

### *The ability of young learners to construct word meaning in context*

Yuko Goto Butler

University of Pennsylvania, Philadelphia, USA

<https://orcid.org/0000-0002-9531-3469>

[ybutler@upenn.edu](mailto:ybutler@upenn.edu)

#### Abstract

This study examines young English readers' ability to infer word meanings in context and to use metacognitive knowledge for constructing word meanings in relation to their reading performance. The participants were 61 fourth-grade students in the United States, comprising 24 monolingual English-speaking (ME) students and 37 English-as-a-second-language (L2) students; each group was also divided into strong and emergent readers in English. Participants were asked to read aloud paragraphs containing words unfamiliar to them in two different contextual conditions (i.e., explicit and implicit conditions), to guess the unfamiliar word meanings, and to tell a teacher how they arrived at the inferred meanings. Quantitative analyses found significant differences between strong and emergent readers in their oral fluency as well as in their ability to infer word meanings and articulate their use of metacognitive knowledge. Although significant differences were found in the ability to infer word meanings and the use of metacognitive reasoning between ME and L2 students, such differences disappeared after controlling for the size of students' receptive vocabulary. Qualitative analyses also revealed differences in the kinds of knowledge and strategies that strong and emergent readers relied on when constructing the meaning of unknown words in both explicit and implicit contexts.

*Keywords:* young learners; vocabulary; strategies; meaning construction; context

## 1. Introduction

It is well known that vocabulary plays a critical role in reading comprehension (Marulis & Neuman, 2010; Moghadam, Zainal, & Ghaderpour, 2012; Takanishi & Menestrel, 2017), although the precise mechanism linking learners' vocabulary knowledge and reading comprehension is not totally clear (Cain, Oakhill, & Lemmon, 2004). Many studies have reported positive relationships between learners' vocabulary size (i.e., how many words they know, also referred to as *vocabulary breadth*) and their reading comprehension (Duke & Carlisle, 2011; National Institute of Health, 2000; Tannenbaum, Torgesen, & Wagner, 2006), but vocabulary size is only one aspect of learners' vocabulary knowledge (Nation, 2013). To better understand the relationship between young readers' vocabulary knowledge and reading comprehension, their *depth of vocabulary knowledge*, or "richness of word understandings" (Proctor, Silverman, Haring, & Montecillo, 2012, p. 1636) and their *ability to construct word meanings in a given context* (Duke & Carlisle, 2011; Walters, 2004) must be examined as well. Unfortunately, however, there is a paucity of research looking into young readers' depth of vocabulary, their ability to infer meanings in context, and how such lexical abilities relate to reading comprehension. Research in this area is particularly limited when it comes to young learners of a second language (L2). This study, therefore, focuses on young English readers (both monolingual students and L2 learners at the upper-elementary school level) and aims to provide insights into: (a) their ability to infer meaning and define words in context, (b) their use of metacognitive knowledge for constructing word meanings, and, finally, (c) how such ability and knowledge relate to their reading performance. Information gained from this study is likely to be useful for developing effective strategy instruction for young L2 readers as well for emergent readers. In this paper, "young learners" and "young readers" are used synonymously and are defined as school-age children (both monolingual and L2-learning children) up to the age of 12.

## 2. Literature review

### 2.1. Knowledge and strategies to infer and construct word meaning in context

Vocabulary size, receptive vocabulary size in particular, is often reported to be highly correlated with reading comprehension among both L1 and L2 readers. For example, it is well known that adult English L2 learners need to be familiar with at least 95% to 98% of the words in a text in order to comprehend the text independently (without getting help from teachers, dictionaries, and other means; e.g., Laufer, 1997; Schmitt, Jiang, & Grabe, 2011). Recent studies on adult L2 readers, however,

have uncovered complex relationships between vocabulary size and reading comprehension. For example, Masrai (2019) found that among high-proficient adult English L2 readers, vocabulary size of mid-frequent words contributed most to their reading comprehension, rather than vocabulary size of most-frequent words or low-frequent words. Simply counting the number of known words, therefore, tells us only a partial story of the mechanism of reading comprehension.

Vocabulary depth is a complex notion, and researchers have proposed numerous ways to conceptualize it (Schmitt, 2014). Some researchers have looked at word depth on a continuum from partial to precise knowledge, proceeding from never having seen a word, through the middle stages of having a general sense of the meaning, to knowing it well (Dale, 1965). Cronbach (1942) discussed vocabulary depth in relation to ability levels, ranging from *generalization* (the ability to define a word) to *availability* (the ability to use a word in production). Other researchers have proposed a *receptive-productive* knowledge continuum, although it is not easy to determine where on the continuum a word becomes available for productive use (Read, 2000). Still others have conceptualized vocabulary depth as a composition of multiple elements, such as knowledge of spoken form (e.g., pronunciation), written form (e.g., spelling), meaning, grammar, collocation, register, frequency, and associations (for detailed discussions, refer to Nation, 2013, and Schmitt, 2014).

