Annals of Dyslexia (2020) 70:200–219
https://doi.org/10.1007/s11881-020-00193-y

Predicting Arabic word reading: A cross-classified generalized random-effects analysis showing the critical role of morphology

Sana Tibi[1] · Ashley A. Edwards[2] · Christopher Schatschneider[2] · John R. Kirby[3]

Received: 12 February 2019 / Accepted: 21 February 2020 / Published online: 1 May 2020

Abstract
The distinctive features of the Arabic language and orthography offer opportunities to investigate multiple word characteristics at the item level. The aim of this paper was to model differences in word reading at the item level among 3rd grade native Arabic-speaking children ($n = 303$) using cross-classified generalized random-effects (CCGRE) analysis. The participants read 80 vowelized words that varied in multiple elements that may contribute to their decodability: number of letters, number of syllables, number of morphemes, ligaturing (connectivity), semantics (concrete vs. abstract), orthographic frequency, root type frequency, and part of speech. Morphological awareness (MA) was included as a person-level predictor. Results of individual models showed that MA, number of letters, number of syllables, number of morphemes, number of ligatures, orthographic frequency, and part of speech were significantly related to the probability of a correct response. However, when all predictors were entered simultaneously, only MA and number of morphemes remained significant. These results underscore the important role of morphology in the lexical structure of Arabic words and in Arabic word reading. Discussion focuses on the role of morphology in Arabic reading and the implications for intervention to improve word recognition in children learning to read Arabic.

Keywords Arabic · Cross-classified generalized random-effects · Morphemes · Morphological awareness · Reading

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s11881-020-00193-y) contains supplementary material, which is available to authorized users.

✉ Sana Tibi
sana.tibi@cci.fsu.edu

[1] School of Communication Science and Disorders, Florida State University, Tallahassee, FL, USA

[2] Department of Psychology, Florida State University, Tallahassee, FL, USA

[3] Faculty of Education, Queen's University, Ontario, Canada

Word reading is at the heart of the reading process, and serves as a key element in reading comprehension. For example, the *simple view of reading* (SVR) (Gough & Tunmer, 1986; Kirby & Savage, 2008) proposes that reading comprehension is the product of word reading and linguistic comprehension. Multiple theories have been put forth to explain how word reading develops. For example, Ehri's (2005, 2018; Ehri and McCormick, 1998) *phase theory* describes five phases of word reading development: pre-alphabetic, partial alphabetic, full alphabetic, consolidated alphabetic, and automatic alphabetic; children go through this developmental sequence until word recognition becomes automatized as sight word reading. Morphology appears in the fourth (consolidated alphabetic) phase of this model, which allows for the accumulation of knowledge about the regularities in the writing system (e.g., grapho-phonemic and grapho-syllabic units), which, in turn, leads to growth in the mental lexicon and facilitates word reading ability.

Furthermore, *connectionist models* (Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989) have distributed pathways that mediate the interaction of orthography, phonology, and semantics of lexical items. The connectionist models explain the dynamic cognitive processes of word reading in terms of weighted connections encoding different units (orthographic, phonologic, or semantic knowledge) in a cooperative and competitive manner (Plaut, 2005). Kirby and Bowers (2017, 2018) have suggested that morphology binds phonology, orthography, and semantics because morphemes contain cues to pronunciation, spelling, and meaning.

Predictors of word reading

Skilled reading in English and other languages relies on a constellation of linguistic and cognitive processes including phonological awareness (PA), naming speed (NS), orthographic processing (OP), and morphological awareness (MA), and each of these cognitive processes focuses on different aspects of the reading process (Arabic: Abu-Rabia, Share, & Mansour, 2003; Tibi & Kirby, 2018a, 2019; Chinese: McBride-Chang et al., 2005; Cross-linguistic: Caravolas, Lervåg, Defior, Málková, & Hulme, 2013; Landerl et al., 2019; English: Carlisle, 2000; Castles, Rastle, & Nation, 2018; National Reading Panel, 2000; Roman, Kirby, Parrila, Wade-Woolley, & Deacon, 2009; Finnish: Georgiou, Papadopoulos, Fella, & Parrila, 2012; Silvén, Poskiparta, Niemi, & Voeten, 2007; German: Landerl & Wimmer, 2000; Greek: Georgiou, Torppa, Manolitsis, Lyytinen, & Parrila, 2012; Hebrew: Ravid & Malenky, 2001; Schiff, Raveh, & Fighel, 2011; Japanese: Muroya et al., 2017; Turkish: Babayiğit & Stainthorp, 2007). Empirical evidence has shown that variability in word reading among individuals (skilled readers vs. dyslexics or poor readers) is related to differences in these cognitive processes. For example, skilled readers perform better at grapheme-phoneme decoding, phonemic awareness, and naming speed (Compton, DeFries, & Olson, 2001; Kirby, Parrila, & Pfeiffer, 2003; Landerl et al., 2013, 2019; Parrila & Protopapas, 2018; Wolf & Bowers, 2000). Most current research on word reading has focused on child-level predictors (e.g., decoding abilities, PA, NS, knowledge of words' specific meaning, and MA), rather than word-level predictors such as word regularity, morphological transparency, orthographic frequency, and number of letters (Gilbert, Compton, & Kearns, 2011; Kearns, 2015; Kearns & Al Ghanem, 2019; Kearns et al., 2016). Adding word-level predictors offers the potential to account for variance not accounted for by child-level predictors. For example, Kearns (2015) and Kearns and Al Ghanem (2019) showed that there were significant separate child and word sources of variance (log-odds > 1.00 in an unconditional model). When Kearns and colleagues examined word and child characteristics in the same models, they found significant effects for both word- and child-levels.

Therefore, it is important to acknowledge that accurate reading depends on both person and word factors (Gilbert et al., 2011; Kearns, 2015; Kearns et al., 2016; Steacy, Elleman, Lovett, & Compton, 2016; Steacy et al., 2018). There is growing evidence showing various characteristics of English words contribute to variability in reading. For example, words that contain vowel diagraphs (Gilbert et al., 2011) or consonant clusters (Olson, Forsberg, Wise, & Rack, 1994) are more difficult to read than those that do not.

Orthographic regularity is another word feature that has been shown to affect recognition of low frequency words (Balota & Ferraro, 1993; Waters, Seidenberg, & Bruck, 1984). Regularity describes the level to which words adhere to grapheme-phoneme correspondence (GPC) rules; regularity varies across orthographies, from very regular ones such as Finnish to more opaque ones such as English (Seymour, Aro, & Erskine, 2003). Vowelized Arabic orthography, in spite of its high orthographic regularity, may have other features (allography and nonlinear morphology) that could contribute to word reading difficulty (Tibi & Kirby, 2018a).

Semantic characteristics of words such as concreteness and imageability have also been shown to affect word reading accuracy (Coltheart, Laxon, & Keating, 1988; Keenan, Betjemann, & Olson 2008; for a contrary view, see Duff & Hulme, 2012). For example, performance on word reading is slower when reading irregular words with low imageability (Strain & Herdman, 1999). There is also evidence from connectionist modeling that the addition of a semantic processor improves both nonword and irregular word reading (Plaut et al., 1996).

Another word-specific feature that has been found to affect reading accuracy is length. Longer words, as measured by number of letters (De Luca, Barca, Burani, & Zoccolotti, 2008) or polysyllabic (Muncer & Knight, 2012) or polysyllabic and polymorphemic words (Kearns, 2015) are more challenging to read. Steacy, Elleman, and Compton (2017) argued that such words require readers to employ inductive learning to build orthographic and semantic knowledge taking into account the statistical probabilities of word components. Kearns et al. (2016) investigated item-specific, child-level, and word-level knowledge among fifth graders with either early- or late-emerging reading difficulties who were asked to read polymorphemic words. They found that at the word level, word frequency and root word family frequency were significant predictors.

Previous research has shown that high frequency words are easier or faster to read than low frequency words (Jared & Seidenberg, 1990; Treiman, Goswami, & Bruck, 1990; Yap & Balota, 2009). Frequency has been assessed with respect to different units. For example, some research reports frequency as a count of phonological units such as the final vowel-consonant unit (the rime unit) in monosyllabic words (Booth & Perfetti, 2002; Treiman et al., 1990), whereas others (Calhoon & Leslie, 2002) reported word frequency and rime-neighborhood size as factors.

