Running head: LEXICAL BASIS OF L2 READING COMPREHENSION

The Lexical Basis of Second Language (L2) Reading Comprehension: From (Sub-)Lexical Knowledge to Processing Efficiency

Mona G. Alshehri [1, 2] and Dongbo Zhang [2]

[1] Department of Education, Taif University

[2] Graduate School of Education, University of Exeter

Author Note

Mona G. Alshehri, Department of Education, Taif University, Kingdom of Saudi Arabia; and Graduate School of Education, University of Exeter, United Kingdom.

Dongbo Zhang, Graduate School of Education, University of Exeter, United Kingdom.

This study was part of the project for the first author’s PhD dissertation completed at the University of Exeter. The first author wishes to thank Taif University for sponsoring her study in the United Kingdom. We thank Laura Ciaccio, Theres Grüter (*Language Learning* Associate Editor), Emma Marsden (*Language Learning* Editor), and an anonymous reviewer for their comments and suggestions that helped improve this paper. Any errors are our own.

Correspondence concerning this article should be sent to Mona G. Alshehri, Department of Education, Taif University, Al Hawiyah, Taif 26571, Kingdom of Saudi Arabia. Email: mona.sh@tu.edu.sa

Abstract

This study compared how distinct lexical competences, including lexical knowledge as well as processing skills at both word/lexical and sub-lexical/morphological levels, collectively and relatively predict reading comprehension in adult learners of English as a Foreign Language (EFL). Participants were 220 Arabic-speaking EFL learners in a Saudi university. A battery of paper- and computer-based tests was administered to measure the participants' lexical competences, reading comprehension ability, and working memory. Hierarchical regression analyses revealed that over and above working memory, both lexical and sub-lexical knowledge were significant and unique predictors of reading comprehension; and sub-lexical processing efficiency, as opposed to lexical processing efficiency, predicted reading comprehension significantly. Additionally, among the measured lexical competences, lexical knowledge was the strongest predictor; and the two knowledge variables collectively had a far greater influence on reading comprehension than the two processing efficiency variables. These findings are discussed in light of the lexical basis of text comprehension.

Keywords: Reading comprehension, English as a Foreign Language, lexical quality, lexical knowledge, lexical processing

The Lexical Basis of Second Language (L2) Reading Comprehension: From (Sub-)Lexical Knowledge to Processing Efficiency

The Lexical Quality Hypothesis (Perfetti, 2007) contends that high-quality representations of lexical and sub-lexical features are fundamental for efficient word recognition and word-to-text integration and consequently text comprehension. It underscores the important role of diverse lexical processes and accordingly readers' lexical competences in reading comprehension. Essentially, efficient text comprehension necessitates not only rich knowledge of word meanings but also an ability to process printed words and access their meanings rapidly (i.e., lexical processing or word recognition efficiency).

Previous studies on diverse groups of second language (L2) readers have confirmed the importance of lexical knowledge, notably vocabulary size/breadth, in reading comprehension (Choi & Zhang, 2021; Grabe, 2009; Zhang, 2012). Yet, less is known about the role of knowledge of sub-lexical features, notably morphological knowledge, which, in light of the Lexical Quality Hypothesis, should also play an important role in L2 reading comprehension. Limited research has concurrently considered both lexical and sub-lexical knowledge in adult L2 readers of English (see Zhang & Koda, 2012 for an exception). More importantly, the literature on L2 reading comprehension has paid little attention to lexical, and sub-lexical, processing efficiency, which theoretically should also be fundamentally important given efficient comprehension is a goal of reading (Grabe, 2009; Koda, 2005). Further research is thus warranted on how diverse lexical competences, which are defined in this study to include not only lexical and sub-lexical knowledge but also lexical and sub-lexical processing efficiency, contribute to L2 reading comprehension. To this end, this study measured distinct lexical competences, using a battery of paper- and computer-based tests, in a large group of Arabic-speaking learners of English as a Foreign Language (EFL) in a Saudi university; and compared how the measured competences collectively and relatively predict the learners' L2 reading comprehension.

(Sub-)Lexical Knowledge in Reading Comprehension

Reading comprehension can be understood as "the process of simultaneously extracting and constructing meaning through interaction and involvement with written language" (RAND Reading Study Group, 2002, p. 11). The Construction-Integration Model (Kintsch, 1988) contends that the process of text comprehension starts with the reader accessing and integrating word meanings for establishing a text model, and then the reader building a situation model through activation of background knowledge and various inferencing processes. Reading comprehension thus arguably necessitates various linguistic processes, including, notably, lexical processes, which also underpin the Verbal Efficiency Theory (Perfetti, 1985) and later the Lexical Quality Hypothesis (Perfetti, 2007).

The Lexical Quality Hypothesis places lexical representations and processes at the center of a Reading Systems Framework (Perfetti & Stafura, 2014), and posits that high-quality representations of lexical and sub-lexical features are fundamental to text comprehension (Perfetti, 2007). These representations involve the features of four constituents of word identity: orthography, phonology, semantics, and morphosyntax (Perfetti, 2007). Together, the quality of these four features and the coherence among them facilitate the rapid, low-resource retrieval of lexical word identities and their integration into a mental model of the text (Perfetti, 2007; Perfetti & Stafura, 2014).

The above theoretical outlining of lexical underpinnings of reading comprehension is largely situated in the first language (L1) context but should pertain to L2 reading comprehension as well (Grabe, 2009). Words are the building blocks of texts. To comprehend a text, L2 readers need to know the meanings of the words that make up the text. The knowledge of word meanings, defined in this study as lexical knowledge, should thus play a critically important role in text comprehension. This “instrumentalist” view (Anderson & Freebody, 1981) on the importance of lexical knowledge in text comprehension can be well understood from a strand of L2 research that focuses on lexical coverage and adequate comprehension of texts (e.g., Hu & Nation, 2000; Schmitt, Jiang, & Grabe, 2011). Other studies on L2 readers of English have revealed strong positive correlations between vocabulary size/breadth knowledge (i.e., the number of words whose meanings are known) and reading comprehension ability (e.g., Farran, Bingham, & Matthews, 2012; Qian, 1999; Zhang, 2012). Grabe (2009) highlighted that the correlation could be as high as over .90. Jeon and Yamashita’s (2014) meta-analysis, for example, showed that vocabulary knowledge is one of the strongest correlates of L2 reading comprehension (only next to grammatical knowledge) (on average $r = .79$).

Compared to the wide recognition of and strong empirical evidence on the importance of lexical knowledge (i.e., knowledge of word meanings in the context of this study) in L2 reading comprehension, the attention is limited in the L2 literature to the important role of knowledge of sub-lexical features encapsulated in the Lexical Quality Hypothesis (Perfetti, 2007). As mentioned earlier, high-quality representations of lexical and sub-lexical features are fundamental to efficient word recognition and word-to-text integration (Perfetti, 2007). Additionally, the binding of constituent features also plays an essential role. Morphology, in particular, has been underscored by some scholars as an important constituent binding mechanism; and morphological representations have a strong implication for reading acquisition (e.g., Kirby & Bowers, 2017). For example, English derivation, in addition to modifying the part of speech and meaning of the base word to which a suffix is added, is often characterized by phonological and/or orthographic change to the base word as well (e.g., *apply* → *applicable*). Theoretically, morphological knowledge (and processing, which is discussed in the next section) should also play an important role in the comprehension of English texts, where multimorphemic words are prevalent (Nagy & Anderson, 1984).

In fact, Jeon and Yamashita's (2014) meta-analysis revealed that on average (the number of effect sizes/correlation coefficients meta-analyzed $k = 6$, which was notably smaller than vocabulary knowledge, for which $k = 31$), morphological knowledge had a correlation of .61 with L2 reading comprehension. Thus, even though morphological knowledge, like orthographic knowledge ($r = .51$) and phonological awareness ($r = .48$), was categorized in the meta-analysis as a low-evidence predictor of reading comprehension because of the small number of correlations retrieved from the literature, the moderate average correlation does seem to lend clear support to the importance of morphology in L2 reading comprehension. The issues that wait to be further explored in the L2 literature, however, are often *how* morphological knowledge is important for reading comprehension and whether it predicts L2 reading comprehension over and above lexical knowledge.

Theoretically, morphological knowledge, such as knowledge of roots and affixes, can contribute to text comprehension, independent of lexical knowledge, through at least two major mechanisms. On the one hand, the reader can apply morphological knowledge for more accurate and rapid recognition of morphologically complex words in a text by, for example, dividing those words into their morphemic constituents; on the other hand, morphological knowledge serves as a reliable strategy for the reader to unlock meanings of unknown words in textual reading, that is, instantaneous resolution of vocabulary gaps during reading or "on the spot vocabulary learning” (Nagy, 2007, p. 64).

The empirical literature, however, has produced inconsistent findings. Zhang and Koda (2013), for example, found young Chinese-speaking EFL learners’ English morphological awareness, which covered both derivation and compounding, predicted their reading comprehension, over and above vocabulary knowledge (or lexical knowledge as defined for the purpose of the present study). Similar findings were also reported in some studies on young bilingual readers (e.g., Kieffer & Lesaux, 2008; Zhang, 2017). Yet, a significant, unique effect did not surface in Farran et al.’s (2012) study on grades 3 and 5 Arabic-speaking bilingual readers of English in Canada. English morphological awareness barely explained any additional amout of variance in English reading comprheneison after vocablary knoweldge was also in the regression model (vocabulary knowledge was actually the strongest predictor of reading comprehension; see Table 7, p. 2175). Likewise, in a study on adult Chinese- and Korean-speaking learners of English in Canada, Qian (1999) found that morphological knowledge – knowledge of English affixes and stems, which was intended to be one of the measures for vocabulary depth – did not uniquely and significantly predict reading comprehension. In Zhang & Koda’s (2012) study on adult Chinese-speaking EFL learners, derivational knowledge did not surface as a unique and significant predictor of reading comprehension controlling for vocabulary knowledge.