Because young readers are in the midst of rapid development of word knowledge, the ability to infer and construct word meaning in context appears to be critical for their reading comprehension. Indeed, average young readers have to deal with many unfamiliar words when they read. Nagy and Herman (1987) estimated that the average English-speaking fifth grader encounters at least 10,000 different unknown words in a year. Because words “often have more than one meaning,” even if young readers “know” the word, they still need to identify the most suitable meaning for a given context (Nagy & Scott, 2000, p. 72). In reality, while young readers may have clearer representations of certain words, they may only have partial knowledge of others, and as such they must constantly revise or add meanings to their vocabulary knowledge to make sense of texts.

To successfully infer word meanings in context, what kind of knowledge is necessary? Nagy (1995) proposed that three types of knowledge influence readers’ contextual inferencing: linguistic knowledge, world knowledge, and strategic knowledge. *Linguistic knowledge* includes readers’ vocabulary size, knowledge about possible word meanings (what Nagy called “word schemas”), and syntactic knowledge. *World knowledge* refers to readers’ concepts about the world. *Strategic knowledge* is readers’ metacognitive control over their cognitive resources when reading.

With respect to linguistic knowledge, L2 researchers have found that internal lexical representations (i.e., readers’ knowledge about phonological, orthographical and morphological information about the word in question), as well as

knowledge of syntactical and semantic relations among words and discourse knowledge, play a significant role in the inference of word meaning (e.g., de Bot, Paribakht, & Wesche, 1997; Haynes, 1993; Ke & Koda, 2019; Paribakht & Wesche, 1999). Decoding accuracy appears to be associated with word meaning inference in L2 as well (Prior et al., 2014). In L1 research, it is well documented that children as young as two, if not younger, are able to rely on parts of speech and other syntactic information to arrive at the meanings of new words (Westermann & Mani, 2018).

The roles of world knowledge and strategic knowledge in inferring word meanings in context are much less understood. Nassaji (2003) found that adult L2 readers relied most heavily on world knowledge when making lexical inferences in reading. In an L1 context, Diakidoy (1998) found that US sixth graders' prior knowledge about the content of passages predicted their comprehension, which in turn influenced their learning of unknown words in the passages. The value of strategic knowledge is thought to be important from an instructional point of view. Nassaji also reported that, among adult English learners, repeating the word or sections of the text was the most popular strategy for lexical inference, followed by analogy (i.e., guessing based on other phonologically or orthographically similar words), and verifying (i.e., checking appropriateness against the wider textual context).

Huckin and Block (1993) proposed a cognitive processing model of L2 lexical inference that involves cognitive and metalinguistic control processing. Cognitive processing is a quick and dynamic process that takes place in various components, or what they referred to as *modules*, including vocabulary knowledge, morphosyntax knowledge, world knowledge, knowledge about textual discourse patterns, and so forth. Metalinguistic processing is both a linear and parallel process in which hypotheses about word meanings are generated and tested in context. The model predicts that, with these two processes activating together, a reader can infer meaning by strategically deciding how to use various knowledge types while interacting with contextual sources.

## 2.2. The role of context in constructing word meaning

Constructing meaning in context may not always happen effortlessly, but conscious and strategic use of textual context facilitates students' ability to infer the meaning of unknown words. Researchers have conceptualized *context* differently (Walters, 2004). Some have classified it based on the explicitness or implicitness of cues, others have paid attention to the location of cues for inference (i.e., local vs. global cues), and still others have focused on the types of knowledge required for making inferences. For example, Carton (1971) argues that context is composed of *intralingual context* (arising from knowledge about the target language), *interlingual context* (cues from other languages, including loan

words and cognates), and *extralingual context* (arising from world knowledge and the target culture). Similarly, but somewhat more simply, Pressley, Levin, and McDaniel (1987) distinguished external context from internal context, defining *external context* as “surrounding sentence and paragraph cues” and *internal context* as “inspection of word parts” (p. 121). This distinction is worth noting because the influence of internal context has often been neglected or underestimated in studies due to the use of pseudowords as test stimuli. Even in the absence of many external cues, lexical/morphological knowledge (internal context) could provide enough cues to help readers guess the meaning of some words. Beck, McKeown, and McGaslin (1983) distinguished pedagogical contexts from natural contexts. *Pedagogical contexts* are specially designed for instructional purposes, and they contain explicit information to help students guess the meaning of unknown words. In contrast, *natural contexts* do not contain such explicit cues. Importantly, the majority of words in authentic texts appear without explicit contextual information.