A type of frequency related to morphology is "family size" (Baayen, Lieber, & Schreuder, 1997; Schreuder & Baayen, 1997), which refers to the frequency or type count of the number of words that share the same root. The more words are derived or formed from a root, the more productive the root is, and words from larger families are read more accurately (Carlisle & Katz, 2006; Carlisle & Stone, 2005; Deacon, Whalen, & Kirby, 2011; Kearns et al., 2016).

The distributional properties of words, morphemes, and rimes have been shown to affect word and pseudoword recognition (accuracy and latency) in lexical decision tasks across several languages. For example, Dutch productive noun stems have been shown to elicit higher and faster frequency ratings than less productive noun stems with smaller family sizes (Baayen et al., 1997). Family size of the root (i.e., the number of words sharing a base morpheme) has been shown to predict accuracy and latency of word recognition tasks in Hebrew (Moscoso Del Prado Martin et al., 2005), a language with nonlinear morphology. In developmental studies on Hebrew, Schiff et al. (2011) investigated the effect of morphology on priming words among 4th and 7th graders. Their findings revealed that, with young readers, priming occurred when the prime and target words were morphologically and semantically related. However, among the more skilled readers, strong morphological priming occurred regardless of semantic connection. Ravid (2003) has also provided evidence on the prominent role of the root in the lexical acquisition of Hebrew-speaking young children. Arabic, a language similar to Hebrew in its morphological structure that combines roots with word patterns nonlinearly, has also shown an effect of root productivity in lexical decision tasks (Boudelaa & Marslen-Wilson, 2011). Interestingly, Boudelaa and Marslen-Wilson found that priming occurred only when the consonantal roots were productive, regardless of the productivity of the word pattern. Accordingly, they concluded that the distributional property of the Arabic root dominates the lexical access process in Arabic.

The distinctive characteristics of Arabic orthography and morphology call for analysis of the effects of word characteristics, especially in children learning to read. Findings from research on English and other orthographies may not be generalizable to Arabic (see Share, 2008 on the Anglocentricity of research on reading) due to the fact that Arabic morphology and orthography have distinct features from other languages (Tibi & Kirby, 2017, 2018a, 2019; Wofford & Tibi, 2018). Before we describe the current study, we provide a brief description of the Arabic language and its orthography.

Arabic language and orthography

Arabic, like Hebrew, is a Semitic language in its typology (Owens, 2013) and is read right to left. Both are characterized by nonlinear morphology and the primacy of consonantal roots, placing them in stark contrast to Indo-European languages.

Vowelization Arabic orthography is comprised of 28 letters, including three long vowels (/a:/, /u:/, and /i:/) that are represented orthographically as letters, whereas the three corresponding short vowels (a, i, u) are represented as diacritic marks (e.g., ـْ ,ـُ ,ـِ ,ـَ) and may or may not be added to the words, depending on the type of text; for example, these diacritics are mostly seen in children’s books and religious texts, but are not present in adult novels or newspapers. Although these short vowels are not represented orthographically in all types of texts, their role is quite significant because they provide phonemic and morpho-syntactic information. Abu-Rabia (1996, 1997, 1999, 2007) emphasized that the presence of short vowels enhances reading accuracy for all readers and across all ages and reading skills. Short vowels contribute to the transparency of Arabic orthography allowing for close grapheme-phoneme correspondences. Abu-Rabia (2007) maintained that both beginning and advanced Arabic readers rely on short vowelization for correct pronunciation. He also argued that because of homography and the role of vowelization of words’ final letters in determining grammar, word recognition is highly dependent on vowelization. Abu-Rabia (2007) also reported that roots of words and short vowelization were essential factors for reading accuracy in highly skilled adult Arabic readers.

Ligaturing Arabic orthography is also characterized by ligaturing (connectivity), which requires certain letters (22 out of 28) to be connected to other letters. The process of ligaturing changes the shape of the letters (termed allographs), depending on their location in the word (isolated, initial, medial, or final), and the change for some letters is substantial; for example, the letter ‘ه’, roughly equivalent to “h,” has the following different shapes: ه , ـه , ـهـ , هـ. Ligaturing contributes to orthographic depth (Share & Daniels, 2015; Tibi & Kirby, 2018a, 2019). Several researchers (Khateb, Taha, Elias, & Ibrahim, 2013; Taha, Ibrahim, & Khateb, 2013) have found that connected words (due to compulsory ligaturing) were easier to process than non-connected words for skilled readers and dyslexics. Taha and Khateeb (2018) showed that first graders were more likely to accept ligatured words than unconnected words as real words and that kindergarteners similarly accepted ligatured pseudowords; they argued that this was due to the frequent occurrence of connected words in Arabic. Dai, Ibrahim, and Share (2013) found that grade 3 children read ligatured pseudowords more slowly than non-ligatured pseudowords, but in a second experiment, diacritics had even more effect. In the end, they concluded that “ligatures do not create any special burden for Arabic readers” (p. 204).

Morphology A third characteristic of Arabic lexical structure is its morphology, both linear and nonlinear (Boudelaa & Marslen-Wilson, 2001; Tibi & Kirby, 2017, 2018a, 2019). In Arabic linear morphology (used for inflections), morphemes are added to words as prefixes and/or suffixes (as in English) and provide information about the grammatical functions of a word such as gender, number (singular, dual, and plural), person, and time. for example, /rassaaam/ “painter” +/a:n/ “dual inflectional marker” = /rassa:ma:n/ ‘رَسّامان’ which means “two painters.” In contrast, in nonlinear morphology (used for derivations), words are formed by interleaving two abstract bound morphemes: the consonantal root (providing meaning) and the mostly vocalic word pattern (providing morphosyntactic functions). For example, from the consonantal root /r.s.m./, we can derive /**r**a**s**a**m**a/ ‘رَسّام’ “to draw,” /ma**rs**a**m**/‘مَرْسَم’ “studio/atelier,” /**r**a**ss**a:**m**/ ‘رَسَم’ “painter/artist,” /ma**rs**u:**m**/ ‘مَرْسو م’ “was drawn/painted,” and many more. In the preceding examples, /a.a.a/, /ma..a/, /a.aa./, and /ma..u./ constitute the word patterns. Because of this compulsory nonlinear process in Arabic word derivation, the vast majority of Arabic words are at least bimorphemic (Tibi & Kirby, 2017). It is important to note that the root and the word pattern are both bound morphemes, and only when they are combined is a meaningful word created. Also, because ligaturing in Arabic orthography allows for affixes and clitics[1] to be connected to derived words, some Arabic words may contain the conjunction “and” as a clitic, and subject, verb, and object (e.g., /fajarsumu:hʊm/ ‘فيَرْسُمو هُم’ “and they draw/paint them”). Parsing such a one-word sentence into its constituents (/fa/ ‘ف’ “and,” /jarsum/ يرسم’ “paint/draw,” /u:/ “suffix indicating ‘they’” as the masculine plural, and /hum/ هم ‘ ’ them”) is probably a challenging task for beginning Arabic learners, although we are not aware that this has been demonstrated empirically. We know though that root awareness (RA) plays a central role in multiple reading outcomes (word reading accuracy,

[1] Clitics are morphemes that may attach to the word as unstressed prefixes or suffixes and more than one can co-occur within the same word, resulting in one-word phrases and clauses.

pseudoword reading accuracy, word reading fluency, text reading fluency, and maze reading comprehension) among Arabic-speaking third graders (Tibi, Tock, & Kirby, 2019). Using structural equation modeling, Tibi et al. (2019) reported that RA accounted for 43.30% of the variance in reading achievement as the latent variable, after accounting for vocabulary. In a number of studies, Abu-Rabia (2007), Abu-Rabia and Abu-Rahmoun (2012), Abu-Rabia et al. (2003) have found that (a) all readers (adults and normal and dyslexic children) performed better on vowelized measures as compared with unvowelized measures; (b) made use of the information provided by the root when reading pseudowords with real roots; and (c) the dyslexic readers tend to rely on roots and short vowelization in reading words whereas normal readers tend to rely on roots only if words are unvowelized. Abu-Rabia and Abu-Rahmoun explained that typically-achieving readers have richer morphological lexicons, whereas dyslexic readers utilize root identification as a compensatory strategy to make up for their poor phonological decoding skills. Finally, Boudelaa and Marslen-Wilson (2005, 2011) demonstrated that priming among adult students was determined entirely by the root, not the word pattern, irrespective of semantic transparency or root productivity. Boudelaa and Marslen-Wilson (2011) concluded that "the basic processes of morphemic segmentation in Arabic are organized around the root" (p. 641).