(Sub-)Lexical Processing Efficiency and Reading Comprehension

While it is essential that readers possess diverse linguistic knowledge for text comprehension, comprehension would be hampered if lower-level linguistic processes are not automatized. Comprehension requires simultaneous orchestration or execution of a number of processes (Perfetti, 1999); yet working memory capacity is limited (Baddeley, 2007). A lack of automatized lower-level processes would constrain the participation of higher-order processes such as textual inferencing for effective construction of a mental model. From a lexical perspective, because words are intended for use in the real world, including text reading, knowing a word should not be simply about an ability to “recognize it in connected speech or in print” and “to access its meaning” but should entail the competence “to do these things within a fraction of a second” (Nagy & Scott, 2000, p. 273). The Lexical Quality Hypothesis (Perfetti, 2007), and its predecessor the Verbal Efficiency Theory, embodies “a capacity theory of comprehension” (Just & Carpenter, 1992). It underscores high-quality representations of (sub-)lexical features because they are fundamental to rapid recognition of printed words and word-to-text integration processes. (Sub-)lexical processing efficiency is an essential element of reading comprehension process (Perfetti & Stafura, 2014). In the L1 English reading literature, particularly studies on school children or developing readers, sight word recognition efficiency and word decoding fluency are critical determinants of reading comprehension (Garcia & Cain, 2014).

Theoretically, the above emphasis on efficient lexical and sub-lexical processing should not pertain to L1 or monolingual readers only. In fact, word recognition efficiency, that is, accurate and rapid recognition of printed words, has been recognized as essential to L2 reading comprehension (Grabe, 2009; Koda, 2005). Empirically, however, compared to the L1 reading literature, research that considered fluency-related lexical competences is much less in the literature on L2 English reading; and the existing body of research often approached the issue from diverse perspectives and generated mixed findings.

On the one hand, some studies on young ESL learners or bilingual children, like those on monolingual children, considered the contribution of word decoding fluency to reading comprehension. Proctor, Carlo, August, and Snow (2005), for example, found that after oral vocabulary was controlled for, English decoding fluency was not a unique and significant predictor of fourth-grade Spanish-speaking ESL learners' reading comprehension in the US. Yet, in Pasquarella, Gottardo, and Grant's (2012) study on adolescent L2 readers of English in Canada, real and pseudoword decoding fluency, after controlling for vocabulary knowledge, significantly predicted reading comprehension.

On the other hand, there were a small number of studies, mostly on foreign language learners of English, that approached the issue of lexical processing efficiency in light of readers' rapid lexical/semantic decision. As part of the NELSON project, van Gelderen et al. (2004), for example, measured adolescent Dutch-speaking EFL readers' "speed of word recognition" with a lexical decision task, that is, a task that asked learners to decide as fast as they could whether a letter string presented on a computer screen was an existing word (see also Harrington, 2018 where lexical decision tasks were intended to measure L2 lexical facility). Reaction times (RTs) and accuracy of responses were both recorded. Among the five concurrent predictors of English reading comprehension, only vocabulary knowledge, in addition to metacognitive knowledge, uniquely and significantly predicted reading comprehension. A significant, unique effect did not surface of the RTs or the word recognition speed. Yamashita's (2013) study on Japanese-speaking university EFL learners, on the other hand, found that reading comprehension was significantly predicted by learners' efficiency of "decoding" (judgement on whether a nonce word could be "read as an English word") and lexical meaning access (judgement on whether words in a pair were antonyms) measured with a paper-based, timed Yes/No decision task. Note, however, that Yamashita, unlike van Gelderen et al. (2004), did not concurrently consider the students' lexical knowledge. It thus remains unclear whether the significant effect identified of the processing efficiency measures would remain, had a lexical knowledge measure been included.

To date, very little research has aimed to test whether sub-lexical processing efficiency, particularly morphological processing efficiency, would be a dimension of lexical competence that may uniquely predict L2 reading comprehension, along with other dimensions (lexical vs. sub-lexical/morphological knowledge on the one hand and lexical processing efficiency on the other). Overall, despite increasing interests in morphological knowledge and L2 reading comprehension (e.g., Kieffer & Lesaux, 2008; Zhang & Koda, 2012) and L2 morphological processing and lexical representation (Clahsen, Felser, Neubauer, Sato, & Silva, 2010; see also Ciaccio & Clahsen, 2020), little effort aimed to combine the two lines of research and examine how morphological processing efficiency may have a unique role to play during text reading. Logic suggests that if morphological knowledge is important for lexical inferencing and/or word decoding fluency during text comprehension, as some L2 studies suggested (e.g., Zhang & Koda, 2012), the unitization of or access to this knowledge must be in a rapid manner for comprehension to be smooth and efficient. Zhang and Ke (2020) underscored morphological decoding fluency in L2 reading comprehension. If efficient morphological processing, which entails quick access to morphological features such as morphological structure and meanings of morphemic constituents, is not in place, fluent morphological decoding would not be possible. In other words, morphological knowledge is necessary but insufficient for efficient processing or recognition of multimorphemic words in print. Empirically, as in the case of *lexical* knowledge vs. *lexical* processing efficiency, it is warranted to study *morphological* processing efficiency in conjunction with *morphological* knowledge to explore their hypothetically unique contribution to L2 reading comprehension.

The Present Study

This study set out to address the aforementioned gaps and explore the lexical basis of L2 reading comprehension in light of the Lexical Quality Hypothesis. The overarching question to be answered is: How do distinct lexical competences collectively and relatively predict L2 reading comprehension? Three sets of questions were further posed to guide this study. The first set examined the contribution of lexical vs. sub-lexical predictors; the second set the contribution of knowledge vs. processing efficiency predictors; and the last one the collective and relative contributions of the four lexical competences.

1. How does lexical vs. sub-lexical knowledge on the one hand, and lexical vs. sub-lexical processing efficiency on the other, relatively predict L2 reading comprehension? How does lexical-level competence (knowledge and processing efficiency) vs. sub-lexical competence (knowledge and processing efficiency) relatively predict L2 reading comprehension?

2. How does lexical knowledge vs. processing efficiency on the one hand, and sub-lexical knowledge vs. processing efficiency on the other, relatively predict L2 reading comprehension? How does knowledge (lexical and sub-lexical) vs. processing efficiency (lexical and sub-lexical) relatively predict L2 reading comprehension?

3. How do the four lexical competences – lexical and sub-lexical on the one hand and knowledge and processing efficiency on the other – collectively and relatively predict L2 reading comprehension?

Method

Participants

This study was conducted at a women's university in Saudi Arabia. The participants were 268 Arabic-speaking first-year students in the university. The analyses for this study, however, were based only on those who attended all the testing sessions described below ($N = 220$). For various random reasons such as absence from class or schedule conflict, 48 students missed one or more of the testing sessions. The students' age ranged between 17 and 22 years old (mean age = 20 years old). The background questionnaire showed that a large majority of them started learning English when they were about 12 years old.

The participants represented a range of undergraduate majors offered by the Saudi university, including, for example, media, English, chemistry, nutrition, and computer science. English is generally the medium of instruction in Saudi universities, particularly for science and engineering majors. Before proceeding to their disciplinary learning in English, which typically starts from the second year, Saudi university students need to go through a whole year of intensive English learning to harness their English proficiency, particularly English for academic purposes. This was the case for the participants of this study, who were first-year students. They all participated on a voluntary basis.

Measures

A battery of paper- and computer-based tests, which is described in detail below, was administered, on a group or individual basis, to measure the participants' distinct lexical competences, reading comprehension, as well as working memory. All instruments were piloted on 30 other first-year students who studied at the same university but did not participate later in the study. Modifications were subsequently made to some instruments; and feedback was also collected from some students to help with the modification process. All tasks except the reading comprehension task are available on https://www.iris-database.org.

All measures showed a fair to high level of reliability (see Table 1) (Brown, 2014). In light of recent discussions on and recommendations for instrument reliability based on internal consistency in the literature on psychoeducational assessment and applied linguistics (e.g., McNeish, 2018; Plonsky & Derrick, 2016), we reported McDonald's omega (ω), which does not assume Tau-equivalence and was calculated using the Structural Equation Modeling method and following Hayes and Coutts (2020). We also reported Cronbach's α, considering its wide, albeit increasingly contested, use in the literature.

Reading comprehension. Reading comprehension was measured with a standardized reading test, namely, Gates-MacGinitie Reading Tests, Fourth Edition (Form S) (MacGinitie, MacGinitie, & Maria, 2000). This test was selected because it considers different types of texts and assesses literal as well as inferential comprehension. Another consideration was that this test, as opposed to any retired standardized tests that target non-native speakers of English (e.g., IELTS), would unlikely have been taken by a participant. Although this test is more commonly used for L1 populations, it has been widely used to measure L2 readers' comprehension as well (e.g., Akamatsu, 2003; Li & Kirby, 2014). Four short reading passages were selected (mean length about 120 words) from Level 5, based on the authors' expert knowledge about local students' reading proficiency and the pilot study. Deliberately, of the four passages selected, two were informational and the other two narrative. Each passage was accompanied by five or six multiple-choice questions, with a total of 21 questions across the four passages. This test was paper-based and administrated in two class sessions, two passages in each session and each session about 25 minutes. Participants were instructed to read the passages silently and circle an answer for each question.

Lexical knowledge. Lexical knowledge was narrowly defined as learners' knowledge of meanings of individual words. It was intended to represent learners' vocabulary breadth and was measured with a modified version of the Vocabulary Levels Test (VLT) (Schmitt, Schmitt, &

Clapham, 2001; Webb, Sasao, & Ballance, 2017). The test for this study covered only four levels of word frequency: 1000, 2000, 3000, and 5000. For each frequency level, six items were randomly sampled. Each item consisted of a list of six words and three meaning choices. Different from the original VLT, the three meaning choices were translated and presented in Arabic (the native language of the participants). Participants were asked to select a word to match each meaning choice. This test was administered in one class session; participants were given 20 minutes to complete it.

Sub-lexical knowledge. We were particularly interested in learners' sub-lexical knowledge that pertains to morphology or more specifically derivation. Participants' morphological knowledge was measured with a researcher-developed task modeled on the Word Part Levels Test (Sasao & Webb, 2017). While the format of the original test, the instructions, and the scoring method were the same, some items were redesigned with consideration of the local students' English learning experience and knowledge of English prefixes and suffixes. The test consisted of the following three sections, assessing the knowledge of form, meaning, and use/function of English affixes (e.g., -*less* and *super*-), respectively.

The first section consisted of 12 items that measured the knowledge of the correct written form of common English prefixes and suffixes. Participants were presented with four orthographically resembling options, only one of which was a correct affix and should thus be circled (e.g., *multi*-, *mul*-, *mlt*-, *tui*-). The second section, which consisted of 10 items, measured the knowledge of meanings of affixes. Participants were asked to select a simple English word, out of four choices, that conveys the meaning of a target prefix or suffix. For each affix, such as *un*-, two word examples (e.g., *unhappy* and *unfair*) were given to contextualize its use; additionally, the Arabic translation of the four English word choices (e.g., for *un*-: *again*, *no*, *back*, and *new*) was also provided. Finally, the last section, which included 10 items, measured the knowledge of how an affix indicates the part-of-speech of a derivational word (i.e., the syntactic properties of affixes). For each item, a prefix or suffix (e.g., *-ish*) was presented together with a derivational word (e.g., *selfish*) to show its use. Participants were asked to select noun, verb, adjective, or adverb to demonstrate of an understanding about how the target affix indicates the part-of-speech of a word to which it is attached. The Arabic translation of the words "noun," "verb," "adjective," and "adverb" was also provided. The test was administered to the participants in a separate class session of about 20 minutes.

