magnitude of this relationship to reading and spelling may be mediated by orthographic depth. Based on the literature in Arabic, it was hypothesized that PA would be the strongest predictor of reading and spelling. It was also hypothesized that control measures such as visual concentration and vocabulary may play an important role in Arabic due to Arabic's complex visual characteristics and the diglossic phenomenon. Lastly, it was hypothesized that morphological awareness would play a central role in reading and spelling in Arabic, given the rich and complex morphological structure in the language. The next section presents the methodology employed to investigate these aims. Study 1a used a within-subject design in a sample of bilingual Arabic L1 English L2, third, fourth, and fifth graders. Study 2b employed a between-subject design by comparing the bilinguals in Study 2a to the monolingual third graders in Study 2b.

Study 2a Method

This research employed both a within-subject and a between-subject design to address the aims. In Study 2a, a cross-sectional design was employed, in which the sample of bilinguals was divided into younger and older groups in order to investigate whether there was a developmental trend in the variables under examination. In addition, this allowed for the investigation of the influence of the transition from fully vowelized transparent script to the unvowelized opaque script. Two types of analyses were used: correlation and multiple regressions.

Participants

As demonstrated in the literature, children reading in opaque orthographies are shown to have a slower rate of reading development. For example, children's decoding skills are at ceiling by the end of first grade in transparent orthographies, whereas it takes nearly twice as long in opaque orthographies (Seymour et al., 2003). Furthermore, Arabic reading, even when fully vowelized (i.e., transparent), represents a challenge to the novice reader. Thus, a sample of third, fourth, and fifth graders was selected for this study.

Eighty-six bilingual students took part in the study: 28 third graders (16 female; mean age 8.60 ± .47), 31 fourth graders (15 female; mean age 9.75 ± .49), and 27 fifth graders (16 female; mean age 10.96 ± .42).

All bilingual participants were native Arabic speakers attending an American curriculum international school. The participants were selected from two campuses of the same school, located in the upper-middle-class suburbs of Cairo. The two campuses have the same tuition fees and thus cater to children of the same socio-economic status (SES). Both campuses deliver the same American curriculum, and all teachers are native English speakers with the exception of those teaching Arabic (which is taught daily in a 40-minute class), religion (a 40-minute class once a week), and social studies (one 40-minute class once a week). Teachers for Arabic, religion and social studies are native Arabic speakers. All other curriculum subjects are taught in English. Participants were randomly selected from each grade level. Potential participants were excluded from the study if they had any sensory or cognitive impairments.

Materials

The materials covered background assessments of nonverbal reasoning, attention, print exposure, language abilities, phonological processes, orthographic processing, and reading and spelling. This section describes the measures that were used first in Arabic, followed by the English measures.

Every effort was made to ensure that all measures in Arabic and English were parallel, having the same number of items and the same administration and scoring procedures. The measures in Table 5 are divided into background non-linguistic measures, Arabic phonological processing measures, Arabic orthographic measures, Arabic oral language measures, and Arabic reading and spelling measures. This is followed by the comparable English measures. All Arabic measures were provided by Dr. Nadia Taibah from the Department of Special Education at King Abdulaziz University. All Arabic assessment tools provided were not yet standardized. The majority of the Arabic tools were developed based on standardized English assessments using the same administration and scoring protocols, and they are discussed, where applicable, below.

Nonverbal Control Measures

Two tests were used as control measures: 1) the Raven's-Educational (UK Edition) Standard Progressive Matrices-Plus version (SPM+ Raven, 1998), and 2) the Finger Windows subtest of the Wide Range Assessment of Memory and Learning, Second Edition (WRAML-2; Sheslow & Adams, 2003).

The SPM+ was administered as a measure of nonverbal general ability. Booklets and answer record sheets were distributed to each participant. The booklet comprises five sets, each containing 12 items. The items increase in difficulty within each set and from set to set (e.g., set A has items from A1 to A12). Each item has a pattern with a missing piece. The participant had to determine which one of the four to six choices completes a given pattern and fill in the answer on the answer record sheet. The participants were given 40 minutes to complete this task. The raw score was calculated by summing up the number of correct responses in each set (maximum of 12) and then summing the scores of all sets (maximum of 60). The reported reliability is .94 for ages 7 to 18.

The Finger Windows subtest of WRAML-2 was administered as a measure of short-term visual attention capacity. This task was administered to assess visual attention, which is believed to be relevant to reading in Arabic. The participant is presented with a rigid plastic rectangle piece with cut-out holes ("windows"). The examiner inserts the tip of a pencil approximately one inch into a series of holes, one at a time in a prescribed order. The participant is then asked to insert his or her finger in the same windows and in the same order as the examiner. The order and number of windows increases in difficulty, and the test is terminated after three consecutive errors. This task was individually administered. The raw score was calculated by summing the number of correct responses and the six points below the basal (participants under age 8 are administered first six items). The maximum score for this test is 30. This subtest has a reported reliability of .81 (for age range 8.0 to 8.11) and .82 (for age range 9.0 to 13.11).

Arabic Control Measures

The Arabic Control measures include receptive vocabulary, two measures of verbal short-term memory, working memory, and the Arabic Title Recognition Test that was described in Chapter 2, to assess print exposure. These are all described next.

Receptive Vocabulary. Vocabulary is essential to reading (Kirby et al., 2008). The Arabic receptive vocabulary task is modeled on the commonly used English receptive vocabulary measure, The Peabody Picture Vocabulary Test (PPVT III; Dunn & Dunn, 1997). This test was administered as a measure of children’s vocabulary. The test was administered individually. The examiner asked the examinee to point to one of the four pictures that would best convey the word that the examiner had given orally. There are four practice items and 103 test items. All items were administered, as there is no discontinuation rule. Feedback and correction was given only for practice items. The total score is the sum of correct responses. The maximum score for this test is 103. The reported reliability for this measure is .96 for kindergarten to sixth grade (Mahfoudi et al., in press).

Verbal Short-term Memory. The temporary storage until subsequent units are decoded and finally assembled is an ability crucial to reading because the sequence of sounds that have been decoded need to be retained until the remaining sounds are decoded (Baddeley, 2012; Catts et al., 2006; Wagner et al., 1997). To assess phonological verbal STM, two commonly used tasks were administered. The Arabic verbal STM assessment included two tasks: nonword repetition and digit span forward. These are modeled after the subtests of nonword repetition and memory for digits of the Comprehensive Test of Phonological Processing (CTOPP; Wagner et al., 1999). The two subtests were recorded by a speech and language pathologist to ensure consistency in test administration (Taibah, et al., in press).

The digit span task is administered to assess participants' verbal STM. In this task, the participant is instructed to listen to an audio-recorded series of digits and then repeat the digits verbatim. This test had three practice items of two to three digits and 13 test items. The number of digits for recall increased gradually. The ceiling for this subtest was reached once the participant made three consecutive errors, at which point the test was terminated. The score was determined by the number of correct items for a maximum score of 13. This subtest has a reported reliability of .80 for kindergarten to sixth-grade students.

The Nonword Repetition task, like the Digit Span Forward task, is an additional measure of participants' verbal STM. The participant is asked to repeat audio-recorded nonwords of increasing difficulty. There were two practice items and 20 test items. The score was computed by summing the total number of correct responses. This subtest was terminated after the examinee made three consecutive errors. The score was determined by the number of correct items, for a maximum score of 13. It has a reported reliability of .83 for kindergarten to sixth grade.

Working Memory. While the term working memory is used interchangeably with verbal STM, Baddeley (2012) draws a distinction between the two; working memory is where information is temporarily stored and manipulated. The Digits Backward subset is also administered as a measure of participants' verbal STM; however, this task is more challenging because the participant is required to repeat the digits heard via an audio-recording backwards. There are three practice items of two to three digits each and 13 test items ranging from two digits to six digits in length. The score was determined by the number of correct items, for a maximum score of 13. The test was terminated after three consecutive errors. The reported reliability of this subtest is .74 for the first to sixth grades.

Print Exposure. The Arabic Title Recognition Test (Arabic TRT) was administered as a measure of Arabic print exposure. Participants were presented with a list of 30 Arabic book titles. They were told that some of the book titles are real and some are fictitious. They were instructed to circle the ones they know to be real book titles. Scoring is determined by the number of real titles minus the number of foils selected to correct for guessing, for a maximum score of 20. The Cronbach's α reliability is .79.

Arabic Reading-related Measures

Based on the literature of reading-related skills, the following were assessed: PA, RAN, visual/orthographic processing, and morphological awareness.

Phonological Awareness*.* The ability to identify and manipulate sounds in spoken words is important to reading (Wagner et al., 1997; Manis et al., 2000; Mutter et al., 2004; Ehri, 2005a, 2005b). In order to assess this ability, two Arabic PA tasks, modeled after the elision and blending words subtests of CTOPP (Wagner et al., 1999), were used. Phonological awareness measures such as elision and blending are considered more appropriate for measuring phonological awareness and are less susceptible to ceiling effects (Ziegler et al., 2010). In the Sound Deletion task the participant had to repeat a word, then delete a sound from a word and produce the new word. For example, the tester would say “Say cat. Now say cat without /k/”. There are three practice items composed of compound words. The participant is asked to delete one word, followed by three test items. The second part has two practice items. The participant is asked to delete a sound instead. This test is composed of 20 items and was terminated when the participant made three consecutive errors. The raw score was determined by summing the number of correct items, for a maximum score of 20. The reported reliability for this task is .93 for the first to sixth grades.

The Sound Blending task requires the participant to blend sounds to produce a word. For example, the participant hears sounds from an audio recording, /k/, /a/, /t/, and is required to blend the sounds and produce the word /cat/. This task begins with three practice items in which the participant must blend syllables. Feedback and correction are given on these items, followed by four test items of multisyllabic words. No feedback is given on test items. The second part consists of two practice items. The participant is required to blend sounds to form a given word, followed by 16 test items. This task consists of 20 test items in total. The score is calculated by summing the number of correct items. Testing is terminated if the participant makes three consecutive errors. The items in this task were recorded by a speech and language pathologist, to ensure consistency in test administration. The reported reliability for this task is .81 for kindergarten to sixth grade (Taibah et al., in press).

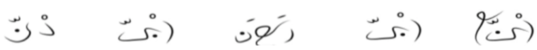
RAN. Naming speed as assessed through the rapid naming of letters, numbers, objects, or colours, has also been linked to reading development in various orthographies (Manis et al., 1999; de Jong & van der Leij, 1999; Kirby et al., 2003; Georgiou et al., 2016; Georgiou et al., 2021). Two alphanumeric RAN tasks were used in this study, RAN-Digit and RAN-Letter, both of which are modeled after the RAN subtests of CTOPP (Wagner et al., 1999). For the RAN-Digit subtest the participant was presented with a card on which were printed four rows containing nine single digits. Participants were asked to name the digits as accurately and as fast as possible. On the practice items, there are six different digits, the same digits that appeared in the test items in random order. The participant had to correctly name all digits on the practice items in order for the test to be administered. Two forms, A and B, were administered, and scoring was determined by summing the number of seconds required to name the digits on both forms. An additional second was added for each error. The reported reliability on this task is .92 for kindergarten to sixth grade (Taibah et al., in press).

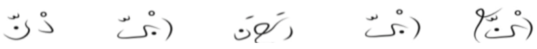
The RAN-Letter subtest is identical to the RAN-Digit task except that it contains letters instead of digits. There are also two forms to this task, using the same administration and scoring procedures as the RAN-Digit task. The participant was instructed to say the letter name and not letter sound during practice items. There were 12 letters presented during practice, and the same 12 letters appear on both forms in random order. The test was not administered if the participant could not name all of the letters on the practice items. Two forms, A and B, were administered, and scoring was determined by summing the number of seconds required to name the letters on both forms. An additional second was added for each error. This subtest has a reported reliability of .88 for kindergarten to sixth grade (Taibah et al., in press).

Visual/Orthographic Processing Measures. The ability to store and retrieve orthographic representations contributes to reading development. Thus, four tasks were administered to assess this ability: Orthographic Memory, Orthographic Matching, Pseudo-letter Memory, and Pseudo-letter Matching (Haynes et al., in press).

The Orthographic Memory task was administered as a measure of participants’ orthographic processing ability. The participant is presented with a letter or letters on a computer screen for two seconds. The target letter/letters then disappear, and three to four letter/letters (target(s) and distractors) appear in a row. For example, the participant would see the target (ـخ ق) for two seconds followed by a row of target and distractors (e.g. ـق ر, ـحـ ت, ـخ ق, ف ـح) . The participant is asked to point to the correct answer. There were three practice items and 30 test items in increasing difficulty, targets beginning with a single letter and ending with seven-letter strings. The participant was given correction and feedback on the practice items but not for the 30 test items. The total score was calculated by summing the number of correct responses on the test items. The maximum score for this task is 30. The reported reliability for this task is .88 for the second to sixth grades.

The Orthographic Matching task was also administered as a measure of participants' orthographic processing ability; however, this task was timed and is considered as a measure of orthographic fluency. The participant was presented with a row of letters and instructed to put a slash through the two items that are the same in each of the given rows. There are three practice items and 25 test items presented in increasing difficulty, beginning with two-letter clusters (e.g., ف ح   ح ت   خ ق   ق ر   خ ق) and ending with a cluster of seven-letter strings (e.g., قغن لآ   م علت   قغن لآ   دمن كت   ظّم و). Each row had five items. After the practice items were administered, the examiner set a timer for two minutes. The total was calculated by summing the number of correct responses. If a participant completed the 25 items before the two-minute period, the time was recorded. The maximum score for this task is 25. The reported reliability for this measure is .83 for the second to sixth grades.

The Pseudo-Letter Memory task is similar to the Orthographic Memory task; however the items are pseudo-letters that resemble Arabic letters and diacritical marks (symbols that cannot be decoded), which is thought to reduce verbal mediation and to rely more on visual processes. The participant was presented with a pseudo-letter or a combination of pseudo-letters on a computer screen for two seconds (e.g., ), followed by four choices displayed in a single row (e.g., ). The participant was required to choose the one that had appeared earlier. There are three practice items, followed by 30 test items. Only the test items are scored, for a maximum score of 30. Correction and feedback was given only for the practice items. The reported reliability of this task is .88 for the second to sixth grades.

The Pseudo-Letter Matching task is similar to the orthographic matching task; however, instead of letters the items are pseudo-letters that resemble Arabic letters. The participant is presented with a sheet with three practice items displayed in a row and told to put a slash through the two items that are the same (e.g., ). Correction and feedback were given for the practice items, followed by 25 test items. The participant was timed for two minutes, and the score was calculated by summing the number of correct items recorded in two minutes. The time was recorded if the participant completed the 25 items in less than two minutes. The maximum score for this task is 25. The reported reliability for this measure is .78 for the second to sixth grades.

Morphological Awareness. Morphological awareness has been linked to reading outcomes (Bowers, Kirby, & Deacon, 2010; Tibi et al., 2020). The morphological production task was developed based on the subtask of the English Clinical Evaluation of Language Fundamentals (CELF-5; Wiig et al., 2013). This task was administered as a measure of morphological awareness. Each slide had two pictures side by side. The first picture was described by the examiner. The examinee was expected to describe the second picture by changing from singular to plural, feminine to masculine, etc. This measure contains four practice items and 39 test items. The total was calculated by summing the number of correct responses for a maximum score of 39. The reported reliability of this task is .94 for kindergarten to sixth grade (Mahfoudi et al., in press).

Arabic Reading and Spelling Outcome Measures

Two Arabic reading measures were administered: Word Recognition and Nonword Reading Fluency. A spelling to dictation test was administered to assess spelling ability in Arabic. Administration and scoring are described below.

Word Recognition*.* The word recognition task was modeled after the English Word Identification subtest of the Woodcock Reading Mastery Test Third Edition (WRMT-III; Woodcock, 2011) (Alsudairi et al., in press). This widely used task was administered as a measure of reading accuracy. This task consists of 50 items of increasing difficulty. The participant is first shown three practice items. Correction and feedback were given during the practice items. The instructions given by the examiner were as follows: “Now you will read some more words. Read all words on the first row (the examiner pointed to the right corner of the first row) and then move on to the second row and so on. Read all words on each row, paying close attention to diacritical marks. Try not to miss any words.” These instructions were read verbatim in Arabic to each participant. If a participant was unable to read an item after five seconds had lapsed, he/she was instructed to move on to the next item. For each item read correctly a score of 1 is given. The total raw score is the total of items read correctly, including proper pronunciation of diacritical marks. The test was terminated after five consecutive errors. The maximum score for this test is 50. The reported reliability of this task is .90 for third and fourth graders and .88 for the fifth and sixth graders.

Nonword Reading Fluency. The nonword reading fluency task was modeled after the Phonemic Decoding Efficiency (PDE) subset of The Test of Word Reading Efficiency Second Edition (TOWRE-2; Torgesen et al., 2012). This commonly used task was administered as a measure of decoding efficiency. The nonword fluency task required the participant to read Arabic nonwords as accurately as possible in one minute. The items were presented in columns in increasing difficulty. Because the items are nonwords, reading them with the correct diacritical marks was essential. The examiner presented the participant with three practice nonwords and gave the following instructions: “These are not real words. Watch how I read the first word.” The examiner then instructed the participant to read the nonwords and gave corrections and feedback. The test items were then presented, and the participant was instructed as follows: “These are some more nonwords. Start here” (examiner points to the first item in the top right column). If the participant was unable to read an item within three seconds, he/she was instructed to move on to the next item. The test was terminated after one minute. The total raw score was derived by summing all nonwords read correctly in one minute, for a maximum score of 35. The reported reliability for this test is .93 for the second to six grades (Alsudairi et al., in press).

Arabic Single-word Spelling*.* In order to assess spelling ability, a spelling to dictation task was administered. The spelling to dictation task consisted of 38 items. Each word was first read aloud by the examiner, followed by embedding it in a sentence and finally in isolation once more. The participants were instructed to wait until the word was said the third time before spelling it. This test was group administered, and all 38 words were dictated to the participants. The score was determined by the number of items spelled correctly, for a maximum score of 38. The reported reliability for this test is .89 for the third to sixth grades (Alsudairi et al., in press).

English Control Measures

The English control measures included receptive vocabulary, two measures of verbal STM, and the TRT in English that was outlined in Chapter 2, to assess print exposure.

Receptive Vocabulary. The Peabody Picture Vocabulary Test-Fourth Edition (PPVT-4; Dunn & Dunn, 2007) is administered individually. This test was administered as a measure of vocabulary. The participant is presented with an easel with four pictures and then asked to point to one of the four pictures that corresponds with the word just heard. Each set contains 12 items, and a basal is established when a participant has zero to one error on a given set. The ceiling is also established by eight or more errors in a given set. A score of 1 is given for all items below the basal item and summed all the way through to the ceiling item. The score is determined by summing up the correct items, for a maximum score of 228. The average reported reliability of this test is .97 for ages 2.6 to over 81 years old (Dunn & Dunn, 2007).

Verbal STM. The Memory for Digits subtest from CTOPP-2 was administered (Wagner et al., 2013). This test was administered as a measure of verbal STM. Memory for Digits requires the participant to listen to a series of audio-recorded numbers, presented at the rate of two numbers per second, ranging in length from two to eight numbers in total. The participant was asked to repeat the series of digits verbatim in the order in which it was heard. The ceiling for this subtest was reached after three consecutive errors, at which point the test was terminated. The task has 28 test items, and the raw score is computed by summing the number of correct responses, for a maximum score of 28. This subtest has an average reported reliability score of .81 for ages 4 to 24.

Similarly, The Nonword Repetition subtest from CTOPP-2 (Wagner et al., 2013) was administered as a measure of verbal STM. The Nonword Repetition task requires the participant to listen to nonwords on an audio-recording. The nonwords range from 3 to 15 phonemes in length. The participant was asked to repeat verbatim the nonword just heard. This test was terminated once the participant made three consecutive errors. The task has 30 items of increasing length. The raw score is computed by summing the number of correct responses, for a maximum score of 30.It has an average reported reliability of .77 for ages 4 to 24.

Print Exposure. The English Title Recognition Test (English TRT) was administered as a measure of English print exposure. Participants are presented with a list of 30 English book titles. They are told that some of the book titles are real and some are fictitious. They are to circle the ones that they know to be real book titles. Scoring is determined by the number of real titles minus the number of foils selected, to correct for guessing, for a maximum score of 20. The Cronbach's α reliability is .81.

English Reading-related Measures

Phonological Awareness. Two subtests of the CTOPP-2 (Wagner et al., 2013) Form A were administered as a measure of PA: Elision and Blending Words. The elision task requires the participant to first hear a compound word, "basketball" and then to repeat the word and delete one word from the compound word (e.g., "now say basketball without ball"). The items increase in difficulty. The participant was asked to delete a phoneme from beginning, medial, or final position. For example, “say bold without /b/.” The test has 34 items and was discontinued once the participant made three consecutive errors. The raw score is calculated by summing the number of correct responses, for a maximum score of 34. This subtest has an average reported reliability of .91 for ages 4 to 24.

The blending words task requires the participant to listen to a series of audio-recorded separate words in compound words, such as “what word do these sounds make: /tooth/, /brush/”? The participant blends the compound word to produce “toothbrush”. The items increase in difficulty. The participant was then required to blend the individual sounds to produce the word. This task contains 32 items and is discontinued after three consecutive errors. The raw score is calculated by summing the number of correct responses, for a maximum score of 32. This subtest and has an average reported reliability of .86 for ages 4 to 24.

RAN. Two subtests of CTOPP-2 (Wagner et al., 2013) Form A were administered: RAN digit and RAN letter. These tests were administered as measures of naming speed. The Rapid Digit Naming test requires the participant to name a series of numbers listed in rows. Participants are first asked to name all the digits on a practice sheet, to ensure that they are familiar with these numbers. The test cannot be administered if the participant cannot name all six numbers presented on the practice card. The participant is then given a card with the same numbers appearing in random order in four rows and instructed to read each row from left to right. The score is derived from the total number of seconds taken to name all the digits. The participant is instructed to read the digits in each row quickly and accurately. The time in seconds is recorded. The participant’s score is determined by the number of seconds taken to name all the digits. One second is deducted for each digit named incorrectly. This task has an average reported reliability of .85 for ages 4 to 24.

The Rapid Letter Naming test requires the participant to name a series of letters listed in rows. Participants are first asked to name six randomly arranged letters appearing on a practice card to ensure knowledge of these letters. If the participant is able to name all the letters correctly, he/she is then presented with a card with four rows of the same letters from the practice items appearing in random order on each row. The score is derived from the total number of seconds taken to name all the letters. The examinee is instructed to read the letters in each row quickly and accurately. The time in seconds is recorded, and the score is determined by the number of seconds taken to name all the digits. One second is deducted for each digit named incorrectly. Like the Rapid Digit Naming task, this task has an average reported reliability of .87 for ages 4 to 24.

Orthographic Measure. The Orthographic Matching subtest of the Test of Orthographic Competence (Mather et al., 2008) was used to assess participants' orthographic processing ability. This subtest is standardized for ages six to seven years, but since the Arabic Orthographic Matching task was developed based on this subtest, it was decided to modify the task so that the number of items was equal to that in the Arabic version. This resulted in deleting the first 13 rows, so that there were 25 rows. The participant is presented with a sheet containing rows of five orthographic patterns, two of which are the same (e.g., oa ao co oa oc). The participant is asked to put a slash through the two that are the same. This test is timed for two minutes, and the score is calculated by summing the number of correct responses.

English Reading and Spelling Outcome Measures

English Reading Accuracy. English word recognition was assessed using the Word Identification subtest Form A of the Woodcock Reading Mastery Test Third Edition (WRMT-III; Woodcock, 2011). The participant was presented with a stimulus book and instructed to read the words presented. This test contains 46 items in increasing difficulty. The test was terminated once the participant made four consecutive errors. The raw score was calculated by summing all the items below the participant's basal and all words read correctly, for a maximum score of 46. The reported reliability of this subtest is .91 for third graders, .94 for fourth graders, and .91 for fifth graders.

Fluency Measures. The Decoding Efficiency (PDE) subset of The Test of Word Reading Efficiency Second Edition (TOWRE-2; Torgesen et al., 2012) Form A was administered as a measure of decoding efficiency. Words must be read accurately within 45 seconds. The participant was first presented with practice items. The participant must be able to read at least one practice item to administer the test. The participant was then presented with a card containing the test items and told to read them as accurately and as fast as possible. There is no stopping rule for this test. The raw score is calculated by summing the number of items read correctly in 45 seconds. One second is deducted for each item read incorrectly. The overall reported reliability of TOWRE-2 PDE is .92.

English Single Word Spelling. The Single Word Spelling Test (SWST; Sacre & Masterson, 2000) was administered as a measure of single word spelling in English. The spelling to dictation task consists of 50 items. The word is first read aloud by the examiner, followed by embedding it in a sentence, and finally in isolation once more. The participants were instructed to wait until the word is said the third time before spelling it on their response sheets. This test was group administered, and all 50 words were dictated to the participants. The score is determined by summing up the number of items spelled correctly, for a maximum score of 50.

Ethical Considerations

Ethical committee approval was granted by the Institute of Education, University College London, to conduct this research. Permission to conduct the study at the school was obtained in response to a request sent to the school principal (Appendix E). Upon receipt of approval from the university and the school principal, letters of parental/carer consent were sent to potential participants' parents/carers (Appendix F). The procedures for the protection of human participants were scrupulously followed. All participants remain and will remain in perpetuity anonymous to all but this researcher. Only this researcher had or has access to the data entered by the participants and used for data analysis. Participants' names were not revealed in any of the documents nor to other researchers. A random numeric identifier was assigned to each participant to ensure the anonymity of responses throughout and beyond the research process. Parental/carer consent forms were sent home with students at the participating school. The letter included brief information about the researcher's background, the aim of the study, the method of gathering data, and a request for their child's participation (Appendix F). Parents/carers were assured that participation was entirely voluntary and that their child's participation or lack thereof would not reflect in any way on his/her grades. Parents/carers were encouraged to discuss the researcher’s request with their child. Classroom teachers distributed the parental/carer consent forms to students and collected them a week from the day on which they were distributed and returned them to the researcher. At every stage of data collection, each student was reminded of the right to withdraw from the study at any point.

Further, participating students were told that the data collected would only be used for the purposes of the research and would not reflect on their grades. Great care was taken to ensure that the participants and parents/carers fully understood the nature of the study and that participation was voluntary. No information regarding participation of any individual was communicated to the school or teachers. Confidentiality of data was maintained at all times. These conditions were communicated to all participants at the start of the research.

Procedure

Once ethical approval was obtained and consent forms were collected, data collection began. The first data collection took place in February 2016 and included 13 third graders, 11 fourth graders and 7 fifth graders. In May 2016, 20 fourth graders and 20 fifth graders were assessed from the other campus. Finally, an additional 17 third graders were assessed in November 2016. Bilingual participants were seen during four sessions lasting 40 minutes each. Each student was seen individually for three sessions and once for a group session. Two of the individual sessions were for Arabic assessments, and one individual session was for English assessments. The administration of English and Arabic assessment sessions was counterbalanced. Each student either started with a session of English assessment followed by the two Arabic assessments, or two Arabic sessions followed by a session of English assessments. It was deemed necessary to separate English and Arabic assessments because of reported switching costs observed with bilinguals (Thomas, & Allport, 2000; Peeters et al., 2014). Participants were seen in a group session of no more than six for the Raven's matrices measure and English spelling. The first Arabic session included Arabic phonological measures, orthographic measures, and reading measures. The second session included Arabic vocabulary measure, morphological production task, Arabic spelling, and Arabic print exposure. Each session lasted for 40 minutes. The English session included all phonological measures, orthographic measure (orthographic matching), reading measures, receptive vocabulary, and English print exposure. Table 5 describes the assessments used in Arabic and English for the bilingual sample.

Table 5

Arabic and English Measures Administered to Bilinguals

Areas/processes assessed	Arabic measures	English measures
Nonverbal control measures	**Nonverbal reasoning** Raven's-Educational (UK Edition) Standard Progressive Matrices-Plus version (max = 60) **Visual concentration** Finger Windows subtest of the Wide Range Assessment of Memory and Learning (max = 30)	**Nonverbal reasoning** Raven's-Educational (UK Edition) Standard Progressive Matrices-Plus version (max = 60) **Visual concentration** Finger Windows subtest of the Wide Range Assessment of Memory and Learning (max = 30)
Control measures	**Print exposure** (max 20): Arabic Title Recognition Test **Receptive Vocabulary** (max = 103)**:**	**Print exposure** (max = 20): English Title Recognition Test

	Arabic version of The Peabody Picture Vocabulary Test **Verbal short-term memory**: Nonword repetition (max = 13) and digit span forward (max = 13) subsets of the Arabic version of Comprehensive Test of Phonological Processing **Working memory:** Digit span backward (max = 13)	**Receptive Vocabulary**: The Peabody Picture Vocabulary Test (max = 228) **Verbal short-term memory**: Nonword repetition (max = 30) and digit span forward (max = 28) subsets of the Comprehensive Test of Phonological Processing
PA	**Elision (**max= 20), **blending words (**max = 20): Arabic version of Comprehensive Test of Phonological Processing	**Elision (**max = 34), **blending words (**max = 33): Comprehensive Test of Phonological Processing
RAN	**RAN-digit, RAN-letter** (timed): subsets of the Arabic version of Comprehensive Test of Phonological Processing	**RAN-digit, RAN-letter** (timed): subsets of the Comprehensive Test of Phonological Processing
Orthographic processing	**Orthographic matching:** subtest of the Arabic version of the Test of Orthographic Competence (max = 25) **Pseudo-letter matching:** Experimental task (max = 25) **Orthographic memory:** Experimental task (max = 30) **Pseudo-letter memory:** Experimental task (max = 30)	**Orthographic matching:** subtest of the Test of Orthographic Competence (max = 25) **Pseudo-letter matching:** Not assessed in English **Orthographic memory:** Not assessed in English **Pseudo-letter memory:** Not assessed in English
Morphological measure	**Morphological production:** Subtest of the Arabic version of the Clinical Evaluation of Language Fundamentals (max = 29)	**Morphological production:** Not assessed in English
Reading measures	**Word recognition:** Arabic version of the Word Identification subtest of the Woodcock Reading Mastery Test (max = 50) **Nonword reading fluency:** Arabic version of the Phonemic Decoding Efficiency subset of The Test of Word Reading Efficiency (max = 35)	**Word recognition:** Word Identification subtest of the Woodcock Reading Mastery Test (max = 46) **Nonword reading fluency:** Phonemic Decoding Efficiency subset of The Test of Word Reading Efficiency (max = 108)

Spelling measures	**Spelling to dictation** Arabic version of The Single Word Spelling Test (max = 38)	**Spelling to dictation** The Single Word Spelling Test (max = 50)

Note. All Arabic measures were administered to the monolingual participants, with the exception of orthographic memory and pseudo-letter memory.

Study 2a Results

This section presents the results for the bilingual participants in Study 2a. The bilingual sample was divided into younger and older groups to explore developmental trends. Descriptive statistics for the Arabic measures are presented, followed by correlations between Arabic measures and hierarchical regressions exploring predictors of reading and spelling in Arabic (within-language) for the bilingual sample in Study 2a. Descriptive statistics for the English measures are then presented, followed by correlations and a series of hierarchical regressions exploring predictors of reading and spelling in English (within-language) for the bilingual sample in Study 2a.

Exploratory Data Analysis

Prior to analysis, EDA was carried out to determine if the variables met the assumption of normality. Normality was assessed via the Shapiro Wilk statistic, significance values above .05 considered normal, skewness and kurtosis ratio values with value within +/-3 considered normal, and finally through an examination of histograms. Across all analyses mild violations of normality were not considered problematic, as the sample size of 86 is above the value of 30 required for consideration of normality within the central limit theorem.

Inspection of the bilingual results indicated that most variables were normal; however, some deviated from normality and required adjustment in outlier removal or transformation. Arabic pseudo-letter memory had one low-end outlier, which when removed resulted in only a mild violation of normality based on Shapiro Wilk, $p = .016$, and was therefore considered normal. Arabic RAN digit had two high-end outliers, which when removed resulted in a still significant Shapiro Wilk value; however, inspection of the histogram indicated only a mild violation of normality. Arabic RAN letter had three high-end outliers, which when removed resulted in only a mild violation of normality. English word recognition had two low-end outliers, which when removed result in a Shapiro Wilk significance of $p = .133$. English elision had two low-end outliers; however, when removed there was still significant departure from normality, resulting in a negative skew. In order to address this, the variable was reverse-coded (so a high score became a low score, in order to create a positively skewed variable). Log10 transformation was applied resulting in a normally distributed variable for use in subsequent analysis.

Within-language Predictors of Reading and Spelling for the Bilingual Group

The following sections present predictors of reading and spelling for Arabic and English separately. First, preliminary analysis is presented, and then regression analyses. Because the bilingual sample was drawn from two campuses at two times, independent samples *t*-tests were conducted to determine whether the performance of third-, fourth-, and fifth-grade students differed across the two locations. The findings revealed that fourth and fifth graders from the first campus did not differ from the participants from the second campus on any measure. However, the findings for third-grade students revealed that the students did differ significantly in two of the English measures. Third-grade students in the first campus had significantly higher English PPVT (p = .046) and NR (p = .001) scores than did the students in the second campus.

Arabic Descriptive Statistics and Preliminary Analysis for the Bilingual Group

One of the aims of this study was to explore developmental trends in reading and spelling using a cross-sectional design; however, the sample size was small for each individual grade. Splitting the sample into older and younger participants increased the sample size without compromising the cross-sectional data. The bilingual sample was divided into two groups on the basis of their chronological age (Abu-Rabia, 1995). This resulted in 46 participants in the younger group (age < 9.11 years), and 40 in the older group (age =/+ 9.11). The younger group had a mean age of 8.61± .50, and the older group had a mean age of 10.27 ± .41. Table 6 presents descriptive statistics for all variables before data reduction for the younger and older groups.

Table 6

Descriptive Statistics for Arabic Measures by Age Group for the Bilingual Group in Study 2a

Measures	Younger group (N=46)		Older group (N=40)		Differences between the groups	
	Mean	SD	Mean	SD	*t*	*P*

Nonverbal control measures						
Nonverbal ability (max = 60)	24.3	6.3	27.1	5.1	-2.3	.02
Visual concentration (max = 30)	12.7	2.5	14.1	3.1	-2.23	.03
Arabic control measures						
Nonword repetition (max = 13)	12.1	3.6	12.3	3.3	-0.32	.75
Digits forward (max = 13)	7.4	1.8	8.1	1.8	-1.66	.10
Digits backward (max = 13)	5.8	1.9	7	1.8	-2.82	.006
Receptive vocabulary (max = 103)	51.1	11.1	55	11.3	-1.63	.10
TRT (max = 20)	2.7	3	4.6	3.7	-2.65	.01
Arabic reading-related measures						
Elision (max = 20)	13.7	3	15.9	2.8	-3.53	.001
Blending (max = 20)	8	3	9.1	2.6	-1.74	.09

RAN digit (seconds)	63.9	29.9	44.3	14.8	3.76	<.001
RAN letter (seconds)	62.9	16.5	61.6	20.8	0.33	.74
Orthographic memory (max = 30)	20.5	2.8	21.7	2.5	-2.06	.04
Orthographic matching (max = 25)	12.3	4.2	15.4	3.7	-3.57	.001
Pseudo-letter memory (max = 30)	23.5	2.7	24.5	2.6	-1.73	.09
Pseudo-letter matching (max = 25)	12.9	3.1	15.3	2.8	-3.59	.001
Morphological production (max = 29)	13.8	5.3	16.7	5.2	-2.58	.01
Arabic reading and spelling measures						
Word recognition (max = 50)	19.2	8.8	25.8	9.9	-3.27	.002
Nonword reading fluency (max = 35)	8.9	5.6	13.3	6.6	-3.32	.001

Single-word spelling (max = 38)	16	6.2	19.8	5.6	-2.96	.004

Note. All scores are raw scores. SD=Standard Deviation

To reduce the number of variables, composite measures were created for PA, verbal STM and RAN. The composite measure for PA comprised the scores from the elision and blending tasks, for verbal STM it comprised the scores from the nonword repetition and digits forward tasks, and for RAN it comprised the scores from the RAN letter and RAN digit tasks. The composites were derived by converting the raw scores to *z* scores and averaging the two. All three composite measures were used in the subsequent analyses. Pearson correlations between all measures were carried out for both the younger and older groups (see Tables 7 and 8, respectively).

Arabic Correlations for Younger Group of Bilinguals. Nonverbal ability was correlated to word recognition (p = .004), nonword reading fluency (p = .001), and spelling (p = .007). Verbal STM was correlated to word recognition (p = .012), nonword reading fluency (p < .001), and spelling (p = .004). Digits backward was correlated to word recognition (p < .001), nonword reading fluency (p < .001), and spelling (p = .007). Receptive vocabulary was correlated to word recognition, (p = .003) nonword reading fluency (p < .001), and spelling, (p = .005).

PA was correlated to word recognition (p < .001), nonword reading fluency (p < .001), and spelling, (p < .001). RAN was negatively correlated to nonword reading fluency (p < .001) and spelling, (p = .003). Orthographic memory was correlated to word recognition (p = .005), nonword reading fluency ($p = .003$), and spelling ($p = .042$). Orthographic matching was correlated to word recognition ($p = .047$), nonword reading fluency ($p = .009$), and spelling ($p = .032$). Pseudo-letter matching was correlated to word recognition ($p < .001$), nonword reading fluency ($p = .001$), and spelling ($p < .001$). Morphological production was correlated to word recognition ($p < .001$), nonword reading fluency ($r = .59$, $p < .001$), and spelling ($p < .001$).