There are very few empirical studies investigating the role of context in inferring word meaning among young learners. In Cain, Oakhill, and Lemmon (2004), L1 readers aged 9-10 with less developed reading comprehension skills found it more challenging than their more skilled counterparts to infer the meaning of unknown words (pseudowords were used) in contexts that required more processing demands (i.e., the cues for the word meaning did not appear in the immediate context). From a different perspective, Cain and Oakhill (2014) examined the impact of young L1 readers’ word knowledge (both size and depth) on two types of inferences: local cohesion inferences (i.e., making connections between propositions by relying on synonyms or mapping related lexical items) and global coherence inferences (i.e., using vocabulary knowledge or general world knowledge to fill in gaps). Both types of inferences were associated with reading comprehension, but the students’ vocabulary knowledge, their depth of knowledge in particular (measured by defining word meaning and judging similarities in the meaning of pairs of words), was more important for global coherence inferences than for local cohesion inferences.

### 2.3. Variables associated with young learners’ reading comprehension

In addition to word-meaning inference abilities, many other possible variables influence reading comprehension. Concerning young L1 readers, research consistently reports strong relationships among word recognition (including phonological awareness), reading fluency, and reading comprehension (e.g., National Institute of Health, 2000). Interestingly, various oral language abilities (e.g., listening comprehension, syntactic complexity in spoken language, etc.) show increasingly

stronger associations with reading comprehension from the early to later elementary school years, perhaps at least in part because the growing complexity of texts for upper-grade students requires more sophisticated language knowledge in general (Duke & Carlisle, 2011). Toward the middle and upper elementary school years, some children start exhibiting difficulties in reading, a phenomenon known as the *fourth-grade slump* (Chall, 1987). Importantly, the major challenges experienced by students going through the fourth-grade slump appear to differ from the major challenges experienced by students who start having reading difficulties at earlier grades. Rather than struggling with word recognition, students who face challenges at upper primary grades seem to have difficulties with texts that contain increasingly more unfamiliar and abstract words and that pose heavier conceptual demands (Kucan & Palincsar, 2011). Working memory also influences their performance (Cain et al., 2004). The timing of the onset of these challenges coincides with the drastic development of their metacognitive abilities, including metalinguistic knowledge such as knowledge of morphosyntax (Anglin, Miller, & Wakefield, 1993). In addition, various home factors, such as the frequency of storybook reading at home, are found to influence reading comprehension throughout kindergarten to upper elementary school (Sénéchal, 2006). Importantly, such home factors interact with school factors. Various instructional approaches and techniques (e.g., types of questions students are asked, teachers' interactive styles of engagement, the amount of time spent on reading tasks, etc.) all appear to influence reading comprehension (Duke & Carlisle, 2011).

#### 2.4. Reading among young L2 readers

Compared with the large body of research on young L1 readers, empirical information on young L2 readers is relatively limited. The variables addressed in the previous section, such as oral language proficiency, reading fluency, decoding skills, word recognition in the target language, as well as working memory, socioeconomic status (SES), and literacy environment at home, also by and large apply to young L2 readers (e.g., August & Shanahan, 2006). However, research has also shown that, compared with monolingual children, young L2 readers have different types of linguistic knowledge (including vocabulary knowledge), background knowledge, and metacognitive reading strategies, depending on their unique bilingual experiences (e.g., García, 1991; Jimenez, García, & Pearson, 1996; Peregoy & Boyle, 2000; Verhoeven, 2011). Young L2 readers' L1 decoding skills and vocabulary knowledge can have some positive effects on their L2 reading, but the effects appear to depend on their L1-L2 typological combination and their chance of receiving high-quality L1 instruction (e.g., Proctor et

al., 2012). While the importance of bilingualism is increasingly recognized in education, language-minority children worldwide are often denied access to reading instruction (and schooling) in their home language.

Identifying unique characteristics of young L2 readers can offer useful information for improving instruction, but comparative studies of L1 and L2 readers need to be conducted with careful consideration since researchers can easily overlook L2 readers' home language resources and unique bilingual experiences. For example, it is frequently reported that young L2 learners have smaller vocabulary sizes both in their L1 and L2 compared with their respective monolingual counterparts (e.g., Bialystok, Luk, Peets, & Yang, 2010; Carlo et al., 2004; Mancilla-Martinez & Lesaux, 2011). However, L2 learners are usually exposed to L1 and L2 words in different contexts with different purposes, and their L1 and L2 vocabularies largely do not overlap (e.g., Peña, Bedore, & Zlatic-Giunta, 2002). In studies where L2 readers' vocabulary in their home language was taken into account, their overall vocabulary size (L1 and L2 combined) was compatible with that of their monolingual peers (e.g., Butler, 2019; De Houwer, 2009; Goodrich & Lonigan, 2018).