These characteristics of Arabic language and orthography (transparency of vowelized Arabic, ligaturing, and complex morphology) invite the question: How might these characteristics of Arabic orthography affect word reading? In spite of the transparency of vowelized Arabic orthography, there are good reasons to regard Arabic as having orthographic depth (Share & Daniels, 2015; Tibi & Kirby, 2018a). First, the graphemes themselves are more complex due to allography and ligaturing. Second, some Arabic orthographic units (e.g., roots and derivational morphemes) may be more difficult to learn because they do not occur contiguously (because of the interweaving of roots and word patterns). These features of Arabic suggest that a number of word-level characteristics may contribute to word reading difficulties, in particular the number of morphemes in the words. They also suggest that individual differences in morphological awareness (Carlisle & Katz, 2006; Carlisle & Stone, 2005; Kearns et al., 2016) will be important.

Although there is a growing research literature on the predictors of Arabic reading (for review, see Tibi & Kirby, 2019), no Arabic research has addressed the multiple word-level features in crossed-random effects models. To paint a more complete picture, we also need to examine the characteristics of words that make reading more difficult for beginning readers.

The current study

The aim of the current study was to investigate the effects of word-level characteristics on Arabic word reading using Cross-Classified Generalized Random-Effects (CCGRE) analysis. CCGRE allows for the simultaneous estimation of multiple word-level predictors in addition to person-level predictors. Although prior research on Arabic has investigated the predictors of Arabic word reading (e.g., Abu Ahmad, Ibrahim, & Share, 2014; Abu-Rabia et al., 2003; Asaad & Eviatar, 2014; Asadi & Khateb, 2017; Asadi, Khateb, Ibrahim, & Taha, 2017; Layes, Lalonde, & Rebaï, 2017; Taibah & Haynes, 2011; Tibi & Kirby, 2017, 2018a, 2019), none of the existing studies has explored multiple word factors in one model using CCGRE. Utilizing CCGRE allows for random effects models in which word characteristics and person factors are investigated simultaneously. The current study investigated the effects of number of letters, number of syllables, number of morphemes, ligaturing, orthographic frequency, root type frequency, concreteness, and part of speech as word-level characteristics and morphological awareness (MA) as a person-level predictor of word reading accuracy. We hypothesized, based on the research in other languages reviewed earlier (e.g., Carlisle & Katz, 2006; Carlisle & Stone, 2005; Kearns, 2015; Kearns et al., 2016), that each of these word characteristics (except for ligaturing) will have a negative effect on Arabic word reading. More specifically and based on the predictive validity of MA in reading Arabic (Tibi & Kirby, 2017, 2019) as well as the primacy of the root in children's Arabic reading (Tibi & Kirby, 2019; Tibi et al., 2019; Abu-Rabia & Abu-Rahmoun, 2012) and adults' lexical word recognition (Abu-Rabia, 2007; Boudelaa & Marslen-Wilson, 2001, 2004, 2005), the productivity of the root in priming (Boudelaa & Marslen-Wilson, 2011), and the importance of MA in Arabic reading (e.g., Abu-Rabia & Abu-Rahmoun, 2012; Tibi & Kirby, 2017, 2019; Tibi et al., 2019), we hypothesized that the morphological variables (number of morphemes, root type frequency, and, at the person level, MA) will predict reading accuracy. Following Taha and Khateeb (2018) and Taha et al. (2013), we did not expect ligaturing to influence word reading accuracy.

Method

Participants

The participants were 303 grade three children from two studies conducted by Tibi and Kirby (2017, 2018a). The first sample included 102 Arabic-speaking children (51 male; mean age = 104 months, SD = 5.7 months) who came from a broad range of public female and male schools in Abu-Dhabi, representing a range of socioeconomic backgrounds (Tibi & Kirby, 2017), as described by Abu-Dhabi Council of Education officials. The second sample comprised 201 students (101 male; mean age = 97 months, SD = 5.4) who were randomly selected from a group of public schools provided by the Ministry of Education in Dubai (Tibi & Kirby, 2018a). The criteria for inclusion in both samples were informed parental consent and being native Arabic-speaking Emirati children with both parents as native speakers of Arabic. None of the participants showed signs of hearing, visual, or language impairment.

Word reading measure

A set of 80 fully vowelized words (see Appendix) was administered to both groups of participants as part of a battery of tests (Tibi & Kirby, 2017, 2018a), with scores ranging from 0 to 80 words correct (mean = 49.61, SD = 23.23). All words were shown with all vowels. We did this (a) to increase the chances of accurate reading because the participants may not have mastered reading yet and may still be relying on decoding (Tibi & Kirby, 2018a), (b) to eliminate the homographic effect, that is, the possibility of multiple accurate readings of some words (e.g., some verbs could be read either in the passive or active form if not vowelized, e.g., /jastaxdim/ 'يَسْتَخْدِم' "to use" or /justaxdam/ 'يُسْتَخْدَم' "was used," and (c) because this is the form most familiar to children. Words were presented on a laminated A-4 paper, in a series of rows with four words in each row.

The words were selected based on the frequency count from two sources each representing a different corpus. One source was the lexical database of Modern Standard Arabic known as

Table 1 Description of word level predictors

Variable	Mean	SD	Min	Max
Number of letters	5.76	1.66	2	10
Number of syllables	3.49	1.07	1	6
Number of morphemes	2.86	1.14	1	5
Ligaturing	3.09	1.42	1	7
Orthographic frequency	0.58	3.63	0.03	32,189.29
Root type frequency	16.35	10.24	1	50
Pronoun*	0.05	0.22		
Verb*	0.29	0.46		
Noun*	0.56	0.50		
Adjective*	0.05	0.22		
Other*	0.05	0.22		
Concreteness*	0.51	0.50		
Morphological awareness	13.37	4.21	0	20
Word reading	49.61	23.23	0	84

For the parts of speech and concreteness (indicated by *), the mean represents the proportion of those words in the test, and the SD is the product of the probability of a 1 and the probability of a 0 on that characteristic

Aralex (Boudelaa & Marslen-Wilson, 2010). Aralex consists of 40 million word tokens primarily drawn from unvowelized Arabic newspapers available online and includes information about orthographic forms, vowelized stems, unvowelized stems, roots, and word patterns. Aralex also provides statistical information about words and morphemes. The second source was a corpus of 147,527 word tokens compiled from Arabic textbooks used in primary schools (Grades 1–6) from two different Arab countries (Libya and UAE) by Belkhouche, Harmain, Al Najjar, Taha, and Tibi (2010). The following is a brief description of each of the word characteristics (see also Table 1).

Orthographic frequency Words on the reading test ranged in orthographic frequency from most frequent (the preposition /fi:/ ‘فـي’ (in) with a frequency of 32,189.29 in Aralex and 4100 in the textbook corpus) to the lowest frequency of .03 in Aralex (and 1 in the textbook corpus) for the one-word phrase / azi:matuka/ ‘عزيمَتُكَ’ (your determination/will). We used the Aralex frequencies divided by 1000 to keep the predictors on similar scales. The frequencies indicate the number of times each distinct orthographic form occurs in the 40-million-word corpus. Boudelaa and Marslen-Wilson (2010) defined orthographic form as “the graphic entity that occurs with white space on either side of it” (p. 484). There were two words for which Aralex did not provide orthographic frequencies, although it provided their respective root type frequencies. For these two words, we used ArabiCorpus (Parkinson, n.d.), which is based on adult corpora and provides the frequency for the surface form only. Because ArabiCorpus provides the frequency per 100,000 words, these were multiplied by 10.