Table 7

Correlations between Predictors and Reading Outcomes for the Younger Bilingual Group (N = 46) in Study 2a

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. NVA															
2. VC	.26														
3. VSTM	.22	.05													
4. DB	.45**	0	.08												
5. A-RV	.05	.23	.48**	.11											
6. A-TRT	.14	.10	.32*	.17	.36*										
7. APA	.30*	-.06	.53***	.43**	.25	.16									
8. A-RAN	-.09	.05	-.21	-.33*	-.003	-.10	-.44**								
9. A-OMem	.27	.03	.24	.57***	.09	.12	.46**	-.23							
10. A-OM	.31*	.56***	.36*	.08	.31*	.23	.13	-.02	.10						
11. A-PMem	.25	.30*	.11	.24	.06	-.02	.22	-.14	.36*	.38*					
12. A-PM	.38**	.43**	.30*	.33*	.34*	.18	.31*	-.10	.27	.67***	.47**				
13. A-MP	.17	.22	.50***	.28	.75***	.18	.54***	-.31*	.22	.26	.18	.33*			
14. A-WR	.42**	.24	.37*	.54***	.44**	.09	.55***	-.26	.41**	.30*	.27	.50***	.56***		
15. A-NWRF	.46**	.13	.52***	.52***	.50***	.25	.63***	-.32*	.42**	.38**	.11	.46**	.59***	.76***	
16. A-SWS	.40**	.08	.42**	.40**	.41**	-.01	.71***	-.38*	.30*	.32*	.23	.51***	.61***	.77***	.75***

Note. NVA=Nonverbal ability, VC=Visual concentration, VSM=Verbal short-term memory, DB=Digits backward, A-RV=Receptive vocabulary, A-TRT=Title recognition test, A-PA=Phonological awareness, A-NWRF=Nonword reading fluency, A-RAN=Rapid automatized naming, A-OMem=Orthographic memory, A-OM=Orthographic matching, A-PMem=Pseudo-letter memory, A-PM=Pseudo-letter matching, A-MP=Morphological production A-WR=Word recognition, A-SWS=Single word spelling. A=Arabic measures. *$*p < .05$, $**p < .01$, $***p < .001$*

Arabic Correlations for Older Group of Bilinguals. Visual concentration was correlated to word recognition (*p* = .002) and nonword reading fluency (*p* = .021) but not spelling. Verbal STM was correlated to word recognition (*p* < .001), nonword reading fluency (*p* = .003), and spelling, (*p* = .002). Digits backward was correlated to word recognition (*p* = .001) nonword reading fluency (*p* = .013), and spelling (*p* = .003). Receptive vocabulary was correlated to word recognition (*p* = .015).

PA was correlated to word recognition (*p* < .001), nonword reading fluency (*p* = .001), and spelling (*p* < .001). RAN was negatively correlated word recognition (*p* < .001) and to nonword reading fluency (*p* = .003). Orthographic memory was correlated to word recognition (*p* = .001) and nonword reading fluency (*p* = .003), and spelling, (*p* = .032). Orthographic memory fluency was correlated to word recognition (*p* = .015) and nonword reading fluency (*p* = .022) but not spelling. Pseudo-letter memory was correlated to spelling (*p* = .048). Morphological production was correlated to word recognition (*p* < .001), nonword reading fluency, (*p* = .014), and spelling, (*r* = .50, *p* = .001).

Table 8

Correlations between Predictors and Reading Outcomes for the Older Bilingual Group (N=40) in Study 2a

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. NVA														
2. VC	-.12													
3. VSTM	-.06	.30												
4. DB	.08	.04	.46**											
5. A-RV	-.24	.45**	.21	.01										
6. A-TRT	-.34*	.03	.004	.02	.22									
7. APA	.04	.37*	.68***	.41**	.24	-.05								
8. A-RAN	-.01	-.30	-.29	-.46**	-.12	.12	-.40*							
9. A-OM	.19	.44**	.41**	.31*	.06	-.02	.53***	-.54***						
10. A-OMF	.02	.32*	.10	.03	-.02	.18	.24	-.32*	.35*					
11. A-PM	.18	.30	.19	.14	.20	-.32*	.35*	-.31	.25	.09				
12. A-PMF	.34*	.30	.03	-.12	-.04	-.09	.18	-.32*	.42**	.69***	0.3			
13. A-MP	-.09	.56**	.38*	.25	.71***	.05	.53***	-.29	.22	.01	.33*	-.04		
14. A-WR	-.05	.47**	.61***	.49**	.38*	.14	.73***	-.58***	.50**	.38*	.14	.09	.63***	
15. A-NWRF	-.03	.37*	.46**	.39*	.20	.11	.51**	-.65***	.46**	.36*	.15	.20	.39*	.81***
16. A-SWS	.25	.26	.47**	.45**	.21	.03	.57***	-.31	.34*	-0.02	.32*	.02	.50**	.56***

Note. NVA=Nonverbal ability, VC=Visual concentration, VSM=Verbal short-term memory, DB=Digits backward, A-RV=Receptive vocabulary, A-TRT=Title recognition test, A-PA=Phonological awareness, A-NWRF=Nonword reading fluency, A-RAN=Rapid automatized naming, A-OMem=Orthographic memory, A-OM=Orthographic matching, A-

PMem=Pseudo-letter memory, A-PM=Pseudo-letter matching, A-MP=Morphological production A-WR=Word recognition, A-SWS=Single word spelling. A=Arabic measures.
*p < .05, **p <.01, ***p <.001

Summary of Arabic Correlations for Younger and Older Bilinguals. For the younger group, nearly all measures were correlated to word recognition, nonword reading fluency, and spelling, with the exception of visual concentration, TRT, and pseudo-letter memory. RAN was only correlated to nonword reading fluency and spelling but not to word recognition. The strongest correlations were observed for PA, word recognition, nonword reading fluency, and spelling.

For the older group, visual concentration, RAN, and orthographic matching correlated to word recognition and nonword reading fluency but not spelling. Receptive vocabulary was only correlated to word recognition. Pseudo-letter memory was only weakly correlated to spelling. Verbal STM, digits backward, PA, orthographic memory, and morphological production were all correlated to word recognition, nonword reading fluency, and spelling.

Predictors of Arabic Reading and Spelling for the Bilingual Group

Regression analyses were conducted to determine the significant predictors of Arabic word recognition and spelling. Arabic measures that failed to correlate to reading and spelling were not included in the regression analyses. As mentioned, the orthographic memory task was removed from all further analyses. It was observed during administration of this task that children would decode the letters, which allowed them to easily identify the target letters from distractors. Principal Component Analysis (PCA) confirmed this observation. The orthographic memory task loaded on the phonological tasks rather than the visual and orthographic tasks. Thus, orthographic memory scores were not included in any further analyses. Age, however, was correlated to orthographic matching, pseudo-letter matching, and spelling for the younger bilinguals, and RAN for the older group of bilinguals. Thus, age was included in the subsequent analyses.

Predictors of Arabic Word Recognition. A six-step hierarchical regression analysis was carried out to investigate predictors of word recognition in Arabic. Age was entered in the first step. In the second step, control variables were entered (nonverbal ability, visual concentration, verbal STM, digits backward, and receptive vocabulary). In steps 3 through 6, PA, RAN, orthographic matching, and morphological production were entered, respectively. Results are presented in Table 9.

Table 9

Summary of Hierarchical Regression Analyses with Arabic Word Recognition as the Dependent Variable for the Bilingual Group in Study 2a

		Younger group (N=46)			Older group (N=40)		
Step	Predictor	Ba	βb	ΔR2	βa	βb	ΔR2
1	Age	.16	-.07	.03	.26	.02	.07
2	Nonverbal ability	.21	.19	.51***	.004	-.04	.50***
	Visual concentration	.18	.23		.24	-.02	
	Verbal STM	.13	-.02		.34*	.16	
	Digits backward	.43**	.29		.32*	.14	
	Receptive vocabulary	.26	.30		.18	.04	

3	PA	.33*	.32	.06*	.46**	.26	.10**
4	RAN	-.04	-.04	.001	-.27*	-.19	.04*
5	Orthographic matching	-.03	-.04	0	.20	.24*	.03
6	Morphological production	-.03	-.03	0	.31	.31	.03

Note. β[a] =Standardized beta coefficient when first entered. β[b]= Standardized beta coefficient entered in last step.
$*p < .05$, $**p <.01$, $***p <.001$

For the younger group, digits backward was the only control measure that contributed to word recognition. The control variables explained 51% of the variance in word recognition. PA was the only reading-related measure that predicted word recognition. Overall, the model explained 60% of the variance in word recognition. For the older group, verbal STM and digits backward were the only two control variables that predicted word recognition. Control variables accounted for 50% of the variance in word recognition. PA explained 10% of the variance in word recognition. RAN was a marginally significant predictor of word recognition (p = .05). Overall, the model explained 76% of the variance in word recognition.

Predictors of Arabic Nonword Reading Fluency. A six-step hierarchical regression analysis was carried out to investigate predictors of nonword reading fluency in Arabic. Age was entered in the first step. In the second step, control variables were entered (nonverbal ability, visual concentration, verbal STM, digits backward, and receptive vocabulary). In steps 3 through 6, PA, RAN, orthographic matching, and morphological production were entered, respectively. Results are presented in Table 10.

Table 10

Summary of Hierarchical Regression Analyses with Arabic Nonword Reading Fluency as the Dependent Variable for the Bilingual Group

		Younger group (N=46)			Older group (N=40)		
Step	Predictor	βa	βb	ΔR2	βa	βb	ΔR2
1	Age	.30*	.02	.09*	.21	-.12	.04
2	Nonverbal ability	.20	.19	.52***	-.01	.001	.29*
	Visual concentration	.01	-.04		.24	.04	
	Verbal short-term memory	.31*	.09		.26	.19	
	Digits backward	.36**	.20		.26	.002	
	Receptive vocabulary	.26*	.35		.001	.02	
3	PA	.32*	.35*	.05*	.26	.08	.03
4	RAN	-.07	-.09	.003	-.57**	-.52**	.18**

5	Orthographic matching	.14	.13	.01	.16	.17	.02
6	Morphological production	-.08	-.08	.001	.10	.10	.003

β[a] =Standardized beta coefficient when first entered. β[b]= Standardized beta coefficient entered in last step.
**p < .05, **p <.01, ***p <.001*

For the younger group verbal STM, digits backward, and receptive vocabulary were the only control variables that predicted nonword reading fluency. The control measures explained 52% of the variance. PA accounted for 5% of unique variance in nonword reading fluency and continued to be a significant contributor even in the last step. Overall, the final model accounted for 68% of the variance. RAN was the one and only predictor of nonword reading fluency for the older group. It accounted for 18% of unique variance in nonword reading fluency. Overall, the final model explained 56% of the variance in nonword reading fluency.

Predictors of Arabic Spelling. A six-step hierarchical regression analysis was carried out to investigate predictors of spelling in Arabic. Age was entered in the first step. In the second step, control variables were entered (nonverbal ability, visual concentration, verbal STM, digits backward, and receptive vocabulary). In step 3, PA, and nonword reading fluency were entered. In steps 3 through 6, PA, RAN, orthographic matching, and morphological production were entered, respectively. Results are presented in Table 11.

Table 11

Summary of Hierarchical Regression Analyses with Arabic Spelling as the Dependent Variable for the Bilingual Group in Study 2a

		Younger group (N=46)			Older group (N=40)		
Step	Predictor	Ba	βb	ΔR2	βa	βb	ΔR2
1	Age	.43**	.16	.19**	-.004	-.16	0
2	Nonverbal ability	.19	.17	.28**	.32*	.28*	.43**
	Visual concentration	-.01	-.04		.17	.07	
	Verbal short-term memory	.23	-.10		.26	.09	
	Digit backward	.23	-.04		.30*	.16	
	Vocabulary	.18	.19		.20	.003	
3	Phonological awareness	.58***	.52**	.18***	.35	.28	.06
4	RAN	-.12	-.11	.01	-.10	-.11	.01
5	Orthographic matching	.07	.09	.002	-.16	-.12	.02
6	Morphological production	.10	.10	.002	.26	.26	.02

Note. β^a =Standardized beta coefficient when first entered. β^b= Standardized beta coefficient entered in last step.
**p < .05, **p <.01, ***p <.001*

Only PA predicted spelling for the younger group, even in the last step. PA accounted for 17.5% of the variance in spelling ability for the younger group. Overall, the model accounted for 66% of the variance. Nonverbal ability and digits backward were the only marginally significant (*p*

= .05) predictors of spelling ability for the older group. Overall, the model explained 53% of the variance.

English Descriptive Statistics and Preliminary Analysis for the Bilingual Group

Descriptive statistics for all English variables before data reduction are presented in Table 12.

Table 12

Descriptive Statistics for English Measures by Age Group for the Bilingual Group in Study 2a

	Younger group (N = 46)		Older group (N =40)		Differences between the groups	
Measures	Mean	SD	Mean	SD	*t*	*p*
Nonverbal control measures						
Nonverbal ability (max =60)	24.3	6.3	27.1	5.1	-2.3	.02
Visual concentration (max = 30)	12.7	2.5	14.1	3.1	-2.23	.03
English control measures						
Nonword repetition (max = 30)	12.9	3.5	13.2	3.6	-0.72	.48
Digit span forward (max = 28)	17.2	2.7	17.6	3.4	-0.34	.74

Vocabulary (max = 228)	129.2	22.2	146.5	19.5	-3.8	< .001
TRT (max = 20)	7.7	2.4	9.3	1.8	-3.45	.001
English reading-related measures						
Elision (max = 34)	29.2	4.9	30.5	3.1	1.43	.16
Blending (max = 33)	21.8	3.6	22.8	4.3	-1.12	.27
RAN digit (seconds)	15.1	3	13.6	3.1	2.2	.03
RAN letter (seconds)	18	3.8	16.3	3.1	2.36	.02
Orthographic matching (max = 25)	16.5	3.9	18.7	3.5	-2.8	.006
English reading and spelling measures						
Word recognition (max = 46)	28.1	4.1	31.6	4.3	-3.87	< .001
Nonword reading fluency (max = 108)	40.5	10.9	45.3	10.4	-2.09	.04

Single-word spelling (max = 50)	31.1	10.2	36.1	9.4	-2.36	.02

Note. All reported scores are raw scores. SD=Standard Deviation.

English Correlations for Younger Group of Bilinguals. As was done with the Arabic variables, composite scores were created for the PA measures (from elision and blending words), the verbal STM measures (from nonword repetition and digit span forward), and for RAN (RAN letter and RAN digit). This was done by converting the raw scores into *z* scores and averaging the two. All three composite measures were used in subsequent analyses. Pearson correlations between all measures for the younger bilingual group (N=46) are presented in Table 13.

Table 13

Correlations between English Predictors and English Reading and Spelling Outcomes in the Younger Bilingual Group in Study 2a

	1	2	3	4	5	6	7	8	9	10
1. NVA										
2. VC	0.26									
3. E-VSTM	0.26	0.11								
4. E-RV	.41**	-0.1	0.22							
5. E-TRT	.49**	.37*	0.1	.30*						
6. E-PA	.32*	0.13	.37*	0.05	0.12					
7. E-RAN	-0.08	-0.04	-.37*	-0.06	0.02	-0.26				
8. E-OMF	0.27	0.18	0.1	0.24	.32*	0.1	-0.24			
9. E-WR	.54***	0.14	.46**	.66***	.42**	0.21	-.31*	.58***		
10. E-NWRF	.44**	0.11	.53***	.46**	.41**	0.23	-.60***	.52***	.78***	
11. E-SWS	.53***	0.14	.49**	.33*	.37*	0.22	-0.29	.47**	.67***	.69***

NVA=Nonverbal ability, VC=Visual concentration, E-RV=Receptive vocabulary, E-TRT=Title Recognition Test, E-PA=Phonological awareness, E-VSM=Verbal short-term memory, E-RAN=Rapid automatized naming, E-OMF=Orthographic memory fluency, fluency, E-WR=Word recognition, E-NWRF=Nonword reading fluency, E-SWS=Single word spelling. E=English measures.
**p < .05, **p <.01, ***p <.001*

English Correlations for Older Group of Bilinguals. Nonverbal ability was correlated to word recognition (*p* < .001), nonword reading fluency (*p* = .002), and spelling (*p* < .001). Verbal STM was correlated to word recognition (*p* = .002), nonword reading fluency (*p* < .001), and spelling (*p* = .001). Receptive vocabulary was correlated to word recognition (*p* < .001), nonword reading fluency (*p* = .001), and spelling (*p* = .025). The E-TRT was correlated to word recognition (*p* = .004), nonword reading fluency (*p* = .005), and spelling (*p* = .012).

RAN was negatively correlated to word recognition (*p* = .037), and nonword reading fluency (*p* < .001). Orthographic matching fluency was correlated to word recognition (*p* < .001), nonword reading fluency (*p* < .001), and spelling (*p* = .001).

Pearson correlations between all measures for the older group (N=40) are presented in Table 14.

Table 14

Correlations between English Predictors and English Reading and Spelling Outcomes in the Older Bilingual Group in 2a

	1	2	3	4	5	6	7	8	9	10
1. NVA										
2. VC	-0.12									
3. E-VSTM	0.23	0.003								
4. E-RV	.32*	0.02	.43**							
5. E-TRT	0.01	-0.25	-0.24	-0.15						
6. E-PA	.38*	0.15	.45**	.39*	-0.23					
7. E-RAN	0.07	0.003	-.34*	-0.18	-0.29	-0.22				
8. E-OM	-0.17	.35*	0.3	0.16	0.14	0.05	-0.29			
9. E-WR	0.19	0.04	.37*	.45**	0.04	.45**	-.37*	0.25		
10. E-NWRF	0.02	-0.02	0.25	0.12	0.23	0.29	-.63***	0.27	.65***	
11. E-SWS	0.1	0.06	0.24	0.11	0.22	.41**	-.52***	.43**	.61***	.75***

NVA=Nonverbal ability, VC=Visual concentration, E-RV=Receptive vocabulary, E-TRT=Title Recognition Test, E-PA=Phonological awareness, E-VSM=Verbal short-term memory, E-RAN=Rapid automatized naming, E-OMF=Orthographic memory fluency, fluency, E-WR=Word recognition, E-NWRF=Nonword reading fluency, E-SWS=Single word spelling. E=English measures.

**p < .05, **p <.01, ***p <.001*

Summary of English Correlations for Younger and Older Group of Bilinguals. Verbal STM was correlated to word recognition (*p* = .02) as was receptive vocabulary (*p* = .004). PA was correlated to word recognition (*p* = .004) and spelling (*p* < .01). RAN was negatively correlated to word recognition (*p* = .02), nonword reading fluency (*p* < .001), and spelling (*p* = .001). Orthographic matching fluency was correlated to spelling (*p* = .01).

Predictors of English Reading and Spelling for the Bilingual Group

Regression analyses were conducted to examine predictors of English word recognition, nonword reading fluency, and spelling. Visual concentration did not correlate to any of the outcome measures for both groups and was not included in the subsequent analyses. Age, however, was correlated to vocabulary, orthographic matching, and nonword reading fluency for the younger bilinguals, and visual concentration for the older group of bilinguals. Thus, age was included in the analyses.

Predictors of English Word Recognition for the Bilinguals. A five-step hierarchical regression analysis was carried out to investigate predictors of word recognition in English. Age was entered in the first step. In the second step, control variables were entered (nonverbal ability, verbal STM, receptive vocabulary, and title recognition test). In step 3, PA and nonword reading fluency were entered. In steps 3 through 5, PA, RAN, and orthographic matching were entered respectively. Results are presented in Table 15.

Table 15

Summary of Hierarchical Regression Analyses Predicting English Word Recognition in the Bilingual Group in Study 2a

		Younger group (N=46)			Older group (N=40)		
Step	Predictor	Ba	βb	ΔR2	βa	βb	ΔR2
1	Age	.40**	.05	.16**	.042	-.01	.002
2	Nonverbal ability	.16	.14	.48***	-.003	-.02	.28*
	Verbal short-term memory	.32**	.29**		.25	.14	
	Vocabulary	.41***	.41***		.39*	.33	
	TRT	.13	.07		.14	.09	
3	Phonological awareness	.003	.002	0	.32	.21	.07
4	RAN	.12	.03	.01	-.26	-.244	.04
5	Orthographic matching	.37***	.37***	.10***	.04	.04	.001

Note. β^a =Standardized beta coefficient when first entered. β^b= Standardized beta coefficient entered in last step.
**p < .05, **p <.01, ***p <.001*

Of all the control measures, only verbal STM and receptive vocabulary predicted word recognition for the younger group. The control measure accounted for approximately 48% of the variance in word recognition.

Orthographic matching predicted word recognition and accounted for an additional 10% of the variance when entered in the final step. Overall, the model accounted for 75% of the variance in word recognition. For the older group, control measures accounted for 28% of the variance in word recognition, and the majority of this contribution is accounted for by receptive vocabulary, which was the only variable that reached significance. No other variable predicted word recognition for the older group. Overall, the model accounted for 39% of the variance in word recognition.

Predictors of English Nonword Reading Fluency for the Bilingual Group. A five-step hierarchical regression analyses was carried out to investigate predictors of nonword reading fluency in English. Age was entered in the first step. In the second step, control variables were entered (nonverbal ability, verbal STM, receptive vocabulary, and title recognition test). In step 3, PA and nonword reading fluency were entered. In steps 3 through 5, PA, RAN, and orthographic matching were entered respectively. Results are presented in Table 16.

Table 16

Summary of Hierarchical Regression Analyses Predicting English Nonword Fluency in the Bilingual Group in Study 2a

		Younger group (N=46)			Older group (N=40)		
Step	Predictor	Ba	βb	ΔR2	βa	βb	ΔR2
1	Age	.35*	.10	.12*	.23	.09	.05
2	Nonverbal ability	.10	.08	.39***	-.12	-.13	.17
	Verbal short-term memory	.45***	.43***		.33	.25	
	Receptive vocabulary	.16	.16		.11	.06	
	TRT	.21	.16		.29	.22	
3	Phonological awareness	-.004	-.01	0	.31	.13	.07
4	RAN	.09	.02	.01	-.37	-.38	.08
5	Orthographic matching	.33*	.33*	.08*	-.03	-.03	.001

Note. β^a =Standardized beta coefficient when first entered. β^b= Standardized beta coefficient entered in last step.
**p < .05, **p <.01, ***p <.001*

For the younger group, control measures accounted for approximately 43% of the unique variance in nonword reading fluency. Of the control measures, verbal STM was the sole significant contributor to nonword reading fluency. Orthographic matching predicted nonword reading fluency and explained 8% of the variance. Overall, the model accounted for 60% of the variability in nonword reading fluency. None of the measures predicted nonword reading fluency for the older group. RAN was the only measure that came close to reaching significance. Overall, the model accounted for 36% of the variability in nonword reading fluency.

Predictors of English Spelling for the Bilingual Group. A five-step hierarchical regression analysis was carried out to investigate predictors of spelling in English. Age was entered in the first step. In the second step, control variables were entered (nonverbal ability, verbal STM, receptive vocabulary, and title recognition test). In steps 3 through 5, PA, RAN, and orthographic matching were entered respectively. Results are presented in Table 17.

Table 17

Summary of Hierarchical Regression Analyses Predicting English Spelling in the Bilingual Group in Study 2a

		Younger group (N=46)			Older group (N=40)		
Step	Predictor	Ba	βb	ΔR2	βa	βb	ΔR2
1	Age	.12	-.16	.01	.22	-.004	.05
2	Nonverbal ability	.35*	.33	.42***	-.02	.03	.15
	Verbal short-term memory	.37**	.37**		.32	.07	
	Vocabulary	.06	.06		.07	-.07	
	TRT	.15	.10		.27	.18	
3	Phonological awareness	-.07	-.05	.004	.46*	.32	.14*

4	RAN	.15	.08	.02	-.38*	-.30	.08*
5	Orthographic matching	.36*	.36*	.09*	.27	.27	.05

Note. β[a] =Standardized beta coefficient when first entered. β[b]= Standardized beta coefficient entered in last step.
**p < .05, **p <.01, ***p <.001*

For the younger group, control measures accounted for approximately 39% of the variance in spelling. Of the control measures, nonverbal ability and verbal STM were the only significant contributors of spelling. Orthographic matching was the only reading-related measure that predicted spelling and accounted for approximately 8% of the variance. Overall, the model accounted for 54% of the variability in spelling.

For the older group, PA and RAN were the only predictors of spelling. PA accounted for 14% of the variance in spelling. RAN accounted for approximately 8% of unique variance in spelling. Overall, the model accounted for 47% of the variance in spelling.

Summary of Within-language Predictors of Reading and Spelling for the Bilinguals

For the younger bilingual group, digits backward and PA predicted Arabic word recognition. A similar pattern emerged for the older bilingual group: only verbal STM, digits backward, and PA were significant predictors. Arabic nonword reading fluency was predicted by verbal STM, digits backward, and vocabulary for the younger group. Only RAN predicted nonword reading fluency for the older group. PA was the only predictor of Arabic spelling for the younger group.

For the younger bilingual group, only verbal STM and orthographic matching predicted English word recognition. Only receptive vocabulary predicted English word recognition for the older bilingual group. English nonword reading fluency was predicted by verbal STM and orthographic matching for the younger bilingual group. Just as for Arabic, the only measure that came close to reaching significance for the older bilingual group was RAN. Similarly, verbal STM and orthographic matching predicted English spelling for the younger bilingual group, as it did for nonword reading fluency. For the older bilingual group, RAN and PA predicted English spelling.

Study 2b Method

This study explored predictors of Arabic reading and spelling in Arabic. In Study 2b a between-participant design was also employed to explore differences between monolinguals and bilinguals from study 2a. Two types of analyses were used: correlation and multiple regressions.

Participants

One hundred sixteen monolingual students took part in study 2b (58 female; mean age = 8.10, SD ± .45). The monolingual participants were native Arabic speakers attending a private school located in a middle-class neighborhood in Cairo. Unlike international schools, private schools offer a national curriculum that is delivered almost entirely in Arabic; the students receive one 40-minute class of ESL instruction per day. The selection of monolinguals from a private school rather than a public school was deemed appropriate to control for the homogeneity of the sample in SES and other environmental factors. Participants were randomly selected from four third-grade classrooms. Potential participants were excluded from the study if they had any sensory or cognitive impairment.

Materials

All Arabic measures discussed in Study 2a were administered to the monolingual sample, with the exception of two orthographic measures: orthographic memory and pseudo-letter memory. The orthographic memory measure was excluded because it was found to tap phonological skills rather than orthographic skills for the bilingual sample. The pseudo-letter memory assessment was not administered to the monolinguals because it did not correlate with reading and spelling measures in the bilingual sample.

Procedure

Data collection for the monolingual students began in October 2019. The monolinguals were only assessed on Arabic measures, excluding the orthographic memory task and pseudo-letter memory task. Monolingual participants were seen individually in two one-on-one sessions and one group session. In the first session, phonological processing measures, orthographic measures, and reading measures were administered. During the second session, the participants were assessed on measures of receptive vocabulary, morphological awareness, spelling, visual concentration, and print exposure. Finally, participants were seen in a group session for administration of Raven's Matrices.

Study 2b Results

The following section will present the results for the monolingual sample and the results of comparisons between subsamples of monolinguals and subsamples of bilinguals from Study 2a. The bilinguals and monolinguals were matched first for word recognition, then nonword reading fluency, then spelling. For each comparison, three analyses were conducted: predictors of word recognition, predictors of nonword reading fluency, and predictors of spelling. The rationale for the matching on each of the reading and spelling measures allowed for the investigation of the underlying cognitive and linguistic abilities that predict reading and spelling for bilinguals and monolinguals that have the same reading and spelling level. The results for the monolingual Arabic speaking group are presented first: descriptive statistics for the measures, followed by correlations between measures and hierarchical regressions determining predictors of reading and spelling. Next, this section presents comparisons of predictors of reading and spelling between subsamples of monolinguals and bilinguals who were matched on word recognition, nonword reading fluency, and spelling. For each comparison, descriptive statistics are presented first, followed by correlations and regressions. The chapter concludes by presenting a summary of the findings.

Exploratory Data Analysis

Prior to analysis, EDA was carried out to determine if the variables met the assumption of normality. Normality was assessed via the Shapiro Wilk statistic, significance values above .05 considered normal, skewness and kurtosis ratio values with value within +/-3 considered normal, and finally through an examination of histograms. Inspection of the monolingual results indicated that most variables were normally distributed, and thus no adjustments were required.

Predictors of Reading and Spelling for the Monolingual Group

Analyses were conducted to examine the cognitive and linguistic abilities that were associated with reading and spelling in the monolingual children. Descriptive statistics and correlations between the variables are presented first, followed by the results of regression analyses.

Table 18

Descriptive Statistics for the Monolingual Group in Study 2b

Measures	Max. possible score	Mean	SD	Minimum	Maximum
Nonverbal control measures					
Nonverbal ability	60	17.1	7.3	5	33
Visual Concentration		11.4	2.2	6	16
Control measures					
Nonword repetition	20	9.5	3.6	2	19
Digit span forward	13	7.4	1.7	4	11
Digits backward	13	4.7	1.5	0	9
Vocabulary	103	55.4	10.9	29	79
TRT	20	4.1	3	-3	13
Reading-related measures					
Elision	20	12.7	3.9	2	19
Blending	20	5.6	2	0	12
RAN digit (seconds)		18.6	3.3	11	32
RAN letter (seconds)		24.1	5.8	15	46
Orthographic matching	25	11.7	3.4	4	20

Pseudo-letter matching	25	10.5	3	1	16
Morphological production	29	15	4.7	5	25
Reading and spelling measures					
Word recognition	50	23.2	11.3	1	46
Nonword reading fluency	35	14.6	7.3	0	31
Spelling	38	20.6	6.9	5	32

Note. SD = Standard deviation.

Table 19

Correlations between Measures for the Monolingual Group (N = 116) in Study 2b

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. NVA													
2. VC	.41***												
3. VSTM	.24**	0.07											
4. DB	0.17	0.06	.19*										
5. RV	.26**	0.12	.36***	.35***									
6. TRT	-0.07	0.03	0.07	0.05	0.1								
7. PA	.38***	0.14	.40***	.35***	.40***	0.04							
8.RAN	-.23*	-.23*	-0.1	-.19*	-.21*	0.01	-.38***						
9. OM	.36***	.34***	0.05	.24**	.19*	-0.04	.25**	-.33***					
10. PM	.32***	.36***	0.05	.25**	0.11	-0.14	.25**	-.33***	.53***				
11. MP	.38***	.23*	.44***	.22*	.62***	0.023	.39***	-.25**	.21*	.29**			
12. WR	.37***	.23*	.33***	.36***	.37***	-0.02	.64***	-.44***	.34***	.46***	.48***		
13. NWRF	.31**	.26**	.27**	.41***	.35***	-0.01	.62***	-.46***	.33***	.43***	.44***	.83***	
14. SWS	.33***	.20*	.36***	.34***	.36***	-0.02	.59***	-.46***	.33***	.41***	.45***	.79***	.75***

Note: NVA=Nonverbal ability, VC=Visual concentration, VSTM=Verbal short-term memory, DB=Digit backward, RV=Receptive vocabulary, TRT=Title recognition test, PA=phonological awareness, RAN=Rapid automatized naming, OM=orthographic matching, PM=Pseudo-letter matching, MP=Morphological production, WR=Word recognition, NWRF=Nonword reading fluency, SWS=Single word spelling. **p < .05, **p <.01, ***p <.001*

Nonverbal ability was correlated with word recognition ($p < .001$), nonword reading fluency ($p = .001$), and spelling ($p < .001$). Visual concentration was correlated to word recognition ($p = .013$), nonword reading fluency ($p = .005$), and spelling ($p = .033$). Verbal STM was correlated to word recognition ($p < .001$), nonword reading fluency ($p = .004$), and spelling ($p < .001$). Digits backward was correlated to word recognition ($p < .001$), nonword reading fluency, ($p < .001$), and spelling ($p < .001$). Receptive vocabulary was correlated to word recognition ($p < .001$), nonword reading fluency ($p < .001$), and spelling ($p < .001$). Scores on the title recognition test did not correlate with any measure and were not included in further analyses.

PA was correlated to word recognition ($p < .001$), nonword reading fluency ($p < .001$), and spelling ($p < .001$). RAN was negatively correlated to word recognition ($p < .001$), nonword reading fluency, ($p < .001$), and spelling, ($p < .001$). Orthographic matching was correlated to word recognition ($p < .001$), nonword reading fluency ($p < .001$), and spelling ($p < .001$). Pseudo-letter matching was correlated to word recognition ($p < .001$), nonword reading fluency ($p < .001$), and spelling ($p = .04$). Morphological production was correlated to word recognition ($p < .001$), nonword reading fluency ($p < .001$), and spelling ($p < .001$).

Three separate seven-step hierarchical regression analyses were performed for each outcome measure (word recognition, nonword reading fluency, and spelling) and are presented in Table 20. Because age was correlated to orthographic matching, age was entered in the first step. In the second step, nonverbal control variables were entered (nonverbal ability and visual concentration). In the third step, control measures were entered:

verbal STM, digits backward, and receptive vocabulary. In steps 4 through 7, PA, RAN, orthographic measures (orthographic matching and pseudo-letter matching), and morphological production were entered respectively.

Table 20
Summary of Hierarchical Regressions for the Reading and Spelling Outcome Measures for the Monolingual Group in Study 2b

		Word recognition			Nonword reading fluency			Spelling		
Step	Predictor	β^a	β^b	ΔR2	β^a	β^b	ΔR2	β^a	β^b	ΔR2
1	Age	0.17	0.06	0.03	0.15	0.04	0.02	0.12	0.002	0.02
2	Nonverbal ability	.31**	0.01	.12**	.23*	-0.06	.10**	.29**	0.01	.11**
	Visual concentration	0.09	0.002		0.15	0.06		0.07	-0.02	
3	Verbal short-term memory	.26*	0.01	.12**	0.21	-0.06	.10**	.43**	0.21	.15***
	Digits backward	-0.1	0.05		-0.1	0.06		-0.25	-0.11	
	Vocabulary	.25**	0.03		.26**	0.04		.21*	0.03	
4	PA	.52***	.43***	.19***	.55***	.45***	.21***	.44***	.34***	.14***
5	RAN	-.19*	-0.13	.03*	-.22**	-.17*	.04**	-.25**	-.20*	.05**
6	Orthographic matching	-0.02	-0.01	.06**	-0.01	0.004	.04*	0.02	0.03	.04*
	Pseudo-letter matching	.28**	.25**		.24**	.21*		.22*	.20*	
7	Morphological production	0.16	0.16	0.01	0.14	0.14	0.01	0.12	0.12	0.01

Note. β^a =Standardized beta coefficient when first entered. β^b= Standardized beta coefficient entered in last step.
$^*p < .05$, $^{**}p <.01$, $^{***}p <.001$

Summary of Predictors of Reading and Spelling for the Monolingual Group

Nonverbal ability consistently predicted word recognition, nonword reading fluency, and spelling, explaining between 10 and 12% of the variance in each of the outcome measures. Similarly, receptive vocabulary consistently predicted word recognition, nonword reading fluency, and spelling, whereas verbal STM predicted word recognition and spelling but not nonword reading fluency. PA explained 19% of the variance in word recognition scores, 21% in nonword reading fluency, and14% in spelling. RAN was also a consistent predictor of word recognition, nonword reading fluency, and spelling. It explained 3% of variance in word recognition, 4% in nonword reading fluency, and 5% in spelling. Pseudo-letter matching predicted word recognition, nonword reading fluency and spelling. It explained between 6% and 4% of the variance in the three outcome measures.

Comparing Monolingual and Bilingual Groups

One aim of this study was to attempt to separate universal from script-dependent predictors of single word reading. This was done in two ways: first, by examining predictors of single word reading and spelling in Arabic and in English for the bilingual children discussed in Study 2a, and second, by comparing results for this group to those of the Arabic-speaking monolingual group in Study 2b. The following section will present the results for the monolingual sample and the results of comparisons between subsamples of monolinguals and subsamples of bilinguals from Study 2a. The bilinguals and monolinguals were matched first for word recognition then nonword reading fluency, and then spelling. For each comparison, three analyses were conducted: predictors of word recognition, predictors of nonword reading fluency, and predictors of spelling. The rationale for the matching on each of the reading and spelling measures allowed for the investigation of the underlying cognitive and linguistic abilities that predict reading and spelling for bilinguals and monolinguals that have the same reading and spelling level.

Comparing Monolingual and Bilingual Groups Matched on Word Recognition Ability

Analyses were conducted to determine whether the results regarding predictors of reading and spelling in the bilingual children differed from those for monolingual children matched on word recognition ability. A subsample of bilingual children was matched on word recognition with a subsample of third-grade monolingual children. The selection process involved matching all bilinguals and monolinguals on scores from the word recognition test. When there were more monolinguals than bilinguals (e.g., 10 bilinguals with a score of 8 and 14 monolinguals with a score of 8), the additional monolinguals were randomly deleted in order to have an equal number in both groups (e.g., 10 bilinguals and 10 monolinguals with a score of 8 on word recognition). This was done for all possible word recognition scores and resulted in 51 bilinguals with mean age of 9.6 (SD = .86) and 51 monolinguals with mean age of 8.1 (SD = .45) matched on word recognition.

Descriptive Statistics and Preliminary Analysis of Bilingual and Monolingual Groups Matched on Word Recognition. Descriptive statistics for all measures in Study 2b are provided in Table 21.