### 3. Research questions

Unfortunately, many young L2 learners around the world receive school instruction only in their target language. Improving our understanding of the relationship between young readers' – particularly young L2 readers' – ability to construct word meaning in context and their reading comprehension, has great potential to inform the development of useful strategy instruction for young readers in need. Therefore, the present study, as part of a larger study investigating the role of lexical abilities in young learners' reading comprehension, explores the following questions related to English reading:

- RQ1: Among fourth-grade students (both monolingual and bilingual) who have received academic instruction only in English, are there any differences in their performance, in relation to their reading comprehension levels (*reading* – strong readers vs. emergent readers) and language backgrounds (*language* – monolingual English-speaking students vs. L2 students), in the following areas?
  - vocabulary size (receptive domain);
  - oral reading fluency;
  - inference of meaning of unfamiliar words;
  - metacognitive reasoning to arrive at meaning.

- RQ2: Does context (explicit context vs. implicit context) influence the students' performance, in relation to their reading levels and language backgrounds, when inferring/constructing the meaning of unfamiliar words?
- RQ3: How does context influence the kinds of knowledge and strategies the students use when inferring/constructing the meaning of unfamiliar words?

The study focuses on fourth graders because this grade is a critical time for children's reading development (recall the *fourth-grade slump*, Chall, 1987) and because it is when children start drastically improving their metacognitive skills (Anglin et al., 1993).

With respect to *context*, this study adopts the conceptualizations of Pressley et al. (1987) and Beck et al. (1983) mentioned above, and compares two context conditions: explicit and implicit contexts. The former contains more external information "surrounding sentence and paragraph cues" (Pressley et al., 1987, p. 121), while the latter has very little such external information. Thus, learners must rely more on "inspection of word parts" (p. 121). The explicit context is designed to be more pedagogically friendly because it provides learners with more information that they can use for making lexical inferences.

It should be emphasized that this study focuses on how students infer and explain word meanings for reading comprehension rather than for vocabulary learning (a distinction suggested by Nation, 2013). Even if a reader can successfully infer word meanings in a given context, it does not necessarily mean that the same individual could learn such words from the context. Moreover, in examining learners' abilities to infer word meaning, some researchers make a distinction between *source of knowledge* (e.g., linguistic knowledge, world knowledge, etc.) and *strategies* (e.g., analyzing, monitoring, etc.; e.g., Nassaji, 2003). However, our data among children did not always allow us to make such a distinction reliably; consequently, the present study combines these elements and refers to them as *inference knowledge and strategies* hereafter.

## 4. Method

### 4.1. Participants

The participants were 61 fourth-grade students (aged 9-10) in the United States from a single school district. Twenty-four of them were monolingual English-speaking students (ME students), and 37 were L2 learners of English from either Spanish- or Vietnamese-speaking homes.<sup>1</sup> In an effort to minimize variations in the amount and type of formal English instruction participants had previously

---

<sup>1</sup> There were 14 Spanish-speaking students and 23 Vietnamese-speaking students in the study.



received, we included only participants (both ME and L2 students) who had been enrolled in the same schools since kindergarten.<sup>2</sup> The L2 students all received English language development (ELD) instruction (i.e., they were pulled out of regular classes and received individual or small group ELD instruction) but had received no special support in their home language at school. All participants came from Title I schools, meaning that they came from middle to lower socioeconomic status (SES) backgrounds. Responses to questionnaires distributed to the participants' parents as part of this study also indicated that there were no major differences in SES (measured by parental educational levels and occupation, the number of books at home, habits of reading to children, etc.) across groups.<sup>3</sup>

Both ME and L2 students were further categorized as either strong or emergent readers based on their reading levels in English. The reading level for L2 students was gauged by the following measurements: (a) a standardized reading test (reading performance on the Stanford Achievement Test; using normal curve equivalent [NCE] scores, students scoring 40 and lower were grouped as emergent readers and students scoring 60 and higher were grouped as strong readers); (b) San Diego Quick (a reading diagnostic test); (c) a district-administered running record (a teacher-based assessment in which miscues and self-corrections during oral reading were examined by teachers), and a recommendation from district ELD teachers. The Stanford Achievement Test was the only measurement used to determine reading levels for ME students. The same criterion based on NCE described above was used to group ME students into strong and emergent readers. The average NCEs for Stanford Achievement Test reading scores were 70.5 for strong ME readers (ME+), 25.3 for emergent ME readers (ME-), 70.2 for strong L2 readers (L2+), and 30.7 for emergent L2 readers (L2-).<sup>4</sup>

All the L2 participants were originally judged as English-language learners (ELLs) who needed special ELD assistance when they entered kindergarten. At the time of their participation in this study (while they were enrolled in the fourth

---

<sup>2</sup> The precise age of first intensive exposure to English among L2 students was not known, but it was earlier than 5 years of age.