Lexical processing efficiency. While the lexical knowledge measure described earlier aimed to assess how many words one knows or written vocabulary size, lexical processing efficiency, in the context of this study, was about visual word recognition efficiency, that is, how rapidly learners can recognize a printed word that they know. To measure lexical processing efficiency, a computer-based lexical decision task was adopted.

The lexical decision task consisted of 40 real words as well as 20 decodable pseudowords (e.g., *toag*) as fillers. The real words were randomly selected from the 1,000 level of the most frequent words based on the Corpus of Contemporary American English (www.wordfrequency.info) and should thus be known to the participants. The order of those real words and pseudowords was randomized. Participants were asked to indicate whether they knew a word (i.e., knowing its [partial] meaning) on the screen by pressing as quickly as possible the "yes" or "no" key marked on the keyboard (cf. the literature on measuring efficiency of vocabulary recognition using computer-based, Yes/No tests such as Harrington, 2018; Pellicer-Sánchez & Schmitt, 2012). Both RTs and Yes/No responses were recorded. Details on the testing procedure are provided later in the Data Collection Procedure section.

Sub-lexical processing efficiency. The sub-lexical processing efficiency measure focused on morphological processing. Two computer-based tasks were included. In the separability task, following Koda (2000), participants were asked to decide, as quickly as possible, whether a word presented in the center of the computer screen can be separated into two or more meaningful components (i.e., stem and affix). There were 30 stimulus words that were assumed to be known to participants. Those words were selected from an initial list created by the first author based on her many years of teaching in the university and familiarity with the students' curriculum. They were also checked by the English teachers of the participants and later piloted on a separate group of students, as mentioned earlier. Fifteen of the stimuli were actual derivational words, such as *government* and *disappear*, which can be segmented into *govern* and -*ment*, and *dis*- and *appear*, respectively. The other 15 words were monomorphemic words that included a letter or a string of letters resembling an English affix, such as *pow<u>er</u>* and *kitch<u>en</u>*. Conversely, the combinability task asked the participants to decide, as quickly as possible, whether the two word-parts presented on the computer screen can be combined to make a meaningful "bigger" English word. There were 24 items in this task, including 12 items that were combinable, such as *fear* and *less*, and 12 items that were not (e.g., *un* and *home*).

Working memory. Text comprehension necessitates the execution of a number of processes, the efficiency of which depends heavily on readers' mental capacity (Just & Carpenter, 1992) Working memory capacity, in particular, is a significant correlate of L2 reading comprehension (Harrington & Sawyer, 1992). (Sub-)lexical processing itself also depends on working memory (Tokowicz, 2014). Subsequently, to obtain a more accurate understanding of the effect of lexical competences, particularly, that of processing efficiency, learners' working memory was also measured and later included as a covariate when reading comprehension was predicted by different lexical competences in regression analysis.

Working memory was measured with a computerized digit span task, which is one of the most widely used tests (Richardson, 2007). The test for this study consisted of 20 numerical sequences – 10 for forward span and 10 backward span – assessing short-term storage of the stimulus sequences. For the forward span items, participants were asked to decide, as quickly as possible, whether a digit sequence presented on the computer screen was the one they saw earlier and in the given order. Likewise, for the backward span items, they were to decide whether a digit sequence was the one that they saw earlier but had the order reversed. For both types of span, there were five sets of random numerical digits increasing in number or the length of sequence (it started with two-digit sequences and ended with six-digit sequences). Each set consisted of two items: one with the order matched and the other with the wrong order.

Data Collection Procedure

For the paper-based lexical and sub-lexical knowledge measures and the reading comprehension test, the first author negotiated with the English language teachers of the participants to administer them on a group basis in 4-6 class sessions (each session about 20-30 minutes) based on the convenience of the classes. The working memory and the lexical and sub-lexical processing efficiency measures were administered individually on a laptop computer and run on PsychoPy Version 3.0 (Peirce et al., 2019). The computer-based testing was conducted in a quiet space on the university campus in one session, which lasted about 15 minutes. Data collection was completed over a period of four weeks. Task instructions were given in Arabic (or Arabic and English) to ensure participants' full understanding.

For the computer-based testing, the Working Memory test was first administered, followed by the (sub-)lexical processing efficiency measures. For all measures, testing began with an on-screen instruction and some practice items. Participants were asked to give a response for an item presented in the center of the computer screen by pressing as quickly as possible "yes" (the left arrow key) or "no" (the right arrow key) marked with stickers on the keyboard. Both RTs and Yes/No responses were recorded. RT was calculated as the interval between the onset of an item and the time of Yes/No being pressed. For the working memory test, participants began the test by seeing a digit sequence for a fixed rate of 1000 ms. Upon the offset of the stimulus sequence, a question, " Is this (a digit sequence) the number you saw in the given / reverse order," appeared on the screen. For all the computer-based tests, the pressing of a key automatically activated the next item. If no key was pressed for an item or no response was detected after a certain period of time, the item would automatically disappear and the next item would appear. The time assigned for an item to be answered before disappearing ranged from 1000 to 2000 milliseconds, with the baseline time for each item estimated on the basis of the pilot study.

Scoring and Missing Values

For the paper-based measures, one point was awarded for each correct answer; an incorrect answer or a missing response did not receive any point. The maximum score possible was thus 21 for the reading comprehension test and 72 for the lexical knowledge test. For the three sections of the sub-lexical or morphological knowledge test, the maximum score possible was 12, 10, and 10, respectively.

The scoring and handling of missing data for the computer-based tasks was less straightforward. There are no consistent methods for handling the data of decision tasks like those in the present study (Jiang, 2012). In the literature on L2 reading comprehension, while some studies considered both accuracy and RTs of responses, such as Cremer and Schoonen

(2013), others incorporated only RTs (e.g., van Gelderen et al., 2004). In the present study, for all computerized measures, we adopted RTs for our analysis. This choice was also in line with our purpose to compare (sub-)lexical *knowledge* against *processing efficiency* for which it was of greater interest to examine participants' response or decision latency. We, however, accommodated accuracy rate in the final RT calculation for all computerized measures (please see the discussion below on Inverse Efficiency Score or IES).

For the lexical processing efficiency task, we only focused on the 40 real words; and for those words, we relied on the RTs of correctly answered items. To calculate the right RTs for analysis, for each participant, we recoded the RT of a No decision on a real word as missing; and the RT of a missing decision was also coded as missing. Then we calculated the mean RT for each item. A RT that was above or below the item mean by two or more standard deviations was subsequently considered to be an outlier and further recoded as missing. This was followed by computing the mean RT of correctly answered items for each participant. Finally, to accommodate the rate of correct responses, a raw RT was replaced by an inversed value (Ratcliff, 1993). For each participant, the Inverse Efficiency Score (IES) was calculated by having the raw mean RT divided by the percentage of correct responses (Townsend & Ashby, 1983). In this respect, participants with a low RT but a low accuracy rate as well would be penalized for the low accuracy. The same procedure was largely followed for calculating the adjusted RTs for the two sub-lexical/morphological processing efficiency tasks as well. The only exception was that, unlike the filler items or pseudo words in the lexical decision task, the monomorphemic items for the separability task and the items not combinable for the combinability task were not excluded for RT calculation. This was because a No decision on those items was considered to also show the participants' attention to morphological features and thus reflect their morphological processing efficiency.

Results

Descriptive Statistics, Reliabilities, and Normality

All statistical analyses, unless stated otherwise, were performed on SPSS 26. The means, standard deviations, reliabilities (McDonald's ω and Cronbach's α), and skewness and kurtosis values of all measured competences are presented in Table 1 (Appendix S1 presents elaborated descriptive statistics that include the minimum and maximum scores and the range for each variable). The accuracy rate, raw RT, as well as IES/adjusted RT are shown for the (sub-)lexical processing efficiency measures, although for the reason we mentioned earlier only IES RTs were used for the subsequent bivariate correlation and regression analyses. For most measures, the skewness and kurtosis estimates were generally below the rule-of-thumb values for univariate normality (e.g., ±2 for both skewness and kurtosis) as well as the critical values that may result in significant deviation from multivariate normality (e.g., ±2 and ±7 for skewness and kurtosis, respectively; Curran, West, & Finch, 1996). The kurtosis of the IES RTs was higher than that of the raw RTs and the paper-based measures. Nonetheless, it was within the acceptable range for multivariate normality. Normality of residuals was checked and confirmed through the examination of histograms and P-P plots (cf. Gelman & Hill, 2007).

Insert Table 1 Here

Bivariate Correlations

Table 2 shows the bivariate correlations between all the variables. Alpha was set at .05

for all correlations and subsequent regression analyses. To highlight, reading comprehension correlated positively and significantly with all knowledge variables, notably lexical knowledge ($r = .643$, $p < .001$), which also produced the highest correlation with reading comprehension. Reading comprehension also correlated negatively and significantly with working memory ($r = -.169$, $p = .012$) and the two sub-lexical processing efficiency tasks ($r = -.183$, $p = .007$ and $r = -.193$, $p = .004$ respectively for the separability and combinability tasks). The correlation between reading comprehension and lexical processing efficiency was negative as well ($r = -.084$, $p = .213$); however, it did not achieve the significance level.

It is also important to note that the three measures of sub-lexical knowledge were all significantly correlated with each other. Knowledge of affix forms significantly correlated with knowledge of affix meanings ($r = .463$, $p < .001$) and knowledge of affix function ($r = .506$, $p < .001$); and knowledge of affix meaning and knowledge of affix function also showed a significant correlation ($r = .518$, $p < .001$). All three sub-lexical knowledge measures also significantly correlated with lexical knowledge, $r = .405$, $r = .447$, and $r = .557$ (all *p*s < .001), respectively, for the affix form, meaning, and function tasks.

Finally, all the (sub-)lexical knowledge measures negatively and significantly correlated with all the (sub-)lexical processing efficiency measures, which makes sense because the processing efficiency measures had a focus on speed (that is, the lower the value, the higher the speed). The two sub-lexical processing efficiency measures and working memory were also positively and significantly correlated. The correlation between lexical processing efficiency and working memory was also positive but not statistically significant ($r = .087$, $p = .197$).