Word length The numbers of letters and syllables in each word were counted. The words varied in their number of letters, ranging from 2 (“in” ‘في’) to 10 letters (“the hospitals” ‘المُسْتَشْفَيات’). The mean number of letters was 5.76 (SD = 1.66). The words also varied in their number of syllables from one to six syllables with a mean of 3.49 (SD = 1.07).

Number of morphemes Words ranged in number of morphemes from one (e.g., the pronoun “he” or the preposition “in”) to five morphemes (e.g., “the contestants” / almutasa:bequ:n/) ‘الْمُتَسابِقون’ with a mean of 2.86 (SD = 1.14). Each word was given a score based on the number of morphemes (derivational, inflectional, and clitics) in it and the number of morphological transformations the word had undergone. For example, the word “the book” ‘الكِتاب’ / al-kita:b/ received 3 points: 1 point for its root [k.t.b], another point for being transformed into a noun, and another point for the particle “the” / al/. Another example of a polymorphemic word that underwent multiple morphological transformations and included an inflectional morpheme is “the contestants” / almutasa:biqu:n/‘مُتَسابِقون’ which received a score of 5 as follows: 1 point for the root /s.b.q/, another point for the derived verb /tasa:baqa/ ‘مُتَسابق’, a third point for the agentive noun /mutasa:biq/ '‘المُتَسابِقون’ that is derived from the verb, a fourth point for the inflectional suffix /oun/ to indicate plural masculine added to the noun /mutasa:bq**u:n**/ '‘المُتَسابِقون’ and a fifth point for the addition of the definite article “the” / al/.

Ligaturing Words varied in the number of ligatured letters, ranging from one (“he” هو) to seven (e.g., “the hospitals” ‘ال/م/س/ت/ش/ف/ي/ات – ’المُستَشفيات) with a mean of 3.09 (SD = 1.42).

Root type frequency According to Boudelaa and Marslen-Wilson (2010) root type frequencies are “raw counts and are extracted from the dictionary. Specifically, the type frequency (or the morphological family size) for a particular root is the number of stems containing that particular root” (p. 484). Roots of the words in the current study spanned a range of frequencies from a minimum of one (“spider”—‘‘عنْكَبوت / ankabu:t/) to a maximum of 50 (the root for the word “circle,” i.e., to form a circle or a round shape), with a mean of 16.35 (SD = 10.24). It should be noted that 5 out of 80 words did not have a root type frequency in Aralex. Accordingly, each was assigned as value of 1, as recommended by Boudelaa (personal communication, January 23, 2019).

Concreteness Words were assessed as either concrete or abstract. Of the 80 words, 41 (51%) were concrete words (e.g., vegetables, hospitals, spider, bed) and 39 (49%) abstract (what, was, private, contemplation, remember). Words deemed concrete are the ones that are mentally more imageable (Paivio, Yuille, & Madigan, 1968).

Part of speech The words covered a range of different parts of speech: 45 nouns (56%), 23 verbs (29%), 4 adjectives (5%), 4 personal pronouns (5%), and 4 other words including 1 preposition (1.25%), two relative pronouns (2.5%), and 1 question word (1.25%). A group of contrasts between adjectives and the other categories was used to assess this with adjectives as the reference group; thus, the contrasts were nouns vs. adjectives, verbs vs. adjectives, etc.

Morphological awareness The person-level characteristic MA was assessed using the sentence completion task (Tibi & Kirby, 2017), which was based on Carlisle’s (2000) measure. Participants were presented with 20 written sentences each preceded by a clue word, and they were asked to complete each sentence in writing with the correct form of the word to complete each sentence. The participant's score was the number of items correct. Scores ranged from 0 to 20 with a mean of 13.37 (SD = 4.21). Some of the items required inflections (e.g.,ساحَةُ المَدرَسَةِ ---------------- (كبير).) and others involved derivations (e.g يُساعِدُنا رَجُلُ ------------- الإطفاءِ على إخمادِ . (حَرَق)).

Procedure

Both the word reading and MA tests were administered by the first author who is a native speaker of Arabic. Participants were instructed to read the words on the word reading test as accurately as possible. The test was administered individually and was not timed. Testing began with three practice items and was discontinued after seven consecutive errors to spare children repeated failure. The termination rule was justified because the reading test was previously piloted (Tibi & Kirby, 2017) and deemed to be ordered from easy to difficult. The MA test was preceded by four practice items to ensure children's understanding of the task. Instruction was provided in the Emirati dialect, which was the children's spoken dialect.

Results

Descriptive statistics for all item-level predictors, MA score, and the word reading measure are provided in Table 1. The number of correct words read by the participants ranged from 0 to 79 with a mean of 49.61 (SD = 23.23). The correlation between MA and word reading was 0.63. The correlations among the word-level predictors appear in Table 2. While there were many nonsignificant relations among the predictors, a strong subset of correlations were observed for ligaturing, number of letters, number of syllables, and number of morphemes with the correlations among these measures ranging above .70.

Cross-classified generalized random-effects models

Item responses for all 80 words were modeled using a series of cross-classified generalized random-effects (CCGRE) models. These models allow for the estimation of variability in item responses between students as well as the variance between words. Because all the responses were binary (correct/incorrect) we used a binary distribution with a logit link function to predict the probability of getting an item correct based upon the set of item characteristics. In these models, persons are crossed with items and both persons and items were allowed to be random factors. For the conditional models, one person level predictor (morphological awareness) and eight item level measures (number of letters, number of syllables, number of morphemes, amount of ligaturing, orthographic frequency, root type frequency, part of speech (whether the word was a noun, pronoun, verb, adjective, or other), and concreteness) were included as predictors. To estimate the zero-order relations of these predictors to the probability of getting an item correct, we fit a separate model for each predictor; to investigate which predictors were uniquely significant after controlling for the other predictors in the model, we fit one model with all the predictors entered simultaneously.

Table 2 Correlations

	Number of letters	Number of syllables	Number of morphemes	Ligaturing	Orthographic frequency	Root type frequency	Pronoun	Verb	Noun	Adjective	Other	Concrete semantics
Number of letters	1	.81*	.80*	.74*	− .32*	0.16	− .28*	−0.13	.30*	0.14	− .28*	−0.03
Number of syllables	.81*	1	.80*	.70*	− .30*	.25*	−0.16	0.05	0.07	0.16	− .27*	−0.16
Number of morphemes	.80*	.80*	1	.64*	− .25*	.34*	− .27*	0.05	0.11	.23*	− .33*	−0.12
Ligaturing	.74*	.70*	.64*	1	− .22*	0.13	− .26*	0.04	0.14	0.07	−0.22	−0.17
Orthographic frequency	− .32*	− .30*	− .25*	− .22*	1	−0.18	0.03	−0.08	−0.16	−0.03	.54*	0.13
Root type frequency	0.16	.25*	.34*	0.13	−0.18	1	−0.12	0.09	0.1	0.03	− .32*	−0.05
Pronoun	− .28*	−0.16	− .27*	− .26*	0.03	−0.12	1	−0.15	− .26*	−0.05	−0.05	.22*
Verb	−0.13	0.05	0.05	0.04	−0.08	0.09	−0.15	1	− .72*	−0.15	−0.15	− .49*
Noun	.30*	0.07	0.11	0.14	−0.16	0.1	− .26*	− .72*	1	− .26*	− .26*	.30*
Adjective	0.14	0.16	.23*	0.07	−0.03	0.03	−0.05	−0.15	−0.26	1	−0.05	−0.01
Other	− .28*	− .27*	− .33*	−0.22	.54*	−0.32	−0.05	−0.15	− .26*	−0.05	1	0.11
Concrete semantics	−0.03	−0.16	−0.12	−0.17	−0.13	−0.05	.22*	− .49*	.30*	−0.01	0.11	1

*Significant at $p < .05$

Table 3 Unconditional model

Parameter	Logit	*z* value	*p* value	Odds ratio
Intercept	0.85	3.56	< 0.001	2.34

The first CCGRE model fit was an unconditional model (see Table 3). This model estimates the variance attributable to students, the variance attributable to words, and a grand mean (intercept) that yields the probability of getting an item correct (in logits). The results of this model revealed the variability due to students (SD = 2.34) and the variability due to items (SD = 1.76), indicating that there is enough variability in item responses to try and predict this variability with our predictors.