<sup>3</sup> There was no statistically significant difference in the years of schooling among parents with different L1: 14.4, 13.0, and 11.5 years for English, Vietnamese, and Spanish speakers, respectively. All the L2 students' parents indicated that both English and the students' home language (either Vietnamese or Spanish) were used at home, except one Vietnamese family who reported that Vietnamese was almost exclusively used at home.

<sup>4</sup> For L2 students, there was no significant difference in the average scores of Stanford Achievement Test by their home languages. For the L2+ students, the average scores were 70.3 and 69.9 for Vietnamese and Spanish groups, respectively. For the L2- students, the average scores were 27.8 and 32.6 for Vietnamese and Spanish, respectively. Because the numbers of each language group were small and there was no home language support at school among the L2 students, systematic analyses by home language were not performed in this study.

grade), all L2+ readers were classified as fluent English proficient (FEP) students and were no longer classified as ELLs by the district, while all L2-readers were still classified as ELLs. However, all L2 students from both reading groups had acquired sufficient oral skills in English based on a standardized test (the IDEA Oral Language Proficiency Test, IPT), which was also used by the district as a redesignation criterion. Namely, L2-students in the study were still classified as ELLs because they had not yet met the districts' criteria in reading and writing for redesignation, even though they had already acquired sufficient oral skills in English.

Participants were randomly selected from the students at the participating schools who met the preceding criteria.

#### 4.2. Instruments and procedures

Two vocabulary assessments were employed in this study. First, to assess the size of students' English receptive vocabularies, the *peabody picture vocabulary test-revised* (PPVT-R, Dunn & Dunn, 1981, referred to as PPVT hereafter) was individually administered. Second, in order to assess students' ability to infer and explain word meanings in context, a vocabulary assessment (referred to as the *inference assessment* hereafter) was developed and administered. This assessment had two components. First, students were asked to read aloud short paragraphs that contained words that were unfamiliar to them, and then they were asked to guess the meanings of the unfamiliar words in question. This was designed to assess their *reading fluency* and ability to *infer and explain the meaning* of specified words in context. The students were allowed to read the paragraphs silently after reading them aloud if they wished to do so. They were also allowed to ask the administrator the meanings of any unfamiliar words (if any) besides the words in question. Second, immediately after inferring the meaning of each word, students were asked to explain how they arrived at the meanings of the words as they defined them. This question was designed to elicit their use of *metacognitive reasoning* for determining the meanings of the words in question.

The inference assessment consisted of 20 items that were assumed to be unfamiliar<sup>5</sup> to fourth graders, and two versions of this test (each version with the same set of 20 items) were prepared. Half of the words in each version were presented with explicit external contextual information and the other half were presented

---

<sup>5</sup> Unfamiliarity was originally judged by two ELD specialists and one fourth-grade teacher at one of the participating schools. The words identified by the teachers as being potentially "unfamiliar" were pilot tested with fourth graders at a nonparticipating school within the same district. Six ME students and 10 L2 students who were considered to be strong readers were tested in the pilot study. The final words used in the inference assessment were composed of words that none of the pilot study participants had known.

with implicit contextual information. As is evident from the following example, the paragraphs were short, and every effort was made to keep the syntax simple:

Explicit context	The teacher left Ricardo to watch the class while she went to make copies. She told him to be responsible and make sure we kept working on our project. She wanted him to <u>oversee</u> us.
Implicit context	The teacher left Ricardo with the class while she went to make copies. She wanted him to <u>oversee</u> us.

The items were counterbalanced, with half of the students in each group taking Version 1 and the other half taking Version 2 of the test. The items were selected from a larger pool of items for which four research assistants with ELD teaching experience had graded the explicitness of contextual information on a scale from 1 to 5, with 1 being least explicit and 5 being most explicit. Only the pairs whose difference in average rating was more than 2.5 scores apart were selected as final test items. The average score for items with explicit context was 4.3; for items with implicit context the average score was 1.3. Note that we controlled only for external contextual information; the degree of explicitness was judged based on the availability of external contextual cues (e.g., restatements of the word meaning, examples, and synonyms) but not on the availability of internal contextual cues (e.g., affixes and compounds). All the words in question were actual nouns, verbs, and adjectives; we did not use pseudowords (see the appendix for the list of words used in the inference assessment). The assessment was conducted individually by a district ELD coordinator with more than 10 years of teaching experience. She was blind to the backgrounds of the participating students.

#### 4.3. Data analyses

All of the students' responses as well as their interactions with the teacher during the inference assessment were audio-recorded and transcribed. Several different coding systems were employed to analyze the data, as explained below. For each coding system, 12 students (three students from each group) were randomly chosen, and their responses were independently coded by two researchers. The disagreements were discussed (accounting for approximately 15% to 20% of the items for each coding system), and the rest of the data were independently coded again. Intraclass correlation coefficients between the two raters for the entire data set were calculated for each coding system, as reported below. The transcribed data were also analyzed qualitatively in order to understand the students' use of metacognitive knowledge.