Insert Table 2 Here

Contribution of Lexical Competences to Reading Comprehension

A series of hierarchical regression analyses (Jeon, 2015) was performed to examine how different dimensions of lexical competences – lexical vs. sub-lexical; knowledge vs. processing efficiency – collectively and relatively contributed to L2 reading comprehension over and above working memory. For all analyses, working memory was entered first into the regression equation as a covariate (it explained about 2.9% of the variance in reading comprehension), followed by different lexical competences entered individually or as a block. The three sub-lexical knowledge measures were always entered as a block; likewise, the RTs for the morphological separability and combinability tasks were also entered as a block to represent sub-lexical processing efficiency. The order of entry was also switched for different predictors to test, and compare, their unique contribution to reading comprehension. Multicollinearity was diagnosed for multiple regression analysis through Variable Inflation Factors (VIF), which ranged from 1.108 to 1.864 and was smaller than the lowest bound of rule-of-thumb values (that is, 2.5) for indicating presence of multicollinearity (Allison, 1999).

Comparing lexical and sub-lexical predictors. The first research question sought to compare lexical and sub-lexical predictors of reading comprehension. Three sets of regression analyses were conducted for this purpose. We first examined how lexical knowledge and sub-lexical knowledge predictors relatively contributed to reading comprehension; and then analyzed how lexical processing efficiency and sub-lexical processing efficiency relatively contributed to reading comprehension. Finally, we compared how the two lexical-level competences (i.e., knowledge and processing efficiency together) and the two sub-lexical competences (also knowledge and processing efficiency together) relatively predicted reading comprehension.

Estimates of the regression coefficients in the final regression model can be seen in Appendix S2 for each set of analyses.

As shown in the upper panel of Table 3, controlling for working memory, lexical knowledge additionally explained 38.6% of the variance in reading comprehension ($p < .001$). Over and above working memory and lexical knowledge, sub-lexical knowledge also significantly predicted reading comprehension ($p = .001$); it, however, only additionally explained 4.4% of the variance. When sub-lexical knowledge was entered into the regression equation as the second step, it added 27.1% to the variance explained ($p < .001$). The unique effect of lexical knowledge remained significant: over and above working memory and sub-lexical knowledge, it explained 15.9% of the variance in reading comprehension. It can thus be concluded that lexical knowledge explained a far greater amount of unique variance than did sub-lexical knowledge, although the unique effect of both predictors was significant.

The middle panel of Table 3 shows the unique contribution of lexical vs. sub-lexical processing efficiency. After controlling for working memory, lexical processing efficiency did not predict reading comprehension significantly, whether it was entered before or after sub-lexical processing efficiency. It barely explained any additional variance in reading comprehension when sub-lexical processing efficiency was already in the model. On the other hand, sub-lexical lexical processing efficiency uniquely explained a small yet significant proportion of variance in reading comprehension. Specifically, with working memory and lexical processing efficiency were in the regression model, sub-lexical lexical processing efficiency additionally explained 3.4% of the variance ($p = .021$).

Lastly, we compared the effects of the two lexical predictors with those of the two sub-lexical predictors. As shown in the bottom panel of Table 3, the lexical predictors (knowledge and processing entered as a block) had a far greater unique effect on reading comprehension than did the sub-lexical predictors, although the unique effect of both was significant. Specifically, over and above working memory and the lexical predictors, the sub-lexical predictors additionally explained about 4.9% of the variance in reading comprehension ($p = .002$). On the other hand, the lexical predictors, when entered into the regression model at the last step, significantly explained about 16.7% of the variance in reading comprehension ($p < .001$).

Taken together, the findings suggested that lexical-level competences overall had a stronger effect on reading comprehension than did sub-lexical competences; and this advantage seemed to be attributed to the large effect of lexical knowledge. With respect to processing efficiency, the effect at the sub-lexical level, though small, was greater.

Insert Table 3 Here

Comparing knowledge and processing efficiency predictors. The second research question aimed to compare the effects of knowledge and processing efficiency predictors. Three sets of regression analyses again were conducted. We first compared these two types of competence at the lexical level, and then at the sub-lexical level. Lastly, we compared the effects of lexical and sub-lexical knowledge (i.e., the two levels together) and those of lexical and sub-lexical processing efficiency. Estimates of the regression coefficients can be seen in Appendix S3 for each set of analyses.

The upper panel of Table 4 shows the results of the first comparison. Controlling for working memory and lexical processing efficiency, lexical knowledge significantly explained a unique proportion of variance in reading comprehension (about 38.5%; $p < .001$). Conversely, however, a unique effect did not surface for lexical processing efficiency when it was entered lastly into the model (p = 0.214); and minimal additional variance was explained of reading comprehension (ΔR^2 =.004).

The middle panel of Table 4 presents the results of the second comparison. Sub-lexical knowledge, whether entered in the model before and after sub-lexical processing efficiency, significantly predicted reading comprehension (both ps < .001). As the last predictor entered in the model, sub-lexical knowledge uniquely explained about 23.2% of the variance of reading comprehension. Conversely, although controlling for working memory, sub-lexical processing efficiency significantly predicted reading comprehension (ΔR^2 =.039, p = .012), it failed to significantly predict reading comprehension when sub-lexical knowledge was also in the model (ΔR^2= .000, p = .997).

Finally, as shown in the bottom panel of Table 4, the two knowledge measures (lexical and sub-lexical combined) collectively and uniquely explained about 40% of the variance in reading comprehension (p < .001) when working memory and the two processing efficiency measures (lexical and sub-lexical) were also in the model. Conversely, entered after working memory and the two knowledge predictors, the two processing efficiency measures, however, barely explained any additional variance in reading comprehension (ΔR^2 = .009, p =.341).

Taken together, the above findings seem to suggest that knowledge was a far stronger predictor of reading comprehension than processing efficiency, which was true for both the lexical and the sub-lexical level or disregarding the level of competence.

Insert Table 4 Here

Unique contribution of each predictor. Distinct from the first two questions, the last research question focused on the unique and relative contribution of each predictor. Regression coefficient estimates can be seen in the bottom panel of either Appendix S2 or S3. Table 5 shows that the four lexical competences collectively explained over 40% of the variance in reading comprehension. The top section of the table shows the results on the unique contribution of sub-lexical knowledge, and sub-lexical processing efficiency, when all the other predictors (working memory included) were in the model. The unique contribution was significant for sub-lexical knowledge ($\Delta R^2 = .040$, $p = .001$), but not for sub-lexical processing efficiency ($\Delta R^2 = .001$, $p = .866$). Likewise, the bottom section of Table 5 shows the unique contribution of lexical knowledge, and lexical processing efficiency, when all the other predictors were in the model. The unique contribution of lexical knowledge was significant ($\Delta R^2 = .165$, $p < .001$); yet a significant, unique effect did not surface of lexical processing efficiency ($\Delta R^2 = .008$; $p = .068$).

Based on the unique proportion of variance explained of reading comprehension (i.e., ΔR^2), lexical knowledge appeared to be the strongest unique predictor, followed by sub-lexical knowledge. With the presence of the knowledge predictors and working memory in the model, lexical and sub-lexical processing efficiency barely contributed to reading comprehension.

Insert Table 5 Here

Discussion

The present study set out to investigate how four distinct dimensions of lexical competence – lexical vs. sub-lexical on the one hand and knowledge vs. processing efficiency on the other – collectively and relatively contributed to reading comprehension in adult learners of

English so as to shed light on the lexical basis of L2 reading comprehension. To answer the research questions, the four lexical competences collectively explained over 40% of the variance in the participants' reading comprehension. Compared to the processing efficiency predictors, the knowledge predictors had a predominant influence on reading comprehension. In fact, when the effects of the knowledge predictors were taken into consideration, those of the processing efficiency predictors were no longer significant. Additionally, the lexical predictors collectively had a greater effect on reading comprehension than did the sub-lexical predictors; yet this overall effect did not seem to hold specifically for processing efficiency in that sub-lexical processing efficiency seemed to have a larger effect on reading comprehension (nonetheless, the effect of both processing efficiency predictors was very small). Finally, among the four lexical competences, lexical knowledge was the strongest predictor, followed by sub-lexical knowledge and processing efficiency predictors.

Lexical vs. Sub-lexical Knowledge in Reading Comprehension

The Lexical Quality Hypothesis (Perfetti, 2007) contends that high-quality representations of lexical and sub-lexical features are fundamentally important for text comprehension. The lexical basis of reading comprehension it underscores (Perfetti & Hart, 2001) has been largely supported in the L2 (as well as L1) reading comprehension literature. Notably, a strong association has been consistently found between vocabulary knowledge and reading comprehension (Choi & Zhang, 2021; Grabe, 2009; Jeon & Yamashita, 2014). This relationship was confirmed in the present study. The lexical knowledge measure, which targeted vocabulary size, explained nearly 40% of the variance in reading comprehension (when the effects of working memory and sub-lexical knowledge were concurrently considered; see Table 4). Considering that the previous findings were derived largely from speakers of languages other than Arabic (e.g., Japanese, Chinese, Spanish), the present finding seems to suggest that disregarding learners' L1 background, lexical knowledge or knowledge of word meanings is fundamentally important for L2 reading comprehension.

An issue under-studied in the literature pertains to the (unique) importance of knowledge of sub-lexical features encapsulated in the Lexical Quality Hypothesis. In the present study, we focused on morphological features, because morphology could serve to bind other sub-lexical features, including orthography, phonology, semantics, and grammar (see Kirby & Bowers, 2017). In fact, this study attended to several aspects of morphological knowledge that touched on orthography (the affix form measure), semantics (the affix meaning measure), and grammar (the affix function measure). In the L2 literature, despite an increasing interest in the role of morphology in reading comprehension, the attention is overall limited and most existing studies focused on young EFL learners or bilingual children (e.g., Kieffer & Lesaux, 2008; Zhang & Koda, 2013). Few studies have attended to this issue in adult learners of English (see Zhang & Koda, 2012 for an exception). In the present study, which focused on adult Arabic-speaking EFL learners, all three measures of morphological knowledge significantly correlated with reading comprehension; and collectively, they significantly predicted reading comprehension over and above lexical knowledge (i.e., vocabulary size), even though the unique effect was much smaller than that of lexical knowledge (see Table 3). This finding thus lends support to the highlight of the Lexical Quality Hypothesis on the importance of sub-lexical representations for text comprehension.

The finding also suggests that morphological knowledge is uniquely important for reading comprehension independent of lexical knowledge in adult learners of English. Yet, it seems to differ from the findings of two previous studies that also focused on adult learners.