The next set of CCGRE models estimated the effect of each predictor alone to predict the odds of getting an item correct. For MA, number of letters, syllables, ligatures, and morphemes, we grand-centered the predictor so that the intercept could be interpreted as the odds of getting an item correct when the score on the predictor was average in the sample. For the other predictors, we dummy-coded the predictors and left them uncentered because they have a natural zero. The intercept then becomes the estimate for the referent category. The results of these models are in Table 4. Because we were fitting nine models (with a total of 12 parameters because of the 5 parts of speech), we controlled for Type I error using the Benjamini-Hochberg linear step up procedure (Benjamini & Hochberg, 1995). Ignoring the *p* values for the intercepts, there were 12 *p* values we needed Type I error control over. After performing the linear step-up procedure, the critical *p* value was determined to be .038, and therefore only MA, number of letters, number of syllables,

Table 4 Results from the nine separate models predicting word reading

Parameter	Logit	*z* value	*p* value	Odds ratio
Intercept	0.85	4.11	< 0.001	2.34
Number of letters	− 0.64	− 6.74	< 0.001*	0.53
Intercept	0.85	4.05	< 0.001	2.34
Number of syllables	− 0.96	− 6.32	< 0.001*	0.38
Intercept	0.85	4.29	< 0.001	2.34
Number of morphemes	− 1.04	− 8.11	< 0.001*	0.35
Intercept	0.85	3.78	< 0.001	2.34
ligaturing	− 0.51	− 4.00	< 0.001*	0.60
Intercept	0.85	3.76	< 0.001	2.34
Orthographic frequency	0.19	3.76	< 0.001*	1.20
Intercept	0.85	3.58	< 0.001	2.34
Root type frequency	− 0.02	− 1.25	0.212	0.98
Intercept	− 0.86	− 1.13	0.258	0.42
Pronoun vs adjective	3.48	3.25	0.001*	32.46
Verb vs adjective	1.39	1.71	0.086	4.01
Noun vs adjective	1.63	2.09	0.037*	5.10
Other vs adjective	4.32	4.04	< 0.001*	75.19
Intercept	0.55	1.78	0.076	1.73
Concrete vs abstract	0.59	1.51	0.130	1.80
Intercept	0.85	3.83	< 0.001	2.34
Morphological awareness	0.35	13.86	< 0.001*	1.42

*Significant after Benjamini-Hochberg linear step-up correction (critical *p* value = .038)

Table 5 Results from the simultaneous prediction model of word reading

Parameter	Logit	*z* value	*p* value	Odds ratio
Intercept	0.17	0.27	0.787	1.19
Number of letters	− 0.23	− 1.26	0.209	0.79
Number of syllables	− 0.07	− 0.26	0.797	0.93
Number of morphemes	− 0.79	− 3.33	< 0.001*	0.45
Ligaturing	0.22	1.51	0.132	1.25
Orthographic frequency	0.10	2.09	0.037	1.11
Root type frequency	0.02	1.55	0.122	1.02
Pronoun vs adjective	1.13	1.23	0.219	3.10
Verb vs adjective	0.34	0.52	0.607	1.40
Noun vs adjective	0.69	1.07	0.284	1.99
Other vs adjective	1.17	1.17	0.241	3.22
Concrete vs abstract	0.16	1.51	0.611	1.17
Morphological awareness	0.35	13.86	< 0.001*	1.42

*Significant after Benjamini-Hochberg linear step-up correction (critical p value = .008)

number of morphemes, ligaturing, orthographic frequency, the difference in probability between pronouns and adjectives, nouns and adjectives, and adjectives and other were statistically significant.

The final CCGRE model entered the predictors simultaneously into the model, and the parameter estimates are reported in Table 5. After conducting the linear step-up procedure (critical p value = 0.008), only number of morphemes and MA remained significant. The odds ratio for number of morphemes was .45, meaning that for every increase in the number of morphemes, the odds of reading this word correctly dropped by more than half.

Discussion

The purpose of the current study was to investigate the effects of word-level characteristics and individual differences in morphological awareness on word reading accuracy. We used a novel statistical method in Arabic research, cross-classified generalized random-effects modeling, which allowed the effects of person and word characteristics on word reading to be estimated simultaneously. Of the word characteristics hypothesized to be related to word reading accuracy, all but root type frequency and semantics were significant (Table 4). Contrary to what would be expected based on Taha and Khateeb (2018) and Taha et al. (2013), the number of ligatures was negatively related to word reading. This negative effect may indicate that words with more ligatures tend to also have other features (length, infrequency, etc.) that are related to word reading (see Table 1). When these other features are controlled in the combined model (Table 5), the ligaturing effect becomes positive but nonsignificant.

When all features were combined in one model, the number of morphemes and morphological awareness were the only significant predictors (Table 5). This result partially confirms our hypothesis that the morphological variables (MA, number of morphemes, and root type frequency) would be important predictors and supports the expectation that the morphological complexity of Arabic words presents a serious obstacle for children learning to read. These findings corroborate previous evidence from Arabic regarding person-level predictors (Abu-Rabia & Abu-Rahmoun, 2012; Tibi 2016; Tibi & Kirby, 2017, 2018a, 2019; Tibi et al., 2019) and are consistent with much research in English and other orthographies (e.g., Carlisle &

Katz, 2006; Carlisle & Stone, 2005; Gilbert et al., 2011; Kearns, 2015; Kearns et al., 2016; Steacy et al., 2016; Steacy et al., 2018; Verhoeven & Perfetti, 2011). On the other hand, this result raises the question why root type frequency, a measure of morphological family size, failed to reach significance in both the individual and full models.

It was surprising that root type frequency did not show a significant effect, given what previous studies have found (Boudelaa & Marslen-Wilson, 2011; Carlisle & Katz, 2006; Deacon et al., 2011; Kearns et al., 2016). Several possible explanations can be suggested. First, the participants in the current study were much younger than the university student participants in the studies by Boudelaa and Marslen-Wilson (2011) who would also have more reading experience and more extensive lexicons. The current participants may not have had enough reading experience to have developed mental representations of the roots that were robust enough to help them in reading. These children may still be working at the grapheme-phoneme decoding phase, not having yet developed "consolidated" representations (Ehri, 2005, 2018). This may be particularly true for Arabic in which the root consonants are separated by graphemes from the word patterns (e.g., [d.r.s] "to study" 'دَرَسَ' – 'مَدارِس' "schools"). It may take many more exposures to identify a root in multiple words to develop a solid mental representation when the elements of the root are dispersed rather than contiguous (as they are in English and other alphabetic orthographies). We would expect that as children increase in reading experience, their mental representations of roots will become better established, and that characteristics such as root type and root token frequency will contribute to their successful word reading.

A second explanation is that the Aralex root type frequencies are based on dictionary listings (i.e., the number of words in the dictionary that include each root) and therefore may include many words that children would never have encountered. We had considered using Aralex's root token frequencies (i.e., how frequently each root appeared in the Aralex corpus) but could not do this due to its collinearity with orthographic frequency. The fact that orthographic frequency was a significant predictor in the single-predictor model suggests root token frequency might also have been, but the nonsignificance of orthographic frequency in the full model suggests that this would also have happened to root token frequency. Future research should employ frequencies derived from a corpus based on words children are likely to have seen, for instance those from school textbooks. In addition, including unvowelized words in the word reading measure might change the results in favor of root type frequency. When words are unvowelized, the young reader might resort more to the root extraction process (Boudelaa & Marslen-Wilson, 2005, 2011, 2015) and to identification of larger groups of letters (Ehri, 2005, 2018).

Last, but certainly not least, the absence of the root type frequency effect could be a consequence of instructional practices. It may be the result of a lack of explicit instruction in morphological awareness and morphemic analysis of multimorphemic words and a lack of practice in deriving multiple words from the same root. Alternatively, children may be taught to recognize polymorphemic words as *whole* words or to sound them out as sequences of phonemes rather than as structures composed of morphemes. Although phonological processing is important in Arabic (Tibi & Kirby, 2018a), it needs to be combined with morphological processing to deal with polymorphemic words (a comparable argument can be made in English, e.g., Kirby & Bowers, 2017, 2018). In the absence of direct observations of instructional practices, however, these instructional interpretations remain speculative.