First, *overall fluency* was judged based on students' performance when they read the item paragraphs aloud, and it was coded once for each student. Overall fluency was intended to holistically capture the level of speed and accuracy in oral text reading and was judged using a scale from 0 to 3 (where 0 indicated "not fluent" and 3 indicated "very fluent"). In other words, it was a gross measurement of students' ease in oral reading processing. The intraclass correlation coefficient was .91.

Second, students' ability to infer and define words in context (*defining words*, considered a kind of depth of word knowledge in previous research) was coded for each of the students' responses. A 4-point scale was employed (where 0 was "mentioned completely irrelevant meaning in the given context or no response" and 3 was "could clearly define a relevant meaning in context"). Even if an answer did not exactly match the conventional meaning but the definition perfectly made sense in the given context, students could score a 3 (although such cases were rare). Admittedly, this coding could not unpack students' abilities to construct word meaning and to articulate ideas. What this coding captured, therefore, was their ability to articulate constructed meanings. The intraclass correlation coefficient was .92. In addition, *word category identification* was also coded for each response. This coding aimed to capture to what extent students' explanations of word meanings matched the lexical categories of the target words, such as nouns, verbs, and so forth. This is considered a kind of vocabulary knowledge (depth) that gives us additional information on the accuracy of their inference. A 3-point scale (where 0 was "did not match the target lexical category" and 2 was "perfectly matches") was employed. The intraclass correlation coefficient was .88.

Finally, students' use of metacognitive knowledge was analyzed both quantitatively and qualitatively. For the quantitative analysis, first, the *degree of metacognitive reasoning* was judged holistically using a 4-point scale (0 was "no sign of metacognitive reasoning" and 3 indicated "extensive metacognitive reasoning") for each response. This measure captures how well students could articulate their reasoning in meaning making. To receive a 3 for this measure, for example, students needed to provide an involved explanation or theory for why they gave a particular definition, such as identifying one or more specific sources of cues in context. The intraclass correlation coefficient was .85. To further examine the source of inferencing knowledge and strategies, the following coding scheme was developed inductively while consulting with previous studies of strategies employed by adult learners (e.g., Nassaji, 2003; Paribakht & Wesche, 1999): (a) phonological cues; (b) lexical cues (e.g., use of knowledge of compound and root words, morphological knowledge such as prefixes and suffixes, and use of first language lexical knowledge or cognates); (c) world knowledge; (d) external contextual information (use of contextual cues available in the text); (e) partial memory/knowledge (use

of partial knowledge of the word in question); and (f) unspecified or no response.<sup>6</sup> The first two categories (phonological and lexical cues) concern internal information residing in the target words, whereas the next three (world knowledge, external contextual information, and the partial memory/knowledge) concern information external to the targeted words. Thus, for example, identifying a synonym of the target word in a given paragraph is categorized as the use of external contextual information. Multiple entries were possible for this coding, although it hardly happened.<sup>7</sup> The intraclass correlation coefficient was .91.

## 5. Results

Before answering the research questions, it was necessary to confirm that the two versions of the inference assessment were indeed compatible. First, the reliabilities of the two versions of the assessment were checked. The reliability coefficients (Cronbach's alpha) were .85 for Version 1 and .76 for Version 2. Second, the mean scores of *defining words* of the two versions were compared. The means for versions 1 and 2 were 1.92 and 1.76, respectively, and a one-way ANOVA failed to find a significant difference ( $F(1, 38) = .59, p > .1, \eta_p^2 = .03$ ). Thus, the scores from both versions were combined for the analyses described below.

### 5.1. Students' ability to infer and define words in context and other related abilities by group

First, the students' performance on the PPVT (*receptive vocabulary size*) and the inference assessment (*overall reading fluency, defining words, word category identification, and metacognitive reasoning*) was examined. The descriptive results (means and standard deviations) are indicated in Table 1. With respect to the PPVT, students' performance was normalized with a mean ( $M$ ) of 50 and a standard deviation ( $SD$ ) of 21.06.<sup>8</sup> A two-way ANOVA indicated significant differences

---

<sup>6</sup> One can argue that reading texts silently after the read-aloud (an option given to the students) can be a strategy. In the current study, it was not coded as a strategy. This is because when the students made a pause before responding and it was not clear whether they were reading the text silently or thinking about the meaning of the target word. A systematic analysis on the potential impact of this strategy can be a topic of future investigation.

<sup>7</sup> Out of 10 items for each contextual condition, the average numbers of coding per child were 10.20 (for the explicit context condition) and 10.24 (for the implicit context condition). It was very unlikely that the multiple coding significantly inflated the frequencies of each category of metacognitive knowledge and strategies for inferring word meaning.