Table 21

Descriptive Statistics for Arabic Measures for Monolingual and Bilingual Groups Matched on Word Recognition Scores in Study 2b

	Monolinguals (N=51)				Bilinguals (N=51)			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Nonverbal control Measures								
Nonverbal ability	18.3	7.6	5	33	26.5	5.2	14	36
Visual concentration	11.7	2.2	7	16	13.7	3.1	6	21
Control measures								
Nonword repetition	9.7	3.3	5	17	12.2	3.3	5	20
Digit span forward	7	1.7	4	11	8.2	1.7	4	11
Digits backward	4.8	1.6	0	9	6.8	1.9	3	11
Vocabulary	56.5	10.6	29	79	54.3	12	31	83
TRT	3.8	3.2	-3	12	3.9	3.6	-3	14
Reading-related measures								
Elision	13.7	3.3	6	19	15.4	2.6	9	20
Blending	6	2.1	2	12	8.9	2.9	2	19
RAN digit	18.2	3.1	13	26	55.9	21.5	25	123
RAN letter	22.8	5.1	15	37	58.1	13.7	35	93
Orthographic matching	12	3.1	4	18	14.5	4.2	2	22

Pseudo-letter matching	10.8	2.8	4	16	24.4	2.7	19	30
Morphological production	15.7	3.8	9	25	16.2	5.2	8	33
Reading and spelling outcome measures								
Word recognition	25.1	8.7	12	47	25.1	8.7	12	47
Nonword reading fluency	16.2	6	5	30	12.9	6.4	3	31
Single word spelling	22	5.1	8	31	19.5	5.8	10	32

Note. SD = Standard deviation

In order to examine group differences on the measures, independent samples *t*-tests were used. There was a significant difference t (100) = -6.4, $p < .001$ for nonverbal ability and visual concentration, t (100) = -3.4, p = .03, in favor of the bilingual group. There was also a significant difference for both the verbal STM measures (nonword repetition t (100) = -3.6, $p < .001$ and digits forward t (100) = -3.7, $p < .001$) in favor of the bilinguals. There was a significant difference t (100) = -5.9, $p < .001$ on the digits backward task between the groups, also in favor of the bilinguals. Although the monolinguals outperformed the bilinguals on the receptive vocabulary measure, this difference did not reach significance. There was no group difference for the title recognition test. There was a significant difference for both PA measures (elision *t* (100) = -2.9, *p* = .005 and blending words *t* (100) = -5.9, *p* < .001), in favor of the bilinguals. There was a significant difference for the RAN measures (RAN digit *t* (100) = -2.9, *p* = .005 and RAN letter *t* (100) = -5.9, *p* < .001) in favor of the monolinguals. There was a significant difference in favor of the bilinguals on orthographic matching fluency, *t* (100) = -3.3, *p* = .001 and pseudo-letter matching, *t* (100) = -25.1, *p* < .001. There was no significant difference in performance on the morphological production task between the groups. Finally, there was a significant difference on nonword reading fluency, *t* (100) = 2.7, *p* = .008, and spelling, *t* (100) = 2.3, *p* =.02, in favor of the monolinguals.

In order to reduce the number of variables, as was done in Study 2a, composite scores were created for PA, comprising the scores for the elision and blending tasks, for verbal STM comprising scores for the nonword repetition and digits forward tasks, and for RAN, comprising scores for the RAN letter and RAN digit tasks. Composite scores were obtained by converting the raw scores into *z* scores and averaging the two. The three composite measures were used in all subsequent analyses.

Table 22

Correlations between Variables for the Monolingual Group (N = 51) Matched on Word Recognition in Study 2b

	1	2	3	4	5	6	7	8	9	10	11	12	
1. NVA													
2. VC	.44**												
3. VSTM	.36**	.22											
4. DB	.23	.004	.15										
5. A-RV	.31*	.19	.38**	.42**									
6. A-TRT	-.07	-.01	-.05	.17	.05								
7. A-PA	.37**	.17	.33*	.29*	.36*	-.05							
8. A-RAN	-.003	.08	.04	-.14	-.01	-.002	-.15						
9. A-OM	.35*	.26	.14	.28*	.10	-.03	.30*	-.47***					
10. A-PM	.27	.25	.01	.19	-.09	-.15	.15	-.17	.44**				
11. A-MP	.48***	.29*	.56***	.34*	.61***	.13	.45**	.05	.21	.09			
12. A-WR	.33*	.21	.07	.47***	.25	.09	.58***	-.25	.36**	.30*	.33*		
13. A-NWRF	.16	.28*	.02	.49***	.27	.08	.58***	-.24	.32*	.17	.23	.76***	
14. A-SWS	.27	.18	.22	.37**	.17	.10	.38**	-.14	.32*	.20	.41**	.60***	.

Note. NVA=Nonverbal ability, VC=Visual concentration, A-RV=Receptive vocabulary, DB=Digits backward, A-TRT=Title Recognition Test, A-MP=Morphological production, A-PA=Phonological awareness, A-VSM=Verbal short-term memory, A-RAN=Rapid automatized naming, A-OMF=Orthographic matching, A-PMF=Pseudo-letter matching, A-WR=Word recognition, A-NWRF=Nonword reading fluency, A-SWS=Single word spelling. A=Arabic measures. *$p < .05$, **$p < .01$, ***$p < .001$

For the monolinguals, nonverbal ability was correlated to word recognition (*r* = .33, *p* = .02). Visual concentration was correlated to nonword reading fluency (*r* = .28, *p* = .05). Digits backward was correlated to word recognition (*r* = .47, *p* < .001), nonword reading fluency (*r* = .49, *p* < .001), and spelling, (*r* = .37, *p* = .007).

PA was correlated to word recognition (*r* = .58, *p* < .001), nonword reading fluency (*r* = .58, *p* < .001), and spelling (*r* = .38, *p* = .006). Orthographic matching fluency was correlated to word recognition (*r* = .36, *p* = .01), nonword reading fluency (*r* = .32, *p* = .02), and spelling (*r* = .32, *p* = .02). Pseudo-letter matching fluency was only correlated to word recognition (*r* = .30, *p* = .03). Finally, morphological production was correlated to word recognition (*r* = .33, *p* = .02), and spelling (*r* = .41, *p* = .003).

Table 23

Correlations between Variables for the Bilingual Group (N = 51) Matched on Word Recognition in Study 2b

	1	2	3	4	5	6	7	8	9	10	11	12	
1. NVA													
2. VC	-.03												
3. VSTM	.04	.29*											
4. DB	.23	-.07	.18										
5. A-RV	-.16	.37**	.57***	-.07									
6. A-TRT	-.25	.08	.18	.01	.29*								
7. A-PA	.20	.19	.50***	.50***	.23	.04							
8. A-RAN	-.10	-.15	-.21	-.46**	-.2	-.05	-.55***						
9. A-OM	.10	.43**	.26	-.14	.20	.27	.13	-.13					
10. A-PM	.33*	.32*	.12	.14	.15	-.11	.32*	-.18	.31*				
11. A-MP	.02	.44**	.60***	.17	.71***	.11	.61***	-.42**	.19	.28*			
12. A-WR	.15	.38**	.52***	.38**	.43**	.13	.66***	-.55***	.31*	.23	.63***		
13. A-NWRF	.15	.38**	.45**	.35*	.38**	.23	.55***	-.60***	.41**	.19	.51***	.80***	
14. A-SWS	.29*	.24	.31*	.39**	.33*	.01	.58***	-.48***	.15	.23	.62***	.64***	.4

Note. NVA=Nonverbal ability, VC=Visual concentration, A-RV=Receptive vocabulary, DB=Digits backward, A-TRT=Title Recognition Test, A-MP=Morphological production, A-PA=Phonological awareness, A-VSM=Verbal short-term memory, A-RAN=Rapid automatized naming, A-OMF=Orthographic matching, A-PMF=Pseudo-letter matching, A-WR=Word recognition, A-NWRF=Nonword reading fluency, A-SWS=Single word spelling. A=Arabic measures.
**p < .05, **p <.01, ***p <.001*

Nonverbal ability was correlated to spelling ($r = .28$, $p = .04$). Visual concentration was correlated to word recognition ($r = .38$, $p = .007$) and nonword reading fluency ($r = .38$, $p = .007$). Verbal STM was correlated to word recognition ($r = .52$, $p < .001$), nonword reading fluency, ($r = .45$, $p = .001$), and spelling, ($r = .31$, $p = .03$). Digits backward was correlated to word recognition ($r = .38$, $p = .006$), nonword reading fluency, ($r = .35$, $p = .01$), and spelling ($r = .39$, $p = .005$). Receptive vocabulary was correlated to word recognition ($r = .43$, $p = .002$), nonword reading fluency ($r = .38$, $p = .006$), and spelling ($r = .33$, $p = .02$). Arabic TRT was correlated to nonword reading fluency ($r = .22$, $p = .049$).

PA was correlated to word recognition ($r = .66$, $p < .001$), nonword reading fluency ($r = .55$, $p < .001$), and spelling ($r = .58$, $p < .001$). RAN was negatively correlated to word recognition ($r = -.55$, $p < .001$), nonword reading fluency ($r = -.60$, $p < .001$), and spelling ($r = -.48$, $p < .001$). Orthographic matching fluency was correlated to word recognition ($r = .31$, $p = .03$) and nonword reading fluency ($r = .41$, $p = .003$) but not to spelling. Finally, morphological production was correlated to word recognition ($r = .63$, $p < .001$), nonword reading fluency ($r = .51$, $p < .001$), and spelling, ($r = .62$, $p < .001$).

Predictors of Word Recognition for the Monolingual and Bilingual Groups Matched on Word Recognition. Based on these correlations, hierarchical regression analyses were conducted to determine predictors of word recognition among the monolingual and bilingual groups. Scores in the TRT did not correlate to any of the outcome variables and was excluded from the analyses. Similarly, pseudo-letter matching was only weakly correlated to word recognition, and only in the monolingual sample, and thus was excluded. Because age was correlated to TRT, PA, RAN, orthographic matching, and nonword reading fluency for the bilinguals, age was entered in the first step, followed by the control measures in step 2 (nonverbal ability, visual concentration, verbal STM, and receptive vocabulary). In steps 3 through 6, PA, RAN, orthographic matching, and morphological production were entered respectively. Results are presented in Table 24.

Table 24

Summary of Hierarchical Regression Analyses Predicting Arabic Word Recognition for Monolingual and Bilingual Groups Matched on Word Recognition in Study 2b

		Monolingual (N=51)			Bilingual (N=51)		
Step	Predictor	βa	βb	ΔR2	βa	Bb	ΔR2
1	Age	.26	.15	.07	.34*	-.03	.12*
2	Nonverbal ability	.13	.01	.26*	.10	.04	.37***
	Visual concentration	.18	.17		.22	.11	
	Verbal STM	-.13	-.23		.25	.10	
	Digits backward	.38*	.30*		.31*	.14	
	Receptive vocabulary	.09	-.04		.22	.12	
3	PA	.50***	.46***	.19***	.40**	.28	.09**
4	RAN	-.11	-.12	.01	-.22	-.20	.03
5	Orthographic matching	.01	.002	0	.14	.15	.01
6	Morphological production	.12	.12	.01	.12	.12	.004

Note. β[a] =Standardized beta coefficient when first entered. β[b]= Standardized beta coefficient entered in last step.
**p < .05, **p <.01, ***p <.001*

Of all the control measures, only digits backward predicted word recognition for both monolinguals and bilinguals. The control measures explained approximately 26% of the variance in word recognition for the monolingual group and 37% for the bilingual group. PA was the strongest predictor of word recognition for the monolingual and bilingual groups alike. It explained 19% of the variance in word recognition for the monolingual group and 8.5% of the variance in word recognition for the bilinguals. Overall, the model explained 53% of the variance in word recognition for the monolinguals and 62% of the variance in word recognition for the bilinguals.

Comparing Monolinguals to Bilinguals Matched on Nonword Reading Fluency Ability

A subsample of bilinguals was matched on nonword reading fluency with a subsample of third-grade monolinguals. The selection process involved matching all bilinguals and monolinguals on scores in the nonword reading fluency test. The same matching procedure for word recognition was employed for matching on nonword reading fluency. This was done for all possible scores and resulted in 63 bilinguals (mean age = 9.3, SD = .94) and 63 monolinguals (mean age = 8.2, SD = .48) matched on nonword reading fluency. The following section presents preliminary analysis followed by regression analysis.

Descriptive Statistics and Preliminary Analysis of Bilingual and Monolingual Groups Matched on Nonword Reading Fluency. Descriptive statistics for all measures administered in Study 2b are provided in Table 25.

Table 25

Descriptive Statistics for Arabic Measures for Monolingual and Bilingual Groups Matched on Nonword Reading Fluency Scores in Study 2b

	Monolinguals (N=63)				Bilinguals (N=63)			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Nonverbal control measures								
Nonverbal ability	17.3	7.3	6	33	25.7	5.9	10	36
Visual concentration	11	2.1	6	15	13.6	2.8	9	21
Control measures								
Nonword repetition	9.4	3.9	2	19	12.6	3.2	5	20
Digit span forward	7.2	1.8	4	11	7.8	1.8	4	11
Digits backward	4.4	1.5	0	8	6.3	1.9	3	10
Receptive vocabulary	53.5	10.7	32	75	54	11.6	31	83
TRT	4	3	-1	13	3.7	3.5	-5	14
Reading-related measures								
Elision	11.9	3.9	2	19	15.1	2.6	9	20
Blending	5.4	2	0	12	8.7	2.5	4	19
RAN digit	19	3.7	11	32	54.3	25.9	0	123
RAN letter	24.4	6.3	15	46	60.5	17.4	35	131

Orthographic matching	11.4	3.3	4	20	14	4	6	24
Pseudo-letter matching	10.3	3.1	1	16	24	2.4	17	30
Morphological Production	14.5	4.9	5	25	16	5.4	7	33
Reading and spelling outcome measures								
Word recognition	20.7	11.2	1	46	23.5	10.2	0	47
Nonword reading fluency	12	6.3	0	31	12	6.3	0	31
Single word spelling	19.2	6.9	5	32	18.7	6	5	32

Note. SD = Standard deviation

To assess for differences in performance between the monolinguals and bilinguals, independent samples *t*-tests were used. There was a significant difference, t (124) = -7.1, $p < .001$, on nonverbal ability and visual concentration, t (124) = -5.9 $p = .03$, in favor of the bilinguals. There was a significant difference on the verbal STM tasks (nonword repetition t (124) = -5.0, $p < .001$ and digit span forward t (124) = -2.0, $p = .05$) tasks between the two groups in favor of the bilinguals. There was a significant difference, t (124) = -6.3, $p < .001$ on digits backward, also in favor of the bilinguals. There was no difference between the monolinguals and the bilinguals on the receptive vocabulary measure. There was no difference in the means on the title recognition test between the two groups. There was a significant difference on the PA tasks (elision t (124) = -5.4, p < .001 and blending words t (124) = -8.2, p < .001) in favor of the bilinguals. There was a significant difference on RAN (RAN digit t (124) = -10.7, p < .001 and RAN letter t (124) = -15.5, p < .001) tasks in favor of the monolinguals. There was a significant difference in favor of the bilinguals on orthographic matching fluency, t (124) = -4.0, p < .001, and pseudo-letter matching, t (124) = -27.7, p < .001. There was no significant group difference in performance for morphological production, word recognition or spelling.

In order to reduce the variables, three composite scores were created for PA (from elision and blending task scores), verbal STM measures (from nonword repetition and digits forward scores), and RAN (RAN letter and RAN digit). This was done by converting the raw scores into z scores and averaging the two for each of the three measures. The three sets of composite scores were used in subsequent analyses.

Table 26

Correlations between Variables for the Monolingual Group (N = 63) Matched on Nonword Reading Fluency in Study 2b

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. NVA													
2. VC	.38**												
3. VSTM	.35**	.05											
4. DB	.26*	-.03	.24										
5. A-RV	.33**	.11	.45**	.43**									
6. A-TRT	.11	.17	.03	.001	.03								
7. A-PA	.50***	.13	.47***	.34**	.52***	.01							
8. A-RAN	-.38**	-.30*	-.12	-0.17	-.24	.02	-.37**						
9. A-OM	.47***	.43**	.02	0.16	.19	-.15	.25*	-.39**					
10. A-PM	.32**	.35**	.07	0.08	.12	-.24	.26*	-.38**	.45**				
11. A-MP	.50***	.34**	.56***	.30*	.63***	-.003	.48**	-.36**	.28*	.35**			
12. A-WR	.41**	.19	.33**	0.2	.38**	-.09	.61***	-.51***	.38**	.48***	.45**		
13. A-NWRF	.42**	.23	.33**	.27*	.36**	-.05	.56**	-.54***	.36**	.49***	.47**	.89***	
14. A-SWS	.42**	.13	.35**	.29*	.36**	-.02	.54***	-.49***	.31*	.40**	.43**	.78***	.81***

Note. NVA=Nonverbal ability, VC=Visual concentration, A-RV=Receptive vocabulary, DB=Digits backward, A-TRT=Title Recognition Test, A-MP=Morphological production, A-PA=Phonological awareness, A-VSM=Verbal short-term memory, A-RAN=Rapid automatized naming, A-OMF=Orthographic matching, A-PMF=Pseudo-letter matching, A-WR=Word recognition, A-NWRF=Nonword reading fluency, A-SWS=Single word spelling. A=Arabic measures. **p < .05, **p <.01, ***p <.001*

For the monolinguals, nonverbal ability was correlated to word recognition (*r* = .41, *p* = .001), nonword reading fluency (*r* = .42, *p* =.001), and spelling (*r* = .42, *p* = .001). Verbal STM was correlated to word recognition (*r* = .33, *p* = .008), nonword reading fluency (*r* = .33, *p* = .008), and spelling (*r* = .35, *p* = .005). Digits backward was correlated to nonword reading fluency (*r* = .27, *p* = .03) and spelling (*r* = .29, *p* = .02). Receptive vocabulary was correlated to word recognition (*r* = .38, *p* = .002), nonword reading fluency (*r* = .36, *p* = .004), and spelling (*r* = .36, *p* = .004).

PA was correlated to word recognition (*r* = .61, *p* < .001), nonword reading fluency (*r* = .56, *p* < .001), and spelling (*r* = .54, *p* < .001). RAN was correlated to word recognition (*r* = .51, *p* < .001), nonword reading fluency (*r* = .54, *p* < .001), and spelling (*r* = .49, *p* < .001). Orthographic matching fluency was correlated to word recognition (*r* = .38, *p* = .002), nonword reading fluency (*r* = .36, *p* = .004), and spelling (*r* = .31, *p* = .01). Pseudo-letter matching fluency was only correlated to word recognition (*r* = .48, *p* < .001), nonword reading fluency (*r* = .49, *p* < .001), and spelling (*r* = .40, *p* = .001). Finally, morphological production was correlated to word recognition (*r* = .45, *p* < .001), nonword reading fluency (*r* = .47, *p* < .001), and spelling (*r* = .43, *p* < .001).

Table 27

Correlations between Variables for the Bilingual Group (N = 63) Matched on Nonword Reading Fluency in Study 2b

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. NVA													
2. VC	.02												
3. VSTM	.14	.26*											
4. DB	.41**	.20	.25										
5. A-RV	-.05	.35**	.48**	.17									
6. A-TRT	.04	.18	.27*	.18	.30*								
7. A-PA	.29	.27*	.53**	.49***	.27*	.14							
8. A-RAN	-.11	-.16	-.21	-.44**	-.15	-.10	-.45**						
9. A-OM	.32*	.44**	.25*	.09	.13	.25*	.20	-.16					
10. A-PM	.18	.37**	.09	.12	.22	-.15	.21	-.03	.21				
11. A-MP	.10	.43**	.52***	.40**	.74**	.19	.56***	-.43**	.14	.25*			
12. A-WR	.30*	.48***	.47***	.56***	.44**	.18	.63***	-.44**	.38**	.23	.65***		
13. A-NWRF	.34**	.36**	.46**	.51***	.39**	.29*	.61***	-.60***	.47***	.15	.59***	.81***	
14. A-SWS	.40**	.25*	.43**	.45**	.39**	.12	.61***	-.40**	0.22	.12	.61***	.68***	.66***

Note. NVA=Nonverbal ability, VC=Visual concentration, A-RV=Receptive vocabulary, DB=Digits backward, A-TRT=Title Recognition Test, A-MP=Morphological production, A-PA=Phonological awareness, A-VSM=Verbal short-term memory, A-RAN=Rapid automatized naming, A-OMF=Orthographic matching, A-PMF=Pseudo-letter matching, A-WR=Word recognition, A-NWRF=Nonword reading fluency, A-SWS=Single word spelling. A=Arabic measures. **p < .05, **p <.01, ***p <.001*

For the bilingual group, nonverbal ability was correlated to word recognition (r = .30, p = .02), nonword reading fluency (r = .34, p = .007), and spelling (r = .40, p = .001). Visual concentration was correlated to word recognition (r = .48, p < .001), nonword reading fluency (r = .36, p = .003), and spelling (r = .25, p = .05). Verbal STM was correlated to word recognition (r = .47, p < .001), nonword reading fluency (r = .46, p < .001), and spelling (r = .43, p < .001). Digits backward was correlated to word recognition (r = .56, p < .001), nonword reading fluency (r = .51, p < .001), and spelling (r = .45, p < .001). Receptive vocabulary was correlated to word recognition (r = .44, p < .001), nonword reading fluency (r = .39, p =), and spelling (r = .39, p = .001). Arabic TRT was correlated to nonword reading fluency (r = .29, p = .02)

PA was correlated to word recognition (r = .63, p < .001), nonword reading fluency (r = .61, p < .001), and spelling (r = .61, p < .001). RAN was negatively correlated to word recognition (r = -.44, p < .001), nonword reading fluency (r = -.60, p < .001), and spelling, (r = -.40, p =.001). Orthographic matching fluency was correlated to word recognition (r = .38, p = .002) and nonword reading fluency (r = .47, p < .001) but not to spelling. Finally, morphological production was correlated to word recognition (r = .65, p < .001), nonword reading fluency (r = .59, p < .001), and spelling, (r = .61, p < .001).

Predictors of Nonword Reading Fluency for Monolinguals and Bilinguals. Based on these correlations, hierarchical regression was conducted to determine predictors of nonword reading fluency among monolinguals and bilinguals. TRT did not correlate to any of the outcome variables and was excluded from analyses. Because age was correlated to orthographic matching, pseudo-letter memory, word recognition, and nonword reading fluency for the monolinguals; and was nearly correlated to every variable for the bilinguals, age was entered in step 1, followed by control measures in step 2 (nonverbal ability, visual concentration, verbal STM, and receptive vocabulary). PA and nonword reading fluency was entered at step 3, and RAN was entered in step 4. In step 5, orthographic matching and pseudo-letter matching were entered. Finally, in step 6, morphological production was entered. Results are presented in Table 28.

Table 28

Summary of Hierarchical Regression Analyses Predicting Arabic Nonword Reading Fluency for Monolingual and Bilingual groups in Study 2b

		Monolingual (N=63)			Bilingual (N=63)		
Step	Predictor	βa	βb	ΔR2	βa	βb	ΔR2
1	Age	.31*	.17	.09*	.53***	.14	.28***
2	Nonverbal ability	.22	.002	.23**	.08	.08	.28***
	Visual concentration	.05	-.05		.06	-.004	
	Verbal STM	.14	.09		.28*	.10	
	Digits backward	.05	.02		.28**	.14	
	Vocabulary	.22	.06		.12	.11	
3	PA	.40**	.28*	.09**	.23	.16	.03
4	RAN	-.34**	-.28*	.09**	-.33***	-.29**	.08***
5	Orthographic matching	-.01	-.01	.051	.22*	.24*	.03

	Pseudo-letter matching	.26*	.25*		-.05	-.06	
6	Morphological production	.06	.06	.001	.11	.11	.003

Note. β[a] =Standardized beta coefficient when first entered. β[b]= Standardized beta coefficient entered in last step.
$*p < .05$, $**p <.01$, $***p <.001$

None of the control measures predicted nonword reading fluency for the monolinguals; however, verbal STM and digits backward predicted nonword reading fluency for the bilinguals. The control measures explained 23% of the variance in nonword reading fluency for the monolinguals and 28% for the bilinguals. PA was the strongest predictor of nonword reading fluency for the monolinguals only. RAN predicted nonword reading fluency for both the monolinguals and bilinguals. RAN explained approximately 8.5% of the variance in nonword reading fluency for the monolinguals and 7.5% for the bilinguals. Pseudo-letter matching was a significant predictor of nonword reading fluency for the monolinguals, whereas orthographic matching was a significant predictor of nonword reading fluency for the bilinguals. Overall, the model explained 55% of the variance in nonword reading fluency for the monolinguals and 69% of the variance in nonword reading fluency for the bilinguals.

Comparing Monolinguals to Bilinguals Matched on Spelling Ability

This section presents preliminary analysis followed by regression analyses for monolingual and bilingual children matched on single word spelling scores. The same matching procedure for word recognition and nonword reading fluency was employed for matching on spelling. This resulted in 62 bilinguals (mean age = 9.6, SD = .95) and 62 monolinguals (mean age = 8.09, SD = .43) matched on spelling.

Descriptive Statistics and Preliminary Analysis for Bilingual and Monolingual groups Matched on Spelling. Descriptive statistics for all measures are provided in Table 29.

Table 29

Descriptive Statistics of Arabic Measures for Monolingual and Bilingual Groups Matched on Spelling Scores in Study 2b

	Mean	SD	Min	Max	Mean	SD	Min	Max
Nonverbal control measures								
Nonverbal ability	16.35	6.905	6	33	26.4	6.026	10	36
Visual concentration	11.39	2.377	6	16	13.5	2.672	9	21
Control measures								
Nonword repetition	8.79	3.173	2	19	12.68	3.458	6	20
Digit span forward	7.08	1.711	4	11	7.77	1.92	4	11
Digits backward	4.31	1.5	0	7	6.4	1.987	3	11
Vocabulary	53.48	11.322	29	74	52.87	11.223	30	83
TRT	4.06	3.156	-3	13	3.42	3.481	-5	14
Reading-related measures								
Elision	12.21	3.846	3	19	15.11	3.178	5	20
Blending	5.05	1.97	0	12	8.82	2.973	1	19

RAN digit	18.68	3.57	11	32	51.85	24.657	0	113
RAN letter	24.77	5.849	15	46	62.73	20.081	35	131
Orthographic matching	11.52	3.425	4	20	13.39	4.066	6	24
Pseudo-letter matching	10.42	3.247	1	16	23.97	2.858	15	30
Morphological production	14.55	4.858	5	25	15.82	5.632	5	33
Reading and spelling outcome measures								
Word recognition	21.19	10.17	4	44	23.52	10.381	0	47
Nonword reading Fluency	13.73	6.962	1	31	11.63	6.686	0	31
Single word spelling	19.1	6.518	5	32	19.1	6.518	5	32

Note. SD = Standard deviation

To compare performance for this subsample of monolinguals and bilinguals, independent samples *t*-tests were used. There was a significant difference, t (122) = -8.6, $p < .001$, on nonverbal ability and visual concentration, t (122) = -4.7, $p < .001$, between the two groups in favor of the bilinguals. There was a significant difference in the means on the verbal STM tasks (nonword repetition, t (122) = -6.5, $p < .001$, and digit span forward, t (122) = -2.1, $p = .04$), tasks between the two groups in favor of the bilinguals. There was a significant difference, t (122) = -6.6, p < .001, on digits backward, also in favor of the bilinguals. Although the monolinguals outperformed the bilinguals on the receptive vocabulary measure, this difference did not reach significance. There was no difference in the means on TRT. There was a significant difference in the means on PA (elision t (122) = -4.6, p < .001 and blending words, t (122) = -8.3, p < .001), in favour of the bilinguals. There was a significant difference for RAN (RAN digit t (122) = -10.5, p < .001 and RAN letter, t (122) = -14.3, p < .001), in favor of the monolingual group. There was a significant difference in favour of the bilinguals on orthographic matching fluency, t (122) = -2.8, p = .006, and pseudo-letter matching, t (122) = -24.6, p < .001. There was no significant difference in morphological production, nonword reading fluency or word recognition between the two groups.

Table 30

Correlations between Variables for the Monolingual Group (N = 62) Matched on Spelling Scores in Study 2b

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. NVA													
2. VC	.56**												
3. VSTM	.29*	.13											
4. DB	.10	.11	.20										
5. RV	.07	.02	.38**	.37**									
6. TRT	-.04	-.06	.02	.15	.11								
7. PA	.40**	.26*	.28*	.31*	.39**	.11							
8.RAN	-.14	-.32*	.10	-.25*	-.19	-.06	-.28*						
9. OM	.41**	.40**	.19	.18	.19	-.03	.41**	-.36**					
10. PM	.18	.38**	-.07	.15	-.06	-.24	.25	-.35**	.56**				
11. MP	.36**	.29*	.47**	.23	.62**	-.03	.47**	-.37**	.30*	.26*			
12. WR	.28*	.27*	.31*	.33**	.34**	-.01	.60**	-.48**	.44**	.39**	.58**		
13. NWRF	.26*	.27*	.26*	.41**	.29*	-.01	.62**	-.49**	.45**	.33**	.46**	.82**	
14. SWS	.27*	.22	.28*	.40**	.31*	-.10	.53**	-.42**	.44**	.35**	.54**	.81**	.72**

Note. NVA=Nonverbal ability, VC=Visual concentration, RV=Receptive vocabulary, DB=Digits backward, TRT=Title Recognition Test, MP=Morphological production, PA=Phonological awareness, VSM=Verbal short-term memory, RAN=Rapid automatized naming, OM=Orthographic matching, PM=Pseudo-letter matching. WR=Word recognition, NWRF=Nonword reading fluency, SWS=Single word spelling. *$p < .05$, **$p < .01$, ***$p < .001$*

For the monolingual group, nonverbal ability was correlated to word recognition (*r* = .28, *p* = .03), nonword reading fluency (*r* = .26, *p* = .04), and spelling (*r* = .27, *p* = .04). Visual concentration was correlated to word recognition (*r* = .27, *p* = .04) and nonword reading fluency (*r* = .27, *p* = .04). Verbal STM was correlated to word recognition (*r* = .31, *p* = .02), nonword reading fluency (*r* = .26, *p* = .04), and spelling, (*r* = .28, *p* = .03). Digits backward was correlated to word recognition (*r* = .33, *p* = .009), nonword reading fluency (*r* = .41, *p* = .001), and spelling (*r* = .40, *p* = .001). Receptive vocabulary was correlated to word recognition (*r* = .34, *p* = .007), nonword reading fluency (*r* = .29, *p* = .02), and spelling, (*r* = .31, *p* = .01).

PA was correlated to word recognition (*r* = .60, *p* < .001), nonword reading fluency (*r* = .62, *p* < .001), and spelling (*r* = .53, *p* < .001). RAN was correlated to word recognition (*r* = -.48, *p* < .001), nonword reading fluency (*r* = -.49, *p* < .001), and spelling (*r* = -.42, *p* = .001). Orthographic matching fluency was correlated to word recognition (*r* = .44, *p* < .001), nonword reading fluency (*r* = .45, *p* < .001), and spelling (*r* = .44 *p* < .001). Pseudo-letter matching fluency was correlated to word recognition (*r* = .39, *p* = .002), nonword reading fluency (*r* = .33, *p* = .008), and spelling (*r* = .35, *p* = .006). Finally, morphological production was correlated to word recognition (*r* = .58, *p* < .001), nonword reading fluency, (*r* = .46, *p* < .001), and spelling (*r* = .54, *p* < .001).

In order to reduce the variables, composites were created for PA, verbal STM, and RAN as before. This was done by converting the raw scores for the two tasks for each measure into *z* scores and averaging the two. The composite scores are used in the analyses.

Table 31

Correlations between Variables for the Bilingual Group (N = 62) Matched on Spelling Scores in Study 2b

	1	2	3	4	5	6	7	8	9	10	11	12
1. NVA												
2. VC	.19											
3. VSTM	.13	.30*										
4. DB	.37**	.26*	.33**									
5. RV	.003	.29*	.48**	.25								
6. TRT	-.02	.11	.22	.23	.27*							
7. PA	.20	.24	.63**	.57**	.37**	.20						
8.RAN	-.05	-.15	-.36**	-.41**	-.19	-.06	-.51**					
9. OM	.32*	.50**	.41**	.24	.24	.31*	.44**	-.27*				
10. PM	.31*	.40**	.17	.26*	.27*	-.11	.35**	-.29*	.30*			
11. MP	.07	.35**	.54**	.41**	.78**	.18	.60**	-.39**	.27*	.37**		
12. WR	.21	.46**	.51**	.63**	.52**	.22	.68**	-.45**	.50**	.27*	.60**	
13. NWRF	.25*	.38**	.53**	.56**	.55**	.29*	.62**	-.45**	.56**	.20	.56**	.82**
14. SWS	.36**	.25*	.44**	.54**	.45**	.16	.73**	-.47**	.42**	.33*	.64**	.73**

Note. NVA=Nonverbal ability, VC=Visual concentration, RV=Receptive vocabulary, DB=Digits backward, TRT=Title Recognition Test, MP=Morphological production, PA=Phonological awareness, VSM=Verbal short-term memory, RAN=Rapid automatized naming, OM=Orthographic matching, PM=Pseudo-letter matching. WR=Word recognition, NWRF=Nonword reading fluency, SWS=Single word spelling. **p < .05, **p <.01, ***p <.001*

For the bilingual group, nonverbal ability was correlated to nonword reading fluency ($r = .25$, $p = .047$) and spelling ($r = .36$, $p = .004$). Visual concentration was correlated to word recognition ($r = .46$, $p < .001$), nonword reading fluency ($r = .38$, $p = .002$), and spelling ($r = .25$, $p = .047$). Verbal STM was correlated to word recognition ($r = .51$, $p < .001$), nonword reading fluency ($r = .53$, $p < .001$), and spelling ($r = .44$, $p < .001$). Digits backward was correlated to word recognition ($r = .63$, $p < .001$), nonword reading fluency ($r = .56$, $p < .001$), and spelling, ($r = .54$, $p < .001$). Receptive vocabulary was correlated to word recognition ($r = .52$, $p < .001$), nonword reading fluency ($r = .55$ $p = .001$), and spelling, ($r = .45$, $p < .001$). Arabic TRT was correlated to nonword reading fluency ($r = .29$, $p = .02$).

PA was correlated to word recognition ($r = .68$, $p < .001$), nonword reading fluency ($r = .62$, $p < .001$), and spelling, ($r = .73$, $p < .001$). RAN was negatively correlated to word recognition ($r = -.45$, $p < .001$), nonword reading fluency ($r = -.45$, $p < .001$), and spelling ($r = -.47$, $p < .001$). Orthographic matching fluency was correlated to word recognition ($r = .50$, $p < .03$), nonword reading fluency ($r = .56$, $p < .001$), and spelling ($r = .42$, $p = .001$). Pseudo-letter matching fluency was correlated to word recognition ($r = .27$, $p = .04$) and spelling ($r = .33$, $p = .01$). Finally, morphological production was correlated to word recognition ($r = .60$, $p < .001$), nonword reading fluency ($r = .56$, $p < .001$), and spelling, ($r = .64$, $p < .001$).

Predictors of Spelling for Monolingual and Bilingual Groups Matched on Spelling. A hierarchical regression was conducted with single word spelling as the outcome measure. Age was entered in step 1, followed by control measures in step 2 (nonverbal ability, visual concentration, verbal STM, and receptive vocabulary). In steps 3 and 4, PA and RAN were entered respectively. In step 5, orthographic matching and pseudo-letter matching were entered. Finally, morphological production was entered in step 6. The results are presented for both the monolinguals and bilinguals in Table 32.

Table 32

Summary of Hierarchical Regression Analyses of Predictors of Spelling for Monolingual and Bilingual Children in Study 2b

		Monolinguals (N=62)			Bilinguals (N=62)		
Step	Predictor	βa	Bb	ΔR2	Ba	Bb	ΔR2
1	Age	.25*	.11	.07*	.59***	.18	.35***
2	Nonverbal ability	.07	-.06	.25**	.15	.23	.21***
	Visual concentration	.09	-.08		-.05	-.02	
	Verbal short-term memory	.13	.05		.14	-.15	
	Digits backward	.29*	.26*		.25*	.02	
	Vocabulary	.18	-.15		.25	.13	
3	PA	.36**	.24	.09**	.55***	.47***	.12***
4	RAN	-.24*	-.13	.05*	-.11	-0.1	.01
5	Orthographic matching	.09	.17	.03	-.02	.02	.004
	Pseudo-letter matching	.15	.04		-.07	-.08	
6	Morphological production	.36*	.36*	.04*	.23	.23	.01

Note. β^a =Standardized beta coefficient when first entered. β^b= Standardized beta coefficient entered in last step.
$^*p < .05$, $^{**}p < .01$, $^{***}p < .001$

Of the control measures, digits backward predicted spelling for both bilingual and monolingual groups. PA predicted spelling for both the monolinguals and bilinguals. The influence of PA on spelling was greater in the bilingual sample and explained 12% of the variance in spelling. RAN was also a predictor of spelling, but only for the monolingual group. The strongest predictor of spelling among the monolinguals was morphological awareness, explaining an additional 4% of the variance in spelling when entered in the last step. Overall, the model explained 52% of the variance in spelling for the monolingual group and 71% of the variance for the bilingual group.

Summary of Predictors of Reading and Spelling in Arabic and English for Bilinguals and Monolinguals

Cross-linguistic analysis of predictors of reading and spelling for the bilingual children revealed some key differences between Arabic and English. Arabic digits backward and PA emerged as the strongest predictors of reading for the younger bilinguals, whereas Arabic RAN was a consistent predictor of reading for the older group of bilinguals. Arabic PA was the only predictor of spelling and that was only for the younger bilinguals.

A very different pattern emerged for English reading and spelling. Verbal STM, receptive vocabulary, and orthographic matching predicted English word recognition for the younger bilinguals. Receptive vocabulary was the only predictor of English word recognition for the older bilinguals.

Verbal STM was the strongest predictor of English nonword reading fluency, followed by orthographic matching for the younger group of bilinguals. None of the variables significantly predicted English nonword reading fluency for the older bilinguals. Verbal STM and orthographic matching predicted English spelling for the younger bilinguals, whereas for the older group, PA and RAN predicted English spelling.

As for the monolinguals, verbal STM, receptive vocabulary, PA, RAN, and orthographic/visual processing predicted Arabic word recognition, nonword reading fluency, and spelling. PA emerged as the strongest predictor of all Arabic outcome measures. Nonverbal ability was a consistent predictor of reading and spelling in Arabic among the monolinguals.

The comparisons between the monolinguals and bilinguals revealed that Arabic PA was the strongest predictor of Arabic word recognition for both monolinguals and bilinguals, followed by digits backward. Arabic RAN was the only predictor shared by both monolinguals and bilinguals for Arabic nonword reading. RAN was the strongest predictor of nonword reading for the bilinguals, whereas PA was the strongest predictor of nonword reading, followed by RAN for the monolinguals. PA was the strongest predictor of Arabic spelling for both the bilinguals and monolinguals, followed by digits backward. Arabic morphological production predicted Arabic spelling for the monolinguals only. The next section discusses these findings in relation to the existing literature.