Implications, limitations, and future directions

The main implication of this research for practice is that children learning to read Arabic need to be taught to recognize and manipulate morphemes in the words they are reading and that particular attention needs to be given to longer, multimorphemic words. We already know that morphological awareness is a predictor of success in reading Arabic (Tibi & Kirby, 2017; Tibi et al., 2019) and that it survives the control of phonological awareness (Tibi & Kirby, 2019). Even if phonological awareness is the stronger predictor during learning to read (Tibi & Kirby, 2019), morphological awareness adds valuable variance and may provide access to decoding for children whose phonological awareness is weak, such as those with phonological dyslexia (Elbro & Arnback, 1996). Morphological processing would allow children to work with larger orthographic units, a necessary step towards fluency. As we noted earlier, Arabic morphology is more complex than that of English, so teaching practices and materials that work in English (e.g., Bowers, Kirby, & Deacon, 2010; Goodwin & Ahn, 2010, 2013; Kirby & Bowers, 2017, 2018) will need to be adapted. Phonologically based instruction can be integrated with morphologically based instruction. There is a need for curriculum development and instructional research in Arabic.

The main implication for future research is that more word-level analyses are required in Arabic, and this research should begin to jointly consider further child-level factors, such as phonological awareness, naming speed, specific semantic knowledge of words, and vocabulary. Moreover, further analyses involving orthographic and phonological shifts are needed in Arabic. For example, in the words "asked" 'سَأَلَ' /sa ala/ and "question" 'سُؤال' /su a:l/, there are orthographic and phonological shifts by virtue of the different word patterns and the spelling convention (the middle consonant-glottal stop changing its shape due to the /u/ vowel preceding it). We know that shift words are more difficult to read in English than transparent words especially among young children (Carlisle, 2000; Carlisle & Stone, 2005; Kearns, 2015). As noted earlier, research in English has found significant interactions between word- and child-level factors (e.g., Gilbert et al., 2011; Kearns, 2015; Kearns et al., 2016; Steacy et al., 2016; Steacy et al., 2018). For example, Kearns (2015) found that students with greater MA skill were better able to read words with shifts. Therefore, such interactions need to be explored in Arabic in models at the item level. It would also be useful to include a measure of children's knowledge of the roots of each word on the reading test (Kearns et al., 2016).

There are some limitations in the present study that should be acknowledged. First, the children in the current study came from one educational system (UAE), so generalizing from these results should be done with caution. The UAE has its own spoken dialect that is different from the dialects spoken elsewhere, and this could have played a role in the children's reading due to some linguistic interference (e.g., vocabulary) from the spoken dialect. Future studies should draw samples from multiple Arabic-speaking countries and educational systems. Second, the orthographic form and root type frequencies were taken from Aralex (Boudelaa & Marslen-Wilson, 2010), which is based on an Arabic dictionary and a corpus primarily related to adults. Frequency counts based on a corpus of children's books may lead to different results. Future efforts should be geared towards developing a large corpus from children's books. Furthermore, future research should include word pattern frequencies as these may also affect word reading. A third limitation is that the findings are tied to the words that were selected in the current study. It is possible that a different set of words that did not include clitics or multiple morpheme suffixes would yield somewhat different results.

In conclusion, this study sought to identify the role of word features in word reading accuracy. Findings showed that number of morphemes and morphological awareness were the strongest factors in children's reading accuracy. These results should encourage future studies using item-level analyses, particularly as a part of intervention studies (e.g., see Steacy et al., 2016). These results also support the explicit inclusion of morphological instruction in early Arabic literacy programs (Tibi & Kirby, 2018b; Wofford & Tibi, 2018). All of this should help children become more aware of the morphological structure in their language and contribute to improved reading development.

Acknowledgments The authors would like to give special thanks to the Ministry of Education in Dubai and the Abu Dhabi Council for Education, United Arab Emirates, who allowed us to work with school children to collect data for this project.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

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). New York, NY: Springer.

Abu-Rabia, S. (1996). The role of vowels and context in reading of highly skilled native Arabic readers. *Journal of Psycholinguistic Research, 25*, 629–641. https://doi.org/10.1007/bf01712413.

Abu-Rabia, S. (1997). Reading in Arabic orthography: The effect of vowels and context on reading accuracy of poor and skilled native Arabic readers. *Reading and Writing: An Interdisciplinary Journal, 9*, 65–78. https://doi.org/10.1023/A:1007962408827.

Abu-Rabia, S. (1999). The effect of Arabic vowels on the reading comprehension of second- and sixth-grade native Arab children. *Journal of Psycholinguistic Research, 28*, 93–101. https://doi.org/10.1023/a:1023291620997.

Abu-Rabia, S. (2007). The role of morphology and short vowelization in reading Arabic among normal and dyslexic readers in grades 3, 6, 9, and 12. *Journal of Psycholinguistic Research, 36*, 89–106. https://doi.org/10.1007/s10936-006-9035-6.

Abu-Rabia, S., & Abu-Rahmoun, N. (2012). The role of phonology and morphology in the development of basic reading skills of dyslexic and normal native Arabic readers. *Creative Education, 13*, 1259–1268. https://doi.org/10.4236/ce.2012.37185.

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: An Interdisciplinary Journal, 16*, 423–442. https://doi.org/10.1023/A:1024237415143.

Asaad, H., & Eviatar, Z. (2014). Learning to read in Arabic: The long and winding road. *Reading and Writing: An Interdisciplinary Journal, 27*, 649–664. https://doi.org/10.1007/s11145-013-9469-9.

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*, 486–505. https://doi.org/10.1080/02702711.2017.1299821.

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: An Interdisciplinary Journal, 30*, 1835–1867. https://doi.org/10.1007/s11145-017-9755-z.

Baayen, R. H., Lieber, R., & Schreuder, R. (1997). The morphological complexity of simplex nouns. *Linguistics, 35*, 861–877. https://doi.org/10.1515/ling.1997.35.5.861.

Babayiğit, S., & Stainthorp, R. (2007). Preliterate phonological awareness and early literacy skills in Turkish. *Journal of Research in Reading, 30*, 394–413. https://doi.org/10.1111/j.1467-9817.2007.00350.x.

Balota, D. A., & Ferraro, F. R. (1993). A dissociation of frequency and regularity effects in pronunciation performance across young-adults, older adults, and individuals with senile dementia of the Alzheimer-type. *Journal of Memory and Language, 32*, 573–592. https://doi.org/10.1006/jmla.1993.1029.

Belkhouche, B., Harmain, M., Al Najjar, L., Taha, H., & Tibi, S. (2010) Analysis of primary school Arabic language textbooks. Paper presented at the 10th Arab conference on information technology Retrieved from http://acit2k.org/ACIT/index.php/proceedings/acit-2010-proceedings

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B., 57*, 289–300 https://www.jstor.org/stable/2346101.

Booth, J. R., & Perfetti, C. A. (2002). Onset and rime structure influences naming but not early word identification in children and adults. *Scientific Studies of Reading, 6*, 1–23. https://doi.org/10.1207/S1532799XSSR0601_01.

Boudelaa, S., & Marslen-Wilson, W. D. (2001). Morphological units in the Arabic mental lexicon. *Cognition, 81*, 65–92. https://doi.org/10.1016/S0010-0277(01)00119-6.

Boudelaa, S., & Marslen-Wilson, W. D. (2004). Allomorphic variation in Arabic: Implications for lexical processing and representation. *Brain and Language, 90*, 106–116. https://doi.org/10.1016/S0093-934X(03)00424-3.

Boudelaa, S., & Marslen-Wilson, W. D. (2005). Discontinuous morphology in time: Incremental masked priming in Arabic. *Language and Cognitive Processes, 20*, 207–260. https://doi.org/10.1080/01690960444000106.

Boudelaa, S., & Marslen-Wilson, W. D. (2010). ARALEX: A lexical database for modern standard Arabic. *Behavior Research Methods, 42*, 481–487. https://doi.org/10.3758/BRM.42.2.481.

Boudelaa, S., & Marslen-Wilson, W. D. (2011). Productivity and priming: Morphemic decomposition in Arabic. *Language and Cognitive Processes, 26*, 624–652. https://doi.org/10.1080/01690965.2010.521022.