<sup>8</sup> The normed scores were used for the PPVT because the normalization was based on a large, nationwide representative sample in the USA, and all the participating students in this

between strong and emergent readers ( $F(1, 57) = 41.22, p < .001, \eta_p^2 = .42$ ) and ME and L2 readers ( $F(1, 57) = 19.25, p < .001, \eta_p^2 = .25$ ). It failed to find a significant interaction effect, however ( $F(1, 57) = .83, p = .37, \eta_p^2 = .01$ ). Note that for L2 students, only their L2 (English) vocabulary size was assessed, without taking their L1 vocabulary knowledge into account; as discussed in the literature review section, because this could be a potential bias towards L2 students, PPVT is used as a covariate in the analyses below.

In comparing the mean scores across groups for the rest of the measures, a series of two-factor ANOVAs were employed, first without controlling for PPVT and then controlling for PPVT (PPVT was used as a covariate). As shown in Table 2, when no covariate was used, significant differences between strong and emergent readers were found in all these measures, and differences between ME and L2 readers were found only for defining words and metacognitive reasoning. None of the variables showed interaction effects. After controlling for PPVT, a significant main effect for *reading level* was found for all the measures (overall fluency, defining words, word category identification, and metacognitive reasoning) while the main effect for *language* in defining words and metacognitive reasoning disappeared. In other words, when we controlled for students' receptive vocabulary size, all the other measures examined in the inference assessment showed differences in performance between strong and emergent readers (strong readers having higher scores) but failed to show any difference in performance between ME and L2 students.

Table 1 Means and standard deviations for overall fluency, defining word meaning, word category identification, and metacognitive reasoning

	Monolingual English-speaking (ME) readers		English as a second language (L2) readers	
	Strong (ME+) (N = 12)	Emergent (ME-) (N = 12)	Strong (L2+) (N = 18)	Emergent (L2-) (N = 19)
PPVT <sup>a</sup>	66.92 (15.30)	38.67 (13.05)	46.50 (14.83)	25.26 (15.15)
Overall fluency (0-3)	2.83 (.58)	1.25 (.87)	2.83 (.38)	1.37 (.68)
Defining words (0-3)	2.33 (.39)	1.61 (.33)	2.07 (.33)	1.45 (.40)
Word category (0-2)	1.86 (.13)	1.56 (.34)	1.79 (.14)	1.50 (.15)
Metacognitive reasoning (0-3)	2.48 (.23)	1.90 (.30)	2.25 (.25) <sup>b</sup>	1.69 (.34)

Notes. Word category identification scores were based on a 3-point scale (from 0 to 2), and the others were based on a 4-point scale (from 0 to 3). Standard deviations are indicated in parentheses. <sup>a</sup>Normalized scores. <sup>b</sup>N = 17 (one student could not offer any reasoning and was thus excluded from the metacognitive reasoning analysis).

study received schooling only in English. Although the PPVT has been used extensively among L2 learners as well, problems with using it with L2 learners have been raised, and potential influence in performance from the students' L1 was reported (e.g., Goriot et al., 2018).

Table 2 Comparison of mean scores across groups

	No covariate			PPVT controlled		
	Main effect reading	Main effect language	Interaction	Main effect reading	Main effect language	Interaction
	Strong vs. emergent	ME vs. L2	Reading vs. language	Strong vs. emergent	ME vs. L2	Reading vs. language
Overall fluency	$F(1, 57) = 84.38^{**}$ ( $\eta_p^2 = .60$ )	$F(1, 57) = .13$ ( $\eta_p^2 = .002$ )	$F(1, 57) = .13$ ( $\eta_p^2 = .002$ )	$F(1, 56) = 39.39^{**}$ ( $\eta_p^2 = .35$ )	$F(1, 56) = 3.48$ ( $\eta_p^2 = .06$ )	$F(1, 56) < .001$ ( $\eta_p^2 < .001$ )
Defining words	$F(1, 57) = 50.41^{**}$ ( $\eta_p^2 = .47$ )	$F(1, 57) = 5.00^*$ ( $\eta_p^2 = .08$ )	$F(1, 57) = .29$ ( $\eta_p^2 = .01$ )	$F(1, 56) = 10.82^*$ ( $\eta_p^2 = .16$ )	$F(1, 56) = .01$ ( $\eta_p^2 < .001$ )	$F(1, 56) < .001$ ( $\eta_p^2 < .001$ )
Word category	$F(1, 57) = 33.43^{**}$ ( $\eta_p^2 = .37$ )	$F(1, 57) = 1.62$ ( $\eta_p^2 = .03$ )	$F(1, 57) = .001$ ( $\eta_p^2 < .001$ )	$F(1, 56) = 9.74^*$ ( $\eta_p^2 = .15$ )	$F(1, 56) < .001$ ( $\eta_p^2 < .001$ )	$F(1, 56) = .05$ ( $\eta_p^2 < .001$ )
Metacognitive reasoning	$F(1, 56) = 55.96^{**}$ ( $\eta_p^2 = .5$ )	$F(1, 56) = 8.39^{**}$ ( $\eta_p^2 = .13$ )	$F(1, 56) = .042$ ( $\eta_p^2 = .84$ )	$F(1, 55) = 13.52^{**}$ ( $\eta_p^2 = .20$ )	$F(1, 55) = .42$ ( $\eta_p^2 = .01$ )	$F(1, 55) = .09$ ( $\eta_p^2 < .001$ )