Discussion

Study 2 involved examining predictors of reading and spelling first, in a sample of bilingual Arabic L1-English L2, third, fourth, and fifth graders;

and second, in a sample of third-grade Arabic speaking monolingual students. In Study 2a, the bilinguals were divided into younger and older groups in order to examine possible developmental trends. In Study 2b, the third-grade Arabic- speaking monolinguals' data was analysed both separately and by matching them to the subsamples of the bilinguals in Study 2a. The results of Study 2a are discussed first, followed by a discussion on the findings from Study 2b. The last section will present a synthesis of the findings along with limitations and conclusions.

Study 2a Discussion

The discussion of Study 2a will outline the differences between predictors of reading and spelling between the younger and older samples of bilinguals that differed considerably within-language and across both languages.

As expected, there were significant differences between the younger and older groups of bilinguals on nearly every Arabic measure.

The younger and older bilinguals were significantly different on nonverbal ability, and this was a significant predictor of Arabic spelling for the older group and of English spelling for the younger group. Nonverbal ability did not predict reading measures in the bilingual children. These findings demonstrated that, although it is possible to distinguish between universal and script-dependent predictors of reading and spelling, the magnitude of these relationships changes across development.

To investigate the effect of visual complexities of Arabic, visual concentration was measured. In all of the analyses for the older and younger bilinguals, visual concentration entered in the second step did not predict any of the Arabic or English reading or spelling measures. Although this finding is not surprising for English, it was unexpected for Arabic.

Verbal STM was measured in both Arabic and English for the bilingual children using digit span forward and nonword repetition. In addition, working memory was assessed in Arabic using digits backward. The Arabic working memory task predicted word recognition for the younger and older bilinguals. It also predicted Arabic nonword reading fluency for the younger group, and Arabic spelling for the older group. Verbal STM played a bigger role in English reading and spelling for the younger bilinguals only. Verbal STM predicted Arabic reading and English reading and spelling for the younger bilinguals, which demonstrates that, for young readers of Arabic, verbal STM predicts Arabic reading and spelling as it would a second language.

Ironically, the two groups of younger and older bilinguals did not differ in receptive vocabulary in Arabic. However, receptive vocabulary in English differed significantly between the age groups, in favor of the older group. Taken together, these results may indicate the long-lasting effects of diglossia on Arabic vocabulary development. While English is the children's second language, vocabulary was found to develop with age. According to the literature, verbal STM tasks are associated with vocabulary; however, research in English has shown that verbal STM ceases to predict vocabulary after the early years of reading instruction. The younger and older groups of bilinguals did not differ significantly on the English digit span or nonword repetition tasks. Therefore, the lack of group differences on the verbal STM tasks may reflect a universal developmental trend rather than a specific linguistic feature of Arabic.

Arabic receptive vocabulary did not predict Arabic reading or spelling for the bilinguals, except for nonword reading fluency for the younger group. The odd finding that vocabulary predicted nonword reading fluency in Arabic may be explained by the consonantal nature of the Arabic script: words are composed of usually three to four consonants, making nonwords only possible in vowels (diacritics). Having a better vocabulary may result in increased decoding speed for the nonwords. In contrast, English vocabulary was the strongest predictor of English word recognition for the younger and older bilingual children. This finding is consistent with findings from opaque orthographies such as English, in which children must rely at least in part on their existing vocabulary knowledge for irregular words. For example, it was observed that many of the children in the bilingual sample would phonologically decode the word “naive” and then pause and then produce the correct pronunciation. Indeed, this is consistent with the way Ehri (2005b) describes how children approach the task of reading irregular words by approximating the correct pronunciation, but that is only possible if the word exists in their oral vocabulary. In contrast, vocabulary in Arabic did not predict word recognition or spelling in both younger and older bilinguals, which may suggest that Arabic-speaking children do not rely on oral language due to diglossia.

The TRTs that were developed in Arabic and English as measures of print exposure did not predict reading or spelling in either language. As discussed in Chapter 2, this finding may reflect the fact that reading for leisure is not fostered in Arabic. Although reading is encouraged in English for the bilingual sample, the print exposure measure may have not been a suitable measure in this context because the majority of test items were selected based on book titles provided by international school libraries which are similar to the school which the bilingual sample was drawn from. Since such measures are a proxy of leisurely reading, the item selection from school libraries may not be related to single word reading and spelling. The children may have been very familiar with the titles because they are found in the school library and they are encouraged to read them. Although the English TRT did not predict English reading or spelling in the regression analyses, there were moderate correlations with word recognition (.42), nonword reading (.41), and spelling (.37), for the younger group. In English, print exposure might be important for children's development of orthographic skills which are critical for reading in the deep orthography of English. In Arabic, development of orthographic processes may be associated with vocabulary, given the diglossic lexical distance (Saiegh-Haddad, 2018).

This is the first study to have assessed Arabic print exposure, and although findings in both English and Arabic suggest that Arabic-speaking children's reading development is not influenced by exposure to print, more studies examining print exposure and using a variety of tools may yield different results, particularly in diglossic Arabic (e.g., Abu-Rabia, 2000).

Based on the literature reviewed regarding PA and reading and spelling as a function of orthographic depth, it was not clear whether PA would be a predictor in transparent, fully vowelized Arabic. The findings revealed that in Arabic, PA was the strongest predictor of reading and spelling for both groups of bilinguals. Arabic PA was the strongest predictor of word recognition, nonword reading fluency, and spelling for the younger bilinguals but only predicted word recognition for the older bilinguals. This suggests that the role of PA in reading and spelling may slowly start to diminish in older grades. Both the Arabic word-recognition task and Arabic nonword- reading fluency tasks were fully vowelized, making them fully transparent in orthographic depth. Thus, this is an inconsistent finding according to the cross-linguistic literature on transparent orthographies discussed in this chapter. In fact, the pattern of results is consistent with very opaque orthographies such as English, in which the contribution of PA to reading continues to the third grade (Seymour et al., 2003). Arabic PA was the strongest predictor of Arabic spelling among the younger bilinguals; however, it made no contribution to spelling for the older bilinguals. In contrast, English PA was the strongest predictor of English spelling for the older bilinguals only. Based on the literature reviewed in this chapter, it was expected that PA would emerge as a strong predictor of reading and spelling in English; however, that was not the outcome observed in the present study.

The only exception to the overwhelming influence of Arabic PA on reading and spelling was RAN. RAN emerged as the strongest predictor of Arabic nonword reading fluency for the older bilinguals. RAN also predicted Arabic word recognition for the older bilinguals. The bilingual sample found Arabic RAN was much more laborious in comparison to English RAN. In fact, bilingual students are not taught the digits in Arabic class, and since all subjects are in English, numbers in Arabic are probably taught before preschool entry and therefore never become automatic. This was evident when children would pause and often confuse visually similar numbers (e.g., ٨, ٧). Although Arabic is mandatory from kindergarten onwards, it is only taught for 45 minutes a day, and just as with digits, it was also observed that bilingual students often confused the names of similar letters. This may have diluted the influence of Arabic RAN on Arabic reading and spelling measures and rendered the alphanumeric RAN task inappropriate for this sample of bilinguals.

Orthographic knowledge was assessed using an orthographic matching task in Arabic and English. The Arabic orthographic matching task entered in the second-last step after background measures, PA, and RAN, did not predict any of the Arabic reading or spelling measures for the younger and older bilinguals. However, the English orthographic matching task was the second strongest predictor of English reading and spelling when entered in the last step of the model for the younger group of bilingual children but failed to predict reading or spelling for the older group. Although the Arabic and English orthographic matching tasks were comparable and the developers of the Arabic task modeled the measure after the English orthographic matching task, choosing this task as a measure of orthographic knowledge in Arabic may have not been appropriate due to the linguistic characteristics of Arabic. This task required the participants to scan rows consisting of five nonwords and put a slash through the two that are identical (e.g., nesp, mesp, wesp, espw, wesp) in 2 minutes. In the Arabic version, the construction of nonwords consisted of letters that were in the wrong shape for their position in the word. There is a difference between the Arabic and English versions in that in English, having an illegal letter combination such as ‘uuns’ renders it impossible to decode phonologically, whereas in Arabic having the wrong letter shape according to position does not violate its phonology. Thus, the orthographic matching task may have not been appropriate for Arabic. Indeed, Tibi et al.

(2021) addressed this problem recently by developing a valid and reliable orthographic measure in Arabic. Indeed, Tibi et al. (2021) addressed this problem by developing a valid and reliable orthographic measure in Arabic.

The pseudo-letter matching task was a predictor of Arabic reading and spelling for the younger bilinguals, even when entered in the second-last step. Unlike the orthographic matching task described above, this measure may have been a purer measure of visual processing because the items could not be decoded (items were constructed with pseudo-letters that resemble Arabic letters). While this task was not correlated to reading and spelling for the older group of bilingual children, it had a strong correlation with reading (r = .50) and spelling (r = .51) for the younger group. Taken together, it seems that this measure taps into a processing skill necessary for reading in Arabic, but only for younger children.

Study 2b Discussion

The aim was to investigate the predictors of Arabic reading and spelling in a large sample of third-grade monolinguals. Subsamples of monolinguals were selected and matched on word recognition, nonword reading fluency, and spelling to subsamples of bilinguals, in order to determine whether they differed in the cognitive and linguistic abilities that predict their reading and spelling. The following sections will present a discussion of predictors of Arabic reading and spelling for the third-grade monolinguals, followed by comparisons of bilinguals and monolinguals.

Predictors of Reading and Spelling for the Monolinguals

Nonverbal ability consistently predicted all three outcome measures of reading and spelling among the third-grade monolinguals. Although these findings are not supported by the literature in the vast majority of languages, they are consistent with literature on Arabic-speaking children (Abu Ahmad et al., 2014). It has been argued that Arabic's visual complexities place higher general cognitive demands on the reader than in other orthographies; however, this was not supported in the current study. In all of the analyses for the third-grade monolinguals, visual concentration entered in the second step did not predict any of the Arabic reading or spelling measures. This result was unexpected; however, these findings may be due to the fact that the particular task employed may not be sensitive enough. Other visual tasks may be more strongly related to reading, such as conjunction search, which seems to tap into an ability that has been shown to be impaired in dyslexics (e.g., Buchholz & McKone, 2004). Many Arabic reading researchers have suggested that the visual complexities of written Arabic are a major source of difficulty for young readers. The fact that nonverbal ability consistently predicted reading and spelling in Arabic for the monolingual sample supports the idea that higher cognitive demands are placed on the reader of Arabic.

Verbal STM played an even bigger role in reading and spelling for the monolinguals. Verbal STM predicted Arabic reading and spelling for the sample of third-grade monolinguals. The results mirror the ones obtained for the younger bilinguals in Study 2a and are consistent with the prediction that diglossia creates a situation akin to learning a foreign language. Vocabulary, in contrast to the results observed for the bilinguals in Study 2a, which only predicted word recognition, predicted Arabic word recognition, nonword reading fluency, and spelling for the sample of monolinguals, again suggesting that reading and spelling in Arabic place higher demand on general cognitive abilities (Abu-Ahmad et al., 2014). The finding that receptive vocabulary is a predictor of reading and spelling in Arabic, like verbal STM, may also reflect the effects of diglossia on vocabulary development and its relationship to reading and spelling (Saiegh-Haddad, 2003, 2005).

Arabic PA was the strongest predictor of Arabic reading and spelling for the third-grade monolinguals, even when entered in step 4. This finding may provide evidence that Arabic, even when fully vowelized, behaves much like what we would expect of an opaque orthography. However, Zeigler et al. (2010) argued that if PA tasks are difficult enough, then the association of PA and reading/spelling could be found in transparent orthographies as well. Indeed, since there were no observed ceiling effects on the PA tasks in the current study—suggesting that the tasks were sufficient in difficulty to pick up on the association between PA and in fully vowelized Arabic reading and spelling—this may provide an alternative interpretation that fully vowelized Arabic behaves like opaque orthographies.

RAN consistently predicted word recognition, nonword reading, and spelling for the monolinguals, even when entered in step 5, after controlling for age, nonverbal ability, verbal STM, working memory, vocabulary, and PA.

The pseudo-letter matching task was a predictor of Arabic reading and spelling for the monolinguals, whereas the orthographic matching task did not predict any of the outcome measures for the monolinguals.

Nonverbal ability consistently predicted word recognition, nonword reading fluency, and spelling, explaining a substantial amount of the variance in each of the outcome measures. Similarly, receptive vocabulary consistently predicted word recognition, nonword reading fluency, and spelling, whereas verbal STM predicted word recognition and spelling. These findings are consistent with the literature in Arabic suggesting that reading in Arabic may require more general cognitive abilities (Abu Ahmad et al., 2014). PA emerged as the strongest predictor explaining a substantial (19%–21%) amount of the variance in reading ability for the monolinguals. Although to a lesser degree, PA also predicted a significant variance (14%) in spelling for the monolinguals. This is consistent with the literature on Arabic (e.g., Taibah & Haynes, 2011; Taha, 2013; Assad & Eviatar, 2014; Asadi & Khateb, 2017; Asadi et al., 2017; Tibi & Kirby, 2018, 2019). Like PA, RAN was also a consistent predictor of word recognition, nonword reading fluency, and spelling. It explained the most variance in spelling (5%) for the monolinguals. Pseudo-letter matching predicted word recognition, nonword reading fluency and spelling and explained between 6% and 4% of the variance in the three outcome measures.

Comparisons between the Monolinguals and Bilinguals Matched on Word Recognition

The bilinguals outperformed the monolinguals on all tasks with a few notable exceptions. The monolinguals were superior on the receptive vocabulary task and morphological production task, but these differences never reached significance. While these results are somewhat surprising, they demonstrate the effects of diglossia and how Arabic-speaking children develop reading and spelling skills in the absence of sufficient oral language.

The monolinguals were also superior on the RAN tasks, and that difference was significant. As discussed, the advantage of the monolinguals on this task may reflect the fact that monolinguals are far more exposed to digits and letters in Arabic than the bilinguals are. In other words, the RAN digit and letters may not have been an appropriate task for the bilinguals due to their lack of automaticity with the stimulus.

Despite the advantage that bilinguals have on nearly all control and reading-related measures, the monolinguals outperformed the bilinguals on nonword reading fluency and spelling, and that difference was significant.

Predictors of Word Recognition for the Sample of Monolinguals and Bilinguals Matched on Word Recognition

The predictors of reading and spelling for the bilinguals and monolinguals matched on word recognition were nearly identical. Digits backward task (a measure of working memory) predicted word recognition. The strongest and only reading-related predictor of word recognition for both groups of bilinguals and monolinguals matched on word recognition was PA, explaining 19% of the variance in word recognition for the monolinguals and 8.5% of the variance in word recognition for the bilinguals.

Comparisons between the Monolinguals and Bilinguals Matched on Nonword Reading Fluency

The comparisons between a subsample of monolinguals and a subsample of bilinguals matched on nonword reading fluency nearly mirrored those of the subsamples matched on word recognition described above. In nearly all the comparisons, the bilinguals outperformed the monolinguals. The exceptions were in the superiority of monolinguals on the RAN tasks. Finally, there was no difference between the bilinguals and monolinguals on word recognition and spelling.

Predictors of Nonword Reading Fluency for the Sample of Monolinguals and Bilinguals Matched on Nonword Fluency

None of the control measures predicted nonword reading fluency for the monolinguals; however, verbal STM and digits backward predicted nonword reading fluency for the bilinguals. Unlike the samples matched on word recognition, PA was the strongest predictor of nonword reading fluency for the monolinguals only, whereas RAN emerged as a predictor of nonword reading fluency for both the monolinguals and bilinguals matched on nonword reading fluency and explained a similar variance for both groups. Pseudo-letter matching was a significant predictor of nonword reading fluency for the monolinguals, whereas orthographic matching was a significant predictor of nonword reading fluency for the bilinguals. Because the pseudo-letter matching and orthographic matching tasks are timed measures, it is possible that processing speed rather than orthographic processing is involved in the nonword-reading fluency timed task.

Comparisons between the Monolinguals and Bilinguals Matched on Spelling

For the subsamples of bilinguals and monolinguals matched on spelling ability, the bilinguals significantly outperformed the monolinguals on nearly all tasks. These results once again mirror those of the comparisons of subsamples of bilinguals and monolinguals matched on word recognition and nonword reading fluency. Like the subsamples matched on word recognition, the monolinguals matched on spelling ability outperformed the bilinguals on the vocabulary task; however, this difference never reached significance. RAN tasks were in favor of the monolinguals, as was the case in the comparisons of samples matched on word recognition and nonword fluency. Finally, there was no statically significant difference on nonword reading fluency and word recognition between the two groups.

Predictors of Spelling for the Sample of Monolinguals and Bilinguals Matched on Spelling

Of the control measures, digits backward predicted spelling for both bilinguals and monolinguals. Although spelling in Arabic is very transparent, it does require holding on to individual phonemes in STM until proceeding letters are transcribed. This ability may even be more crucial in Arabic since some of these phonemes are diglossic phonemes. PA predicted spelling for both the monolinguals and bilinguals; however, the influence of PA on spelling was greater in the bilingual sample and explained 12% of the variance in spelling. It was not surprising for PA to exert more influence on the bilinguals' spelling ability since the English curriculum places great emphasis on PA and phonics instructions. RAN was also a predictor of spelling but only for the monolinguals; however, as discussed, RAN scores for the bilinguals were near floor, and thus RAN may have also contributed to spelling ability for the bilinguals. The strongest predictor of spelling among the monolinguals was morphological awareness, explaining an additional 4% of the variance in spelling when entered in the last step.

Summary of Comparisons between Bilinguals and Monolinguals

In every matching procedure (word recognition, nonword reading fluency, and spelling), the bilinguals were older than the monolingual sample, suggesting that the bilinguals' reading and spelling abilities temporarily lag behind those of the monolinguals. However, the significant predictors of reading and spelling were nearly identical for both groups. Taken together, although the bilinguals scored significantly lower than did the third-grade monolinguals on reading and spelling tasks, the data indicate that both bilinguals and monolinguals are using the same underlying cognitive and linguistic abilities for reading and spelling. Table 34 provides a summary of predictors of Arabic word recognition, nonword reading fluency, and spelling between monolinguals and bilinguals.

General Discussion

Tables 33, 34, 35, and 36 provide a summary of the predictors of reading and spelling from Studies 2a and 2b. Table 33 provides a summary of predictors of English word recognition, nonword reading fluency and spelling for the younger and older bilinguals. Table 34 provides a summary of predictors of Arabic word recognition, nonword reading fluency, and spelling for the younger and older bilinguals. Table 35 provides a summary of predictors of reading and spelling for the monolinguals. Table 36 provides a summary of predictors of reading and spelling for monolinguals and bilinguals.

Table 33

Summary of Predictors of English Word Recognition, Nonword Reading Fluency, and Spelling for the Sample of Younger and Older Bilinguals

	Word recognition		Nonword Reading Fluency		Spelling	
Predictor	Younger	Older	Younger	Older	Younger	Older
Age	✓	—	✓	—	—	—
Nonverbal ability	—	—	—	—	✓	—
Verbal short-term memory	✓	—	✓	—	✓	—
Vocabulary	✓	✓	—	—	—	—
TRT	—	—	—	—	—	—
Phonological awareness	✓	—	✓	—	—	✓
RAN	—	—	—	—	—	✓
Orthographic matching	✓	—	✓	—	✓	—

Table 34

Summary of Predictors of Arabic Word Recognition, Nonword Reading Fluency, and Spelling for the Sample of Younger and Older Bilinguals

	Word Recognition		Nonword Reading Fluency		Spelling	
Predictor	Younger	Older	Younger	Older	Younger	Older
Age	—	—	✓	—	✓	—
Nonverbal ability	—	—	—	—	—	✓
Visual concentration	—	—	—	—	—	—
Verbal STM	—	✓	✓	—	—	—
Digits backward	✓	✓	✓	—	—	—
Receptive vocabulary	—	—	✓	—	—	—
PA	✓	✓	✓	—	✓	—
RAN	—	✓	—	✓	—	—
Orthographic matching	—	✓	—	—	—	—
Morphological production	—	—	—	—	—	—

Table 35

Summary of Predictors of Arabic Word Recognition, Nonword Reading Fluency, and Spelling for the Sample of Monolinguals

Predictor	Word Recognition	Nonword Reading Fluency	Spelling
Age	—	—	—
Nonverbal ability	✓	✓	✓
		✓	
Visual concentration	—	—	—
Verbal short-term memory	✓	—	✓
Digits backward	—	—	—
Vocabulary	✓	✓	✓
PA	✓	✓	✓
RAN	✓	✓	✓
Orthographic matching	—	—	—
Pseudo-letter matching	✓	✓	✓
Morphological production	—	—	—

Table 36

Summary of Predictors of Arabic Word Recognition, Nonword Reading Fluency, and Spelling for Monolinguals and Bilinguals

	Matched on Word Recognition		Matched on Nonword Reading Fluency		Matched on Spelling	
Predictor	Monolingual	Bilingual	Monolingual	Bilingual	Monolingual	Bilingual
Age	—	✓	✓	✓	✓	✓
Nonverbal ability	—	—	—	—	—	—
Visual concentration	—	—	—	—	—	—
Verbal STM	—	—	—	✓	—	—
Digits backward	✓	✓	—	✓	✓	—
Receptive vocabulary	—	—	—	—	—	—
PA	✓	✓	✓	—	✓	✓
RAN	—	—	✓	✓	✓	—
Orthographic matching	—	—	✓	✓	—	—
Morphological production	—	—	—	—	✓	—

The overwhelming influence of PA on reading in fully vowelized Arabic was somewhat surprising. PA consistently emerged as the strongest predictor of reading. The findings in this study demonstrate that our current definition and parameters of orthographic depth may need to be revised in order to be inclusive of the individual characteristics of Arabic and other non-Indo-European languages.

When a subset of monolinguals and a subset of older bilinguals were matched on word recognition, PA was the strongest predictor of word recognition. Similarly, when a subsample of monolinguals and a subsample of older bilinguals were matched on spelling, PA once again emerged as the strongest predictor. This is consistent with Asadi and Khateeb's (2017) finding that PA consistently predicted Arabic spelling from first to sixth grades. These findings underscore the importance of PA in Arabic and its universality in predicting reading and spelling.

The comparisons between subsamples of monolinguals and bilinguals may offer some hints on the RAN-reading relationship. When a subsample of monolinguals and a subsample of bilinguals were matched on Arabic nonword reading fluency, RAN emerged as the strongest predictor of nonword reading fluency for the bilinguals and second-strongest predictor after PA for the monolinguals. RAN also predicted Arabic spelling for the subsample of monolinguals only that were matched on spelling to a subsample of bilinguals. The fact that RAN consistently predicted spelling for the monolinguals (the whole sample and for the subsample) would suggest that RAN taps into orthographic processing (Wolf & Bowers, 2000; Manis et al., 2000), however, the fact that RAN was the strongest predictor of Arabic nonword reading fluency for the bilinguals and second-strongest predictor of nonword reading fluency for the subsample of monolinguals would suggest that RAN taps into phonological processes. Thus, neither the monolingual nor bilingual data offer a definitive resolution to the role of RAN in reading. It is important to note, however, that, despite literature on the use of alphanumeric RAN, this may have not been appropriate for the sample of bilinguals and thus may make it difficult to draw conclusions based on their results. One of the issues regarding Arabic alphanumeric RAN in general, and not only for the bilingual sample, is that digits and letters in Arabic, unlike English, are multisyllabic. The other issue that became very evident when comparing the performance of bilinguals to that of monolinguals on the alphanumeric RAN was that bilinguals were far slower. This was evident when comparing the mean scores of bilinguals and monolinguals on RAN-digits. The RAN-digits mean score in seconds for the bilinguals was 64 for the younger group and 44 for the older group, whereas the mean for the monolinguals was 19 seconds. Similarly, the RAN-letter mean score in seconds for the bilinguals was 63 for the younger group and 62 for the older group, whereas the mean for the monolinguals was 24 seconds. In fact, the similar letter shapes were the source of most errors on the RAN-letters task for both the bilingual and monolingual children alike, although to a lesser degree for the monolinguals (e.g., ض,ص, ذ,د,ر,ز). Indeed, Ibrahim and Hertz-Lazarowitz (2014) found that Arabic speakers were three times slower at reading Arabic text than were Hebrew speakers reading Hebrew text, demonstrating that visual similarities in Arabic letters cause a considerable deal of uncertainty.

Thus, if we only consider the monolingual data, RAN contributed approximately 3% of the variance in word recognition, 4% in nonword reading fluency, and 5% in spelling. At first glance, it may be tempting to conclude that RAN may be tapping into orthographic processing since RAN's strongest contribution was to spelling; however, seeing that RAN's contribution to nonword reading fluency is close to that of spelling casts doubt on that conclusion. Although nonword reading is often considered as a measure of PA, it is important to note that such a task would surely depend on existing orthographic patterns. Coupled with the fact that this task was also timed, the contribution of RAN to the nonword-reading fluency task may be via orthographic skills and the speed of processing. In fact, it has been suggested that since Arabic is an *abjad* writing system consisting of consonants (usually 3 or 4), it is likely that constructing nonwords would result in an identical real word only varying by vowel diacritics (Hansen, 2014). It may be more difficult to construct Arabic nonwords that do not overlap with words that already exist in the lexicon. Taken together, the influence of RAN on spelling, word recognition, and nonword reading fluency could be evidence of RAN tapping into orthographic processing skills. An alternative to this explanation could be that RAN taps into the integration of both phonological and orthographic representations.

It has been proposed by Arabic researchers that the Arabic lexicon may be morphologically organized (Boudalaa, 2014; Frost et al., 2005). While this assumption is beyond the scope of this paper, the role of morphological awareness in Arabic reading and spelling was investigated. The morphological awareness task assessed morphological production and, despite the high correlations between this task and reading and spelling, morphological production did not predict reading or spelling in either the bilinguals or monolinguals. The only time morphological production made a contribution to spelling was for a subsample of monolinguals that were matched on spelling to a subsample of bilinguals. In fact, morphological production predicted spelling when entered in the last step and contributed 1% of the variance in spelling. The finding that morphological awareness is linked to spelling in Arabic is consistent with the findings from Asadi and Khateb (2017) showing that morphological awareness was the second-strongest predictor of spelling.

There are several potential reasons why morphological production failed to contribute to reading and spelling in Arabic in the present study. First, in reading literature, morphological awareness is often associated with reading comprehension rather than with single word reading (Tibi & Kirby, 2019). Second, the rich morphological structure in Arabic may not have been fully accounted for by the task used in this study. Third, the morphological production task was always entered last in the model, after vocabulary, phonological tasks, and orthographic tasks. It has been argued that morphological tasks and vocabulary tap into a related underlying ability (Kirby et al., 2012). Thus, it is possible that the lack of contribution may have been diluted by the contribution made by vocabulary. Indeed, this was observed in the data reduction analyses for both the bilingual sample and monolingual samples with vocabulary and morphological consistently loading onto one factor.

Limitations

There are several important limitations to consider. First, the measures used in this study were not standardized. Currently, there are limited standardized assessments in Arabic (Tibi et al., 2021), which should be addressed in future research. Second, studies in Arabic are difficult to generalize, as there are various spoken varieties across different Arabic-speaking countries and even within the same country. This is indeed the case in Egypt, where there are distinct varieties of Arabic according to region. Another issue that warrants attention pertains to the word recognition task in Arabic. Earlier research in Arabic has often confounded two distinct diacritical marks—phonemic diacritics, which are necessary for lexical access because they represent short vowels necessary for lexical access map—and morpho-syntactic diacritics (appearing at the end of words) to represent abstract grammatical forms which are not necessary for word identification (Saiegh-Haddad, 2018). The word recognition task obtained for this study required participants to pronounce the morph-syntactic diacritics at the end of each word to be considered correct, which may have obscured the results on the word recognition task. Future research should be more aware of this distinction when administering fully vowlelized word recognition tasks. Although this study employed a large number of measures, it would have been preferable to use more measures for each construct. Indeed, the lack of contribution of morphological awareness to reading may be related to the fact that only inflectional morphology was tested. Saiegh-Haddad and Taha (2017) tested PA and derivational morphological awareness and found that PA was the strongest predictor, morphological awareness explaining a small unique variance. For example, only one measure was administered for vocabulary and one measure for morphology. Because of diglossia, having more measures to assess these constructs would have strengthened the conclusions that can be drawn from this study. Finally, this study used a cross-sectional approach to examine predictors of reading and spelling from a developmental perspective. Instead, longitudinal research designs would be more appropriate to address the developmental progression of reading and spelling skills.

To conclude, this study adds to the cross-linguistic literature on predictors of reading. The general aim of this study was to examine cognitive, linguistic, and environmental factors that contribute to reading in Arabic, in order to separate the universal predictors from the language-specific ones. However, much more research in Arabic is required to achieve this ambitious aim. The findings underscore the importance of PA in Arabic and its universality in predicting reading and spelling. The results, however, show that the magnitude of the contribution of PA to reading may be more consistent with opaque orthographies. While this finding is inconsistent with findings in cross-linguistic literature about the diminishing role of PA in transparent orthographies past first grade (Landrl & Wimmer, 2000; Papadopoulos et al., 2009; Furnes & Samuelsson, 2011), it is consistent with findings from literature on Arabic (Taibah & Haynes, 2011; Taha, 2013; Assad & Eviatar, 2014; Asadi & Khateb, 2017; Asadi et al., 2017; Tibi & Kirby, 2018, 2019)). This pattern of results has often been explained by diglossia and the visual complexities of Arabic that force a reliance on phonological detail. Indeed, Arabic's visual complexities (ligature and diacritics) were examined in Chapter 4.

Chapter 4

Study 3: Orthographic Learning in Arabic

Although the study of predictors of reading is a very useful starting point to understanding word recognition, it is clear from Chapter 3 that it is extremely difficult to determine the magnitude of the role these factors play and how they interact with the nature of orthography of a given language (Ehri, 2005a, 2005b; Castles & Nation, 2010). It is also difficult to ascertain whether the predictors identified are aspects of the process of orthographic learning or whether they are outcomes of the reading process. The fast and efficient recognition of words is the hallmark of reading development (Share, 2004). However, what is yet unclear are the mechanisms that underpin the learning process of word recognition. Several researchers have attempted to document this process from either a developmental approach (e.g., Ehri, 1994) or an item-based approach (Share, 1999, 2004). According to Castles and Nation (2010), the process of learning that leads to word recognition is referred to as orthographic learning.

The simulation of connectionist models comprising phonology (sound), orthography (symbol), and semantics (meaning) connected by hidden units (Plaut, 2005) seems to have prompted the research linking these constituents (Harm et al., 2003). The lexical quality hypothesis (Perfetti & Hart, 2002) has clear parallels to connectionist models. According to this hypothesis, a high-quality representation of the word is established when all three constituents are activated. The more exposure an individual has to the word, the stronger the connections become between the constituents, making the word become what appears to be a unitary lexical entry. The quality of word-specific representation defines how well it is identified and recalled (Verhoeven & Perfetti, 2011), which is consistent with Ehri's amalgamation theory (as cited in Ehri & Wilce, 1979). This theory proposed a bonding process in which all word identities (semantic, syntactic, phonological, and orthographic) are learned and then stored as a unitary entry (Ehri & Wilce, 1979). The process of forming word-specific representations during reading is best described by Share's self-teaching hypothesis (as cited in Share, 2004), which stipulates that, once the alphabetic rule is acquired, children will decode unfamiliar words (given that correct decoding occurs) and self-teach implicitly. According to the self-teaching hypothesis, even one encounter with an unknown word activates the orthographic constituent (Share, 2004). Converging evidence in support of this hypothesis comes from several studies (e.g., Cunningham, 2006; Nation et al., 2007; Ricketts et al., 2011; Shahar-Yames & Share, 2008). Indeed, even a single encounter with an unknown word activates the orthographic constituent (Share, 2004). Developmental, computational, and item-based accounts are consistent with statistical learning models. Based on these theoretical frameworks on the acquisition of word-specific representations, the contribution of the orthographic constituent was investigated in the current study.

Literature Review

In order to understand orthographic learning, Castles and Nation (2010) propose that the only way to address the underpinnings of this process is through training studies that use experimental designs to dissect the various components. Experimental paradigms such as Share's self-teaching and paired associate learning may provide insight into how this process happens (Deacon et al., 2012). Such experiments specify the factors that aid in the transition from the laborious alphabetic reading to skilled word recognition (orthographic learning). The following sections will review the literature on the different paradigms used to assess orthographic learning and predictors of orthographic learning.

What is Orthographic Learning?

Orthographic learning is an “umbrella” term that consists of word-specific representations and general knowledge about the orthographic regularities in one’s own writing system (Apel, 2011; Castles et al., 2018). According to Ehri (2005a), this phase is characterized by recognizing words rapidly and associating their spelling to their meaning without resorting to decoding. How children go from novice readers using phonological decoding (full alphabetic phase) to the seemingly effortless and automatic recognition of words has been the subject of extensive research. Phonological decoding during reading (the ability to segment individual graphemes, associate them with phonemes, and finally blend these into a word) is what drives the built-in self-teaching mechanism, according to Share (1995). The very learning of letter-sound-correspondence rests not only on phonological skills, but rather the pairing and mapping of the phonological (sound) onto the orthographic (symbol). Thus, learning the pairing of the auditory and visual stimulus is the gateway to literacy acquisition. To date, Share’s (1995) self-teaching hypothesis is the most prominent account of the process involved in orthographic learning. According to Share, the novice reader will encounter unfamiliar words, and it is through the decoding process at first and then each subsequent encounter that provides an opportunity for orthographic learning. Share's account is consistent with Perfetti and Hart's (2002) lexical quality hypothesis which stipulates that the frequency with which words are encountered in print determines how well specified they are in the mental lexicon. According to their hypothesis, stored mental representations that are well specified (high lexical quality) are the hallmark of word recognition. With exposure, a bonding of the word's form, pronunciation, and meaning creates a high-quality lexical representation. This interaction with exposure to print is considered to shape the word recognition system in what is referred to as lexical tuning (Castles et al., 2007). Exposure to print plays a major role in determining words that are of high- and low-quality lexical representation and thus will be controlled for in this study. The following review of the literature will address the mechanism that underpins this learning process.

Orthographic Learning Paradigms

In perhaps one of the very first studies to examine orthographic learning using a training paradigm, Ehri and Wilce (1979) presented first- and second-grade students with nonwords that were squiggles, initial letter, misspelled, or correctly spelled, during training. Ehri and Wilce were able to show that children recalled nonwords best in the correctly spelled condition, followed by initial letter, as opposed to misspellings and squiggles. In fact, children in the squiggle condition failed to learn 67% of the nonwords. The authors argued that the results indicated that, once the alphabetic phase is achieved, children rely on spellings to store and recall whole-word representations.

The Self-teaching Paradigm

Share (1995) was one of the first to provide an account of the process of orthographic learning which he termed the self-teaching hypothesis. The main premise of the self-teaching hypothesis is that the letter-sound conversions and subsequent decoding skills provide the novice reader an opportunity to establish word-specific orthographic representations. In order to empirically test this hypothesis, Share (1999) devised a naturalistic experiment that mimicked to a great degree how children acquire orthographic representations through independent reading (see page 45 for details). There was clear evidence of orthographic learning on all outcome measures. On the orthographic choice task, the children were more likely to choose the target. On the naming task, reaction times were significantly faster for the targets than for the foils. Similarly, the children were more likely to spell the nonwords as seen in the stories rather than spell a homophone. However, the number of exposures did not have a significant effect on orthographic learning, at least in a transparent orthography such as Hebrew. Share speculated that in transparent orthographies, children do not require many exposures because of the consistency of spelling-to-sound, whereas in opaque orthographies, more exposures would be necessary for orthographic learning to occur. Nonetheless, orthographic learning was achieved in as little as four exposures. More importantly, orthographic learning did not occur under conditions when phonological decoding was minimized.

Share (2004) then conducted a series of experiments to investigate the role of number of exposures and how long children retain the words learned via self-teaching. In order to address these issues, Share manipulated the number of exposures to the nonwords (appearing once, twice, or four times), and post-tests of orthographic choice, naming, and spelling were administered after 3, 7, and 30 days to assess orthographic learning. In experiment 1, orthographic learning as measured by orthographic choice, naming, and spelling was evident regardless of the number of exposures for the sample of third graders. These results show that a single exposure to a novel word is sufficient for orthographic learning. Since Hebrew is a highly transparent language and since successful phonological decoding is the building block of orthographic learning, Share reasoned that even novice readers of Hebrew would demonstrate orthographic learning. Thus, a sample of first graders participated in the second experiment using the same stimuli and procedure as in the first experiment. The first graders decoded 93% of the nonwords correctly, matching that of the third graders demonstrating that, even by first grade, reading in a transparent language is near ceiling after the first year of reading instruction. Despite the high phonological decoding accuracy, orthographic learning was not evident for the first graders. Share reasoned that “competent novices of regular orthographies are capable of decoding nearly any word in a ‘bottom-up’ or ‘surface’ fashion but appear to be relatively insensitive to surface information, that is, to word-specific orthographic detail” (p. 289). In other words, implicit learning may be facilitated by the consistency of spelling-to-sound and the finer granularity used in decoding unfamiliar words (Castles & Nation, 2010; Arciuli, 2018; Elleman et al., 2019). Indeed, findings from other languages reveal the complex relationship between orthographic learning and how it is modulated by orthographic depth.