Zhang and Koda (2012) found morphological knowledge only indirectly contributed to reading comprehension through vocabulary knowledge; when vocabulary knowledge was controlled for, the effect of morphological knowledge was not significant. Likewise, Qian (1999) did not report a unique and significant effect of the morphological knowledge predictor, which was intended to measure an aspect of vocabulary depth knowledge.

One reason for the discrepancy of findings might be that the lexical/vocabulary measures in both Zhang and Koda (2012) and Qian (1999) considered aspects of knowledge beyond that of individual word meanings. Specifically, both studies, in addition to vocabulary size (measured with a Vocabulary Levels Test), concurrently considered word association ability as a vocabulary depth measure, which was not the case in our study. Another reason might be, in contrast to the two previous studies, that our study had a more comprehensive consideration for aspects of morphological knowledge, including form, meaning, as well as function. Notably, the affix function task, which targeted learners' knowledge of the syntactic properties or part-of-speech information of derivational affixes, had the highest correlation with reading comprehension in this study ($r = .519$, $p < .001$; see Table 2). This aspect of knowledge, which was not specifically considered in the two previous studies, is particularly underscored by Nagy (2007) as contributive to sentence parsing and reading comprehension.

Whichever the reason might be, the above discussion suggests that morphological knowledge overall should be an important underpinning of reading comprehension (see also the size of correlation reported in Jeon & Yamashita, 2014). Yet, whether a unique effect can emerge, over and beyond lexical knowledge, may depend on what aspects of morphological knowledge are the focus on the one hand and what aspects of knowledge at the lexical level are the concurrent focus on the other. This issue warrants further research.

(Sub-)Lexical Processing Efficiency in Reading Comprehension

The processing efficiency measures generated a few very intriguing findings. To begin with, overall when working memory and the two knowledge predictors were concurrently in the model, neither lexical nor sub-lexical processing efficiency predicted reading comprehension significantly. This was a surprising finding, because, theoretically, for smooth text comprehension to happen, efficient word recognition and word-to-text integration are essential (Perfetti, 2007). In other words, text comprehension necessitates not only rich knowledge of word meanings and sub-lexical morphological features, which was discussed earlier, but also an ability to efficiently process printed words, including multimorphemic words, and access their meanings during text comprehension. The ability to quickly recognize a word (and word parts) or the ease of accessing word knowledge should have an added value to reading comprehension (Nagy & Scott, 2000; Perfetti & Hart, 2001). Automatized lower-level processing skills are essential to enable effective participation of higher-order processes for constructing mental models during text reading. This is in line with a capacity view of discourse comprehension (Just & Carpenter, 1992), and should pertain to any readers of English, whether English is their native or second language (Grabe, 2009; Koda, 2005).

One interpretation for the lack of a unique and significant effect of the processing efficiency measures, as we speculate, is that this finding may reflect what characterizes lexical involvement at the particular developmental stage of our participants. Although the students had learned English for at least six years (in a foreign language context), their English proficiency tended to be low. This can be partly seen from their low performance on the reading comprehension measure (the average score was about eight out of 21 items; see Table 1): Level 5 of the Gates-MacGinitie Reading Tests, from which the passages and questions were sampled, actually targets 5th graders in an English-speaking context. In other words, for the participants to comprehend the passages, knowledge of word meaning (and knowledge of morphemic meanings for morphologically complex words) should reasonably be a dominant influence. In the L1 reading literature at least, less skilled comprehenders, compared to skilled comprehenders, tended to have problems with word processing or show less immediate use of word meanings in the integration process (Nation & Snowling, 2004; Perfetti & Stafura, 2014). On the other hand, as discussed later in the Limitations section, this finding could also be affected by how RT scores may not adequately represent individual differences in (sub-)lexical processing.

Another factor for attention might be that the comprehension test was not administered in a timed condition. Although the students were asked to complete a test session within a specified period of time, that is, 10-15 minutes per passage, this time restriction might be too relaxed (considering that each passage was only about 120 words long and followed by only five questions) for processing efficiency to make a noticeable difference, particularly when individual differences in working memory were also taken into account. On the other hand, the present finding seemed to corroborate those from the NELSON project on adolescent learners of English in the Netherland (e.g., Fukkink, Hulstijn, & Simis, 2005; van Gelderen et al., 2004). In those studies, word recognition speed was not found to uniquely and significantly predict reading comprehension; additionally, while word recognition training did improve word recognition speed, the effect did not transfer to benefit reading comprehension.

Despite the weak unique effects of the two processing efficiency predictors, their relative contribution shown in Table 3 deserves some attention. Specifically, when lexical processing efficiency was controlled for, sub-lexical processing efficiency had a significant, albeit small, effect on reading comprehension; conversely, however, this significant effect did not surface for lexical processing. We speculate that this gap might be attributed to the psycholinguistic processes that could be differentially involved in the lexical decision task and the morphological processing tasks. Specifically, when learners made a decision on a highly frequent word such as *sweet* and *visit* (in a decontextualized task such as the lexical decision task in this study), they might rely only on orthographic processing with little meaning activation, which would be very different from the processing of those words in an actual text reading situation where access to meanings is essential. In contrast, for the two morphological processing tasks, though also decontextualized, rapid semantic activation or attention to stem and affix meanings (e.g., *inform* and -*ation* for the stimulus word *information*) seemed unavoidable. Consequently, the required meaning activation process that seemed to favor the morphological processing tasks might have resulted in the relatively larger effect of sub-lexical processing efficiency in this study. Such an account seems to be in line with that for Fukkink et al.'s (2005) result as well in that the improvement in the speed for recognizing decontextualized words as a result of the word recognition training might only represent enhanced orthographic (and phonological) processing and not capture the lexical access that is required of reading comprehension.

Limitations and Future Research

A few limitations of this study are noted. To begin with, we only focused on four major types of lexical competence to explore the lexical basis of reading comprehension. Although we considered both lexical and sub-lexical levels and both knowledge and processing efficiency dimensions, and these predictors explained over 40% of the variance in L2 reading comprehension, efficient reading comprehension does not depend solely on these dimensions. There are arguably other lexical knowledge and skills that underpin (L2) reading comprehension. In the L2 literature, there is, for example, an interest in the role of word or semantic association knowledge, which was often studied as a type of vocabulary depth knowledge (Qian, 1999; Zhang, 2012; see Zhang & Koda, 2017 for a review). Cremer and Schoonen (2013) also distinguished between the availability and accessibility of semantic association knowledge, which was more or less equivalent to the knowledge vs. processing efficiency distinction we made in this study. In both L1 and L2 reading literature, there is recently also some attention to the knowledge of connectives (e.g., Crosson & Lesaux, 2013) and knowledge of formulaic language or multi-word lexical units (e.g., Kremmel, Brunfaut, & Alderson, 2017; Martinez & Murphy, 2011). Collectively, these studies and ours contribute to a more comprehensive understanding about the lexical basis of (L2) reading comprehension. Nonetheless, it would seem too ambitious to accommodate all these dimensions into a single study.

The relative contributions of different dimensions of lexical competence to reading comprehension may depend on learners' L2 proficiency. Some researchers split their sample of readers into "proficient" and "less proficient" subgroups and aimed to examine if any relational patterns would differ between the subgroups (e.g., Cremer & Schoonen, 2013; Shiotsu & Weir, 2007). The present study did not perform the *ad hoc* grouping because the participants were literally from the same learner population. Future research, however, might consider recruiting and comparing learners with distinct levels of language proficiency or at distinct developmental stages.

Another limitation pertains to the relatively low reliability found of the Gates-MacGinitie Reading Test (see Table 1) for the present sample (see the meta-analysis of reliability coefficients in L2 research by Plonsky and Derrick [2016] where the median reliability for reading was .86). Although the Gates-MacGinitie Reading Test is primarily intended for native English-speaking readers, it has been popularly used in research on L2 populations as well, such as Li and Kirby (2015) on adolescent Chinese-speaking EFL readers and Akamatsu (2003) on adult ESL readers. We thus did not speculate that the relatively low reliability was attributed to the inappropriateness of the test for adult EFL learners. One reason might be the relatively low number of passages and questions included in this study. We only sampled and administered four passages because the tasks for the purpose of this study already required a commitment of over two hours, not to mention several other tasks we administered for other study purposes on L2 reading. Another reason might be the extreme difficulty of a couple of questions. For example, one question was only answered correctly by 13% of the participants. To keep the test intact, we did not choose to remove those items to augment the reliability for the present sample. Future research should consider adopting more passages with a larger number of questions.

Finally, to accommodate the trade-off between speed and accuracy for computer-based tasks, we followed Townsend and Ashby (1983) to calculate IESs and used those adjusted RTs as predictors of reading comprehension. Nevertheless, we recognize that IES is just one way for RT adjustment; and despite its wide use in the literature, including the language learning literature (e.g., Ke & Koda, 2017), efforts have been taken to explore other methods for better accommodation of the speed-accuracy interaction and debates are not uncommon (e.g., Davison & Martin, 2013; Liesefeld & Janczyk, 2019; Vandierendonck, 2017). Additionally, there have been explorations and debates on the reliability of measuring and representing of individual differences in non-native lexical processing. Schmalz (2020), for example, explored a number of "psycholinguistic marker effects." Word frequency, an important "psycholinguistic marker," notably could modulate RT performance and its representation for individual differences in lexical processing (see also Brysbaert, Mandera, & Keuleers, 2017). In the present study, the lexical decision task's focus on words from the top 1,000 most frequent words (in opposition to the VLT or the lexical knowledge task), in particular, could have reduced its capability for representing participants' individual differences in lexical processing efficiency and consequently its predictive power for reading comprehension. It is beyond the purpose of the present study to directly investigate those methodological issues, and as a result, we could not rule out a possibility that the approach to handling the interaction between speed and accuracy on the one hand and the reliability of the (adjusted) RT-based (sub-)lexical processing measures on the other might have influenced the research findings (e.g., the lack of a significant, unique effect found of the lexical decision task). These can be a direction for future research.

Conclusions

In light of the Lexical Quality Hypothesis, this study explored the lexical basis of L2 reading comprehension in a group of adult Arabic-speaking EFL readers by studying the collective and relative contributions of four distinct lexical competences: lexical vs. sub-lexical and knowledge vs. processing efficiency. Hierarchical regression analyses revealed that the four lexical predictors collectively explained over 40% of the variance in the participants' reading comprehension. Compared to the processing efficiency predictors, the knowledge predictors had a predominant influence on reading comprehension. When the knowledge predictors were not considered, sub-lexical/morphological processing efficiency, as opposed to lexical processing efficiency, significantly predicted reading comprehension, over and above working memory. Overall, among the four lexical competences, lexical knowledge was the strongest predictor, followed by sub-lexical knowledge and the processing efficiency predictors.