Boudelaa, S., & Marslen-Wilson, W. D. (2015). Structure, form, and meaning in the mental lexicon: Evidence from Arabic. *Language, Cognition and Neuroscience, 30*, 955–992. https://doi.org/10.1080/23273798.2015.1048258.

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*, 144–179. https://doi.org/10.3102/0034654309359353.

Calhoon, J. A., & Leslie, L. (2002). A longitudinal study of the effects of word frequency and rime-neighborhood size on beginning readers' rime reading accuracy in words and nonwords. *Journal of Literacy Research, 34*, 39–58. https://doi.org/10.1207/s15548430jlr3401_2.

Caravolas, M., Lervåg, A., Defior, S., Málková, G. S., & Hulme, C. (2013). Different patterns, but equivalent predictors, of growth in reading in consistent and inconsistent orthographies. *Psychological Science, 24*, 1398–1407. https://doi.org/10.1177/0956797612473122.

Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing: An Interdisciplinary Journal, 12*, 169–190. https://doi.org/10.1023/A:1008131926604.

Carlisle, J. F., & Katz, L. A. (2006). Effects of word and morpheme familiarity on reading of derived words. *Reading and Writing: An Interdisciplinary Journal, 19*, 669–693. https://doi.org/10.1007/s11145-005-5766-2.

Carlisle, J. F., & Stone, C. A. (2005). Exploring the role of morphemes in word reading. *Reading Research Quarterly, 40*, 428–449. https://doi.org/10.1598/RRQ.40.4.3.

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.

Coltheart, V., Laxon, V. J., & Keating, C. (1988). Effects of word imageability and age of acquisition on children's reading. *British Journal of Psychology, 79*, 1–12. https://doi.org/10.1111/j.2044-8295.1988.tb02270.x.

Compton, D. L., DeFries, J. C., & Olson, R. K. (2001). Are RAN- and phonological awareness-deficits additive in children with reading disabilities? *Dyslexia, 7*, 125–149. https://doi.org/10.1002/dys.198.

Dai, J., Ibrahim, R., & Share, D. L. (2013). The influence of orthographic structure on printed word learning in Arabic. *Writing Systems Research, 5*, 189–213. https://doi.org/10.1080/17586801.2013.827563.

De Luca, M., Barca, L., Burani, C., & Zoccolotti, P. (2008). The effect of word length and other sublexical, lexical, and semantic variables on developmental reading deficits. *Cognitive and Behavioral Neurology, 21*, 227–235. https://doi.org/10.1097/WNN.0b013e318190d162.

Deacon, S. H., Whalen, R., & Kirby, J. R. (2011). Do children see the danger in dangerous? Grade 4, 6, and 8 children's reading of morphologically complex words. *Applied PsychoLinguistics, 32*, 467–481. https://doi.org/10.1017/S0142716411000166.

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*, 504–525. https://doi.org/10.1080/10888438.2011.598199.

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

Ehri, L. C. (2018). Orthographic mapping and literacy development revisited. In K. Cain, D. Compton, & R. Parrila (Eds.), *Theories of reading development* (pp. 127–146). Amsterdam, the Netherlands: Benjamins.

Ehri, L. C., & McCormick, S. (1998). Phases of word learning: Implications for instruction with delayed and disabled readers. *Reading and Writing Quarterly, 14*, 135–163.

Elbro, C., & Arnback, E. (1996). The role of morpheme recognition and morphological awareness in dyslexia. *Annals of Dyslexia, 46*, 209–240. https://doi.org/10.1007/BF02648177.

Georgiou, G. K., Papadopoulos, T. C., Fella, A., & Parrila, R. (2012). Rapid naming speed components and reading development in a consistent orthography. *Journal of Experimental Child Psychology, 112*, 1–17. https://doi.org/10.1016/j.jecp.2011.11.006.

Georgiou, G. K., Torppa, M., Manolitsis, G., Lyytinen, H., & Parrila, R. (2012). Longitudinal predictors of reading and spelling across languages varying in orthographic consistency. *Reading and Writing: An Interdisciplinary Journal, 25*, 321–346. https://doi.org/10.1007/s11145-010-9271-x.

Gilbert, J. K., Compton, D. L., & Kearns, D. M. (2011). Word and person effects on decoding accuracy: A new look at an old question. *Journal of Educational Psychology, 103*, 489–507. https://doi.org/10.1037/a0023001.

Goodwin, A. P., & Ahn, S. (2010). A meta-analysis of morphological interventions: Effects on literacy achievement of children with literacy difficulties. *Annals of Dyslexia, 60*, 183–208. https://doi.org/10.1007/s11881-010-0041-x.

Goodwin, A. P., & Ahn, S. (2013). A meta-analysis of morphological interventions in English: Effects on literacy outcomes for school-aged children. *Scientific Studies of Reading, 17*, 257–285. https://doi.org/10.1080/10888438.2012.689791.

Gough, P., & Tunmer, W. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*, 6–10. https://doi.org/10.1177/074193258600700104.

Jared, D., & Seidenberg, M. S. (1990). Naming multisyllabic words. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 92–105. https://doi.org/10.1037/0096-1523.16.1.92.

Kearns, D. M. (2015). How elementary-age children read polysyllabic polymorphemic words. *Journal of Educational Psychology, 107*, 364–390. https://doi.org/10.1037/a0037518.

Kearns, D. M., & Al Ghanem, R. (2019). The role of semantic information in children's word reading: Does meaning affect readers' ability to say polysyllabic words aloud? *Journal of Educational Psychology, 111*, 933–956. https://doi.org/10.1037/edu0000316.

Kearns, D. M., Steacy, L. M., Compton, D. L., Gilbert, J. K., Goodwin, A. P., Cho, E., et al. (2016). Modeling polymorphemic word recognition: Exploring differences among children with early-emerging and late-emerging word reading difficulty. *Journal of Learning Disabilities, 49*, 368–394. https://doi.org/10.1177/0022219414554229.

Keenan, J. M., Betjemann R. S., Olson R. K., (2008) Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading, 12*, 281–300. https://doi.org/10.1080/10888430802132279.

Khateb, A., Taha, H. Y., Elias, I., & Ibrahim, R. (2013). The effect of the internal orthographic connectivity of written Arabic words on the process of the visual recognition: A comparison between skilled and dyslexic readers. *Writing Systems Research, 5*, 214–233. https://doi.org/10.1080/17586801.2013.834244.

Kirby, J., Parrila, R., & Pfeiffer, S. (2003). Naming speed and phonological awareness as predictors of reading development. *Journal of Educational Psychology, 95*, 453–464. https://doi.org/10.1037/0022-0663.95.3.452.

Kirby, J. R., & Bowers, P. N. (2017). Morphological instruction and literacy: Binding phonological, orthographic, and semantic features of words. In K. Cain, D. Compton, & R. Parrila (Eds.), *Theories of reading development* (pp. 437–462). Amsterdam, the Netherlands: Benjamins.

Kirby, J. R., & Bowers, P. N. (2018). The effects of morphological instruction on vocabulary learning, reading, and spelling. In R. Berthiaume, D. Daigle, & A. Desrochers (Eds.), *Morphological processing and literacy development: Current issues and research* (pp. 217–243). New York: Routledge.

Kirby, J. R., & Savage, R. S. (2008). Can the simple view deal with the complexities of reading? *Literacy, 42*, 75–82. https://doi.org/10.1111/j.1741-4369.2008.00487.x.

Landerl, K., Freudenthaler, H., Heene, M., De Jong, P. F., Desrochers, A., Manolitsis, G., Parrila, R., & 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*, 220–234. https://doi.org/10.1080/10888438.2018.1510936.

Landerl, K., Ramus, F., Moll, K., Lyytinen, H., Leppänen, P. H. T., & Schulte-Körne, G. (2013). Predictors of developmental dyslexia in European orthographies with varying complexity. *Journal of Child Psychology and Psychiatry, 54*, 686–694. https://doi.org/10.1111/jcpp.12029.

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.