Cunningham et al. (2002) examined orthographic learning in English-speaking second graders. Using the self-teaching paradigm, and the same outcome measures as Share (1999), 34 children read aloud 10 short stories containing 10 pairs of homophonic nonwords (e.g., rupe/roop; beel/beal). After three days, the children were administered a homophonic nonword choice task, a spelling task, and a naming task to assess orthographic learning. Cunningham and colleagues assessed reading, through nonword decoding accuracy, orthographic choice task, cognitive measures, digit span, receptive vocabulary, nonverbal ability, and RAN. They found orthographic learning had occurred as evidenced by selecting targets over homophones three days after exposure. More importantly, they found that successful decoding of the targets was positively related to orthographic learning, thus providing evidence for Share's claim that phonological decoding is the sine qua non of word recognition. Of the control measures, only nonword decoding and orthographic choice predicted orthographic learning. However, the orthographic choice task employed in this study has been criticized for assessing word recognition rather than orthographic knowledge (Castles & Nation, 2010). Furthermore, Bowey and Miller (2007) pointed out that nonword reading accuracy of targets is not a pure measure of phonological decoding since it only shares 27% of the variance with the Word Attack task (a standardized nonword reading task of increasing difficulty). In fact, when Bowey and Miller reanalyzed Cunningham et al.'s data partialing out Word Attack, phonological decoding accuracy was no longer significantly related to orthographic learning. Nonetheless, phonological decoding accuracy was the strongest predictor of orthographic learning.

Nation et al. (2007) were also interested in investigating orthographic learning among English second and third graders. Nine nonwords were either presented in a story context or in isolation, appearing once, twice, or four times. Orthographic learning was assessed 1 day after exposure and 7 days later using only the homophonic choice task (e.g., the target *ferd* was paired with the homophonic *furd*, and distractors *ferp*, and *furp*). They were also interested in examining the relationship between phonological decoding and orthographic learning. Context had no effect on orthographic learning. Unlike in Hebrew (Share, 2004), more exposures to English nonwords significantly increased learning. The authors argued that increased exposures in an opaque orthography such as English gives more opportunities for successful decoding and ultimately more opportunities for orthographic learning. Orthographic learning was better after one day and significantly decreased after a seven-day delay. These results are also inconsistent with those of Share (2004), who demonstrated robust learning even after a 30-day delay. Furthermore, although Nation et al. detected significant levels of orthographic learning (48% after one day, and 40% after seven days), these levels were much lower than those reported by Share (2004). There are two key methodological differences that may account for these differences. Share's homophonic choice task only included the target and a homophone, whereas Nation et al., like Cunningham et al., included an additional two foils. Share also administered a spelling task as an outcome measure of orthographic learning before the homophonic choice task, whereas Nation et al. did not assess spelling, which may have increased exposure to targets in Share's sample and as a result inflated results on the homophonic choice task. However, since Cunningham et al. (2002) used the more stringent homophonic choice task (target and three foils) and found accuracy levels like the ones reported by Share (2004), the methodological differences between Nation et al. and Share may not be the culprit behind these discrepancies. Perhaps these differences between the studies highlight the interaction between orthographic learning and orthographic depth. Indeed, Share (2004) argued that orthographic sensitivity may be more crucial in opaque orthographies since the novice reader must be more attuned to letter-sound representations that are inconsistent, whereas this is not necessary when decoding a transparent script. Ehri (2005b) alluded to the issue by asking whether orthographic learning is as relevant for transparent orthographies as it is for English.

Bowey and Miller (2007) examined orthographic learning during silent reading, using the self-teaching paradigm in a sample of English-speaking third graders. They were also interested in investigating the existing reading and reading-related abilities as potential predictors of orthographic learning. They assessed children on word reading measures and RAN tasks. Children silently read short stories containing six exposures to the target nonword. Orthographic learning was assessed using a homophonic choice task, and list reading that was composed of a set of 10 target words and 10 homophones. After exposure, targets were read faster and were chosen more often than were homophones on the homophonic choice task, indicating orthographic learning had occurred. The authors argue this pattern of results is indicative of phonological decoding during silent reading. “If children had not phonologically recoded the target nonwords as they read them within the stories and retained at least partial phonological representations, then they would not have been able subsequently to read them faster than homophones” (p. 124). Thus, the construction of orthographic representations is achieved during silent reading as well. Phonological decoding was the only predictor of orthographic learning (18%), thus lending additional support that orthographic learning is dependent on phonological decoding. These results are not consistent with findings from Cunningham et al. (2002), where phonological decoding was not consistently associated with the homophonic choice task (orthographic learning). RAN did not predict orthographic learning.

Ricketts et al. (2011) investigated orthographic and semantic aspects of novel word learning in a sample of English-speaking 8-year-old children using Share’s (1999) self-teaching paradigm. They were also interested in predictors of semantic and orthographic learning. Children were exposed to eight nonwords in stories where four nonwords were presented with contextual cues and four nonwords were presented with ambiguous cues. Each nonword appeared four times in the story. After the stories were read aloud and decoding accuracy was recorded, three outcome measures were administered to assess orthographic and semantic learning of the nonwords: a homophonic choice task (target and three foils), a spelling task, and a nonword-picture- matching task (target nonword presented alongside four pictures) to assess semantic learning. Reading, reading comprehension, and expressive vocabulary were assessed to examine predictors of orthographic and semantic learning. The findings in this study replicate the findings described above. There was clear evidence of orthographic learning as indexed by the homophonic choice task and spelling task. Furthermore, there was a correlation between orthographic learning and nonword reading fluency as measured by TOWRE (Torgesen et al., 1999), confirming that phonological decoding is, in part, one of the mechanisms that underlie orthographic learning. Moderate to high correlations were observed between target decoding accuracy and orthographic learning, whereas vocabulary was the strongest predictor of semantic learning. Thus, learning of the orthographic and semantic identities of novel words is separable, as evident from the different predictive patterns.

Apel (2011) posited that spelling tasks assess both mental orthographic representations and orthographic pattern knowledge. Shahar-Yames and Share (2008) reasoned that the analysis of a word's spelling requires more attention to grapheme-phoneme relationships than to decoding and thus would provide superior orthographic learning. These authors tested their reasoning in a sample of 45 Hebrew-speaking third-grade students and found that analysis of a word's spelling provided children with the opportunity for orthographic learning.

Tamura et al. (2017) used a self-teaching paradigm to investigate the process of orthographic learning in 30 native English-speaking children ages 9 to 11. They make a distinction between the initial phases of orthographic learning (lexical configuration) and later phases of lexicality (lexical engagement). The first experiment examined whether prime-lexicality effects (lexical engagement marker) are present after exposure to newly encountered words and if this occurs through deliberate instruction. Lexical configuration was indexed by three outcome measures (orthographic decision, meaning, and spelling), and lexical engagement was indexed by prime-lexicality effects. They divided the children into two groups (read-only aloud group, reading-plus aloud group). The children read stories that contained low-frequency English words. The children in the read-plus read the same stories as the read-only group but engaged with the words by practicing their spellings and meanings (explicit practice of unfamiliar words). Children in both conditions were exposed to the same unfamiliar words, and the number of times the word appeared in the text was the same across conditions. To assess learning after exposure to the words, the authors administered an orthographic decision task, a spelling test, and definition test. In order to assess lexicalization of the words, they administered a masked priming lexical decision. This was administered also pre-exposure, to establish a baseline. In other words, before exposure the words should exhibit a facilitative priming effect, whereas after exposure, this effect should disappear. They predicted that learning of lexical engagement in the reading-plus group would be greater than in the read-only group and that would be evident by the emergence of inhibition as a result of competition between the newly learned words and their neighbors. Learning was evident in both read-only and read-plus groups as measured on all three orthographic configuration measures (orthographic decision, meaning, and spelling). Spelling data pre-exposure confirmed that spelling was significantly better for the unfamiliar words after the exposure phase for both groups, demonstrating orthographic learning. Also, children with higher performance on the reading task (TOWRE) had significantly better performance on the orthographic decision task but not on spelling or meaning. The fact that orthographic decision and reading are related should be interpreted with caution since measures of orthographic learning are contaminated with reading ability (Castles & Nation, 2010). Similarly, lexicalization was evident in both groups after exposure to the unfamiliar words, indicating lexical engagement as indexed by the lexical decision task. These results demonstrate that lexical engagement occurs with the mere exposure to novel words in text. Taken together, the results showed robust, rapid, and seemingly effortless orthographic learning. Based on these results, a second experiment was conducted; however, they eliminated the read-plus condition, and instead they manipulated the number of exposures (4 versus 12) for each unfamiliar word. Similar results were obtained, demonstrating that number of exposures did not have an effect on lexical configuration; however, increased number of exposures did have a significant effect on lexical engagement.

As demonstrated in Chapter 3, Arabic’s linguistic and visual characteristics may impede a seemingly transparent orthography when fully vowelized. Studies examining orthographic learning in Arabic are scant. Dai et al. (2013) examined the role of ligature (connectivity) in Arabic among third-grade students, using the self-teaching hypothesis as a framework (Share, 1999). They constructed nonwords that were either connected or non-connected and hypothesized that decoding the non-connected nonwords would be easier (accuracy and fluency, as measured by RTs), but after exposure (once orthographic learning had occurred), connected items would be faster and more accurate to read. In their Experiment 1, the results were not consistent with their predictions and revealed that connected items were read over half-a-second longer than were unconnected items, showing that connectedness did not facilitate word recognition. Similar results were obtained on the spelling post-test, which also revealed that connected items were more error prone. The authors reasoned that these results might have been obscured by the fact that unconnecting letters have fewer constant diacritics (dots appearing in, above, or below consonants). To test this, they conducted a second experiment and controlled for the number of consonant diacritics. This resulted in three sets of words: those connected with few diacritical marks, those connected with many diacritical marks, and those non-connected. Dai et al. found significant effects of diacritics on accuracy (error rates of 3.3% for a few diacritical marks versus 6.3% for many diacritical marks). The authors concluded that “the apparent effect of connectedness in Experiment 1 was not attributable to ligaturing per se, but to the complex diacritics that often accompany connected letters” (p. 11).

Tamura et al. (2017) make a distinction between the initial phases of orthographic learning (lexical configuration) and later phases of lexicality (lexical engagement). Dai and colleagues reasoned that lexical configuration (initial phases of orthographic learning) would be more challenging for items with connecting letters because it places greater demands on graphemic parsing, but once lexical engagement is achieved, connected letter items would be read faster in whole-unit fashion. As predicted, connected letter items were read slower than were non-connected letter items during the lexical configuration phase. However, contrary to their predictions, connected letter items were still read slower than were non-connected letter items in the lexical engagement phase. Children made significantly more spelling errors on connected letter items, indicating that connectedness in Arabic impedes the creation of an integrated lexical unit.

With consonant diacritics controlled for in both the non-connected and connected letter items, the connected letter items were read faster during the lexical configuration phase and lexical engagement phase, demonstrating that reading in Arabic is facilitated by letter connectedness both at the initial phases of word reading and later word recognition. Indeed, ligatured letters in Arabic occur more frequently than do non-ligatured letters, and this feature is assumed to support reading fluency (Abdelhadi et al., 2011). Items with many diacritics slowed down reading considerably, demonstrating that connected letter items were not the source of difficulty but rather the many inherit consonant diacritics that come with connected letters are. Reading accuracy of the items, whether connected or unconnected, was not a sensitive measure of orthographic learning, owing to Arabic’s near one-to-one grapheme-to-phoneme correspondence; however, reading fluency was negatively impacted by items that contained many diacritics. Similarly, connected letter items with few consonant diacritics were spelled more accurately than were connected items with many diacritics.

The Paired Associate Learning Paradigm

Studies examining self-teaching as the mechanism underlying orthographic learning point to phonological decoding as being crucial and that correct target decoding during the exposure phase is what determines if learning occurs (e.g., Share, 1999, 2004; Cunningham et al., 2002). However, studies using a different paradigm have demonstrated that orthographic learning may not be a conscious process but rather an implicit learning process that is automatically activated when a word's phonological and orthographic identities are encountered (Ricketts et al., 2009; Rosenthal & Ehri, 2008). In fact, the Stroop effect is a classical example demonstrating how word recognition is an automatic unconscious process (Stroop, 1935).

Studies have demonstrated a relationship between paired-associate learning and reading development (Litt & Nation, 2014; Wang et al., 2017). It is proposed that paired-associate learning ability aids in acquiring connections between phonology and orthographic representations of the word. This, in turn, further supports the consolidation of orthographic knowledge via activation of the phonological and orthographic identities of the word (Wang et al., 2017).

Rosenthal and Ehri (2008) reasoned that orthographic mapping to the phonological form (pronunciation) of a word and its semantic information should optimize learning and memory storage of new words. In an experiment involving 20 second graders, Rosenthal and Ehri used visual-verbal paired associated learning (PAL) to assess learning of newly taught words. Unlike self-teaching, no attention was drawn to the word's spelling and children were not required to decode the words during training; instead, presentation of spelling was incidental. The results favored the orthography-present condition in both pronunciation recall and the definition recall of low-frequency words. A second experiment conducted with 32 fifth-grade children mirrored results from their first experiment with the second-grade children, favoring the incidental presence of orthography (Rosenthal & Ehri, 2008). Moreover, the pronunciation post-test demonstrated that learning to pronounce newly taught words renders them more difficult to recall than does the recall of the meaning. Similarly, Ricketts et al. (2009) investigated the role of orthography in vocabulary acquisition in 58 8- to 9-year-old children. The authors used a PAL paradigm like the one used by Rosenthal and Ehri and trained the children to associate 12 nonwords with pictures of novel objects. Participants' learning of the pairings was assessed using spelling and nonword-and-picture-matching tasks. Ricketts et al. demonstrated that stimuli presented with orthography, but without drawing attention to it, resulted in better recall and more accurate spelling. In other words, the presence of orthography improved the children's learning of the newly taught items. Results such as these cast some doubt about the necessity of successful phonological decoding as a prerequisite to orthographic learning since the spelling of the words is incidental and presentation during training is too brief to allow for phonological analysis to occur.

Orthographic facilitation in novel word learning was demonstrated in second language learning (e.g., Hu, 2008; Cerni et al., 2019). Cerni et al. (2019) found that Italian L1 speakers produced pronunciations that were inconsistent with English L2 phonology but consistent with word spellings in the incidental exposure phase; for example, double letters producing longer consonant pronunciation in Italian that does not exist in English and inconsistent with the pronunciation provided during the training phase. Such results underscore the powerful and implicit effects of orthography on novel word learning.

Previous studies demonstrated that the deliberate analysis of a word's orthography and even incidental orthographic presentation facilitates recall of a word's phonological, orthographic, and semantic identities.

Taken together, these studies provide initial support for the theoretical framework that the amalgamation of all of a word's identities produces higher-quality representations of newly taught words. What is unclear are the underlying skills that predict individual differences in orthographic learning.

What Predicts Individual Differences in Orthographic Learning?

Verbal STM has been found to predict new word learning (Gathercole, 2006). The ability to repeat unfamiliar phonological forms has been linked to oral vocabulary acquisition (Gathercole et al., 1999). This relationship has long been established in the literature (Cain, 2010; Vance & Mitchell, 2006). However, the correlation between verbal STM and the ability to learn new phonological forms seems to decrease with age, probably as a result of efficient word recognition that decreases the load on the phonological loop (e.g., Jarrold et al., 2004; Jarrold et al., 2009). Furthermore, verbal STM capacity has been said to overlap with PA and may not uniquely contribute to vocabulary learning (Parilla et al., 2004; Martinez Perez et al., 2012).

Utilizing longitudinal data collected three years previously, Powell et al. (2014) had the opportunity to investigate the persistence of effects of RAN deficits on reading in a group first assessed in second and third grade with poor RAN and controls matched on PA and other cognitive factors. Nearly three years later, the same children were assessed on alphanumeric RAN (digit, letter), PA (elision, blending), print exposure (title recognition test), orthographic choice (e.g., rain/rane), and an orthographic pattern task that required testees to judge which one of two nonwords resembled a real word and that was adapted from Cassar and Treiman (1997). Finally, the children were assessed on reading regular words, irregular words, and nonwords. To assess orthographic learning, eight nonwords were embedded in a naming and lexical decision task where the participants read the items and decided whether it was a real word or a nonword. This was the exposure phase of the learning where target items appeared either once or four times. The exposure phase was followed by a recall phase right after exposure and one week later. Participants were asked to recall the nonwords encountered by selecting the target from an array of four similar nonwords (e.g., ferd, furd, ferp, furp). The low RAN group continued to show impaired RAN compared to controls and had poorer reading skills nearly three years after they were initially identified demonstrating the persistence of RAN deficits. Although it was hypothesized that differences in print exposure would be associated with differences in RAN ability, there was no difference in print exposure between low RAN participants and controls. Thus, the persistence of RAN deficits cannot be caused by low exposure to print, nor can orthographic deficits in low RAN be attributed to lower experience with print. Furthermore, initial analysis revealed that the low RAN group consistently performed poorer on measures of reading, orthographic choice (lexical), and sub-lexical tasks. Once decoding ability was controlled for, the low RAN group did not significantly differ from controls on orthographic knowledge as indexed by the orthographic choice task. However, the orthographic pattern task (sub-lexical) accounted for significant variance for the low RAN group even after controlling for decoding ability. This seems to suggest a specific deficit in implicit statistical learning of commonly occurring orthographic patterns in the low

RAN group. An intriguing finding was the superiority of the low RAN group on the orthographic learning task compared to the control group. While the finding is speculative, the authors argue that, because reading is slow and effortful for the low RAN group, they may have paid more attention to decoding and thus deeper processing of the nonwords. Another possibility is that the low RAN group may have adopted a visual memorization strategy to compensate for their deficiencies.

Tasks of incidental orthographic learning during reading (e.g., Ricketts et al., 2009, Rosenthal & Ehri, 2008) demonstrate the implicit acquisition of word-specific representations after a few exposures. Although the use of paradigms like PAL may be criticized for not being as naturalistic as the self-teaching paradigm, this is not the case in a diglossic situation, where novel word pairing is the norm and not the exception.

The Present Study

Due to the diglossic situation of Arabic, it is important to determine how to facilitate novel word learning in the absence of oral comprehension. Thus, this study examined the facilitative effects of orthography on novel word learning using a paired associate learning task. In addition to diglossia being implicated as the source of literacy difficulties, Arabic's visual complexities have been implicated. Yet, what is not well understood is whether these visual complexities stem from ligature or diacritics. Thus, this study isolated these two potential sources of visual complexities by using nonwords that manipulated conditions of ligature and consonant diacritics. Finally, this study examined participant-level predictors of novel word learning to determine which of children's existing cognitive and linguistic abilities might facilitate novel word learning. Thus, this study addressed the following research questions:

1. What are the cognitive and linguistic abilities that may facilitate the learning of newly taught words?

2. Does incidental exposure of orthography facilitate recall (phonological, orthographic and semantic) identities of newly taught words?

3. Which item condition (non-connected few consonant diacritics, connected few consonant diacritics, connected many consonant diacritics) are facilitated by exposure to orthography?

Method

This study employed a correlational design to investigate the relationships between the variables of interest and an experimental within group repeated measures design. Two types of correlational design were used: a relational one, where all variables, using quantitative methods were analyzed to investigate whether a relationship exists; and a design one that involved having outcome variables, where one or more variables can predict an outcome variable (multiple regression). Both types of correlational analysis were used to address the research questions. The repeated measures design involved two presentation conditions (orthography absent vs. orthography present) and three word conditions (non-connected few consonant diacritics vs. connected few consonant diacritics vs. connected many consonant diacritics) as repeated measures.

Participants

Participants were 116 monolingual children from third grade; (58 females and 58 males; mean age 8.10 ± .45). All participants were native Arabic speakers attending a private school located in a middle-class neighborhood in Cairo. Private schools offer a national curriculum delivered entirely in Arabic and receive only one 40-minute ESL class a day. The selection of monolinguals from a private school rather than a public school was deemed more appropriate to control for homogeneity of the sample in SES and other environmental factors. Participants were randomly selected from four third-grade classrooms. Potential participants were excluded from the study if they had any sensory or cognitive impairments.

Materials

The materials included background assessments of nonverbal reasoning, visual attention, verbal STM, working memory, receptive vocabulary, and print exposure. Reading related skills were assessed though measures of PA, RAN, orthographic/visual processing, and morphological production. Finally, outcome measures of reading (word recognition, nonword reading fluency) and single word spelling were administered. All Arabic measures used in this study are described in detail in Study 2 (see pp. 143–154). This section describes the measures that were used. Table 37 lists the assessments used in this study.

Table 37

Arabic Measures Used in Study 3

Areas/processes assessed	Measures
Nonverbal control measures	Nonverbal reasoning, visual concentration
Control measures	print exposure, receptive vocabulary
PA	Elision, blending words
Verbal STM	Nonword repetition, digit span forward, digits backward
RAN	RAN-digit, RAN-letter
Orthographic processing	orthographic matching, pseudo-letter matching
Morphological measure	Morphological production
Reading measures	Word recognition, nonword reading fluency
Spelling measures	Spelling to dictation

Experimental Stimuli

Stimuli used in this study were developed on the basis of those used by Ricketts et al. (2009) and Dai et al. (2013). The picture referents were adapted from Ricketts et al. (2009) (see Appendix G) and the nonwords were adapted from Dai et al. (2013). The nonwords were designed in accordance with the visual complexities of the Arabic language and were constructed in three groups: four items that are non-connected and have few consonant diacritics (e.g., وزرق), four items that are connected and have few consonant diacritics (e.g. عسلز), and four items that are connected and have many consonant diacritics (e.g. ظحفك) (see Table 32 for all items). The items were designed to be equally matched in length (four constants) and were all vowelized with *fatHa* diacritic (Dai et al., 2013). All orthography conditions and item conditions were counterbalanced.

Table 38

Word Groups in Experimental Stimuli for Study 3

Non-connected few consonant diacritics	Connected few consonant diacritics	Connected many consonant diacritics
زارض	عبيم	ظحفك
زروغ	خسيل	جشيك
دزوم	ضجمل	غشلز
وزرق	عسلز	قثنب

Each word was presented aloud as its accompanying picture (depicting the noun) was displayed. The pictures were presented using PowerPoint. For one set of items, the words were spelled out and presented at the top of the picture, but no attention was drawn to the presence of the word.

The nonwords were split so that each set included six nonwords in the orthography-present condition and six nonwords in the orthography-absent condition. In addition, the item conditions were counterbalanced across the orthography conditions. This split produced six possible combinations, which were counterbalanced across participants so that every six children received each possible combination.

Training

Each child was instructed in Arabic as follows: “In this game, we are going to learn some new words for things that aliens might use. I want you to listen carefully and try to remember which picture goes with which word.” The instructions were adapted and combined from those used by Ricketts et al. (2009) and Hulme et al. (2007).

Using a PAL paradigm, six training sessions were held for each of the sets of items: three repetition sessions and three production sessions. In the repetition session, the experimenter presented a picture, followed by the pronunciation of the word. The child repeated the nonword heard; if pronounced incorrectly, the experimenter provided the correct pronunciation.

During the production session, the child was presented with a picture without orthography and was asked to recall the pronunciation that corresponded with it. The child was instructed as follows: “What word goes with this picture?” The child was given approximately two seconds to respond. The experimenter, irrespective of whether the child pronounced the nonword correctly or incorrectly, pronounced the word, but words in the orthography-present condition were presented with orthography when the experimenter pronounced the word.

In total, each child was exposed to the 12 words six times during the training sessions. In the six sessions, all nonwords were presented for the same amount of time, irrespective of orthography condition (at approximately two-second intervals using PowerPoint).

Post-tests

Three performance post-tests were administered after training to examine whether there were facilitative effects of orthography on the recall of pronunciation (phonological form), recall of spelling (orthographic form), and recall of meaning (semantics) information.

For the pronunciation dependent variable, each child was presented with the referents of the 12 nonwords, one at a time, and asked to provide the corresponding pronunciation. In this task, the experimenter asked, “What word goes with this picture?” as in the production session of the PAL training. This post-test is intended to measure the children’s learning of the phonological identity of the newly taught nonwords.

For the nonword-and-picture-matching dependent variable, each child was given a nonword (oral pronunciation) and asked to choose from one of four pictures presented, following Ricketts et al.’s (2009) process. The target and three nonwords from the same experimental stimuli were presented in a 2×2 grid, and the child had to point to the corresponding picture when the nonword was pronounced by the examiner. This post-test was similar to the receptive vocabulary test The Peabody Picture Vocabulary Test (PPVT III; Dunn & Dunn, 1997). For example, the child was asked, “Can you point to ضجمل?” The presentation of the 12 nonwords was counterbalanced so that each picture appeared an equal number of times.

Finally, for the spelling dependent variable, the 12 nonwords were dictated to each child individually. The child was given paper, pencil, and eraser and was asked to spell each item that the experimenter pronounced.

Procedure

All testing was conducted in a quiet room at the participants' school. Each child was seen in two individual sessions and one group session. Assessment of nonverbal ability was group administered along with spelling. In the first individual session, control measures (visual concentration, verbal STM, digits backward, receptive vocabulary, and TRT), and reading- related measures (PA, RAN, orthographic matching, pseudo-letter matching, and morphological production) were administered. During the second individual session, reading outcome measures (word recognition, nonword reading fluency) were administered, along with the nonword training sessions, the nonword-and-picture-matching post-test, pronunciation post-test, and spelling post-tests in that fixed sequence. The spelling post-tests were administered last, so as to not increase exposure to the nonwords and rule out the possibility of contaminating the performance on the pronunciation and nonword-and-picture-matching post-tests (see Nation et al., 2007).

Results

Descriptive statistics for all measures are presented, followed by correlations between measures and hierarchical regressions exploring predictors of reading and spelling. In order to address research questions regarding orthographic learning, the data were analyzed using analysis of variance, in order to examine the dependent variables' presence or absence of orthography and type of nonword. Three separate repeated measures univariate 2 (orthography present vs. orthography absent) × 3 (non-connected (NC) vs. connected few (CF) vs. connected many (CM) ANOVAs were conducted—one each for the pronunciation, matching, and spelling dependent measures.

Descriptive Statistics and Preliminary Analysis

To reduce the number of variables and simplify further analyses, three composite scores were created for PA measures (elision and blending words), verbal STM measures (nonword repetition and digit span forward), and RAN (letters and digits). This was done by converting the raw scores into *z* scores and averaging the two. All three composite scores were used in subsequent analysis. Descriptive statistics for all 116 children are provided in Table 39.

Table 39

Descriptive Statistics for Control Measures, Reading-related Measures, and Reading and Spelling in Study 3

Measure	Max. Score	Mean	SD	Min	Max
Nonverbal control measures					
NVA	60	17.1	7.3	5	33
Visual Concentration	30	11.4	2.2	6	16
Control measures					

Nonword repetition	20	9.5	3.6	2	19
Digit span forward	13	7.4	1.7	4	11
Digits backward	13	4.7	1.5	0	9
Receptive vocabulary	103	55.4	10.9	29	79
TRT	20	4.1	3	-3	13
Reading-related measures					
Elision	20	12.7	3.9	2	19
Blending	20	5.6	2	0	12
RAN Digit (seconds)		18.6	3.3	11	32
RAN Letter (seconds)		24.1	5.8	15	46
Orthographic matching	25	11.7	3.4	4	20
Pseudo-letter matching	25	10.5	3	1	16
Morphological production	29	15	4.7	5	25
Reading and spelling measures					
Word recognition	50	23.2	11.3	1	46
Nonword reading fluency	35	14.6	7.3	0	31

Spelling	38	20.6	6.9	5	32

Note. SD=Standard Deviation

Descriptive statistics for the scores in the three post-test assessments are given in Table 40. The data presented are for all 116 children in all conditions.

Table 40

Descriptive Statistics for the Orthographic Leaning Post-test Assessments in Study 3

	Pronunciation			Picture-matching			Spelling		
	NC	CF	CM	NC	CF	CM	NC	CF	CM
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
OP	0.69 (.69)	0.84 (.72)	0.45 (.70)	1.22 (.68)	1.34 (.71)	1.14 (.75)	1.34 (.74)	1.44 (.73)	0.91 (.74)
OA	0.16 (.36)	0.27 (.53)	0.02 (.13)	1.16 (.78)	1.09 (.71)	0.97 (.68)	1.04 (.68)	1.1 (.81)	0.66 (.72)

Note. OP=Orthography present, OA=orthography absent, NC=Non-connected, CF=Connected few, CM=Connected many, SD=Standard deviation.

Pearson correlations were conducted between all variables and the pronunciation, matching, and spelling post-test scores in the orthography-present condition.

Table 41

Correlations between All Measures and Orthographic Learning Post-test Assessment Scores in the Orthography Present Condition in Study 3

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. NVA															
2. VC	.41***														
3. VSTM	.24**	.07													
4. DB	.17	.06	.19*												
5. RV	.26**	.12	.36***	.35***											
6. PA	-.07	.03	.07	0.05	.40***										
7. RAN	.38***	.14	.40***	.35***	-.21*	-.38***									
8. OM	-.23*	-.23*	-.10	-.19*	.19*	.25**	-.33***								
9.P M	.32***	.36***	.05	.25**	.11	.25**	-.33***	.53***							
10. MP	.38***	.23*	.44***	.22*	.62***	.39***	-.25**	.21*	.29**						
11. WR	.37***	.23*	.33***	.36***	.37***	.64***	-.44***	.34***	.46***	.48***					
12. NWRF	.31**	.26**	.27**	.41***	.35***	.62***	-.46***	.33***	.43***	.44***	.83***				
13. SWS	.33***	.20*	.36***	.34***	.36***	.59***	-.46***	.33***	.41***	.45***	.79***	.75***			
14. POP	.22*	.22*	.27**	.21*	.17	.27**	-.21*	.04	.15	.24*	.32***	.35***	.29**		
15. MOP	.07	.14	.25**	.16	.13	.19*	-.16	.09	.20*	.15	.26**	.21*	.18*	.56***	
16. SOP	.17	.21*	.26**	.15	.27**	.47***	-.38***	.09	.23*	.31**	.48***	.48***	.54***	.43***	.34***

Note. NVA=Nonverbal ability, VC=Visual concentration, VSTM=Verbal short-term memory, DB=Digit backward, RV=Receptive vocabulary, PA=phonological awareness, RAN=Rapid automatized naming, OM=orthographic matching, PM=Pseudo-letter matching, MP=Morphological production, WR=Word recognition, NWRF=Nonword reading fluency, SWS=Single word spelling, PPT=Pronunciation post-test, MPT=Matching post-test, SPT=Spelling post-test. Orthographic learning in the orthography-present condition for Pronunciation (POP), Matching (MOP), and Spelling (SOP) were computed as overall recall of items independent of orthographic and item conditions. **p < .05, **p <.01, ***p <.001*

NVA was correlated with pronunciation post-test (*r* = .22, *p* = .02). Visual concentration was correlated to pronunciation post-test (*r* = .22, *p* = .02) and spelling post-test (*r* = .21, *p* = .03). Verbal STM was correlated to pronunciation post-test (*r* = .27, *p* = .003), matching post-test (*r* = .25, *p* = .007), and spelling post-test (*r* = .26, *p* = .004). Digits backward was correlated to pronunciation post-test (*r* = .21, *p* = .02). Receptive vocabulary was correlated to spelling post-test (*r* = .27, *p* = .003). Scores on the title recognition test did not correlate with any measure and are not included in this analysis or in any further analyses.

PA was correlated to pronunciation post-test (*r* = .27, *p* = .004), matching (*r* = .19, *p* = .04), and spelling post-test (*r* = .47, *p* < .001). RAN was negatively correlated pronunciation post-test (*r* = -.21, *p* = .02) and spelling post-test (*r* = -.38, *p* < .001). Pseudo-letter matching was correlated to matching post-test (*r* = .20, *p* = .04) and spelling post-test (*r* = .23, *p* = .01). Morphological production was correlated to pronunciation post-test (*r* = .24, *p* = .01) and spelling post-test (*r* = .31, *p* = .001). Orthographic matching was not correlated to any of the orthographic learning post-tests and is excluded from any further analysis.

Word recognition was correlated to pronunciation post-test (*r* = .32, *p* < .001), matching post-test (*r* = .26, *p* = .005), and spelling post-test (*r* = .48, *p* < .001). Nonword reading fluency was correlated to pronunciation post-test (*r* = .35, *p* < .001), matching post-test (*r* = .21, *p* = .03), and spelling post-test (*r* = .48, *p* < .001). Finally, spelling was correlated to pronunciation post-test (*r* = .29, *p* = .002), matching post-test, (*r* = .18, *p* = .049), and spelling post-test (*r* = .54, *p* < .001).

Predictors of Orthographic Learning

The first aim of this study was to investigate predictors of novel word learning. Based on the correlations above, three separate seven-step regression analyses were also performed for each post-test in the orthography-present condition (pronunciation, matching, and spelling) and are presented in Table 40. Age was entered in step 1. In step 2, nonverbal control variables were entered (NVA and visual concentration). In the third step, control measures were entered next (verbal STM, digits backward, and receptive vocabulary). In steps 4 through 7, PA, RAN, orthographic measures (orthographic matching, pseudo-letter matching) and morphological production were entered, respectively.

Table 42

Summary of Hierarchical Regressions for Predictors of Orthographic Learning for the Orthography- Present Condition in Study 3

		Pronunciation			Picture-matching			Spelling		
Step	Predictor	β^a	β^b	ΔR2	β^a	β^b	ΔR2	β^a	β^b	ΔR2
1	Age	.09	.03	.01	.08	.04	.01	-.03	-.11	.001
2	NVA	.07	.03	.13*	-.08	-.11	.09	.01	-.11	.14**
	VC	.16	.14		.14	.10		.18	.12	
	VSTM	.21*	.18		.23*	.23		.18	.09	
	DB	.14	.12		.11	.06		.05	-.05	
	RV	.02	-.05		.02	.03		.16	.03	
3	PA	.12	.09	.01	.09	.05	.01	.42**	.34**	.12***
4	RAN	-.09	-.09	.01	-.09	-.06	.01	-.23**	-.22*	.04*
5	PM	-.01	-.02	0	.13	.14	.01	.07	.07	.01
6	MP	.06	.06	.002	-.04	-.04	.001	.08	.08	.003

Note. NVA=Nonverbal ability, VC=Visual concentration, VSM=Verbal short-term memory, DB=Digits backward, RV=Receptive vocabulary, TRT=Title recognition test, PA=Phonological awareness, RAN=Rapid automatized naming, OM=Orthographic matching, PM=Pseudo-letter matching, MP=Morphological production. Orthographic learning post-tests were computed as recall of items in the orthography-present condition. β^a =Standardized beta coefficient when first entered, β^b= Standardized beta coefficient entered in the last step.
**p < .05, **p <.01, ***p <.001*

The control measures (NVA, visual concentration, verbal STM, digits backward, and receptive vocabulary) explained 13% of the variance in the pronunciation post-test and 14% of the variance in the spelling post-test for the orthography-present items. Verbal STM predicted the pronunciation post-test, and matching post-test, and marginally predicted the spelling post-test measure. PA explained 12% of the variance in spelling post-test scores. Similarly, RAN predicted scores on spelling post-test and explained approximately 4% of the variance in spelling post-test scores for the orthography-present items. Verbal STM was the only measure to predict the matching post-test.

Orthographic Learning as a Function of Word Orthographic Complexity

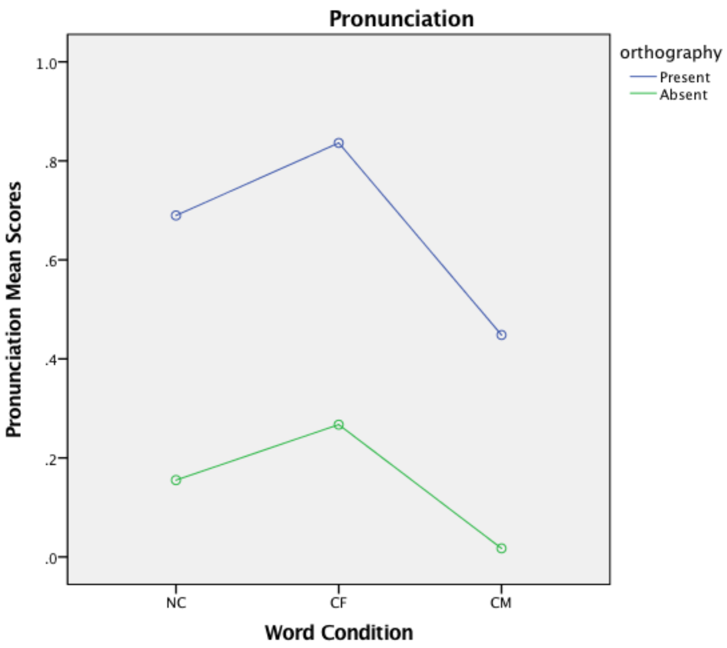
In order to examine the facilitative effects of incidental presentation of orthography on novel word learning and whether orthographic complexity has an effect, a 2×3 repeated measures ANOVA was performed for each of the post-tests.

In order to address the prediction that the recall of nonword phonology would be facilitated in the orthography-present condition for the nonword types ((NC), (CF), and (CM)), a 2×3 repeated measures ANOVA was performed. There was a significant main effect of presence of orthography on recall of pronunciation, F (1, 115) = 99.76, p < .001, r = .68. The effect size was large. There was also a significant main effect of nonword type on recall of pronunciation, F (2, 230) = 28.98, p < .001. Contrasts revealed that pronunciation recall of NC items, F (1, 115) = 8.73, p = .004, was significantly lower than of CF items. Pronunciation recall of NC, F (1, 115) =

24.83, $p < .001$, and CF, F (1, 115) = 51.81, $p < .001$, were significantly higher than of CM items. The NC vs. CF had a small effect size, $r = .27$, a moderate effect size between NC and CM, $r = .42$, and a large effect size between CF and CM, $r = .56$. The interaction between orthography-present or -absent and nonword condition was not significant, F (2, 230) = 1.39, p = .25.