This study confirmed strong lexical involvement in L2 reading comprehension. It underscored the critical importance of knowledge of word meanings that had been found in many previous studies. Yet, it also showed that knowledge of sub-lexical morphological features is important, too. Although the lexical processing efficiency measures did not significantly predict reading comprehension when lexical and sub-lexical knowledge were concurrently in the model, there was emerging evidence that the type of processing where meaning activation is mandated (e.g., judging whether word parts can combine) was also important. To our knowledge, the present study is the first of its kind that concurrently considered both lexical and sub-lexical knowledge and processing efficiency to study reading comprehension in L2 learners. The findings enrich the current understanding about the fundamental role of lexical processes in L2 reading comprehension. They particularly shed light on how morphological knowledge as well as processing skills may have a unique role to play in adult L2 learners of English.

References

Allison, P. D. (1999). *Multiple regression*. Thousand Oaks, CA: Pine Forge Press.

Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.), *Comprehension and teaching: Research reviews* (pp. 77–117). Newark, DE: International Reading Association.

Akamatsu, N. (2003). The effects of first language orthographic features on second language in text. *Language Learning, 53*(2), 207–231. https://doi.org/10.1111/1467-9922.00216

Baddeley, A. (2007). *Working memory, thought, and action*. Oxford: Oxford University Press.

Brown, J. D. (2014). Classical theory reliability. In A. J. Kunnan (Ed.), *The companion to language assessment* (pp. 1165–1181). Oxford, UK: Wiley–Blackwell. https://doi.org/10.1002/9781118411360.wbcla054

Brysbaert, M., Mandera, P., & Keuleers, E. (2017). The word frequency effect in word processing: An update review. *Current Directions in Psychological Science, 27*(1), 45–50. https://doi.org/10.1177/0963721417727521

Ciaccio, L. A., & Clahsen, H. (2020). Variability and consistency in first and second language processing: A masked morphological priming study on prefixation and suffixation. *Language Learning, 70*(1), 103-136. https://doi.org/10.1111/lang.12370

Choi, Y., & Zhang, D. (2021). The relative role of vocabulary and grammatical knowledge in L2 reading comprehension: A systematic review of literature. *International Review of Applied Linguistics in Language Teaching, 59*(1), 1-30. https://doi.org/10.1515/iral-2017-0033

Clahsen, H., Felser, C., Neubauer, K., Sato, M., & Silva, R. (2010). Morphological structure in native and nonnative language processing. *Language Learning, 60*(1), 21–43. https://doi.org/10.1111/j.1467-9922.2009.00550.x

Cremer, M., & Schoonen, R. (2013). The role of accessibility of semantic word knowledge in monolingual and bilingual fifth-grade reading. *Applied Psycholinguistics, 34*(6), 1195–1217. https://doi.org/10.1017/S0142716412000203

Crosson, A. C., & Lesaux, N. K. (2013). Does knowledge of connectives play a unique role in the reading comprehension of English learners and English-only students? *Journal of Research in Reading, 36*(3), 241–260. https://doi.org/10.1111/j.1467-9817.2011.01501.x

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods, 1*(1), 16–29. https://doi.org/10.1037/1082-989X.1.1.16

Davidson, D. J., & Martin, A. E. (2013). Modeling accuracy as a function of response time with the generalized linear mixed effects model. *Acta Psychologia, 144*(1), 83–96. https://doi.org/10.1016/j.actpsy.2013.04.016

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/ hierarchical models*. Cambridge: Cambridge University Press.

Farran, L. K., Bingham, G. E., & Matthews, M. W. (2012). The relationship between language and reading in bilingual English-Arabic children. *Reading and Writing, 25*(9), 2153–2181. https://doi.org/10.1007/s11145-011-9352-5

Fukkink, R. G., Hulstijn, J., & Simis, A. (2005). Does training in Second-Language word recognition skills affect reading comprehension? An experimental study. *Modern Language Journal, 89*(1), 54–75. https://doi.org/10.1111/j.0026-7902.2005.00265.x

García, J. R., & Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research, 84*(1), 74–111. https://doi.org/10.3102/0034654313499616

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York, NY: Cambridge University Press.

Harrington, M. (2018). *Lexical facility: Size, recognition speed and consistency as dimensions of second language vocabulary knowledge*. London: Palgrave MacMillan.

Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition, 14*(1), 25–38. https://doi.org/10.1017/S0272263100010457

Hayes, A. H., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. But… *Communication Methods and Measures, 14*(1), 1–24. https://doi.org/10.1080/19312458.2020.1718629

Hu, H. M., & Nation, P. (2000). What vocabulary size is needed to read unsimplified texts. *Reading in a Foreign Language, 8*(2), 689–696.

Jeon, E. H. (2015). Multiple regression. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 131–158). New York: Routledge.

Jeon, E. H., & Yamashita, J. (2014). L2 reading comprehension and its correlates: A meta-analysis. *Language Learning, 64*(1), 160–212. https://doi.org/10.1111/lang.12034

Jiang, N. (2012). *Conducting reaction time research in second language studies*. New York, NY: Routledge.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122–149. https://doi.org/10.1037/0033-295x.99.1.122

Kieffer, M. J., & Lesaux, N. K. (2008). The role of derivational morphology in the reading comprehension of Spanish-speaking English language learners. *Reading and Writing, 21*(8), 783–804. https://doi.org/10.1007/s11145-007-9092-8

Kintsch, W. (1988). The construction-integration model of text comprehension. *Psychological Review, 95*(2), 163–182. https://doi.org/10.1037/0033-295X.95.2.163

Kirby, J. R. & Bowers, P. N. (2017). Morphological instruction and literacy: Binding phonological, orthographic, and semantic features of words. In K. Cain, D. L. Compton, & R. K. Parrila (Eds.), *Theories of reading development* (pp. 437–462). Amsterdam: John Benjamins Publishing. https://doi.org/0.1075/swll.15.24kir

Koda, K. (2000). Cross-linguistic variations in L2 morphological awareness. *Applied Psycholinguistics, 21*(3), 297–320. https://doi.org/10.1017/S0142716400003015

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. New York, NY: Cambridge University Press.

Kremmel, B., Brunfaut, T., & Alderson, J. C. (2017). Exploring the role of phraseological knowledge in foreign language reading. *Applied Linguistics, 38*(6), 848–870. https://doi.org/10.1093/applin/amv070

Liesefeld, H. R., & Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behavior Research Methods, 51*(1), 40–60. https://doi.org/10.3758/s13428-018-1076-x

Li, M., & Kirby, J. R. (2015). The effects of vocabulary breadth and depth on English reading. *Applied Linguistics, 36*(5), 611–634. https://doi.org/10.1093/applin/amu007

MacGinitie, M., MacGinitie, R., & Maria, K. Dreyer. (2000). *Gates–MacGinitie Reading Tests*. Rolling Meadows, IL: Riverside.

Martinez, R., & Murphy, V. A. (2011). Effect of frequency and idiomaticity on second language reading comprehension. *TESOL Quarterly, 45*(2), 267–290. https://doi.org/10.5054/tq.2011.247708

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods, 23*(2), 412–433. https://doi.org/10.1037/met0000144

Nagy, W. (2007). Metalinguistic awareness and the vocabulary-comprehension connection. In R. K. Wagner, A. E. Muse, & K. R. Tannenbaum (Eds.), *Vocabulary acquisition: Implications for reading comprehension* (pp. 52–77). New York: The Guilford Press.

Nagy, W. E., & Anderson, R. C. (1984). How many words are there in printed school English? *Reading Research Quarterly, 19*(3), 304–330. https://doi.org/10.2307/747823

Nagy, W. E., & Scott, J. A. (2000). Vocabulary processes. In M. L. Kamil, P. B. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research, Vol. 3* (pp. 269–284). Lawrence Erlbaum Associates.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. New York, NY: Cambridge University Press. https://doi.org/10.1017/CBO9781139524759

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading*, *27*(4), 342–356. https://doi.org/10.1111/j.1467-9817.2004.00238.x

Pasquarella, A., Gottardo, A., & Grant, A. (2012). Comparing factors related to reading comprehension in adolescents who speak English as a First (L1) or Second (L2) Language. *Scientific Studies of Reading, 16*(6), 474–503. https://doi.org/10.1080/10888438.2011.593066

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., … Lindeløv, J. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods, 51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pellicer-Sánchez, A., & Schmitt, N. (2012). Scoring Yes-No vocabulary tests: Reaction time vs. nonword approaches. *Language Testing, 29*(4), 489-509. https://doi.org/10.1177/0265532212438053

Perfetti, C. A. (1985). *Reading ability*. Oxford: Oxford university Press.

Perfetti, C. A. (1999). Comprehending written language: A blueprint for the reader. In C. Brown & P. Hagoort (Eds.), *Neurocognition of language* (pp. 167–208). Oxford: Oxford University Press.

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357–383. https://doi.org/10.1080/10888430701530730

Perfetti, C. A., & Hart, L. (2001). The lexical basis of comprehension skill. In D. S. Gorfein (Ed.), *On the consequences of meaning selection: Perspectives on resolving lexical ambiguity* (pp. 67–86). Washington, D.C.: American Psychological Association. https://doi.org/10.1037/10459-004

Perfetti, C. A., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific studies of Reading*, *18*(1), 22–37. https://doi.org/10.1080/10888438.2013.827687

Plonsky, L., & Derrick, D. (2016). A meta-analysis of reliability coefficients in second language research. *The Modern Language Journal, 100*(2), 538-553. https://doi.org/10.1111/modl.12335

Proctor, C. P., Carlo, M., August, D., & Snow, C. (2005). Native Spanish-speaking children reading in English: Toward a model of comprehension. *Journal of Educational Psychology, 97*(2), 246–256. https://doi.org/10.1037/0022-0663.97.2.246

Richardson, J. T. E. (2007). Measures of short-term memory: A historical review. *Cortex, 43*(5), 635–650. https://doi.org/10.1016/S0010-9452(08)70493-3

Qian, D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review, 56*(2), 282–308. https://doi.org/10.3138/cmlr.56.2.282

RAND Reading Study Group. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. Santa Monica, CA: RAND.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114*(3), 510–532. https://doi.org/10.1037/0033-2909.114.3.510

Schmalz, X. (2020). *How (not) to measure individual differences: The case of statistical learning and lexical processing efficiency*. https://doi.org/10.17605/OSF.IO/BZGTR

Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning, 64*(4), 913–951. https://doi.org/10.1111/lang.12077

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal, 95*(1), 26–43. https://doi.org/10.1111/j.1540-4781.2011.01146.x

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing, 18*(1), 55–88. https://doi.org/10.1177/026553220101800103

Shiotsu, T., & Weir, C. J. (2007). The relative significance of syntactic knowledge and vocabulary breadth in the prediction of reading comprehension test performance. *Language Testing, 24*(1), 99–128. https://doi.org/10.1177/0265532207071513

Tokowicz, N. (2014). *Lexical processing and second language acquisition*. New York, NY: Routledge.