Layes, S., Lalonde, R., & Rebaï, M. (2017). Study on morphological awareness and rapid automatized naming through word reading and comprehension in normal and disabled reading Arabic-speaking children. *Reading and Writing Quarterly, 33*, 123–140. https://doi.org/10.1080/10573569.2015.1105763.
McBride-Chang, C., Cho, J.-R., Liu, H., Wagner, R. K., Shu, H., Zhou, A., Cheuk, C. S., & 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.
Moscoso Del Prado Martin, F., Deutsch, A., Frost, R., Schreuder, R., De Jong, N., & Baayen, H. (2005). Changing places: A cross-language perspective on frequency and family size in Dutch and Hebrew. *Journal of Memory and Language, 53*, 496–512. https://doi.org/10.1016/j.jml.2005.07.003.
Muncer, S. J., & Knight, D. C. (2012). The bigram trough hypothesis and the syllable number effect in lexical decision. *Quarterly Journal of Experimental Psychology, 65*, 2221–2230. https://doi.org/10.1080/17470218.2012.697176.
Muroya, N., Inoue, T., Hosokawa, M., Georgiou, G. K., Maekawa, H., & Parrila, R. (2017). The role of morphological awareness in word reading skills in Japanese: A within-language cross-orthographic perspective. *Scientific Studies of Reading, 21*, 449–462. https://doi.org/10.1080/10888438.2017.1323906.
National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction. Reports of subgroups*. Washington, DC: National Institute of Child Health and Human Development.
Olson, R., Forsberg, H., Wise, B., & Rack, J. (1994). Measurement of word recognition, orthographic, and phonological skills. In G. R. Lyon (Ed.), *Frames of reference for the assessment of learning disabilities: New views on measurement issues* (pp. 243–278). Baltimore, MD: Brookes.
Owens, J. (Ed.). (2013). *The Oxford handbook of Arabic linguistics*. Oxford: Oxford University Press.
Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. Journal of Experimental Psychology, 76(1, Pt.2), 1–25. https://doi.org/10.1037/h0025327.
Parkinson, D. (n.d.). ArabiCorpus. http://arabicorpus.byu.edu/.
Parrila, R., & Protopapas, A. (2018). Dyslexia and word reading problems. In K. Cain, D. Compton, & R. Parrila (Eds.), *Theories of reading development* (pp. 333–358). Amsterdam, the Netherlands: 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: Oxford.
Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review, 103*, 56–115. https://doi.org/10.1037/0033-295X.103.1.56.
Ravid, A., & Malenky, A. (2001). Awareness of linear and nonlinear morphology in Hebrew: A developmental study. *First Language, 21*, 025–056.
Ravid, D. (2003). A developmental perspective on root perception in Hebrew and Palestinian Arabic. In J. Shimron (Ed.), *Language processing and acquisition in languages of Semitic, root-based morphology* (pp. 293–319). Amsterdam, the Netherlands: John Benjamins.
Roman, A. A., Kirby, J. R., Parrila, R. K., Wade-Woolley, L., & Deacon, S. H. (2009). Toward a comprehensive view of the skills involved in word reading in grades 4, 6, and 8. *Journal of Experimental Child Psychology, 102*, 96–113. https://doi.org/10.1016/j.jecp.2008.01.004.
Schiff, R., Raveh, M., & Fighel, A. (2011). The development of the Hebrew mental lexicon: When morphological representations become devoid of their meaning. *Scientific Studies of Reading, 16*, 383–403. https://doi.org/10.1080/10888438.2011.571327.
Schreuder, R., & Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language, 37*(1), 118–139. https://doi.org/10.1006/jmla.1997.2510.
Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, 96*(4), 523–568. https://doi.org/10.1037/0033-295X.96.4.523.
Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*, 143–174.
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., & Daniels, P. T. (2015). Aksharas, alphasyllabaries, abugidas, alphabets and orthographic depth: Reflections on Rimzhim, Katz and Fowler (2014). *Writing Systems Research, 1*, 17–31. https://doi.org/10.1080/17586801.2015.1016395.
Silvén, M., Poskiparta, E., Niemi, P., & Voeten, M. (2007). Precursors of reading skill from infancy to first grade in Finnish: Continuity and change in a highly inflected language. *Journal of Educational Psychology, 99*, 516–531. https://doi.org/10.1037/0022-0663.99.3.516.

Steacy, L. M., Compton, D. L., Petscher, Y., Elliott, J. D., Smith, K., Rueckl, J., Sawi, O., Frost, S. J., & Pugh, K. (2018). Development and prediction of context-dependent vowel pronunciation in elementary readers. *Scientific Studies of Reading, 23*(1), 49–63. https://doi.org/10.1080/10888438.2018.1466303.

Steacy, L. M., Elleman, A. M., & Compton, D. L. (2017). Opening the "black box" of learning to read: Inductive learning mechanisms supporting word acquisition development with a focus on children who struggle to read. In K. Cain, D. Compton, & R. Parrila (Eds.), *Theories of reading development* (pp. 99–122). Amsterdam, the Netherlands: Benjamins.

Steacy, L. M., Elleman, A. M., Lovett, M. W., & Compton, D. L. (2016). Exploring differential effects across two decoding treatments on item-level transfer in children with significant word reading difficulties: A new approach for testing intervention elements. *Scientific Studies of Reading, 20*, 283–295. https://doi.org/10.1080/10888438.2016.1178267.

Strain, E., & Herdman, C. M. (1999). Imageability effects in word naming: An individual differences analysis. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale, 53*, 347–359. https://doi.org/10.1037/h0087322.

Taha, H., Ibrahim, R., & Khateb, A. (2013). How does Arabic orthographic connectivity modulate brain activity during visual word recognition: An ERP study. *Brain Topography, 26*, 292–302. https://doi.org/10.1007/s10548-012-0241-2.

Taha, K., & Khateeb, H. (2018). Statistical learning and orthographic preferences among native Arab kindergarten and first graders. *Writing Systems Research, 10*, 15–25. https://doi.org/10.1080/17586801.2018.1473313.

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.

Tibi, S. (2016). Cognitive and linguistic factors in reading Arabic: The role of morphological awareness in reading (Unpublished doctoral dissertation, Queen's University, Canada).

Tibi, S., & Kirby, J. R. (2017). Morphological awareness: Construct and predictive validity in Arabic. *Applied PsychoLinguistics, 38*, 1019–1043. https://doi.org/10.1017/S0142716417000029 .

Tibi, S., & Kirby, J. R. (2018a). Investigating phonological awareness and naming speed as predictors of reading in Arabic. *Scientific Studies of Reading, 22*, 70–84. https://doi.org/10.1080/10888438.2017.1340948.

Tibi, S., & Kirby, J. R. (2018b). Morphology and reading in Arabic. In A. Chekayri (Ed.), *Teaching reading in Arabic: New approaches* (pp. 91–111). Morocco: Ifran [Arabic].

Tibi, S., & Kirby, J. R. (2019). Reading in Arabic: How well does the standard model apply? *Journal of Speech Language and Hearing Research, 4*, 993–1014. https://doi.org/10.1044/2019_JSLHR-L-18-0193.

Tibi, S., Tock, J. L., & Kirby, J. R. (2019). The development of a measure of root awareness to account for reading performance in the Arabic language: A development and validation study. *Applied PsychoLinguistics, 40*, 303–322. https://doi.org/10.1017/S0142716418000589.

Treiman, R., Goswami, U., & Bruck, M. (1990). Not all nonwords are alike: Implications for reading development and theory. *Memory & Cognition, 18*, 559–567.

Verhoeven, L., & Perfetti, C. A. (2011). Morphological processing in reading acquisition: A cross-linguistic perspective. *Applied PsychoLinguistics, 32*, 457–466. https://doi.org/10.1017/S0142716411000154.

Waters, G. S., Seidenberg, M. S., & Bruck, M. (1984). Children's and adults' use of spelling-sound information in three reading tasks. *Memory & Cognition, 12*, 293–305. https://doi.org/10.3758/BF03197678.

Wofford, M. C., & Tibi, S. (2018). A human right to literacy education: Implications for serving Syrian refugee children. *International Journal of Speech-Language Pathology*, 20, 182–190. https://doi.org/10.1080/17549507.2017.1397746.

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*, 322–324. https://doi.org/10.1177/002221940003300404.

Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language, 60*, 502–529. https://doi.org/10.1016/j.jml.2009.02.001.