Figure 1

Pronunciation Post-test Scores in the Three Nonword Type Conditions

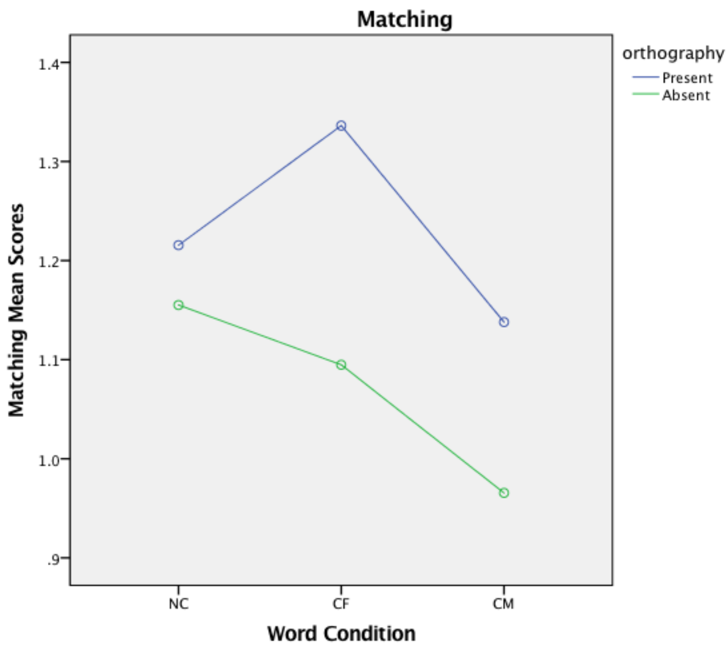


In order to address the prediction that the recall of nonwords' semantic identity would be better in the orthography-present condition than in the orthography-absent condition for the nonword types ((NC), (CF), and (CM)), a 2×3 repeated measures ANOVA was performed. There was a significant main effect for the presence of orthography on semantic recall, F (1, 115) = 10.03, $p = .002$; however, this effect size was small, $r = .28$.

There was a significant main effect of nonword type (NC, CF, CM), *F* (2, 230) = 4.58, *p* = .01. Contrasts revealed that recall of NC, *F* (1, 115) = 5.72, *p* = .02, and CF, *F* (1, 115) = 7.03, *p* = .009, were significantly higher than recall of CM items. Although CF few items enjoyed better recall than did NC items, this difference did not reach significance. There was a small effect size between NC and CM, *r* = .23, and a small effect size between CF and CM, *r* = .24. The interaction between presence of orthography and nonword type was not significant, *F* (2, 230) = 1.18, *p* = .31.

Figure 2

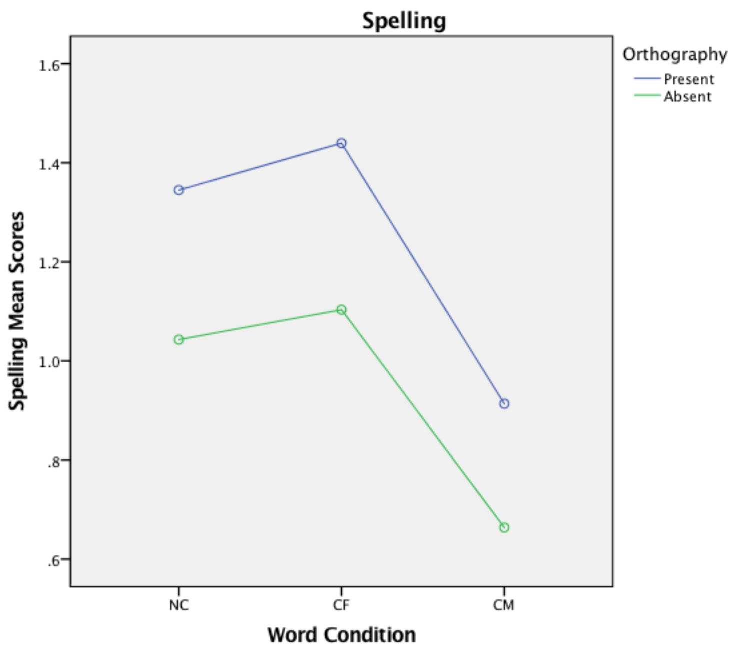
Matching Post-test Scores in the Three Nonword type Conditions



In order to address the prediction that the recall of nonwords' orthographic identity would be better in the orthography-present condition than in the orthography-absent condition for the nonword types ((NC), (CF), and (CM)), a 2×3 repeated measures ANOVA was performed. There was a significant main effect for the orthography (present vs. absent) on the spelling post-test, F (1, 115) = 32.91, p < .001. The presence of orthography had a moderate effect size, r = .47. There was a significant main effect for word condition (NC, CF, CM) on spelling, F (2, 230) = 42.12, p < .001. Contrasts revealed that NC, F (1, 115) = 54.13, p < .001, and CF, F (1, 115) = 73.21, p < .001, were significantly higher than CM. Spelling of CF items was better than of NC items, but this difference did not reach statistical significance. There was a large effect size between NC and CM, r = .57, and an even larger effect size between CF and CM, r = .62. The interaction between orthography-present or -absent and nonword type was not significant, F (2, 230) = .27, p = .76.

Figure 3

Spelling Post-test Scores in the Three Nonword-type Conditions



Summary of Orthographic Learning

Recall of all word identities (phonological, semantic, and orthographic) was enhanced by the mere incidental exposure to its orthographic form. Effect sizes ranged from small to large, the greatest facilitative effect in pronunciation recall. This is not surprising since sound is transient, and even when exposure to the word's orthographic form is incidental it leaves a trace of its form to facilitate recall. Pronunciation recall was the most difficult. Spelling was also facilitated in the orthography-present condition, and its effects were moderate in size. Recall of the picture referent was the least difficult of all three post-test assessments. This finding is consistent with studies on orthographic learning using this paradigm (e.g., Ricketts et al., 2009; Rosenthal & Ehri, 200).

As is evident from Figures 1, 2, and 3, there was no interaction between orthography conditions and nonword conditions. In other words, all item conditions (NC, CF, CM) enjoyed better recall in the orthography-present condition. However, ligature and consonant diacritics conditions affected learning. On all post-tests and irrespective of orthographic condition, connected items with few consonant diacritics were easiest to learn, followed by NC few diacritics. The only exception was the superiority of NC nonword type in the orthography-absent condition on the matching post-test. The most difficult items to learn were the connected with many consonant diacritics. The differences between NC items and CF items only reached statistical significance on pronunciation post-test. On every post-test, NC and CF items were easier to recall than were the CM. The consonant diacritical density appears to be the source of difficulty, because effect sizes were highest between CF and CM. In fact, it appears that connectivity facilitates novel word learning, whereas consonant diacritics hamper it. Although CF nonwords were easier to recall on the matching and spelling post-tests, the contrasts between NC and CF never reached significance.

Discussion

It is widely accepted that reading acquisition and development in Arabic is challenging. Yet, according to our traditional understanding of orthographic depth, vowelized Arabic is considered a transparent orthography that should be acquired with relative ease. This is not the case. As discussed, diglossia is often implicated as a source of this difficulty in reading acquisition. In addition, orthographic and visual factors, such as the density of diacritics, allography (changing letter form depending on the position in the word), and ligature probably contribute to a seemingly transparent orthography. Thus, the broad aim of this study was to examine the facilitative effects of orthography on novel word learning in Arabic. This study hypothesized that children would benefit from incidental exposure to orthography while learning nonsense (novel) words. Based on the lexical quality hypothesis and the use of a learning paradigm of paired associate learning, the study examined the contribution of orthography to the learning of new words. Half of the 12 referents were presented with the nonwords' orthography while the other half were not. No attention was drawn to items presented with orthography. The data provide support for the claim that orthography facilitates the learning of new words in Arabic, which is consistent with the findings in English (e.g., Rosenthal & Ehri, 2008; Ricketts et al., 2009). This study manipulated word characteristics to examine whether the source of difficulty in learning to read in Arabic is due to consonant diacritics or ligature. Another aim of the study was to investigate the predictors of orthographic learning and whether orthographic facilitation in the orthography-present condition is related to individual abilities. This section of the chapter summarizes the findings, identifies the limitations of this study, and discusses the implications.

Predictors of Orthographic Learning

Since the aim of this study was to investigate the mechanisms that underlie successful orthographic learning, only the orthography-present items in the pronunciation post-test (phonological identity), the picture-matching post-test (semantic identity), and spelling post-test (orthographic identity) were used in the regression analyses. Overall, verbal and NVC measures (nonverbal ability, visual concentration, verbal STM, digits backward, and receptive vocabulary) collectively contributed 13% to 14% of the variance in pronunciation and spelling recall in the orthography-present condition. Verbal STM predicted recall of pronunciation (phonological identity), matching (semantic identity), and marginally predicted spelling (orthographic identity) in the orthography present. Verbal STM was the only predictor of the matching post-test. This finding is consistent with findings in the literature linking verbal STM to oral vocabulary acquisition (Gathercole et al., 1999; Gathercole, 2006). Although the literature on verbal STM points to its importance in learning new phonological forms, this relationship seems to diminish with age, probably as a result of efficient word recognition that decreases the load on the phonological loop (e.g., Jarrold et al., 2004; Jarrold et al., 2009). The findings in the present study demonstrate that verbal STM is necessary at the early stages of acquiring new orthographic representations in Arabic. With the exception of spelling post-test (orthographic identity), verbal STM was only a predictor of pronunciation and matching post-tests in the orthography-present condition. Taken together, the results in this study support the claim of an overlap between PA and verbal STM that may account for its contribution to vocabulary learning (Parilla et al., 2004; Martinez Perez et al., 2012).

PA did not predict pronunciation or matching post-tests in the orthography present post-test; however, PA was the strongest predictor of orthographic learning, explaining 12% of the variance in spelling recall for the orthography-present items. This finding underscores the importance of PA in Arabic discussed in Chapter 3. This finding lends support to Share's claim that phonological decoding is the sine qua non of word recognition. However, Share (2004) found that despite the high phonological decoding accuracy, orthographic learning was not evident for his sample of Hebrew-speaking first graders. Share reasoned that beginning competent readers of transparent orthographies rely less on orthographic details and more on bottom-up processes when reading. In Cunningham et al.'s (2002) study, phonological decoding accuracy was the strongest predictor of orthographic learning. However, Bowey and Miller (2007) pointed out that phonological decoding accuracy of targets is not a pure measure of phonological decoding since it only shares 27% of the variance with Word Attack task (a standardized nonword reading task of increasing difficulty). In fact when Bowey and Miller reanalyzed Cunningham et al.'s data partialing out Word Attack, they found that phonological decoding accuracy was no longer significantly related to orthographic learning. This study utilized an incidental learning paradigm, thus decoding accuracy data was not collected. However, the overwhelming contribution of PA to orthographic learning seems to be consistent with opaque orthographies rather than transparent ones. The results in this study demonstrate that our current definition of orthographic depth needs to consider linguistic and visual characteristics that contribute to the rate of reading acquisition (Share & Daniels, 2016).

Verbal STM was a consistent contributor of orthographic learning in the three outcome measures, however; only PA and RAN predicted spelling post-test (orthographic identity). RAN uniquely predicted spelling recall in the orthography-present condition above and beyond the contribution of PA and explained approximately 4% of the variance in spelling post-test scores for the orthography-present items. PA is the strongest predictor of spelling in Arabic as it is in other languages and may reflect a general skill not exclusive to orthographic learning. Indeed, when the data were analyzed for both orthography-present and orthography-absent conditions, PA made the same contribution to spelling (12%) as it did when the outcome was for orthography-present items only. Similarly, RAN made similar contributions to spelling whether the analysis included both absent and present conditions (6.6%), and orthography-present items only (4%).

Facilitative Effects of Orthography on Word Learning

In order to assess the facilitative effects of orthography on novel word learning in Arabic, three post-tests were administered to assess recall of words' identities: phonological, semantic, and orthographic. This section begins with a discussion of the findings regarding the facilitative effects of orthography on recall of these three identities.

The present study adds to existing literature on the value of orthography in pronunciation recall. Based on the lexical quality hypothesis and Ehri's amalgamation theory (as cited in Ehri & Wilce, 1979), it was reasoned that the activation of word-specific identities would enhance learning of newly taught words. Indeed, as the results indicated, the presence of orthography supported the children's ability to recall the pronunciation of the nonwords. There was a large effect size on the recall of pronunciation in the orthography-present condition, thus providing evidence for the facilitative effects of orthography on recall. The findings in this study are consistent with those in several studies that have found that as little as one encounter with an unknown word activates its orthographic constituent (Apel, 2010; Cunningham, 2006; Ricketts et al., 2011; Share, 2004). Indeed, even when the presentation of orthography is incidental, recall of pronunciation was better in the orthography-present condition, and this effect was large. The post-tests were always administered in a fixed order (pronunciation, picture-matching, and spelling) to ensure that exposure was incidental and limited to inhibit explicit decoding of the items. Thus, even when orthographic representation is incidental, learning occurs implicitly. This is consistent with models of statistical learning demonstrating the implicit acquisition of word-specific representations after few exposures.

It was hypothesized that children would benefit from the presence of orthography in recalling semantics information as assessed by the picture-matching task. The presence of orthography facilitated semantic recall; however, unlike pronunciation, this effect was small. The results from the present study are inconsistent with results in studies in English (Ricketts et al., 2009; Rosenthal & Ehri, 2008), and in Chinese (e.g., Hu, 2008), where performance on this task was close to ceiling and thus did not yield a significant effect of orthography. For example, on their matching-words-to-sentences post-test, Rosenthal and Ehri (2008) found the benefits of the presence of orthography did not reach statistical significance, possibly due to ceiling effects (90 to 96% correct). Similarly, on their nonword-and-picture-matching post-test, Ricketts et al. (2009) found that correct responses were also near ceiling. This was not the case in Arabic. Performance on this task was facilitated by the presence of orthography, demonstrating that indeed, orthography facilitates the recall of semantic information. This is the first study to examine the facilitative effects of orthography on semantic learning in Arabic. The facilitative effect on semantic recall may reflect a more accurate account of how Arabic-speaking children learn vocabulary that is often absent from their spoken vernacular.

It was hypothesized that orthography would facilitate spelling recall. This hypothesis was supported by the data. Spelling recall in the orthography-present condition was superior even when the presentation of orthography is incidental. The facilitative effects of orthography on spelling were moderate, which may reflect the transparency in Arabic spelling. Although Arabic is considered highly consistent, diglossic phonemes create a challenge even for skilled readers (Saiegh-Haddad, 2013), and thus the presence of orthography may alleviate this ambiguity during spelling. The moderate effect size could be the product of the training that this finding seems to imply that training (six exposures in total) did not provide the children with enough exposure to establish quality representations, particularly for words that deviate from sound-to-spelling consistency, as is the case for items with diglossic phonemes. Ehri and Wilce (1979) emphasized that “spelling must map to sounds accurately” (p. 36) to facilitate recall. Although the children in the current study heard the nonword pronounced by the examiner, the diglossic phonemes in the nonword may have caused confusion because the speech sound did not map to the orthographic form (Aro & Wimmer, 2003; Ehri & Wilce, 1979; Erdener & Burnham, 2005). Apel (2009) reasoned that less exposure to linguistic rules and patterns, as is probably the case for participants in the present study, will “negatively affect [the] learning of new [words]” (p. 49). The ease of sound-to-spelling in Arabic, coupled with the difficulty and inconsistency of diglossic phoneme could explain why the children in the present study found it most difficult to spell items that should be spelled with relative ease with or without the presence of orthography. Finally, the interaction between orthography-present or -absent and nonword condition was not significant.

Orthographic Learning as a Function of Word Orthographic Complexity

The facilitative effects of incidental presentation of orthography on novel word learning and whether orthographic complexity has an effect was examined by varying ligature and consonant diacritics in the following conditions: NC, CF, and CM. Recall of pronunciation, picture-matching, and spelling was superior for CF items than for NC and CM items, demonstrating that connectivity facilitates orthographic learning in Arabic. The rate of recall of items in the CM condition was the lowest, indicating that consonant diacritics are a source of difficulty in Arabic. Similarly, on the picture-matching post-test, CM items were the most difficult to recall, while

CF enjoyed the highest level of recall. However, like the pronunciation post-test, the interaction between presence of orthography and nonword type was not significant. Finally, the results for the spelling post-test mirrored those of pronunciation and picture-matching, where CM items were the most difficult to spell, whereas the CF items were spelled the most accurately. The largest effect size was observed between the CF and the CM conditions. These results are consistent with those of Dai et al. (2013), showing that items with many consonant diacritics slowed reading.

Researchers have proposed that, since consonant diacritics are used to distinguish several letters having the same basic shape, this feature in Arabic causes confusion and impedes fluency. This linguistic feature of Arabic was investigated neurologically by Eviatar et al. (2004). Indeed, there was no right hemispheric involvement when processing the letters that have the same basic shape but differ in the number of consonant diacritics. The authors take this finding to demonstrate that the right hemisphere is not involved in word recognition in adult readers of Arabic. Ibrahim and Eviatar (2012) showed that, although the right hemisphere is not involved when processing Arabic letters, it contributes to letter identification in both Hebrew and English for the same participants. These results confirm behavioral data in Urdu and Hindi that showed the same participants read Urdu words (Urdu uses Arabic letters) more slowly than they do Hindi words (Rao et al., 2011).

Such studies demonstrate that orthographic depth, as currently defined, does not account for the disparity between the one-to-many (base grapheme and many consonant diacritics) that exists in Arabic, which appears to be at the core of the difficulties experienced by readers of

Arabic. Ligature in Arabic has often been implicated as a source of difficulty. This is not surprising considering that most languages are not cursive. This connectedness appears to facilitate reading fluency and not impede it, as previously proposed. The results of Dia et al. (2013) and the current study confirm this. Although reading speed of connected letter items and NC letter items in Dai et al.'s study was the same, spelling was superior for connected letter items. Similar results were obtained in the current study, where spelling of CF items was better than for NC items; this difference, however, did not reach statistical significance. Taken together, there is some support that connected letter items facilitate orthographic learning in Arabic. Finally, orthographic learning as measured by reading fluency was superior for seen words than for words that were not seen. However, there was no difference in the spelling post-test of items, whether seen or not.

All the interactions between orthography-present or -absent and nonword conditions on the pronunciation, picture-matching, and spelling were not significant. The lack of interaction between orthography conditions and word type may be a reflection of the transparent nature of vowelized Arabic orthography. It has been suggested that the relationship between implicit statistical learning and reading may be weaker in transparent orthographies such as Arabic because regularities are explicitly conveyed (Arciuli, 2018).

The paradigm employed in the present study allows for an alternative explanation to that put forth by Dai and colleagues that the visual demands, and/or the competition from similar-looking letters is the source of difficulty in Arabic. This study employed a paired associate learning task, and the nonwords in any given word condition were pronounced by the examiner and not decoded by the children. Therefore, it is not the visual density of the consonant diacritics that is at the root of the problem but perhaps a phonetic (oral) feature accompanied by letters that contain many diacritics. Upon inspection of the nonwords in the CM condition, three out of the four nonwords (قثنب, جشيك, ظحفك) that were selected from Dai et al.'s stimuli contained diglossic phonemes, whereas the CF items contained no diglossic phonemes. Thus, the apparent difficulty with items with many consonant diacritics is due to difficulty in recalling items with phonemes that do not exist in the children's spoken vernacular.

The effects of diglossia on PA and reading, was first demonstrated by Saiegh-Haddad (2003). Kindergartners had more difficulty isolating diglossic phonemes than they did isolating spoken phonemes. Similarly, first graders read nonwords that adhered to the spoken structure more accurately than they did nonwords designed to mimic the literary syllabic structure. Saiegh-Haddad concluded that the linguistic distance between the spoken and literary varieties of Arabic impedes the development of PA among Arabic-speaking children. The results of this study support Saiegh-Haddad's conclusions and extend them to show that diglossic phonemes also obstruct initial word learning. Indeed, similar results are obtained by languages that have sounds in the dialect that differ from the written form (e.g., Akshara). The inconsistency in sound-to-spelling linguistic features makes it difficult to spell, even when the script is transparent (Nag, 2011). Spelling in Arabic requires segmenting spoken words into their respective phonemes. Since the mappings of the orthographic units and phonological units are near one to one, the standard expectation would be that spelling for Arabic children should develop with relative ease (Caravolas, 2004). This is not the case. According to Arabic researchers, spelling development is affected by diglossic phonemes (e.g., Abu Rabia & Taha, 2006). The influence of dialect on spelling points to a greater impact of phonological processing than of orthographic processing, which is consistent with the findings of this study and the second study. These findings underscore the challenges that Arabic speakers encounter in literacy acquisition that are rooted in diglossia rather than in visual and orthographic demands. Thus, the effects of diglossia may contribute to orthographic depth that is not accounted for in the current definition (Share & Daniels, 2016).

Summary of Word Type and Orthography Condition on Recall

Consonant diacritics and ligature in Arabic affected the learning of the different word identities in various ways. In the recall of pronunciation (phonological identity), the picture-matching (semantic identity), and spelling (orthographic identity), the CM items proved to be most difficult, regardless of whether seen or unseen. At first glance, this finding is consistent with the conclusions of Dai and colleagues that consonant diacritics are both visually and orthographically demanding. However, the lack of interaction between word type and orthography condition in this study points to a different conclusion.

Limitations

It is important to tease through and disentangle the relationships between literacy skills while considering the complexities of the changing relationships of skills in the course of development. Methodological differences make it even more challenging to determine the exact mechanisms underpinning the relationship between orthography and vocabulary. Although the results of this study provide further supporting evidence for the lexical quality hypothesis, generalizability is limited by the fact that the training and post-tests were conducted during a single session. Future research should examine this facilitative effect on delayed recall. Rosenthal and Ehri (2008) delayed their post-tests by one day; thus, it would be informative to replicate such an experiment with post-tests prolonged to several days in order to examine decay, which seems to play an important role in learning new words (Nation et al., 2007). This warrants further investigation. It would be interesting to see if any facilitative effects of orthography continue to be evident after a delay in relationship to word condition in Arabic. Unlike studies that have examined the course of lexicaliztation, RTs were not measured in the current study, due to limited resources. Examining latencies in relationship to word type would have added another level of analysis to the present findings.

Another limitation pertains to the nature of the training itself: it should be more naturalistic (vocabulary learned in the classroom) in order to generalize these results and others (e.g., Apel, 2010). While the use of an incidental exposure offers support to statistical learning models, the training was too brief to establish quality representations. More exposures to a word seems to strengthen the connections among all three constituents and have proved to improve word learning, particularly for opaque languages such as English (Nation et al., 2007; Share, 2004) and perhaps Arabic as well. Furthermore, since the orthographic information was presented incidentally, without the opportunity to analyze the words via decoding (i.e., self-teaching) or spelling the word (Shahar-Yames & Share, 2008), this probably resulted in incomplete representations according to the lexical quality hypothesis. Thus, presenting orthographic information for deliberate analysis during word learning may provide stronger orthographic facilitative effects. Finally, the stimuli used in the present study were concrete, picture-depicted nonwords. Thus, the logical next step would be to replicate this study using abstract words, to determine whether orthography assists with both types of word learning. Words with high imageability (e.g., /butter/) have been found to produce an advantage in learning, unlike words with low imageability (e.g., /better/) (Duff & Hulme, 2012; Walker & Hulme, 1999).

Conclusions and Implications

The findings thus far are intriguing and carry great implications for vocabulary instruction in Arabic. Teachers and parents should emphasize print exposure by encouraging children to read, thereby giving them the opportunity to implicitly self-teach and reap the benefits of orthography. This is probably most important in a diglossic situation. Emphasis should be placed on explicit spelling practice of words that contain diglossic phonemes. It is common classroom practice for Arabic teachers to regard a pronunciation as correct when produced with diglossic phonemes. The findings here suggest that children should be encouraged to pronounce Arabic words correctly and should be discouraged from the use of diglossic phonemes in their pronunciations. Judging from the children’s difficulty in recalling items that had many consonant diacritics, and the fact that letters with many consonant diacritics are diglossic phonemes, more emphasis on spelling should be placed during instructions on words that contain such phonemes. Teachers should consider the use of a word's spelling and proper pronunciation as a powerful tool in literacy instruction, particularly in a diglossic situation such as Arabic.

The theoretical rationale for the contribution of orthography has been available for many years (e.g., statistical learning, lexical quality hypothesis, self-teaching hypothesis, and amalgamation theory), and empirical evidence (Cunningham, 2006; Ricketts et al., 2009; Rosenthal & Ehri, 2008; Share, 2004) has increasingly surfaced in support of these theories. Two issues remain unsettled. 1) What are the exact mechanisms that are involved in orthographic learning? 2) How does orthographic learning happen, and how is it shaped by the characteristics of the language? This is the first study to examine the facilitative effects of orthography on novel word learning in Arabic in order to address these two issues.

Verbal STM was only a predictor of pronunciation (phonological identity) and picture-matching (semantic identity) post-tests in the orthography-present condition, while PA and RAN predicted spelling (orthographic identity). The fact that PA is the strongest predictor of spelling is consistent with the findings in the literature on the strong role of PA on spelling and may reflect a general skill not exclusive to orthographic learning.

The current study, like others, has provided initial support for the claim that orthography facilitates the learning of new vocabulary; however, as noted, much more research in this area is needed before such findings can be generalized to the classroom. The findings of the present study confirm that incidental exposure to orthography facilitates the learning of newly taught words while the presence of orthography also facilitated the recall of pronunciation, semantic information, and spelling of newly taught words. Orthography, even when incidental, exerted a strong effect on the recall, and this result was not dependent on word condition, as all items benefited from training with orthography. These results are particularly relevant for Arabic speakers because they demonstrate that implicit learning is possible even when exposure to print is limited and incidental. What was particularly intriguing in this study is the role that diglossic phonemes rather than consonant diacritics or ligature played in children's learning of new words and the difficulty of items containing diglossic phonemes imposed on these young learners. These results underscore that diglossia impedes literacy acquisition and is the source of difficulty in reading in Arabic.

Chapter 5

General Discussion

The primary purpose of this research was to examine orthographic learning in Arabic-speaking primary school children. The instant and rapid recognition of words is the hallmark of reading (Share, 2004), yet the underpinnings of how this process develops and how it is modulated by orthographic depth are still unclear. With the exception of Gough and Tunmer's (1986) SVoR, many reading theories focus on a particular aspect of the reading process (e.g., Ehri's phases of development of sight-word learning; Share's self-teaching, Connectionist models). These theories are not necessarily competing but rather attempt to explain a narrow piece of the same complex puzzle. There is no doubt that reading development is the product of experience, and nearly all theories of reading take into account how reading is shaped by experience with print. However, despite this underlying knowledge, much of the research on reading does not control for print exposure. This was addressed in the current study for the obvious reason stated, and even more importantly, in a diglossic situation where spoken language is different from the written one. Thus, two TRTs were developed in Arabic L1 and English L2 using the same development procedures as Stanovich and Cunningham's (1992). While experience might be implicit in reading theories, the nature of orthography is not. Orthographic depth has long been recognized as a determinant of reading development (Katz & Frost, 1992; Ziegler & Goswami, 2005). However, our current understanding of qualifying orthographic depth is incomplete and does not take into account orthographic features of non-European languages. Just as reading development in Arabic is constrained by diglossia, other orthographic characteristics also place tremendous demands on the novice reader. Thus, the second aim of this research was to examine predictors of single word reading and spelling in both Arabic L1 and English L2 in order to address the issue of universality of the skills and abilities that have often been implicated by reading research. Also, predictors of single word reading and spelling were examined in a sample of Arabic-speaking monolinguals, to allow for comparisons between bilinguals and monolinguals. The comparisons between bilinguals and monolinguals were carried out to further probe into environmental aspects that may shape reading development. Determining predictors of reading and spelling is of paramount importance yet tells us nothing about the mechanisms underlying fast and efficient word recognition. Thus, the final and primary aim of this research was to examine orthographic learning in Arabic and investigate the skills that are associated with creating new lexical representations. The findings from this research converge on the need to reassess our current definition of orthographic depth to include other orthographic characteristics that play a role in reading development and ultimately in the fast and seemingly effortless instant recognition of words. More importantly, this study provides evidence that diglossia and particularly diglossic phonemes are the root cause of word learning difficulties in Arabic.

The speed at which word recognition develops has long intrigued reading researchers. Beginning readers of alphabetic scripts must first learn the mappings between letters (graphemes) and their corresponding sounds (phonemes). This is thought to be the vehicle in which efficient word recognition is eventually achieved. Once these mappings are applied via decoding of unfamiliar words, an amalgamation process occurs that allows for a word to have lexical status. Although there is a consensus amongst reading researchers that phonological decoding is the first step in achieving rapid and efficient word recognition, less is known about the development of subsequent stages between initial decoding and word recognition. As a result, there is currently no encompassing theory or model that describes exactly how this is achieved. Thus, this study attempted to describe the process from phonological decoding to the fast and efficient word recognition (the process of orthographic learning) that happens in Arabic. This entailed addressing the following questions: what are the underlying abilities and skills that predict efficient word reading and spelling, and what are the skills and abilities that are involved in the process of orthographic learning?

According to the lexical quality hypothesis, the fast and efficient word recognition system is the product of well-specified representations that occur as a result of multiple exposures (Perfetti & Hart, 2002). The final result is the amalgamation of the word’s phonology, orthography, and meaning into what is instantly identified as a unitary lexical entry. Once a word has achieved high-lexical status, instant recognition becomes involuntary (Ehri, 2005b). Indeed, this involuntary activation is demonstrated by the current study and studies using incidental exposure paradigms. Unlike Ehri’s phases that are achieved consequently one after the other, Share (1995) offers an item-based account, which is consistent with the lexical quality hypothesis in which some words enjoy high lexical quality, whereas some words are of low lexical quality. Whether a connectionist, developmental account, or item-based account, all these accounts, explicitly or implicitly, point to the role of exposure that determines orthographic learning. Indeed, statistical learning models offer a way of bridging the gap between these theories and models. Thus, it was an aim of this study to control for the role of print exposure. What is yet unclear is what contributes to the processing of words in a laborious manner to the effortless recognition of words. In order to address this question, we first need to address the skills and abilities that determine the successful development of word recognition. This has been no easy feat for reading researchers for several reasons. First, the research methods used to investigate the development of the word recognition system are often correlational, making it increasingly difficult to separate cause and effect. Second, orthographies vary considerably in depth (Katz & Frost, 1992), making it difficult to draw generalizable conclusions. As such, the second study explored predictors of reading and spelling while trying to minimize these issues, by first having a cross-sectional design which included third-, fourth-, and fifth-grade children. Second, the study included a comparison within-participants Arabic L1 children learning English as L2, and between-participant comparisons of Arabic-speaking monolinguals, and Arabic L1 English L2 bilinguals. Next, in order to gain a better understanding of how lexical entries are established, we need to document the process of orthographic learning through the use of training studies. Based on Ehri's argument that word recognition is an involuntary process, and based on models of statistical learning, the third and final study demonstrated how orthographic learning occurs incidentally. This study provided intriguing results about the linguistic and orthographic features of Arabic and provided clues to explain why literacy acquisition is challenging in Arabic despite being transparent when fully vowelized.

Clearly, cognitive, linguistic, and environmental factors are intertwined in the course of reading development. The role and magnitude these factors play in reading development is influenced and shaped by the characteristics of the script. Most of the evidence, as mentioned, comes from English and European languages (Share, 2008; Share & Daniels, 2016) that do not reflect the linguistic and orthographic features of many other languages and thus limits our understanding of the abilities that are universal versus those that are the result of the script. Thus, predictors of Arabic and English reading and spelling are addressed in this research, in an attempt to separate the skills that are universal from those that are script dependent.

Despite the evidence of the importance of print exposure in the course of reading development, it is seldom measured or controlled as a variable. Thus, the first study developed a TRT in both Arabic and English and hypothesized that print exposure would be a predictor of reading and spelling in both languages. This was not the case. The difficulties in generating book titles for the development of both the Arabic and English versions of TRT is consistent with the anecdotal evidence that Arabic-speaking children seldom read for leisure. The conclusion that Arabic-speaking children do not read for leisure, although speculative, is consistent with other studies that have investigated reading habits in Arabic-speaking children and have encountered similar challenges (e.g., Feitelson et al., 1993). The reason reading habits are not fostered culturally may be rooted in diglossia. Arabic books are written in MSA, which prior to formal reading instruction, is unfamiliar to young children and may deter parents from cultivating a home literacy environment. Instead, Arabic-speaking parents often prefer oral storytelling in the spoken vernacular, which without a doubt enhances oral competence. However, this oral competence does not facilitate reading acquisition due to the linguistic distance between the two varieties.

Using a simple view of reading, coupled with the knowledge about the lack of print exposure in MSA from Chapter 1, this research asked what the skills and abilities are that predict reading and spelling in Arabic in the absence of such basic prerequisites, such as oral comprehension and print exposure. As described, the literature on vowelized Arabic demonstrates that reading develops in a manner similar to that of opaque languages rather than transparent ones, as one might expect based on evidence from cross-linguistic literature. However, even if we only consider cross-linguistic evidence, there is considerable variability in the role and magnitude of abilities that underpin reading development and how they are shaped by orthographic depth.

Predictors of Reading and Spelling in Arabic

The discussion begins with the predictors of reading and spelling in Arabic. The most important finding of this study is the fact that fully vowelized, transparent Arabic behaves much like what we would expect of an opaque orthography like English. This finding is consistent with findings in previous research in Arabic and puts into question our current classification of orthographic depth. Reading tasks, word recognition, and phonological decoding fluency were fully vowelized, which makes Arabic very transparent with near one-to-one letter-sound correspondence.

However, it has long been recognized that Arabic, even when fully vowelized, represents a challenge to readers.

The purpose of the second study was to investigate predictors of reading and spelling in Arabic. Vowelized Arabic is considered transparent; however, the characteristics of Arabic, such as diglossia, ligature, and allography, were expected to complicate a seemingly transparent orthography which would explain why literacy acquisition and development is challenging in Arabic. These features probably influence the cognitive processes that contribute to reading ability.

Based on the literature reviewed in Arabic and in other languages, it was hypothesized that PA would be the strongest predictor of both reading and spelling. Indeed, despite Arabic's seemingly transparent orthography when fully vowelized, PA was the strongest predictor of reading and spelling. While this finding is consistent with evidence in the existing literature on Arabic (e.g., Taibah & Haynes, 2011; Tibi & Kirby, 2018), it is not consistent with the findings from cross-linguistic literature. Cross-linguistic research has demonstrated that the contribution of PA to reading diminishes in the early years of reading acquisition in transparent orthographies (e.g., Landrl & Wimmer, 2000; Papadopoulos et al., 2009; Furnes & Samuelsson, 2010), whereas the contribution of PA to reading persists in opaque orthographies. Clearly, the contribution of PA to reading is mediated by the depth of the orthography (Wimmer & Goswami, 1994; Melby-Lervag et al., 2012). Tibi and Kirby (2018) proposed an explanation for the strong role of PA to reading in Arabic that may be due to instructional methods that do not emphasize the use of phonics; however, this is speculative and was not investigated in their study. Another explanation could be that Arabic even when fully vowelized should be conceptualized as an opaque orthography. This is consistent with the data that shows PA as a stronger predictor in less consistent orthographies (Georgiou et al., 2008; Ziegler et al., 2010). It has been proposed that Arabic's visual and orthographic features are responsible for the heavy reliance on PA (Asadi and Khateb, 2017).

However, PA is not the only ability that predicts reading. This is best demonstrated with words that are irregular that cannot be read using phonological decoding. Thus, efficient word recognition is driven by other factors. Indeed, differences in orthographic processing skills appear to be related to reading and independent of PA (e.g., Cunningham & Stanovich, 1990). However, it has been argued that some of the orthographic processing tasks are essentially reading tasks (e.g., as in orthographic choice tasks where participants might be asked to select the item which is a real word from the pair <rane>, <rain>). Furthermore, many tasks tapping orthographic processing are contaminated by phonological decoding (Castles & Nation, 2010). These limitations in orthographic processing tasks may even be more pronounced in transparent languages such as Arabic, making it difficult to create a purely orthographic measure that is free of phonological decoding. Unlike English, Arabic has a near one-to-one letter-sound correspondence when fully vowelized; thus, there are no irregular words or word units that must be learned by sight. This feature makes it difficult to create an orthographic task that cannot be carried out without phonological decoding. Indeed, one of the orthographic tasks in this study had items that contained letters that were in the wrong shape based on their position in the word. This task was highly correlated to the PA tasks, and when principal factor analysis was conducted, it loaded on the phonological factor rather than the orthographic factor. Because orthographic measures are often contaminated by reading and phonological decoding measures, it is important to examine predictors of orthographic learning by using training studies. Thus, this study examined predictors of orthographic learning.

Although there exists unequivocal evidence of the relationship between RAN and reading, there is much debate on the underlying cognitive underpinnings of this relationship (de Jong, 2011). The ambiguities concerning the underpinnings of the RAN-reading relationship are further complicated by course of reading development and orthographic depth. What RAN exactly taps into and how this ability is related to reading has been debated. Nonetheless, RAN, like PA, appears to be a universal predictor of reading and spelling (Caravolas et al. 2012); however, it is not clear whether this relationship is modulated by orthographic depth (Ziegler et al., 2010; Moll et al., 2014). Based on the literature in Arabic and literature on transparent orthographies, it was hypothesized that RAN would also contribute uniquely to reading. Findings from this study indicate that RAN is indeed a unique predictor of reading even after controlling for PA. RAN predicted word recognition and nonword reading fluency for the older group of bilinguals but not the younger group, which is consistent with opaque orthographies rather than transparent ones. RAN predicted word recognition, nonword reading fluency, and spelling in the sample of third-grade monolinguals. When bilinguals and monolinguals were matched on word recognition, RAN did not contribute to word recognition. However, when a subset of bilinguals and monolinguals were matched on nonword reading fluency, RAN contributed to nonword reading fluency almost equally in both samples. Taken together, the findings are consistent with the view that RAN taps into the efficiency in which orthographic units and their corresponding phonological codes are paired. RAN has been shown to have a stronger predictive role in consistent orthographies and to be a consistent predictor of reading in transparent orthographies above and beyond PA. This is due to the reduced phonological demands, and thus speed becomes a better predictor in transparent orthographies. Reading accuracy is usually at ceiling by first grade in transparent orthographies. Moll and colleagues (2009) posit that reading fluency in transparent orthographies provides a purer assessment of fluency that is not contaminated by accuracy. Ziegler et al. (2010) argue that RAN's contribution to reading in transparent orthographies is due to insufficient measures of PA that are near ceiling. In contrast to the findings of Moll et al. and Ziegler et al., the meta analysis of Araujo et al. (2015) revealed that the influence of RAN was greater in opaque orthographies.