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge: Cambridge University Press.

Vandierendonck, A. (2017). A comparison of methods to combine speed and accuracy measures of performance: A rejoinder on the binning procedure. *Behavior Research Methods, 49*(2), 653–673. https://doi.org/10.3758/s13428-016-0721-5

Van Gelderen, A., Schoonen, R., De Glopper, K., Hulstijn, J., Simis, A., Snellings, P., & Stevenson, M. (2004). Linguistic knowledge, processing speed, and metacognitive knowledge in First-and Second-Language reading comprehension: A componential analysis. *Journal of Educational Psychology, 96*(1), 19–30. https://doi.org/10.1037/0022-0663.96.1.19

Sasao, Y., & Webb, S. A., & (2017). The Words Parts Levels Test. *Language Teaching Research, 21*(1), 12–30. https://doi.org/10.1177/1362168815586083

Webb, S., Sasao, Y., & Ballance, O. (2017). The updated Vocabulary Levels Test: Developing and validating two new forms of the VLT. *ITL - International Journal of Applied Linguistics, 168*(1), 34-70. https://doi.org/10.1075/itl.168.1.02web

Yamashita, J. (2013). Word recognition subcomponents and passage level reading in a foreign language. *Reading in a Foreign Language, 25*(1), 52–71.

Zhang, D. (2012). Vocabulary and grammar knowledge in second language reading comprehension: A structural equation modeling study. *The Modern Language Journal, 96*(4), 558–575. https://doi.org/10.1111/j.1540-4781.2012.01398.x

Zhang, D. (2017). Derivational morphology in reading comprehension of Chinese-speaking learners of English. *Applied Linguistics, 38*(6), 871–895. https://doi.org/10.1093/applin/amv072

Zhang, D., & Ke, S. (2020). The simple view of reading made complex by morphological decoding fluency in bilingual fourth-grade readers of English. *Reading Research Quarterly, 55*(2)*,* 311–329. https://doi.org/10.1002/rrq.287

Zhang, D., & Koda, K. (2012). Contribution of morphological awareness and lexical inferencing ability to L2 vocabulary knowledge and reading comprehension among advanced EFL learners: Testing direct and indirect effects. *Reading and Writing, 25*(5), 1195–1216. https://doi.org/10.1007/s11145-011-9313-z

Zhang, D., & Koda, K. (2013). Morphological awareness and reading comprehension in a foreign language: A study of young Chinese EFL learners. *System, 41*(4), 901–913. http://doi.org/10.1016/j.system.2013.09.009

Zhang, D., & Koda, K. (2017). Assessing L2 learners' vocabulary depth with word associates format tests: Issues, findings, and suggestions. *Asian-Pacific Journal of Second and Foreign Language Education, 2*(1). https://doi.org/10.1186/s40862-017-0024-0

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix S1. Elaborated Descriptive Statistics and Reliability Estimates

Appendix S2. Parameter Estimates of Regression Models Comparing Lexical vs. Sub-lexical Predictors of Reading Comprehension

Appendix S3. Parameter Estimates of Regression Models Comparing Knowledge vs. Processing Predictors of Reading Comprehension

Table 1

Measures and Descriptive Statistics

	N	*M*	*SD*	95% CI of *M*	Rel. (α)	Rel. (ω)[a]	Skewness		Kurtosis	
							Statistics	SE	Statistics	SE
Reading Comprehension	21	8.15	3.47	[7.69, 8.61]	.630	.641	.479	.154	-.161	.306
Affix Form	12	7.78	2.92	[7.39, 8.17]	.755	.757	-.370	.152	-.603	.303
Affix Meaning	10	6.29	2.57	[5.94, 6.63]	.747	.750	-.392	.152	-.638	.303
Affix Function	10	4.41	2.47	[4.09, 4.74]	.661	.670	-.696	.153	-.102	.306
Separability (accuracy)	30	18.44	5.62	[18.03, 19.47]	.812	.795	-.525	.151	-.203	.302
Separability (raw RT)	–	1580.7	343.2	[1538.6, 1622.8]	–[b]	–	-.593	.152	1.126	.302
Separability (IES)[c]	–	2806.0	1286.9	[2635.1, 2977.0]	–	–	1.994	.164	5.357	.327
Combinability (accuracy)	24	14.01	3.99	[13.55, 14.59]	.684	.689	.087	.153	-.424	.304
Combinability (raw RT)	–	1966.5	423.7	[1910.2, 2023.4]	–	–	-.829	.151	1.038	.302
Combinability (IES)	–	3631.8	1298.6	[3459.2, 3804.3]	–	–	1.413	.164	3.720	.327
Vocabulary Levels	72	38.10	14.58	[36.16, 40.04]	.949	.951	.320	.151	-.490	.300
Lexical Decision (accuracy)	40	30.27	7.73	[29.67, 31.65]	.904	.911	-.882	.151	.182	.301

Lexical Decision (raw RT)	–	973.95	186.5	[951.2, 996.7]	–	–	-.298	.151	.050	.301
Lexical Decision (IES)	–	1358.3	537.1	[1302.1, 1535.0]	–	–	1.720	.165	4.071	.328
Working Memory (accuracy)	20	14.02	3.73	[13.84, 14.76]	.754	.736	-1.110	.151	1.740	.300
Working Memory (raw RT)	–	1988.3	320.7	[1949.0, 2027.6]	–	–	-.079	.152	-.046	.302
Working Memory (IES)	–	3039.8	1278.5	[2870.0, 3209.7]	–	–	2.397	.164	5.671	.327

*Note*s. [a] McDonald’s ω, which does not assume Tau-equivalence, was calculated using the Structural Equation Modeling method in M*plus* 8.0 (see Hayes & Coutts, 2020). [b] Reliability was incalculable of raw RTs for computer-based measures because participants showed diverse patterns of correct, “Yes” reactions across real-word stimuli. It was also incalculable of IES RTs because there was only one holistic score for each participant. [c] IES was calculated by having the raw mean RT divided by the percentage of correct responses for each participant.

N = number of items; *M* = mean; *SD* = standard deviation; CI = Confidence Interval; Rel. (α) = Reliability (Cronbach’s alpha); Rel. (ω) = Reliability (McDonald’s omega); SE = standard error; RT = reaction time; IES = Inverse Efficiency Score; Affix form, affix meaning, and affix function = sub-lexical knowledge; separability and combinability = sub-lexical processing; Vocabulary levels = lexical knowledge; lexical decision = lexical processing.

Table 2

Bivariate Correlations Between All Measured Competences

		1	2	3	4	5	6	7	8	9
1	Reading Comprehension	_								
2	Affix Form	**.391(.000) [.276, .497]**	_							
3	Affix Meaning	**.369(.000) [.261, .462]**	**.463(.000) [.365, .557]**	_						
4	Affix Function	**.519(.000) [.401, .613]**	**.506(.000) [.415, .585]**	**.518(.000) [.428, .592]**	_					
5	Separability	**-.183(.007) [-.290, -.062]**	**-.254(.000) [.058, -.374]**	**-.305(.000) [-.417, -.195]**	**-.291(.000) [-.385, -.197]**	_				
6	Combinability	**-.193(.004) [-.317, -.061]**	**-.220(.001) [-.333, -.101]**	**-.294(.000) [-.402, -.190]**	**-.331(.000) [-.424, -.234]**	**.369(.000) [.252, .490]**	_			
7	Vocabulary Levels	**.643(.000)**	**.405(.000)**	**.447(.000)**	**.547(.000)**	**-.193(.004)**	**-.206(.002)**	_		

		[.549, .713]	**[.286, .514]**	**[.333, .548]**	**[.422, .643]**	**[-.289, -.084]**	**[-.322, -.084]**			
8	Lexical Decision	-.084(.213) [-.330, .054]	**-.137(.043)** **[-.328, -.044]**	**-.228(.001)** **[-.398, -.152]**	**-.218(.001)** **[-.367, -.149]**	**.247(.000)** **[.144, .479]**	**.273(.000)** **[.144, .512]**	**-.227(.001)** **[-.365, -.146]**	–	
9	Working Memory	**-.169(.012)** **[-.260, -.075]**	**-.182(.007)** **[-.324, -.046]**	**-.266(.000)** **[-.393, -.130]**	**-.176(.009)** **[-.273, -.071]**	**.202(.003)** **[.053, .353]**	**.144(.033)** **[.021, .274]**	**-.207(.002)** **[-.311, -.066]**	.087(.197) [.013, .228]	–

Notes. P values in paratheses. 95% Confidence Interval (based on 1,000 bootstrap samples) under each correlation. Bold correlations are significant at $p < .05$. Affix form, affix meaning, and affix function = sub-lexical knowledge; separability and combinability = sub-lexical processing; Vocabulary levels = lexical knowledge; lexical decision = lexical processing.