The results of this study support the findings of Araujo et al.'s (2015) meta analysis and demonstrate once again that Arabic, even when fully vowelized, produces results that are more consistent with opaque orthographies rather than transparent ones. The bilingual data offer support for Moll et al.'s argument that, as reading becomes more proficient, the contribution of RAN increases. The strong contribution of PA coupled with the contribution of RAN in the older group demonstrates that fully vowelized Arabic should be considered an opaque orthography. Ziegler et al.'s (2010) argument that the contribution of RAN to reading is due to insufficiently demanding measures of PA was not supported by this study, since RAN uniquely contributed to reading above and beyond the contribution of PA, and there was no indication of ceiling effects in the PA scores (Kirby et al., 2010).

The data in the present study may offer some insights regarding the RAN-reading relationship. The data do not support the view that RAN is a construct of phonological processing, since RAN and PA in the older bilingual group were only modestly correlated, implying that the overlap is minimal. However, comparisons between subsamples of bilinguals and monolinguals matched on word recognition showed that PA was the only reading-related predictor, whereas the subsamples matched on nonword reading fluency showed that both PA and RAN were predictors of reading fluency for both bilinguals and monolinguals. Such results may suggest that RAN taps into the rate of retrieval of phonological information from long-term memory and thus may offer some support to the argument that RAN is a phonological processing construct (e.g., Wagner et al., 1994; Lervåg & Hulme, 2009; Ziegler et al., 2010). There is, however, some evidence that RAN taps into orthographic processes as well (e.g., Bowers & Wolf, 1993; Wolf & Bowers, 1999). RAN predicted spelling in Arabic, which is often considered an orthographic measure, for the monolinguals and older bilinguals, even after the variance of PA was controlled for. Powell et al. (2014) found that RAN was associated with irregular word spelling in English, which supports an orthographic rather than a phonological role. Finally, the data in this study do not support the claim that RAN is a measure of general cognitive ability because of the very small correlations observed between RAN and verbal and nonverbal general ability measures.

The findings in this study regarding the role of PA and RAN have several important implications. First, the finding of the strong role of PA in Arabic challenges the current one-dimensional conceptualization of orthographic depth and calls for a multidimensional framework that encompasses other linguistic and orthographic features. Second, the finding that RAN predicted reading for the monolingual and older bilinguals suggests that, as decoding becomes more efficient, there is less reliance on PA and more reliance on speed. Thus, this study does not support the claim that RAN is a phonological processing construct. Third, the bilinguals who were matched on word recognition, nonword reading fluency, and spelling were always older (fourth and fifth graders) than were the third-grade monolinguals; however, the predictors of reading were very similar. This indicates that the same underlying cognitive abilities are involved in reading and spelling Arabic for both monolinguals and older bilinguals.

There are some noteworthy results regarding the control measure that deserve mention. First, vocabulary is often used as a control measure of general verbal ability (e.g., Ziegler et al., 2010) and is often related to reading comprehension. However, in this study vocabulary predicted reading and spelling measures for the monolinguals. In the sample of bilinguals, vocabulary predicted word recognition for the older and younger groups alike. While vocabulary's influence on word recognition is intuitive, its influence on nonword reading fluency is surprising. Perhaps monolinguals with higher vocabulary knowledge facilitated the fluency in reading the nonwords, which in Arabic usually differ by one letter, since the majority of words in Arabic are composed of three to four consonants. In other words, when constructing nonwords in Arabic, changing one or two letters would usually result in another real word, and nonwords resemble real words to a great degree. Although speculative, the obvious explanation for the influence of vocabulary in reading is diglossia. Indeed, the mean scores on receptive vocabulary between the groups of younger and older bilinguals did not significantly differ, highlighting the effects of diglossia on the development of vocabulary in Arabic, which is nearly stagnant even after years of reading instruction. This is further confirmed by the expected differences between the younger and older bilinguals on the English receptive vocabulary task, in favour of the older group. Taken together, these results may indicate potential long-lasting effects of diglossia on vocabulary development, and that, even though English is a second language, vocabulary develops at the expected rate.

Next, the digits backward task, which in this study was administered as a working memory measure, predicted word recognition, nonword reading fluency, and spelling in the sample of bilinguals and in the subsample comparisons between bilinguals and monolinguals. Digits backward did not predict reading or spelling for the monolinguals. Of all the nonverbal and verbal control measures, verbal STM emerged as the strongest predictor of reading and spelling for the monolinguals, which is inconsistent with Taibah and Haynes’s (2011) findings although the verbal short-term measures used in this study are the same ones used in Taibah and Haynes’s study. Similar results to those of Taibah and Hayes were obtained by Caravolas et al. (2012) longitudinally in four alphabetic languages (English, Spanish, Slovak, and Czech). Verbal STM did not predict reading or spelling in any of the four languages in kindergarten or in first grade; however, the verbal STM had low reliability. Comparing the third-grade monolinguals in this study with Taibah and Haynes's third-grade sample (N = 40) reveals stark differences on the PA measures. The third-grade sample in Taibah and Haynes had a mean score of 14.5 (SD = 6.3) on the elision measure, whereas the third graders in this study had a mean score of 12.7 (SD = 3.9). The disparity between the samples in both studies on the blending words task is even greater, with a mean score of 5.6 (SD = 2) for this sample, and a mean score of 13.5 (SD = 7.6) for Taibah and Haynes's third-grader sample. The samples of third graders in both studies (this one and Taibah and Haynes) had similar scores on their nonword repetition and digit span. This implies that once a certain level of proficiency in PA is achieved, children's reliance on verbal STM diminishes.

Verbal STM also predicted English reading and spelling for the younger bilinguals. These results could be viewed as demonstrating that verbal STM predicts Arabic reading and spelling as it would a second language. This is consistent with the prediction that, due to diglossia, Arabic is akin to a foreign language for young children. However, the role of verbal STM in reading has been shown to diminish after the first few years of reading instruction. The reason for a diminishing role of verbal short-term memory is that, as reading becomes more fluent, there is less need to hold on to phonological units until the subsequent letters are decoded and blended. Since reading development occurs at a faster rate in transparent orthographies, which usually happens by the end of first grade (Seymour et al., 2003), the influence of verbal STM may be limited to the very beginning of reading instruction. Verbal STM may be more important for transparent orthographies because finer-grained units must be decoded and blended, placing higher demands on the limits of verbal STM. For example, Georgiou et al. (2008) found that verbal STM predicted first-grade nonword reading only in Greek but not English. These findings are consistent with PGST, since in Greek finer-grain phonological information must be retained in STM for blending. The results of the current study regarding verbal STM may suggest that the classification of fully vowelized Arabic as transparent is accurate, and indeed, it is when we consider consistency of grapheme-phoneme conversions as the only criterion of orthographic depth. In other words, the results here support the claim that verbal STM may be more important in consistent orthographies than in opaque ones.

It is unclear why working memory was a better predictor of reading and spelling among the bilinguals, whereas verbal STM was a consistent predictor of reading and spelling among monolinguals. The pattern of results between the monolinguals and bilinguals regarding verbal STM and working memory is intriguing and warrants further investigation. Unlike English, Arabic verbal and nonverbal control measures predicted reading and spelling. In particular, the influence of verbal STM, working memory, receptive vocabulary, and nonverbal ability perhaps reflects the overwhelming cognitive demands placed on readers of Arabic (Abu Ahmad et al., 2014).

As discussed, our understanding of orthographic depth comes from evidence that is based on English and other European languages. The findings of the current study add to the emerging literature on the predictors of Arabic reading and spelling and point to the need to reconceptualize orthographic depth to include additional features that may influence literacy development. The results of the current study regarding verbal STM may suggest that the classification of fully vowelized Arabic as transparent is accurate, and indeed, it is when we only consider consistency of grapheme-phoneme conversions as the only criterion of orthographic depth. In other words, the results here support the claim that verbal STM may be more important in consistent orthographies than in opaque ones. The finding of the strong role of PA in Arabic, coupled with the finding that RAN predicted reading for the monolingual and older bilinguals, suggests that, as decoding becomes more efficient, there is less reliance on PA and more reliance on speed, which is what would be expected of an opaque orthography. These findings challenge the current one-dimensional conceptualization of orthographic depth and calls for a multidimensional framework that encompasses other linguistic and orthographic features. As was concluded in the first chapter, Ziegler and Goswami's PGST offers a great framework for reading development from a cross-linguistic perspective; however, like its predecessors, it falls short of encompassing unique linguistic features like diglossia and other visual and orthographic features that are the source of orthographic depth in Arabic (Share & Daniels, 2016).

Exposure and experience with written language is a logical requisite to orthographic learning. Thus, it was hypothesized that print exposure would predict reading and in particular word recognition. Contrary to this prediction, print exposure did not correlate with reading and spelling in Arabic. This was a surprising finding considering that it was assumed that print exposure would even be more pertinent in Arabic due to diglossia. This leads to the question of how Arabic-speaking children construct word-specific representations while contending with limited exposure to print and an improvised oral language as a result of diglossia.

So far, the discussion has been focused on predictors of reading and spelling in Arabic. However, a serious limitation to the study of predictors of reading is that it does not address the question of how orthographic learning occurs and how we construct word-specific orthographic representations.

As discussed, the predominant account of how orthographic learning develops is Share’s (1995) self-teaching hypothesis. Once children have learned the mappings of letters and sounds, they will then use phonological decoding as way of reading unfamiliar words, which in turn will provide the opportunity for orthographic learning. Through this process, children will begin to construct word-specific orthographic representations, which supports the development of word recognition. This process has been modeled computationally (Ziegler et al., 2014) and is consistent with the lexical quality hypothesis (Perfetti & Hart, 2002), where word-specific orthographic representations that are encountered frequently will be of higher quality. This means that words that enjoy high lexical quality are identified instantly. According to Share, phonological decoding is a precursor to word recognition. However, as demonstrated in this study, even incidental exposure to a word’s orthography can facilitate learning of new words. Thus, this study demonstrated that implicit learning is quite powerful and requires few exposures. Although phonological decoding is seen as a precursor to orthographic learning in the self-teaching hypothesis, this study showed that even when the words are not presented long enough to be decoded, the words are recalled more easily when orthography was present. Similar results were obtained by Share (1999) in Hebrew, where even decoding the word once was enough to facilitate orthographic learning. This study adds to the literature that even incidental exposure to the orthographic identity of a word is quite a powerful and robust tool to activate orthographic learning in third-grade Arabic speaking children. These findings run counter to Share's self-teaching hypothesis that puts phonological decoding as a prerequisite to orthographic learning. Support for the fact that successful phonological decoding may not be necessary for orthographic learning comes from studies demonstrating that incidental exposure to a word's form is sufficient for orthographic learning (e.g., Ricketts et al., 2009). Studies using an incidental learning paradigm demonstrate rapid implicit learning from incidental exposure to orthographic representations and are consistent with statistical models of learning (Pacton et al., 2001; Qi et al., 2019).

Statistical learning in both the auditory and visual modalities captures the interaction of both the cognitive processes (e.g., attention, perception, memory) and environmental influences (e.g., characteristics of the language, instructional strategies, exposure to print) involved in reading and spelling acquisition and development. It was hypothesized that implicit statistical learning may explain how Arabic-speaking children acquire reading skills with limited oral and written vocabularies prior to formal reading instruction. Even if reading is facilitated implicitly through statistical learning, there is a consensus that reading in Arabic is challenging. What is yet unclear is the source of this difficulty. There are many culprits, such as diglossia, allography, and ligature; however, only in recent years have Arabic researchers begun to tease these issues out and to conceptualize orthographic depth as multidimensional. According to Castles and Nation (2010), one way to examine orthographic learning is through training studies such as using self-teaching paradigms, to document the process in which orthographic representations are first formed and the eventual amalgamation of the phonological, orthographic, and semantic identities.

Implicit Orthographic Learning

This study took an associative approach to describe the mechanism that underlies orthographic learning in Arabic. The current study used a paired associate learning paradigm to demonstrate the implicit nature of orthographic learning. At the heart of orthographic learning is the notion that much of our vocabulary is learned through exposure to print; however, this was not the case in Arabic, as was demonstrated in studies one and two of this research. Similarly, oral language competence (oral vocabulary) is crucial to reading development (Ehri, 2005a, 2005b). The partial alphabetic phase is where existing oral vocabulary aids in the successful decoding of words. It is proposed that children in this phase will partially decode and deduce the rest by relying on their oral lexicon. However, unlike readers of English and most other languages, beginning readers of Arabic cannot rely on their oral vocabulary to help them partially decode words due to diglossia (Saiegh-Haddad, 2003; Asaad, & Eviatar, 2014). This is consistent with SVoR and underscores the role of role of oral vocabulary in reading development. Nonetheless, once children crack the alphabetic code, the bulk of vocabulary growth is a consequence of exposure to print. Thus, vocabulary in the auditory and visual modalities is important for facilitating word recognition. It was hypothesized that children would abstract orthographic information even when presentation is incidental. While the use of paradigms like PAL may be criticized for not being as naturalistic as the self-teaching paradigm, this is not the case in a diglossic situation, where novel word pairings is the norm and not the exception.

As demonstrated by the second study of this research, orthographic depth, as currently defined, does not account for the disparity between the one-to-many (base grapheme and many consonant diacritics) that exists in Arabic, which is possibly the source of difficulty experienced by readers of Arabic. Ligature, too, has often been implicated as a source of difficulty.

The present study adds to existing literature on the value of orthography in pronunciation recall. Based on the lexical quality hypothesis (Perfetti and Hart, 2002) and Ehri’s amalgamation theory (as cited in Ehri & Wilce, 1979), it was reasoned that the activation of word-specific identities would enhance the learning of newly taught words. There was a large effect size on the recall of pronunciation in the orthography-present condition, thus providing evidence for the facilitative effects of orthography on recall. The findings in this study are consistent with those of several studies that have found that as little as one encounter with an unknown word activates its orthographic constituent (Apel, 2010; Cunningham, 2006; Ricketts et al., 2011; Share, 2004). These findings are consistent with statistical learning models demonstrating the implicit acquisition of word-specific representations after few exposures.

It was hypothesized that children would benefit from the presence of orthography in recalling semantic information as assessed by the picture-matching task. The presence of orthography facilitated semantic recall; however, unlike pronunciation, this size of the effect was small. Studies in other languages using the same paradigm (e.g., Hu, 2008; Ricketts et al., 2009; Rosenthal & Ehri, 2008) have consistently shown that orthography does not facilitate the recall of semantic information, particularly using the picture- referent task like the one used in this study. This was not the case in Arabic, as performance on this task was facilitated by the presence of orthography, demonstrating that indeed, orthography facilitates the recall of semantic information. This is the first study to examine the facilitative effects of orthography on semantic learning in Arabic. The facilitative effect on semantic recall may reflect a more accurate account of how Arabic-speaking children learn vocabulary that is often absent from their spoken vernacular. Perhaps because Arabic-speaking children lack the oral competence (vocabulary) that is outlined in the SVoR, they are more accustomed to abstracting orthographic information to aid in recall of a novel word's semantic identity. This explanation, however, is speculative and requires further investigation.

Spelling recall in the orthography-present condition was also superior in the orthography-present condition. Interestingly, the facilitative effects of orthography on spelling were moderate, which may reflect the transparency in Arabic spelling. However, the fact that orthography still made a moderate contribution to spelling despite Arabic's high consistency may reflect the inconsistency caused by diglossic phonemes (Taha, 2013) which deviate from sound to spelling. It has been suggested that spelling must map to sounds in a consistent manner in order to facilitate recall (Ehri & Wilce, 1979). Furthermore, Arabic-speaking children are less exposed to MSA as a result of diglossia, which adversely affects the learning of novel word forms (Apel, 2009). The ease of sound to spelling in Arabic, coupled with the difficulty and inconsistency of diglossic phonemes, could explain why the children in the present study found it most difficult to spell items that should be spelled with relative ease with or without the presence of orthography. However, this explanation is speculative because diglossia was not systematically examined in this study.

Next, the facilitative effects of incidental presentation of orthography on novel word learning and whether orthographic complexity has an effect on novel word learning was examined by varying ligature and consonant diacritics. As expected, recall of pronunciation, picture-matching, and spelling was superior for CF items than were NC and CM, demonstrating that connectivity facilitates orthographic learning in Arabic. These results are consistent with Dai et al.'s (2013) results showing that items with many consonant diacritics slowed reading.

The fact that CM items were the most difficult to recall demonstrates that consonant diacritics are indeed the source of difficulty in Arabic.

All the interactions between orthography-present or -absent and nonword conditions on the pronunciation, picture-matching, and spelling were not significant. The largest effect size was observed between CF and CM many. In other words, orthography facilitated all the items in all the word conditions. The lack of interaction between orthography conditions and word type may be a reflection of the transparent nature of vowelized Arabic orthography. It has been suggested that the relationship between implicit statistical learning and reading may be weaker in transparent orthographies such as Arabic because regularities are explicitly conveyed (Arciuli, 2018).

These findings are somewhat surprising and are inconsistent with Dai et al.'s conclusions that the visual demands, and/or the competition from similar-looking letters are the source of difficulty in Arabic. Because this study employed a paired associate learning task and because the nonwords in any given word condition were pronounced by the examiner and not decoded by the children, it is not the visual density of the consonant diacritics that is at the root of the problem but perhaps a phonetic feature accompanied by letters that contain many diacritics. Inspection of the nonwords in the CM condition shows that three out of the four nonwords (قثنب, جشيك, ظحفك) that were selected from Dai et al.'s stimuli contained diglossic phonemes, whereas the CF items contained no diglossic phonemes. Thus, the apparent difficulty with items with many consonant diacritics is due to difficulty in recalling items with phonemes that do not exist in the children's spoken vernacular. Although diglossia was not systematically examined in this study, it appears that the challenges that Arabic speakers encounter in literacy acquisition may be rooted in diglossia rather than in visual and orthographic demands.

Consonant diacritics and ligature in Arabic affected the learning of the different word identities in various ways. In the recall of pronunciation (phonological identity), the picture-matching (semantic identity), and spelling (orthographic identity), the CM items proved to be most difficult, regardless of whether seen or unseen. At first glance this finding is consistent with the conclusions of Dai and colleagues that consonant diacritics are both visually and orthographically demanding; however, the lack of interaction between word type and orthography condition in this study points to a different conclusion. An alternative conclusion is that diglossic phonemes are the root of reading difficulty in Arabic, and while this study supports the findings of Dai and colleagues regarding connectivity and consonant diacritics, the incidental approach in the current study demonstrated that diglossic phonemes warrant closer examination.

Limitations

Research in Arabic is emerging, and this study adds to this growing body of research. However, many more studies are required to begin formulating a theory on how reading develops in Arabic. Reading theories and models are Anglocentric (Share, 2008), and using them as backdrops to studying Arabic may not be appropriate.

Much of the research regarding orthographic learning is largely on children in second and third grades, which undermines their substantial existing orthographic knowledge. Although this study did not examine the role of context in orthographic learning, context may be crucial to reading in Arabic, particularly when diacritics are removed. Arabic readers heavily rely on context when reading unvowelized text, and this should be explored to get a more accurate account of how orthographic learning occurs in Arabic.

Because there was some evidence that diglossic phonemes may impede initial word learning, it is essential to control for such phonemes when creating nonwords, in order to separate the effects of consonant diacritics from diglossic phonemes. Diglossic phonemes are not the same across various Arabic speakers, and thus, researchers should construct nonwords that take this fact into account.

Conclusions and Implications

The most striking result of this study is the powerful influence of PA on reading and spelling in fully vowelized Arabic, which according to orthographic depth is transparent. The findings regarding orthographic learning in Arabic point to diglossic phonemes as the source of the difficulties often documented in Arabic reading development. Taken together, there seems to be a *phonological depth* (availability) caused by diglossia, rather than an orthographic depth that is defined by consistency.

References

Abdelhadi, S., Ibrahim, R., & Eviatar, Z. (2011). Perceptual load in the reading of Arabic: Effects of orthographic visual complexity on detection. *Writing Systems Research*, *3*(2), 117–127.

Abu Ahmad, H., Ibrahim, R., & Share, D. L. (2014). Cognitive predictors of early reading ability in Arabic: A longitudinal study from kindergarten to grade 2. In E. Saiegh-Haddad & M. Joshi (Eds.), *Handbook of Arabic literacy: Insights and perspectives* (pp. 171–194). Springer. https://doi.org/10.1007/978-94-017-8545-7_8

Abu-Rabia, S. (1995). Learning to read in Arabic: Reading, syntactic, orthographic and working memory skills in normally achieving and poor Arabic readers. *Reading Psychology: An International Quarterly*, *16*(4), 351–394.

Abu-Rabia, S. (2000). Effects of exposure to literary Arabic on reading comprehension in a diglossic situation. *Reading and Writing*, *13*(1–2), 147–157.

Abu-Rabia, S. (2003). Cognitive and social factors affecting Arab students learning English as a third language in Israel. *Educational Psychology: An International Journal of Experimental Educational Psychology*, *23*, 347–360. https://doi.org/10.1080/ 01443410303214

Abu-Rabia, S., Share, D., & Mansour, M. S. (2003). Word recognition and basic cognitive processes among reading-disabled and normal readers in Arabic. *Reading and Writing*, *16*(5), 423–442.

Abu-Rabia, S., & Taha, H. (2006). Phonological errors predominate in Arabic spelling across grades 1–9. *Journal of Psycholinguistic Research*, *35*(2), 167.

Adelman, J. S. (Ed.). (2012a). *Visual word recognition: Models and methods, orthography and phonology* (Vol. 1). Psychology Press.

Adelman, J. S. (Ed.). (2012b). *Visual word recognition volume 2: Meaning and context, individuals and development*. Psychology Press.

Aguiar, L., & Brady, S. (1991). Vocabulary acquisition and reading ability. *Reading and Writing*, *3*(3-4), 413–425.

Albirini, A. (2016). *Modern Arabic sociolinguistics: Diglossia, variation, codeswitching, attitudes and identity*. Routledge

Allen, L., Cipielewski, J., & Stanovich, K. E. (1992). Multiple indicators of children's reading habits and attitudes: Construct validity and cognitive correlates. *Journal of Educational Psychology, 84*(4), 489.

Al Ghanem, R., & Kearns, D. M. (2015). Orthographic, phonological, and morphological skills and children's word reading in Arabic: A literature review. *Reading Research Quarterly*, *50*(1), 83–109.

Alsudairi, N., Taibah, N. J., Haynes, C., Mahfoudi, A., Hook, P. (in press). *Standardized literacy assessment battery in Arabic: Test of reading and spelling* (اختبار القراءة والإملاء). King Salman Center for Disability Research.

Anderson, R. C., Wilson, P. T., & Fielding, L. G. (1988). Growth in reading and how children spend their time outside of school. *Reading Research Quarterly*, 285–303.

Apel, K. (2009). The acquisition of mental orthographic representations for reading and spelling development. *Communication Disorders Quarterly*, *31*(1), 42–52.

Apel, K. (2010). Kindergarten children’s initial spoken and written word learning in a storybook context. *Scientific Studies of Reading*, *14*(5), 440–463.

Apel, K. (2011). What is orthographic knowledge? *Language, Speech, and Hearing Services in Schools, 42,* 592–603. https://doi.org/10.1044/0161-1461(2011/10-0085)

Apel, K., Wolter, J. A., & Masterson, J. J. (2006). Effects of phonotactic and orthotactic probabilities during fast mapping on 5-year-olds learning to spell. *Developmental Neuropsychology*, *29*(1), 21–42.

Araújo, S., Reis, A., Petersson, K. M., & Faísca, L. (2015). Rapid automatized naming and reading performance: A meta-analysis. *Journal of Educational Psychology*, *107*(3), 868–883. https://doi.org/10.1037/edu0000006

Arciuli, J. (2018). Reading as statistical learning. *Language, Speech, and Hearing Services in Schools*, *49*(3S), 634–643.

Aro, M., & Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Applied Psycholinguistics*, *24*(4), 621–635.

Asaad, H., & Eviatar, Z. (2014). Learning to read in Arabic: The long and winding road. *Reading and Writing*, *27*(4), 649–664.

Asadi, I. A., Ibrahim, R., & Khateb, A. (2017). What contributes to spelling in Arabic? A cross-sectional study from first to sixth grade. *Writing Systems Research*, *9*(1), 60–81.

Asadi, I. A., & Khateb, A. (2017). Predicting reading in vowelized and unvowelized Arabic script: An investigation of reading in first and second grades. *Reading Psychology*, *38*(5), 486–505.

Asadi, I. A., Khateb, A., Ibrahim, R., & Taha, H. (2017). How do different cognitive and linguistic variables contribute to reading in Arabic? A cross-sectional study from first to sixth grade. *Reading and Writing*, *30*(9), 1835–1867.

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29.

Baddeley, A. (2015). 1. Working memory in second language learning. In W. Zhisheng, M. Borges Mota, & A. McNeill (Eds.), *Working memory in second language acquisition and processing* (pp. 17–28). Multilingual Matters. https://doi.org/10.21832/9781783093595-005

Bar-Kochva, I. (2013). What are the underlying skills of silent reading acquisition? A developmental study from kindergarten to the 2nd grade. *Reading and Writing*, *26*(9), 1417–1436.

Batnini, S., & Uno, A. (2015). Investigation of basic cognitive predictors of reading and spelling abilities in Tunisian third-grade primary school children. *Brain and Development*, *37*(6), 579–591.

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.

Bowers, P. G., & Wolf, M. (1993). Theoretical links among naming speed, precise timing mechanisms and orthographic skill in dyslexia. *Reading and Writing: An Interdisciplinary Journal*, *5*, 69–85. https://doi.org/10.1007/BF01026919

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research*, *80*(2), 144–179.

Bowey, J. A., & Miller, R. (2007). Correlates of orthographic learning in third-grade children's silent reading. *Journal of Research in Reading*, *30*(2), 115–128.

Buchholz, J., & McKone, E. (2004). Adults with dyslexia show deficits on spatial frequency doubling and visual attention tasks. *Dyslexia*, *10*(1), 24–43.

Cain, K. (2010). *Reading development and difficulties*. British Psychological Society and Blackwell Publishing.

Cain, K., & Oakhill, J. (2011). Matthew effects in young readers' reading comprehension and reading experience aid vocabulary development. *Journal of Learning Disabilities, 44*(5), 431–443.

Caravolas, M., Hulme, C., & Snowling, M. J. (2001). The foundations of spelling ability: Evidence from a 3-year longitudinal study. *Journal of memory and language*, *45*(4), 751-774.

Caravolas, M. (2004). Spelling development in alphabetic writing systems: A cross-linguistic perspective. *European Psychologist*, *9*(1), 3–14.

Caravolas, M., Lervåg, A., Mousikou, P., Efrim, C., Litavský, M., Onochie-Quintanilla, E., Salas, N., Schöffelová,M., Defior,S.,

Mikulajová, M., Seidlová-Málková, G., Hulme, C. (2012). Common patterns of prediction of literacy development in different alphabetic orthographies. *Psychological Science*, *23*, 678–686. https://doi.org/10.1177/0956797611434536

Caravolas, M., Volín, J., & Hulme, C. (2005). Phoneme awareness is a key component of alphabetic literacy skills in consistent and inconsistent orthographies: Evidence from Czech and English children. *Journal of Experimental Child Psychology*, *92*(2), 107–139.

Cassar, M., & Treiman, R. (1997). The beginnings of orthographic knowledge: Children's knowledge of double letters in words. *Journal of educational psychology*, *89*(4), 631.

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition*, *91*(1), 77–111.

Castles, A., Davis, C., Cavalot, P., & Forster, K. (2007). Tracking the acquisition of orthographic skills in developing readers: Masked priming effects. *Journal of Experimental Child Psychology*, *97*, 165–182. https://doi.org/10.1016/j.jecp.2007.01.006

Castles, A., & Nation, K. (2010). How does orthographic learning happen? In S. Andrews (Ed.), *From inkmarks to ideas: Current issues in lexical processing* (pp. 181–209). Psychology Press. https://doi.org/104324/9780203841211

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, *19*, 5–51. https://doi.org/10.1177/ 1529100618772271

Catts, H. W., Adolf, S. M, & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research*, *49*, 278–293

Cerni, T., Bassetti, B., & Masterson, J. (2019). Effects of orthographic forms on the acquisition of novel spoken words in a second language. *Frontiers in Communication*, *4*, 31.

Chateau, D., & Jared, D. (2000). Exposure to print and word recognition processes. *Memory & Cognition, 28*(1), 143–153.

Chen, R., & Vellutino, F. R. (1997). Prediction of reading ability: A cross-validation study of the simple view of reading. *Journal of Literacy Research*, *29*, 1–24

Chen, S-Y, & Fang, S-P. (2016). Print exposure of Taiwanese fifth graders: Measurement and prediction. *The Asia-Pacific Education Researcher, 25*(1), 69–78.

Cipielewski, J., & Stanovich, K. E. (1992). Predicting growth in reading ability from children's exposure to print. *Journal of Experimental Child Psychology, 54*(1), 74–89.

Cunningham, A. E. (2006). Accounting for children's orthographic learning while reading text: Do children self-teach? *Journal of Experimental Child Psychology, 95,* 56–77. https://doi/org/10.1016/j.jecp.2006.03.008

Cunningham, A. E., Perry, K. E., Stanovich, K. E., & Share, D. L. (2002). Orthographic learning during reading: Examining the role of self-teaching. *Journal of Experimental Child Psychology*, *82*(3), 185–199.

Cunningham, A. E., & Stanovich, K. E. (1990). Assessing print exposure and orthographic processing skill in children: A quick measure of reading experience. *Journal of Educational Psychology*, *82*(4), 733.

Cunningham, A. E, & Stanovich, K. E. (1993). Children's literacy environments and early word recognition subskills. *Reading and Writing, 5*(2), 193–204.

Cunningham, A. E., & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*(6), 934.

Dai, J., Ibrahim, R., & Share, D. L. (2013). The influence of orthographic structure on printed word learning in Arabic. *Writing Systems Research*, *5*(2), 189–213.

David, C. J. (2010). SOLAR versus SERIOL revisited. *European Journal of Cognitive Psychology, 22*(5), 695-724. https://doi.org/10.1080/09541440903155682

Davidse, N. J., de Jong, M. T, Bus, A. G., Huijbregts, S. C. J., & Swaab, H. (2011). Cognitive and environmental predictors of early literacy skills. *Reading and Writing, 24*(4), 395–412.

Deacon, S. H., Benere, J., & Castles, A. (2012). Chicken or egg? Untangling the relationship between orthographic processing skill and reading accuracy. *Cognition*, *122*, 110–117. https://doi.org/10.1016/ j.cognition.2011.09.003

Deacon, S. H., & Kirby, J. R. (2004). Morphological awareness: Just “more phonological”? The roles of morphological and phonological awareness in reading development. *Applied Psycholinguistics*, *25*(2), 223–238.

de Jong, P. F. (2011). What discrete and serial rapid automatized naming can reveal about reading. *Scientific St udies of Reading*, *15*(4), 314–337.

de Jong, P. F., & Share, D. L. (2007). Orthographic learning during oral and silent reading. *Scientific Studies of Reading*, *11*(1), 55–71.

de Jong, P. F., & Van der Leij, A. (1999). Specific contributions of phonological abilities to early reading acquisition: Results from a Dutch latent variable longitudinal study. *Journal of Educational Psychology*, *91*(3), 450.

Duff, F. J., & Hulme, C. (2012). The role of children's phonological and semantic knowledge in learning to read words. *Scientific Studies of Reading*, *16*(6), 504–525.

Dunn, L. M., & Dunn, L. M. (1997). *Examiner's manual for the Peabody Picture Vocabulary Test* (3rd ed). American Guidance Service.

Dunn, L. M., & Dunn, L. M. (2007). *Examiner's manual for the Peabody Picture Vocabulary Test* (4th ed.). American Guidance Service.

Durgunoğlu, A. Y. (2002). Cross-linguistic transfer in literacy development and implications for language learners. *Annals of Dyslexia*, *52*(1), 189–204.

Ecalle, J., Magnan, A., & Calmus, C. (2009). Lasting effects on literacy skills with a computer-assisted learning using syllabic units in low-progress readers. *Computers & Education, 52*(3), 554–561. https://doi.org/10.1016/j.compedu.2008.10.010

Echols, L. D., West, R. F., Stanovich, K. E, & Zehr, K. S. (1996). Using children's literacy activities to predict growth in verbal cognitive skills: A longitudinal investigation. *Journal of Educational Psychology, 88*(2), 296.

Ehri, L. C. (1994). Development of the ability to read words: Update. *Theoretical Models and Processes of Reading*, *4*, 323–358.

Ehri, L.C. (2005a) Development of sight word reading: Phases and findings. In M.J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook.* Blackwell Publishing

Ehri, L. C. (2005b). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, *9*, 167–188. https://doi/10.1207/s1532799xssr0902_4

Ehri, L. C., Nunes, S. R., Willows, D. M., Schuster, B. V., Yaghoub-Zadeh, Z., & Shanahan, T. (2001). Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. *Reading Research Quarterly*, *36*(3), 250–287.

Ehri, L. C., & Wilce, L. S. (1979). The mnemonic value of orthography among beginning readers. *Journal of Educational Psychology*, *71*(1), 26.

Elleman, A. M., Steacy, L. M., & Compton, D. L. (2019). The role of statistical learning in word reading and spelling development: More questions than answers. *Scientific Studies of Reading, 23*(1), 7. https://doi.10.1080/10888438.2018.1549045

Erdener, V. D., & Burnham, D. K. (2005). The role of audiovisual speech and orthographic information in nonnative speech production. *Language Learning*, *55*(2), 191–228.

Eviatar, Z., Ibrahim, R., & Ganayim, D. (2004). Orthography and the hemispheres: Visual and linguistic aspects of letter processing. *Neuropsychology*, *18*(1), 174.

Farran, L. K., Bingham, G., & Matthews, M. (2016). Language predictors of word reading in bilingual English-Arabic children. *Arab Journal of Applied Linguistics*, *2*(1), 91–116.

Feitelson, D., Goldstein, Z., Iraqi, J., & Share, D. L. (1993). Effects of listening to story reading on aspects of literacy acquisition in a diglossic situation. *Reading Research Quarterly*, 71–79.

Fragman, A. (2013). Spelling proficiency of novel phonemes in Arabic among native Hebrew academic students. *Athens Journal of Philology, 1*(1), 33–44.

Franceschini, S., Gori, S., Ruffino, M., Pedrolli, K., & Facoetti, A. (2012). A causal link between visual spatial attention and reading acquisition. *Current Biology*, *22*(9), 814–819.

Friedmann, N. & Haddad-Hanna, M. (2014). Types of developmental dyslexia in Arabic. In E. Saiegh-Haddad & R. M. Joshi (Eds.), *Handbook of Arabic literacy. Literacy Studies* (Vol. 9). https://doi.org/10.1007/978-94-017-8545-7

Frith, U. (1985). Beneath the surface of developmental dyslexia. *Surface Dyslexia*, 32.

Frost, R., Kugler, T., Deutsch, A., & Forester, K. I. (2005). The orthographic structure versus morphological structure: Principles of lexical organization in a given language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(6), 1293–1326

Fumero, K., & Tibi, S. (2020). The importance of morphological awareness in bilingual language and literacy skills: Clinical implications for speech-language pathologists. *Language, Speech, and Hearing Services in Schools*, *51*(3), 572–588.

Furnes, B., & Samuelsson, S. (2011). Phonological awareness and rapid automatized naming predicting early development in reading and spelling: Results from a cross-linguistic longitudinal study. *Learning and Individual Differences, 21*, 85–95.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics,* 27*,* 513–543. https://doi.org/10.1017/-S0142716406060383

Gathercole, S. E., Service, E., Hitch, G. J., Adams, A. M., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development:

Further evidence on the nature of the relationship. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *13*(1), 65–77.

Georgiou, G. K., Cardoso-Martins, C., Das, J. P., Falcón, A., Hosokawa, M., Inoue, T., ... & Vieira, A. P. A. (2022). Cross-language contributions of rapid automatized naming to reading accuracy and fluency in young adults: Evidence from eight languages representing different writing systems. *Journal of Cultural Cognitive Science*, 1–18.

Georgiou, G. K., Parrila, R., & Papadopoulos, T. C. (2008). Predictors of word decoding and reading fluency across languages varying in orthographic consistency. *Journal of Educational Psychology*, *100*(3), 566.

Georgiou, G. K., Parrila, R., & Papadopoulos, T. C. (2016). The anatomy of the RAN-reading relationship. *Reading and Writing*, *29*(9), 1793–1815.

Goldman, J. G, & Manis, F. R. (2013). Relationships among cortical thickness, reading skill, and print exposure in adults. *Scientific Studies of Reading, 17*(3), 163–176.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, *7*(6), 6–10

Haigh, C. A., Savage, R .S, Erdos, C., & Genesee, F. (2011). The role of phoneme and onset-rime awareness in second language reading acquisition, *Journal of Research in Reading, 34*(1), 94–113.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). Canonical correlation: A supplement to multivariate data analysis. *Multivariate data analysis: A global perspective* (7th ed.). Pearson Prentice Hall Publishing.

Hammer, C. S., Hoff, E., Uchikoshi, Y., Gillanders, C., Castro, D. C., & Sandilos, L. E. (2014). The language and literacy development of young dual language learners: A critical review. *Early Childhood Research Quarterly*, *29*(4), 715–733.

Hansen, G. F. (2014). Word recognition in Arabic: Approaching a language-specific reading model. In E. Saiegh-Haddad & M. Joshi (Eds.), *Handbook of Arabic literacy: Insights and Perspectives* (pp. 55–76). Springer.

Harm, M. W., McCandliss, B. D., & Seidenberg, M. S. (2003). Modeling the successes and failures of interventions for disabled readers. *Scientific Studies of Reading*, *7*(2), 155–182.

Haynes, C., Taibah, N. J., Hook, P., Alsudairi, N., Mahfoudi, A. (in press). *Standardized literacy assessment battery in Arabic: Test of visual/orthographic processing* (اختبار معالجة العمليات البصرية والرسم الإملائي). King Salman Center for Disability Research.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2, 127–160.

Hu, C. F. (2008). Use orthography in L2 auditory word learning: Who benefits? *Reading and Writing*, *21*(8), 823–841

Hulme, C., Goetz, K., Gooch, D., Adams, J., & Snowling, M. J. (2007). Paired-associate learning, phoneme awareness, and learning to read. *Journal of Experimental Child Psychology*, *96*(2), 150–166.

Hulme, C., Hatcher, P. J., Nation, K., Brown, A., Adams, J., & Stuart, G. (2002). Phoneme awareness is a better predictor of early reading skill than onset-rime awareness. *Journal of Experimental Child Psychology*, *82*, 2–28.

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*, *2*, 93.

Ibrahim, R. (2013). Beginning readers in Arabic and the distance between literary and spoken Arabic. *Creative Education*, *4*(05), 307.

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

Ibrahim, R., & Eviatar, Z. (2012). The contribution of the two hemispheres to lexical decision in different languages. *Behavioral and Brain Functions*, *8*(3), 2–15.

Ibrahim, R., Shibel, S., & Hertz-Lazarowitz, R. (2014). The complex nature of text reading difficulties: The case of bilingual children. *Psychology*, *5*(16), 1911.

Inoue, T., Georgiou, G. K., Muroya, N., Hosokawa, M., Maekawa, H., & Parrila, R. (2020). Predicting the early growth of word and nonword reading fluency in a consistent syllabic orthography. *Journal of Research in Reading*, *43*(3), 364–381

Jarrold, C., Baddeley, A. D., Hewes, A. K., Leeke, T. C., & Phillips, C. E. (2004). What links verbal short-term memory performance and vocabulary level? Evidence of changing relationships among individuals with learning disability. *Journal of Memory and Language*, 50, 134–148. https://doi.org/10.1016/-j.jml.2003.10.004

Jarrold, C., Thorn, A. S. C., & Stephens, E. (2009). The relationship among verbal short-term memory, phonological awareness, and new word learning: Evidence from typical development and Down syndrome.

Journal of Experimental Child Psychology, 102*,* 196–218. https://doi.org/10.1016/j.jecp.2008.07.001

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. *Advances in Psychology*, 94, 67-84.

Khateb, A., Khateb-Abdelgani, M., Taha, H. Y., & Ibrahim, R. (2014). The impact of orthographic connectivity on visual word recognition in Arabic: A cross-sectional study. *Reading and Writing*, 1–24.

Kirby, J. R., Deacon, S. H., Bowers, P. N., Izenberg, L., WadeWoolley, L., & Parrila, R. (2012). Children's morphological awareness and reading ability. *Reading and Writing: An Interdisciplinary Journal*, *25*, 389–410. https://doi.org/10.1007/ s11145-010-9276-5

Kirby, J. R., Desrochers, A., Roth, L., & Lai, S. S. V. (2008). Longitudinal predictors of word reading development. *Canadian Psychology*, *49*, 103–110. https://doi.org/10.1037/0708-5591. 49.2.103

Kirby, J. R., Georgiou, G. K., Martinussen, R., & Parrila, R. (2010). Naming speed and reading: From prediction to instruction. *Reading Research Quarterly*, *45*, 341–362. https://doi.org/ 10.1598/RRQ.45.3.4

Kirby, J. R., Parrila, R. K., & Pfeiffer, S. L. (2003). Naming speed and phonological awareness as predictors of reading development. *Journal of Educational Psychology*, *95*(3), 453.

Landerl, K., Freudenthaler, H. H., Heene, M., De Jong, P. F., Desrochers, A., Manolitsis, G., ... & Georgiou, G. K. (2019). Phonological awareness and rapid automatized naming as longitudinal predictors of reading in five alphabetic orthographies with varying degrees of consistency. *Scientific Studies of Reading*, *23*(3), 220–234.

Landerl, K., & Wimmer, H. (2000). Deficits in phoneme segmentation are not the core problem of dyslexia: Evidence from German and English children. *Applied Psycholinguistics*, *21*, 243–262. https://doi.org/10.1017/S0142716400002058

Landerl, K., & Wimmer, H. (2008). Development of word reading fluency and spelling in a consistent orthography: An 8-year follow-up. *Journal of Educational Psychology*, *100*(1), 150.

Leil, A. K. A., Share, D. L., & Ibrahim, R. (2014). How does speed and accuracy in reading relate to reading comprehension in Arabic? *Psicológica*, *35*(2), 251–276.

Lervåg, A., & Hulme, C. (2009). Rapid automatized naming (RAN) taps a mechanism that places constraints on the development of early reading fluency. *Psychological Science*, *20*(8), 1040–1048.

Litt, R. A., & Nation, K. (2014). The nature and specificity of paired associate learning deficits in children with dyslexia. *Journal of Memory and Language*, *71*(1), 71–88.

Mahfoudi, A., Haynes, C., Taibah, N. J., Alsudairi, N., Hook, P., (in press). *Standardized literacy assessment battery in Arabic: Test of oral language* (اختبار اللغة الشفهية). King Salman Center for Disability Research.

Manis, F. R., Doi, L. M., & Bhadha, B. (2000). Naming speed, phonological awareness, and orthographic knowledge in second graders. *Journal of Learning Disabilities*, *33*, 325–333. https://doi.org/10.1177/002221940003300405

Manis, F. R., Seidenberg, M. S., & Doi, L. M (1999). See Dick RAN: Rapid naming and the longitudinal prediction of reading subskills in first and second graders. *Scientific Studies of Reading, 3*(2), 129–157

Martinez, D., Georgiou, G. K., Inoue, T., Falcón, A., & Parrila, R. (2021). How does rapid automatized naming influence orthographic knowledge? *Journal of Experimental Child Psychology*, *204*, 105064.

Masterson, J., & Hayes, M. (2007). Development and data for UK versions of an author and title recognition test for adults. *Journal of Research in Reading, 30*(2), 212–219. https//doi.org/10.1111/j.1467-9817.2006.00320.x

Mather, N., Roberts, R., Hammill, D., & Allen, E. (2008). Test of Orthographic Competence (TOC) Pro-Ed. *Austin, TX*.

McBride-Chang, C., Cho, J.-R., Liu, H., Wagner, R. K., Shu, H., Zhou, A., . . Muse, A. (2005). Changing models across cultures: Associations of phonological awareness and morphological structure awareness with vocabulary and word recognition in second graders from Beijing, Hong Kong, Korea, and the United States. *Journal of Experimental Child Psychology*, *92*, 140–160. https://doi.org/10.1016/j.jecp.2005.03.009

McBride-Chang, C., Manis, F. R., Seidenberg, M. S., Custodio, R. G., & Doi, L. M. (1993). Print exposure as a predictor of word reading and reading comprehension in disabled and nondisabled readers. *Journal of Educational Psychology*, *85*(2), 230.

Melby-Lervåg, M., & Lervåg, A. (2011). Cross-linguistic transfer of oral language, decoding, phonological awareness and reading comprehension: A meta-analysis of the correlational evidence. *Journal of Research in Reading*, *34*(1), 114–135.

Melby-Lervåg, M., Lyster, S. A. H., & Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, *138*(2), 322.

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Lawrence Erlbaum Associates.

Mol, S. E., & Bus, A. G. (2011). To read or not to read: A meta-analysis of print exposure from infancy to early adulthood. *Psychological Bulletin*, *137*(2), 267.

Moll, K., Ramus, F., Bartling, J., Bruder, J., Kunze, S., Neuhoff, N., ... & Landerl, K. (2014). Cognitive mechanisms underlying reading and spelling development in five European orthographies. *Learning and Instruction*, *29*, 65–77.

Muter, V. Hulme, C., Snowling, M. J., & Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: Evidence from a longitudinal study. *Developmental Psychology*, *40*(5), 665–681.

Nag, S. (2011). The *akshara* languages: What do they tell us about children's literacy learning? *Language-cognition: State of the Art*, 291–310.

Nag, S., & Snowling, M. J. (2012). Reading in an alphasyllabary: Implications for a language universal theory of learning to read. *Scientific Studies of Reading*, *16*(5), 404–423.

Nag, S., & Snowling, M. J. (2013). Children's reading development: Learning about sounds, symbols and cross-modal mappings. In. B. R. Kar (Ed.), *Cognition and brain development: Converging evidence from* *various methodologies* (pp. 253–270). American Psychological Association. https://doi.org/10.1037/14043-013

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading*, *27*, 342–356.

Nation, K., Angell, P., & Castles, A (2007). Orthographic learning via self-teaching in children learning to read English: Effects of exposure, durability, and context. *Journal of Experimental Child Psychology,* 96*,* 71–84. https://doi.org/10.1016/-j.jecp.2006.06.004

Ouellette, G. P. (2006). What' meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, *98*(3), 554.

Ouellette, G., & Beers, A. (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story. *Reading and writing*, *23*(2), 189–208.

Pacton, S., Perruchet, P., Fayol, M., & Cleeremans, A. (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General*, *130*(3), 401.

Papadopoulos, T. C., Georgiou, G., & Kendeou, P. (2009). Investigating the double-deficit hypothesis in Greek: Findings from a longitudinal study. *Journal of Learning Disabilities*, *42*, 528–547. https://doi.org/10.1177/0022219409338745

Parrila, R., Kirby, J. R., & McQuarrie, L. (2004). Articulation rate, naming speed, verbal short-term memory, and phonological awareness: Longitudinal predictors of early reading development. *Scientific Studies of Reading*, *8*, 3–26.

Peeters, D., Runnqvist, E., Bertrand, D., & Grainger, J. (2014). Asymmetrical switch costs in bilingual language production induced by reading words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(1), 284.

Peng, P., Barnes, M., Wang, C., Wang, W., Li, S., Swanson, H. L., ... & Tao, S. (2018). A meta-analysis on the relation between reading and working memory. *Psychological Bulletin*, *144*(1), 48.

Perez, T. M., Majerus, S., Mahot, A., & Poncelet, M. (2012). Evidence for a specific impairment of serial order short-term memory in dyslexic children. *Dyslexia*, *18*(2), 94–109.

Perez, T. M., Majerus, S., & Poncelet, M. (2012). The contribution of short-term memory for serial order to early reading acquisition: Evidence from a longitudinal study. *Journal of Experimental Child Psychology*, *111*(4), 708–723.

Perfetti, C. A., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro & P. Reitsma (Eds.), *Precursors of functional literacy* (Vol. 11, pp. 189–214). John Benjamins.

Plaut, D. C. (2005). Connectionist approaches to reading. In M. J. Snowling & C. Hulme (Eds.), *The science of reading: A handbook* (pp. 24–38). Blackwell Publishing.

Pollo, T. C., Treiman, R., & Kessler, B. (2008). Preschoolers use partial letter names to select spellings: Evidence from Portuguese. *Applied Psycholinguistics*, *29*(2), 195–212.

Powell, D., & Atkinson, L. (2021). Unraveling the links between rapid automatized naming (RAN), phonological awareness, and reading. *Journal of Educational Psychology*, *113*(4), 706.

Powell, D., Stainthorp, R., & Stuart, M. (2014). Deficits in orthographic knowledge in children poor at rapid automatized naming (RAN) tasks? *Scientific Studies of Reading*, *18*(3), 192–207.

Qi, Z., Sanchez Araujo, Y., Georgan, W. C., Gabrieli, J. D., & Arciuli, J. (2019). Hearing matters more than seeing: A cross-modality study of statistical learning and reading ability. *Scientific Studies of Reading*, *23*(1), 101–115.

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*, *24*(6), 679–695.

Raven, J. C., John Hugh Court, & Raven, J. E. (1998). *Raven's coloured progressive matrices*. Harcourt Assessment.

Ravid, D., & Schiff, R. (2006). Roots and patterns in Hebrew language development: Evidence from written morphological analogies. *Reading and Writing*, *19*(8), 789–818.

Ricketts, J., Bishop, D. V., & Nation, K. (2009). Orthographic facilitation in oral vocabulary acquisition. *Quarterly Journal of Experimental Psychology*, *62*(10), 1948–1966.

Ricketts, J., Bishop, D. V. M., Pimperton, H., & Nation, K. (2011). The role of self-teaching in learning orthographic and semantic aspects of new words. *Scientific Studies of Reading,* 15*,* 47–70. https://doi.org/10.1080/10888438.2011.536129

Ricketts, J., Nation, K., & Bishop, D. V. (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading*, *11*(3), 235–257.

Rose, A. (2009). Identifying and teaching children and young people with dyslexia and literacy difficulties. *An Independent Report from Sir Jim Rose to the Secretary of State for Children, Schools and Families.*

Rosenthal, J., & Ehri, L. C. (2008). The mnemonic value of orthography for vocabulary learning. *Journal of Educational Psychology*, *100*(1), 175.

Sacre, L., & Masterson, J. (2000). Single word spelling test (SWST). NFER-Nelson.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926-–1928.

Saiegh-Haddad, E. (2003). Linguistic distance and initial reading acquisition: The case of Arabic diglossia. *Applied Psycholinguistics*, *24*(03), 431–451

Saiegh-Haddad, E. (2005). Correlates of reading fluency in Arabic: Diglossic and orthographic factors. *Reading and Writing: An Interdisciplinary Journal*, *18*, 559–582. https://doi.org/10.1007/ s11145-005-3180-4

Saiegh-Haddad, E. (2013). A tale of one letter: Morphological processing in early Arabic spelling. *Writing Systems Research*, *5*(2), 169–188.

Saiegh-Haddad, E. (2018). MAWRID: A model of Arabic word reading in development. *Journal of Learning Disabilities*, *51*(5), 454–462.

Saiegh-Haddad, E., & Geva, E. (2008). Morphological awareness, phonological awareness, and reading in English–Arabic bilingual children. *Reading and Writing*, *21*(5), 481–504.

Saiegh-Haddad, E., & Henkin-Roitfarb, R. (2014).The structure of Arabic language and orthography. In E. Saiegh-Hadad & R. M. Joshi (Eds.), *Handbook of Arabic literacy volume 9* (pp. 3-28*).* Springer Netherlands

Saiegh-Haddad, E., & Lina, H. A. J. (2018). Does phonological distance impact quality of phonological representations? Evidence from Arabic diglossia. *Journal of Child Language*, *45*(6), 1377–1399.

Saiegh-Haddad, E., & Taha, H. (2017). The role of morphological and phonological awareness in the early development of word spelling and reading in typically developing and disabled Arabic readers. *Dyslexia*, *23*(4), 345–371.

Sénéchal, M., & LeFevre, J-A. (2002). Parental involvement in the development of children’s reading skill: A five-year longitudinal study. *Child Development, 73*(2), 445–460.

Seymour, P. H., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*(2), 143–174.

Shahar-Yames, D., & Share, D. L. (2008). Spelling as a self-teaching mechanism in orthographic learning. *Journal of Research in Reading, 31,* 22–39. https://doi.org/10.1111/j.1467-9817.2007.00359.x

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, *55*, 151–218. https://doi.org/10.1016/0010-0277(94)00645-2

Share, D. L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, *72*(2), 95–129.

Share, D. L. (2004). Orthographic learning at a glance: On the time course and developmental onset of self-teaching. *Journal of Experimental Child Psychology,* 87, 267–298. https://doi.org/10.1016/j.jecp.2004.01.001

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an “outlier” orthography. *Psychological Bulletin*, 134, 584–615. https://doi.org/10.1037/0033-2909.134.4.584

Share, D. L. (2014). Alphabetism in reading science. *Frontiers in Psychology*, 5, 1–4, https://doi.org/10.3389/fpsyg.2014.00752

Share, D. L., & Daniels, P. T. (2016). Aksharas, alphasyllabaries, abugidas, alphabets and orthographic depth: Reflections on Rimzhim, Katz and Fowler (2014). *Writing Systems Research*, *8*(1), 17–31.

Sheslow, D., & Adams, W. (2003). *Wide Range Assessment of Memory and Learning Second Edition (WRAML-2).* Wide Range.

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, *117*(1), 34.

Snowling, M. J. (1998). Dyslexia as a phonological deficit: Evidence and implications. *Child Psychology and Psychiatry Review*, *3*(1), 4–11

Snowling, M. J. (2001). From language to reading and dyslexia, *Dyslexia*, 7, 37–46.

Spencer, M., Kaschak, M. P., Jones, J. L., & Lonigan, C. J. (2015). Statistical learning is related to early literacy-related skills. *Reading and Writing*, *28*(4), 467–490.

Stainthorp, R. (1997). A children’s author recognition test: A useful tool in reading research. *Journal of Research in Reading*, *20*(2), 148–158.

Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly,* 21, 360–407.

Stanovich, K. E, & Cunningham, A. E. (1992). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. *Memory & Cognition, 20*(1), 51–68.

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, 402–433.

Stuart, M., Stainthorp, R., & Snowling, M. (2008). Literacy as a complex activity: Deconstructing the simple view of reading. *Literacy*, *42*(2), 59–66.

Taibah, N. J., & Haynes, C. W. (2011). Contributions of phonological processing skills to reading skills in Arabic-speaking children. *Reading and Writing: An Interdisciplinary Journal*, *24*, 1019–1042. https://doi.org/10.1007/s11145-010- 9273-8

Taibah, N. J., Haynes, C., Hook, P., Alsudairi, N., Mahfoudi, A. (in press). *Standardized literacy assessment battery in Arabic: Test of phonological processing* (اختبار معالجة العمليات الصوتية). King Salman Center for Disability Research.

Taha, H. (2016). The development of reading and spelling in Arabic orthography: Two parallel processes? *Reading Psychology*, *37*(8), 1149–1161.

Taha, H. Y. (2013). Reading and spelling in Arabic: Linguistic and orthographic complexity. *Theory and Practice in Language Studies*, *3*(5), 721–727.

Tamura, N., Castles, A., & Nation, K. (2017). Orthographic learning, fast and slow: Lexical competition effects reveal the time course of word learning in developing readers. *Cognition*, *163*, 93–102.

Thomas, M. S., & Allport, A. (2000). Language switching costs in bilingual visual word recognition. *Journal of memory and language*, *43*(1), 44-66.

Tibi, S., Edwards, A. A., Schatschneider, C., & Kirby, J. R. (2020). Predicting Arabic word reading: A cross-classified generalized random-effects analysis showing the critical role of morphology. *Annals of Dyslexia*, *70*(2), 200–219.

Tibi, S., Fitton, L., & McIlraith, A. L. (2021). The development of a measure of orthographic knowledge in the Arabic language: A psychometric evaluation. *Applied Psycholinguistics*, *42*(3), 739–762.

Tibi, S., & Kirby, J. R. (2018). Investigating phonological awareness and naming speed as predictors of reading in Arabic. *Scientific Studies of Reading*, *22*(1), 70–84.

Tibi, S., & Kirby, J. R. (2019). Reading in Arabic: How well does the standard model apply? *Journal of Speech, Language, and Hearing Research*, *62*(4), 993–1014.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A. (2012). *Test of Word Reading Efficiency–Second Edition*. Pro-Ed.

Treiman, R., & Cassar, M. (1997). Spelling acquisition in English. *Learning to spell: Research, theory, and practice across languages*, 61-80.

Vance, M., & Mitchell, J .E. (2006). Short-term memory: Assessment and intervention. In M. J. Snowling & J. Stackhouse (Eds.), *Dyslexia speech and language: A practitioner's handbook.* Whurr Publishers Ltd.

Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., Chen, R. (2007). Components of reading ability: Multivariate evidence for a convergent skills model of reading development, *Scientific Studies of Reading*, *11*(1), 3–32.

Verhoeven, L., & Perfetti, C. A. (2011). Introduction to this special issue: Vocabulary growth and reading skill. *Scientific Studies of Reading, 15,* 1–7. https://doi.org/10.1080/10888438.2011.536124

Verhoeven, L., & van Leeuwe, J. (2012). The simple view of second language reading throughout the primary grades. *Reading and Writing*, *25*(8), 1805–1818.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin. 101*(2), 192.

Wagner, R .K., Torgesen, J. K., & Rashotte, C .A. (1999). *CTOPP: Comprehensive test of phonological processing*. Pro-Ed.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., & Pearson, N. A. (2013). *Comprehensive test of phonological processing–2nd ed. (CTOPP-2)*. Pro-Ed.

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., Donahue, J., & Garon, T. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A 5-year longitudinal study, *Developmental Psychology*, *33*(3), 468–479.

Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(5), 1256.

Wang, H. C., Wass, M., & Castles, A. (2017). Paired-associate learning ability accounts for unique variance in orthographic learning. *Scientific Studies of Reading*, *21*(1), 5–16.

Wiig, E. H., Semel, E., & Secord, W. A. (2013). *Clinical evaluation of language fundamentals–Fifth Edition (CELF-5)*. NCS Pearson.

Wimmer, H., & Goswami, U. (1994). The influence of orthographic consistency on reading development: Word recognition in English and German children. *Cognition*, *51*(1), 91–103.

Woodcock, R. W. (2011). *Woodcock Reading Mastery Tests–Third Edition*. Pearson.

Wolf, M., & Bowers, P. G. (1999). The double-deficit hypothesis for the developmental dyslexias. *Journal of Educational Psychology*, *91*, 415–438. https://doi.org/10.1037/0022-0663.91.3.415

Wolf, M., & Bowers, P. G. (2000). Naming-speed processes and developmental reading disabilities: An introduction to the special issue on the double-deficit hypothesis. *Journal of Learning Disabilities*, *33*(4), 322–324.

Zhang, S. Z., Inoue, T., Shu, H., & Georgiou, G. K. (2020). How does home literacy environment influence reading comprehension in Chinese? Evidence from a 3-year longitudinal study. *Reading and Writing*, *33*(7), 1745–1767.

Ziegler, J. C., Bertrand, D., Tóth, D., Csépe, V., Reis, A., Faísca, L., ... & Blomert, L. (2010). Orthographic depth and its impact on universal predictors of reading: A cross-language investigation. *Psychological Science*, *21*(4), 551–559.

Ziegler, J. C., & Goswami, U. C. (2005). Reading acquisition, developmental dyslexia and skilled reading across languages: A psycholinguistic gain size theory. *Psychological Bulletin*, *131*(1), 3–29.

Ziegler, J. C., Perry, C., & Zorzi, M. (2014). Modelling reading development through phonological decoding and self-teaching: Implications for dyslexia. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634), 20120397.

Endnotes

[1] In the course of recruiting schools for this study, eight international schools were contacted. With the exception of two schools, all declined to participate. Of the two that initially showed interest, one school later declined, stating that the majority of parents did not accept the invitation to participate. Thus, the entire sample for study 1a and study 1b had to be recruited from the same school.

[2] English reading fluency was assessed using two measures: phonological decoding (nonword) reading fluency and sight-word reading fluency. In Arabic there was no comparable assessment for the sight-word reading fluency measure. Thus, only nonword reading fluency was assessed in Arabic.

Appendix A

Letter to the School Principal

Study 1

Dear School Principal:

My name is Rima Balshe. I am a PhD student at the Institute of Education at the University of London. I work with children with various special educational needs at the Learning Resource Center (LRC), where I provide academic support in the form of remedial tutoring sessions.

I am writing to inquire if you would be willing for me to do my research in your school.

The purpose of my research is to examine the effects of print exposure on word recognition processes among native Arabic speakers learning English as a second language. The participants of interest are 3rd, 4th and 5th graders. Data will be gathered in the form of academic and experimental assessments. Participants will be assessed individually in their familiar school setting during school hours.

The participants will remain totally anonymous; the participating students' names will not be revealed in any of the documents nor to other researchers. Furthermore, the information collected will be treated sensitively and will be highly confidential; it will be only used for the purpose of the current research.

Students and their parents will be given a 'Consent Form' to ensure their acceptance of the participation in the research. Students and their parents who wish for their kids to participate in the research will sign the form and return it to the class teacher within a specified period of time.

If you would like to discuss any part of this research, please contact me via e-mail or cellular number below. I will provide the school with a copy of the research once it is completed for interested teachers to learn more about the study.

Thank you for taking the time to consider this request, and I hope that you find my research topic of interest to your school.

Rima Balshe
Cellular #: [redacted]
Email: [redacted]

Appendix B

Parental and Student Consent Form

Study 1

Dear Parent and Student:

My name is Rima Balshe. I am a PhD student at the Institute of Education at the University of London. I work with children with various special educational needs at the Learning Resource Center (LRC), where I provide academic support in the form of remedial tutoring sessions.

I am writing to ask your permission for your child to take part in my research project being carried out at your child's school. The purpose of my project is to look at the effects of print exposure on word reading in 1st to 5th grade Arabic-speaking children learning English as a second language. The project will involve asking the children to take part in tasks where they will see the names of children's books and titles of books and they will be asked to check which ones they are familiar with. They will also be asked about their favourite pastimes. All the assessments will be administered by myself, either seeing the children in groups in the case of Grades 2–6 children, or individually, in the case of Grade 1 children, during school hours.

The school administration has given its approval for this project. Taking part is entirely voluntary and your child's participation or lack thereof will not reflect in any way on his/her grades. If you are happy for your child to take part I would ask you to please discuss the contents of this letter with your child prior to making your decision.

The children taking part will remain anonymous in any reports; your child's name will not be revealed in any of the documents, to the school, nor to other researchers. Furthermore, the information collected will be confidential and it will be only used for the purpose of the current project. The data collected will be summarized anonymously in my PhD dissertation and in subsequent publications.

To ensure your confidentiality, an envelope is provided for you in which to place and seal your reply form, to be returned with your child please by the date specified below. Taking part in the study is entirely voluntary. If you and your child consent to taking part you will be free to withdraw from participation at any time and without giving reason. Any data I have collected for your child would then be destroyed and not used in any subsequent reports. If you would like to discuss any part of this project, you can contact me at the e-mail address below.

Please return the enclosed form by 27, 03, 2016***, if you do not wish for your child to participate. Otherwise, your child maybe randomly selected to participate in the research project***

Thank you for your cooperation
Rima Balshe

E-mail Address: [redacted]

Print exposure project
Consent form

I, parent of ____________________, have read the information letter and do not give consent* to my son/daughter taking part in the research project being conducted at school by Rima Balshe.

Parent's Signature________________________

Student's Signature________________________

Appendix C

Scores for Correct Selection of Targets and Selection of Foils

Study 1a

Frequencies and Percentages for Choosing Arabic Targets (Lists A, B, and C)

Title	*n*	%
صلاح الدين الأيوبي بطل حطين	15	51.7
الأمير الصغير	13	44.8
حكايات الفرعون الصغير	11	37.9
حكايات من شعر أمير الشعراء أحمد شوقي	11	37.9
ويليام شكسبير	8	28.6
كليوباترا سلسلة شخصيات شهيرة	7	25
حكايات الحكيم لقمان	7	25
النقطه السوداء	7	25
الولد الكذاب و الذئب	7	24.1
كن نفسك	7	24.1
ماهاتما غاندي	6	21.4
حكايات الحكيم لقمان	6	20.7
مغامرات ياسمينه	6	20.7
كركر هانم	6	20.7
كلب طيب القلب	6	20.7
الكنز السحري	6	20.7
فرحانه تعلم نورا الشجاعة	6	20.7
بساط الحواديت	5	17.9
مارى كورى	5	17.9
قصص من سكسبير	5	17.9
بيبو و القراصنة	5	17.2
استاذ عطسان	5	17.2
ما معنى الروح الرياضية	5	17.2
انا لونا و أنا رياضية	5	17.2
فرحانه و ملابس السهرة	5	17.2
حكايات عمو محمود	5	17.2
قصة الكهرباء المكهربة	4	14.3
لويس باستير	4	14.3
طيري يا طيارة	4	14.3
فرحانه تستقبل أخا	4	13.8
دودو و النجم القطبى	4	13.8
رد جميل	4	13.8
الاتوبيس السحري	4	13.8
أنا و جدتى	4	13.8
هانز كريستسان اندرسن	3	10.7
نساء صغيرات	3	10.7
أجازة فأرية جدا	3	10.7
خالد بن الوليد بطل اليرموك	3	10.7

Title	*n*	%
فرحانه و عيد ميلاد بابا	3	10.7
فرحانة تبحث عن كتابها	3	10.3
الكلب الشقى	3	10.3
اندروكليز و الاسد	3	10.3
الامبراطور و الكروان	3	10.3
نصف فيزو	3	10.3
فيليب و عمر	3	10.3
فرحانه تحب عيد ميلادها	3	10.3
لا أحد ينام في مدينة الاحلام	2	6.9
فرحانة تحب المشاركة	2	6.9
فرحانه تتعرف على الاشكال	2	6.9
الثور الغضبان	2	6.9
فرحانه تحلم بدور سندريلا	2	6.9
أستطيع أن اقرأ و أنا مغمض العينين	2	6.9
تيمور و الخرافات	2	6.9
حقيبة كبيرة من اللهموم	1	3.6
قبلني قبلة المساء دائما	1	3.6
فرحانه تتعرف على الاشكال	1	3.6
ما يحتاجه العالم الأن	1	3.6
فرحانه تقول الحقيقة	1	3.4
تقبلني دائما كما أنا	1	3.4
الفيلسوف الصغير	0	0

Frequencies and Percentages for Choosing Arabic Foils (N = 86)

Title	*n*	%
قصص من زمن العصر الحجري	28	32.6
الصندوق السحري	25	29.1
القرش و السمكة الذكية	21	24.4
أصدقاء التنين	19	22.1
الاشياء التى تسعدنى	14	16.3
رحلة مع عائلتى	13	15.1
ما تعلمته من جدي	13	15.1
لا أستطيع النوم مبكرا	6	7
اشباح القلعة	6	7
لماذا لم يأتى أحمد الى المدرسة	2	2.3

Frequencies and Percentages for Choosing English Targets (Lists A & B)

Title	*N*	%
Diary of a Wimpy Kid	38	88.4
Matilda	36	83.7
Charlie and the Chocolate Factory	35	81.4
Big Nate	23	53.5
Hunger Games	21	48.8
Dork Diaries	20	46.5
Heroes of Olympus	12	27.9
Witches	9	20.9
The Power of Six	9	20.9
Little Darlings	8	18.6
A Candy Apple	7	16.3
The 39 Clues	7	16.3
The War Horse	7	16.3
I, Hero	6	14
Judy Moody	6	14
Running Wild	6	14
The Host	6	14
The Diamond Girls	6	14
Magic Tree House	6	14
Captain Underpants	6	14
Goosebumps	6	14
The Last Wolf	6	14
Superfudge	5	11.6
The Sand Man and the Turtles	5	11.6
Cookie	4	9.3
The Gallagher Academy	4	9.3
The Red Pyramid	4	9.3
The Cupcake Diaries	3	7
White Mountains	3	7
Origami Yoda	2	4.7
The Breadwinner	2	4.7
I am number four	2	4.7
The Graveyard Book	2	4.7
A Poison Apple	2	4.7
Asterix	2	4.7
Hunger, Lies, Plague	1	2.3
Mud City	1	2.3
Geronimo Stilton	0	0
Confessions of a Teenager	0	0
Heist Society	0	0

Frequencies and Percentages for Choosing English Foils (N = 86)

Title	*N*	%
The Dark Horse	27	31.4
Mountain High School	18	20.9
The Secret Society	17	19.8
The Adventures of a Teenager	15	17.4
Cupcake Factory	13	15.1
The Last Missing Clue	11	12.8
Nerd Academy	10	11.6
Dire Games	6	7.0
Captain Boda	6	7.0
Nerd Generation	5	5.8

Study 1b

Frequencies and Percentages for Choosing Arabic Targets (N = 76)

Title	*n*	%
الأمير الصغير	48	63.2
صلاح الدين الأيوبي بطل حطين	41	53.9
قصص من شكسبير	38	50.0
مغامرات ياسمينه	34	44.7
حكايات الفرعون الصغير	35	42.1
الكنز السحري	31	40.8
كليوباترا سلسلة شخصيات شهيرة	30	39.5
بساط الحواديت	27	35.5
حكايات من شعر أمير الشعراء أحمد شوقي	26	34.2
كركر هانم	25	32.9
ويليام شكسبير	24	31.6
حكايات الحكيم لقمان	24	31.6
كلب طيب القلب	24	31.6
فرحانه تعلم نورا الشجاعة	22	28.9
انا لونا و أنا رياضية	20	26.3
بيبو و القراصنة	32	25.0
كن نفسك	19	25.0
حكايات توفيق الحكيم	19	25.0
النقطه السوداء	18	23.7
ماهاتما غاندي	10	13.2

Frequencies and Percentages for Choosing Arabic Foils (N = 76)

Title	*n*	%
مدينتي نظيفه	33	43.4
نور و الطيور	29	38.2
حقيبتي الصغيره	28	36.8
البيت الازرق الصغير	28	36.8
القبطان سعيد	23	30.3
ما تعلمته من جدي	21	27.6
لا استطيع النوم مبكرا	19	25
لماذا لم ياتي احمد الى المدرسه؟	18	23.7
تذكرنى دائما	17	22.4
الاشياء التى تسعدني	14	18.4

Frequencies and Percentages for Choosing English Targets (N = 76)

Title	*N*	%
Big Nate	46	60.5
Dork Diaries	41	53.9
The War Horse	41	53.9
Captain Underpants	37	48.7
Magic Tree House	35	46.1
Heroes of Olympus	34	44.7
The Last Wolf	33	43.4
The Power of Six	31	40.8
The Sand Man and the Turtles	30	39.5
The Diamond Girls	26	34.2
I, Hero	26	34.2
Little Darlings	24	31.6
Witches	23	30.3
Running Wild	22	28.9
Judy Moody	21	27.6
The 39 Clues	21	27.6
Apple Candy	21	27.6
The Host	18	23.7
Superfudge	18	23.7
Goosebumps	16	21.1

Frequencies and Percentages for Choosing English Foils (N = 76)

Title	*N*	%
The Missing Letter	27	35.5
Milton City	23	30.3
Teenage Society	23	30.3
Nerd Academy	21	27.6
Captain Boda	20	26.3
Dire Games	18	23.7
The Hideaway	14	18.4
It’s My Room	12	15.8
Nerd Generation	11	14.5
Man, Mud, Plague	9	11.8

Appendix D
Reading Habits Questionnaire

Below you will be given a choice between engaging in one of two activities. Please put a check mark next to the one that you prefer. Please mark only one. That is, even if you like both activities, please mark only the one you like better. Similarly, even if you dislike both activities, mark the one that you would prefer to do. For each item, please mark only one choice.

1. I would rather (a) listen to music of my choice, or (b) spend time on my hobbies.
2. I would rather (a) read a book of my choice, or (b) listen to music of my choice.
3. I would rather (a) attend a movie of my choice, or (b) play an outdoor sport of my choice.
4. I would rather (a) watch a television program of my choice, or (b) read a book of my choice.
5. I would rather (a) spend time on my hobbies, or (b) play an outdoor sport of my choice.
6. I would rather (a) listen to music of my choice, or (b) talk with friends of my choice.
7. I would rather (a) read a book of my choice, or (b) play an outdoor sport of my choice.
8. I would rather (a) talk with friends of my choice, or (b) play an outdoor sport of my choice.
9. I would rather (a) attend a movie of my choice, or (b) read a book of my choice.
10. I would rather (a) attend a movie of my choice, or (b) talk with friends of my choice.
11. I would rather (a) spend time on my hobbies, or (b) read a book of my choice.
12. I would rather (a) read a book of my choice, or (b) talk with friends of my choice.

Appendix E

Letter to School Principal

Study 2

Dear School Principal:

My name is Rima Balshe. I am a PhD student at the Institute of Education at University College London. I have extensive experience in working with children with various educational needs.

I am writing to inquire if you would be willing for me to do my research in your school.

The purpose of my research is to examine cognitive and linguistic predictors of single word reading among native Arabic speakers learning English as a second language. The participants of interest are 3rd and 5th graders. Data will be gathered in the form of academic and experimental assessments. Participants will be assessed individually in their familiar school setting during school hours.

The participants will remain totally anonymous; the participating students' names will not be revealed in any of the documents nor to other researchers. Furthermore, the information collected will be treated sensitively and will be highly confidential; it will be only used for the purpose of the current research.

Students and their parents will be given a 'Consent Form' to ensure their acceptance of the participation in the research. Students and their parents who wish for their kids to participate in the research will sign the form and return it to the class teacher within a specified period of time.

If you would like to discuss any part of this research, please contact me via e-mail or cellular number below. I will provide the school with a copy of the research once it is completed for interested teachers to learn more about the study.

Thank you for taking the time to consider this request and I hope that you find my research topic of interest to your school.

Rima Balshe
Cellular #: [redacted]
Email:

Appendix F

Parental and Student Consent Form

Study 2

Dear Parent and Student:

My name is Rima Balshe. I am a PhD student at the Institute of Education at University College London. I have worked extensively with children with various special educational needs.

I am writing to ask your permission for your child to take part in my research project being carried out at your child's school. The purpose of my project is to look at predictors of word reading in Arabic speaking children learning English as a second language. The project will involve asking the children to take part in academic assessments in both Arabic and English. All assessments will be administered by myself, seeing the children individually, on school premises during school hours.

The school administration has given its approval for this project. Taking part is entirely voluntary and your child's participation or lack thereof will not reflect in any way on his/her grades. If you are happy for your child to take part I would ask you to please discuss the contents of this letter with your child prior to making your decision.

The children taking part will remain anonymous in any reports; your child's name will not be revealed in any of the documents, to the school, nor to other researchers. Furthermore, the information collected will be confidential and it will be only used for the purpose of the current project. The data collected will be summarized anonymously in my PhD dissertation and in subsequent publications.

To ensure your confidentiality, an envelope is provided for you in which to place and seal your reply form, to be returned with your child by the date specified below. Taking part in the study is entirely voluntary. If you and your child consent to taking part you will be free to withdraw from participation at any time and without giving reason. Any data I have collected for your child would then be destroyed and not used in any subsequent reports. If you would like to discuss any part of this project, you can contact me at the e-mail address below.

Please return the enclosed form by 20, 11, 2016***, if you do not wish for your child to participate. Otherwise, your child maybe randomly selected to participate in the research project***

Thank you for your cooperation
Rima Balshe
E-mail Address: [redacted]

Predictors of Single Word Reading project
Consent form

I, parent of ____________________, have read the information letter and **do not give consent/give consent** to my son/daughter taking part in the research project being conducted at school by Rima Balshe.

Parent's Signature________________________

Student's Signature________________________

Appendix G

Experimental Stimuli for Study 3