Table 3

Comparing Lexical and Sub-lexical Predictors of Reading Comprehension

Steps	Predictors	R^2 [95% CI]	Adjusted R^2	ΔR^2	*p*
1	Working memory	.029 [.001, .084]	.024	.029	.012
Lexical vs. Sub-lexical knowledge					
2	Lexical knowledge	.414 [.315, .493]	.409	.386	.000
3	Sub-lexical knowledge	.459 [.353, .527]	.446	.044	.001
2	Sub-lexical knowledge	.300 [.193, .380]	.287	.271	.000
3	Lexical knowledge	.459 [.353, .527]	.446	.159	.000
Lexical vs. sub-lexical processing					
2	Lexical processing	.034 [.000, .087]	.025	.005	.297
3	Sub-lexical processing	.068 [.008, .125]	.050	.034	.021
2	Sub-lexical processing	.067 [.011, .130]	.055	.039	.012
3	Lexical processing	.068 [.008, .125]	.050	.000	.876
Lexical vs. sub-lexical (knowledge & processing)					
2	Lexical knowledge & processing	.419 [.316, .494]	.411	.390	.000
3	Sub-lexical knowledge & processing	.467 [.353, .529]	.447	.049	.002
2	Sub-lexical knowledge & processing	.300 [.185, .374]	.280	.271	.000
3	Lexical knowledge &processing	.467 [.353, .529]	.447	.167	.000

Table 4

Comparing Knowledge and Processing Efficiency Predictors of Reading Comprehension

Steps	Predictors	R^2 [95% CI]	Adjusted R^2	ΔR^2	*p*
1	working memory	.029 [.001, .084]	.024	.029	.012
Lexical knowledge vs. Lexical processing					
2	Lexical knowledge	.414 [.315, .493]	.409	.386	.000
3	Lexical processing	.419 [.316, .494]	.411	.004	.214
2	Lexical processing	.034 [.000, .087]	.025	.005	.297
3	Lexical knowledge	.419 [.316, .494]	.411	.385	.000
Sub-lexical knowledge vs. Sub-lexical processing					
2	Sub-lexical knowledge	.300 [.193, .380]	.287	.271	.000
3	Sub-lexical processing	.300 [.185, .374]	.280	.000	.997
2	Sub-lexical processing	.067 [.011, .130]	.055	.039	.012
3	Sub-lexical knowledge	.300 [.185, .374]	.280	.232	.000
Knowledge vs. processing (lexical & sub-lexical)					
2	Lexical & sub-lexical knowledge	.459 [.353, .527]	.446	.430	.000
3	Lexical & sub-lexical processing	.467 [.353, .529]	.447	.009	.341
2	Lexical & sub-lexical processing	.068 [.008, .125]	.050	.039	.032
3	Lexical & sub-lexical knowledge	.467 [.353, .529]	.447	.400	.000

Table 5

The Unique Contribution of Each Predictor of Reading Comprehension

Steps	Predictors	R^2 [95% CI]	Adjusted R^2	ΔR^2	p
1	working memory	.029 [.001, .084]	.024	.029	.012
Unique contribution of sub-lexical knowledge vs. processing					
2	Lexical knowledge	.414 [.315, .493]	.409	.386	.000
3	Lexical processing	.419 [.316, .494]	.411	.004	.214
4	Sub-lexical processing	.427 [.318, .498]	.413	.008	.219
5	Sub-lexical knowledge	.467 [.353, .529]	.447	.040	.001
4	Sub-lexical knowledge	.467 [.358, .532]	.452	.048	.000
5	Sub-lexical processing	.467 [.353, .529]	.447	.001	.866
Unique contribution of lexical knowledge vs. processing					
2	Sub-lexical knowledge	.300 [.193, .380]	.287	.271	.000
3	Sub-lexical processing	.300 [.185, .374]	.280	.000	.997
4	Lexical processing	.302 [.184, .373]	.279	.002	.413
5	Lexical knowledge	.467 [.353, .529]	.447	.165	.000
4	Lexical knowledge	.459 [.346, .523]	.441	.159	.000
5	Lexical processing	.467 [.353, .529]	.447	.008	.068

Supporting Information for: Alshehri, M. G., & Zhang, D. The lexical basis of Second Language (L2) reading comprehension: From (sub-)lexical knowledge to processing efficiency. Article accepted in Language Learning on XXX 2021.

Appendix S1: Elaborated Descriptive Statistics and Reliability Estimates

	N	*M*	*SD*	Min	Max	Range	95% CI of *M*	Rel. (α)	Rel. (ω)[a]	Skewness		Kurtosis	
										Statistics	SE	Statistics	SE
Reading Comprehension	21	8.15	3.47	1.0	18.0	17.0	[7.69, 8.61]	.630	.641	.479	.154	-.161	.306
Affix Form	12	7.78	2.92	0.0	12.0	12.0	[7.39, 8.17]	.755	.757	-.370	.152	-.603	.303
Affix Meaning	10	6.29	2.57	0.0	10.0	10.0	[5.94, 6.63]	.747	.750	-.392	.152	-.638	.303
Affix Function	10	4.41	2.47	0.0	10.0	10.0	[4.09, 4.74]	.661	.670	-.696	.153	-.102	.306
Separability (accuracy)	30	18.44	5.62	1.0	29.0	28.0	[18.03, 19.47]	.812	.795	-.525	.151	-.203	.302
Separability (raw RT)	–	1580.7	343.2	265.4	2370.8	2105.5	[1538.6, 1622.8]	–[b]	–	-.593	.152	1.126	.302
Separability (IES)[c]	–	2806.0	1286.9	1004.2	9624.9	8620.8	[2635.1, 2977.0]	–	–	1.994	.164	5.357	.327
Combinability	24	14.01	3.99	5.0	24.0	19.0	[13.55, 14.59]	.684	.689	.087	.153	-.424	.304

(accuracy)													
Combinability (raw RT)	–	1966.5	423.7	549.1	2805.1	2256.0	[1910.2, 2023.4]	–	–	-.829	.151	1.038	.302
Combinability (IES)	–	3631.8	1298.6	1510.1	10380.0	8869.8	[3459.2, 3804.3]	–	–	1.413	.164	3.720	.327
Vocabulary Levels	72	38.10	14.58	8.0	72.0	64.0	[36.16, 40.04]	.949	.951	.320	.151	-.490	.300
Lexical Decision (accuracy)	40	30.27	7.73	5.0	40.0	35.0	[29.67, 31.65]	.904	.911	-.882	.151	.182	.301
Lexical Decision (raw RT)	–	974.0	186.5	420.2	1415.7	995.5	[951.2, 996.7]	–	–	-.298	.151	.050	.301
Lexical Decision (IES)	–	1358.3	537.1	677.5	3819.5	3141.9	[1302.1, 1535.0]	–	–	1.720	.165	4.071	.328
Working Memory (accuracy)	20	14.02	3.73	3.0	20.0	17.0	[13.84, 14.76]	.754	.736	-1.110	.151	1.740	.300
Working Memory (raw RT)	–	1988.3	320.7	1047.2	2710.0	1662.6	[1949.0, 2027.6]	–	–	-.079	.152	-.046	.302
Working Memory (IES)	–	3039.8	1278.5	1496.1	9846.6	8350.5	[2870.0, 3209.7]	–	–	2.397	.164	5.671	.327

Notes. [a] McDonald's ω, which does not assume Tau-equivalence, was calculated using the Structural Equation Modeling method in Mplus 8.0 (see Hayes & Coutts, 2020). [b] Reliability was incalculable of raw RTs for computer-based measures because participants showed diverse patterns of correct, "Yes" reactions across real-word stimuli. It was also incalculable of IES RTs because there was only one holistic score for each participant. [c] IES was calculated by having the raw mean RT divided by the percentage of correct responses for each participant.

N = number of items; *M* = mean; *SD* = standard deviation; CI = Confidence Interval; Rel. (α) = Reliability (Cronbach's alpha); Rel. (ω) = Reliability (McDonald's omega); SE = standard error; RT = reaction time; IES = Inverse Efficiency Score; Affix form, affix meaning, and affix function = sub-lexical knowledge; separability and combinability = sub-lexical processing; Vocabulary levels = lexical knowledge; lexical decision = lexical processing.

Appendix S2: Parameter Estimates of Regression Models Comparing Lexical vs. Sub-lexical Predictors of Reading Comprehension

Predictors	*B*	*95% CI*	*β*	*t*	*p*
Lexical vs. Sub-lexical knowledge					
Working memory	-.044	[-.325, .236]	-.016	-.311	.756
Lexical knowledge	.117	[.088, .147]	.494	7.924	.000
Sub-lexical knowledge	–	–	–	–	–
- Form	.101	[-.041, .244]	.086	1.401	.163
- Meaning	-.003	[-.170, .165]	-.002	-.031	.975
- Function	.285	[.099, .471]	.203	3.024	.003
Lexical vs. Sub-lexical processing					
Working memory	-.341	[-.708, .013]	-.128	-1.899	.059
Lexical processing	-.043	[-.584, .498]	-.011	-.156	.876
Sub-lexical processing	–	–	–	–	–
- Separability	-.283	[-.669, .103]	-.105	-1.445	.150
- Combinability	-.355	[-.737, .026]	-.133	-1.837	.068
Lexical vs. Sub-lexical (knowledge & processing)					
Working memory	-.039	[-.322, .243]	-.015	-.274	.784
Lexical knowledge & processing	–	–	–	–	–
- Knowledge	.121	[-.091, .150]	.507	8.086	.000
- Processing	.390	[-.029, .810]	.099	1.833	.068
Sub-lexical knowledge & processing	–	–	–	–	–
- Form	.096	[-.047, .239]	.081	1.325	.187
- Meaning	.007	[-.163, .178]	.005	.084	.934
- Function	.288	[.098, .477]	.205	2.987	.003

- Separability	-.046	[-.347, .255]	-.017	-.300	.764
- Combinability	-.054	[-.354, .245]	-.020	-.356	.722

Appendix S3: Parameter Estimates of Regression Models Comparing Knowledge vs. Processing Predictors of Reading Comprehension

Predictors	*B*	*95% CI*	*β*	*t*	*p*
Lexical Knowledge vs. Lexical Processing					
Working memory	-.111	[-.395, .173]	-.041	-.771	.441
Lexical knowledge	.154	[.129, .180]	.649	11.961	.000
Lexical processing	.263	[-.153, .678]	.066	1.246	.214
Sub-lexical Knowledge vs. Sub-lexical Processing					
Working memory	-.142	[-.463, .179]	-.052	-.870	.385
Sub-lexical knowledge	–	–	–	–	–
- Form	.170	[.008, .332]	.143	2.063	.040
- Meaning	.115	[-.076, .306]	.086	1.190	.235
- Function	.553	[.351, .756]	.394	5.383	.000
Sub-lexical processing	–	–	–	–	–
- Separability	.014	[-.327, .354]	.005	.079	.937
- Combinability	-.002	[-.339, .335]	-.001	-.011	.991
Knowledge vs. Processing (Lexical & Sub-lexical)					
Working memory	-.039	[-.322, .243]	-.015	-.274	.784
Lexical and sub-lexical knowledge	–	–	–	–	–
- Lexical knowledge	.121	[-.091, .150]	.507	8.086	.000
- Form	.096	[-.047, .239]	.081	1.325	.187
- Meaning	.007	[-.163, .178]	.005	.084	.934
- Function	.288	[.098, .477]	.205	2.987	.003
Lexical and sub-lexical processing	–	–	–	–	–
- Lexical processing	.390	[-.029, .810]	.099	1.833	.068

- Separability	-.046	[-.347, .255]	-.017	-.300	.764
- Combinability	-.054	[-.354, .245]	-.020	-.356	.722