Notes.  $^{**} p < .001$ ,  $^* p < .05$ . Effect sizes (partial eta-squared,  $\eta_p^2$ ) are indicated in parentheses.

## 5.2. The role of context in inferring word meaning

When examining students' performance in defining words in context, we also compared performance in explicit and implicit contextual conditions across groups (Figure 1). A two-factor ANOVA with a repeated measure (*context* being a within-subject measure) indicated that there were main effects in *reading level* ( $F(1, 57) = 50.00, p < .001, \eta_p^2 = .47$ ), *language* ( $F(1, 57) = 5.08, p = .028, \eta_p^2 = .08$ ), and *context* ( $F(1, 57) = 78.51, p < .001, \eta_p^2 = .58$ ). There was no interaction effect.<sup>9</sup> After controlling for PPVT, there were main effects in *reading level* ( $F(1, 56) = 10.62, p = .002, \eta_p^2 = .16$ ) and *context* ( $F(1, 56) = 21.11, p < .001, \eta_p^2 = .27$ ), whereas the main effect for *language* disappeared ( $F(1, 56) = .01, p = .94, \eta_p^2 = .00$ ). An interaction effect was found between *language* and *context* ( $F(1, 56) = 5.19, p = .027, \eta_p^2 = .09$ ) when controlling for PPVT (although the effect size was small).<sup>10</sup>

Next, students' degree of metacognitive reasoning in explicit and implicit contexts was compared (Figure 2). As explained above, this measure captures how well students can articulate their reasoning for arriving at the meaning of words in context. A two-factor ANOVA with repeated measures showed that there were main effects in *reading level* ( $F(1, 56) = 52.83, p < .001, \eta_p^2 = .49$ ), *language* ( $F(1, 56) = 7.84, p = .007, \eta_p^2 = .12$ ), and *context* ( $F(1, 56) = 16.88, p < .001, \eta_p^2 = .23$ ).

<sup>9</sup> The results of the interactions were as follows: *context* and *language* ( $F(1, 57) = 2.30, p = .14, \eta_p^2 = .04$ ); *context* and *reading* ( $F(1, 57) = 3.06, p = .09, \eta_p^2 = .05$ ); and *context, language, and reading* ( $F(1, 57) = 1.52, p = .22, \eta_p^2 = .03$ ).

<sup>10</sup> Other interactions were not significant: *context* and *reading* ( $F(1, 56) = .02, p = .88, \eta_p^2 = .00$ ); and *context, language, and reading* ( $F(1, 56) = 1.06, p = .31, \eta_p^2 = .02$ ).

There was no interaction effect.<sup>11</sup> After controlling for PPVT, only the main effect in *reading level* was significant ( $F(1, 55) = 12.05, p < .001, \eta_p^2 = .18$ ).<sup>12</sup>

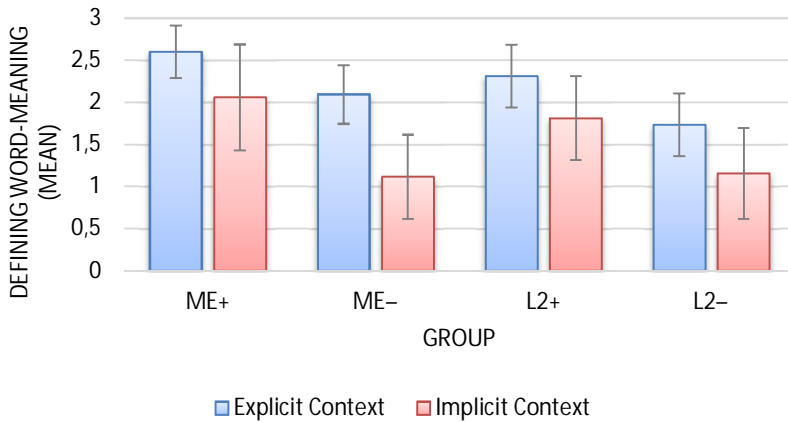


Figure 1 Defining words by context (means and SDs; the black lines indicate SDs for each group).

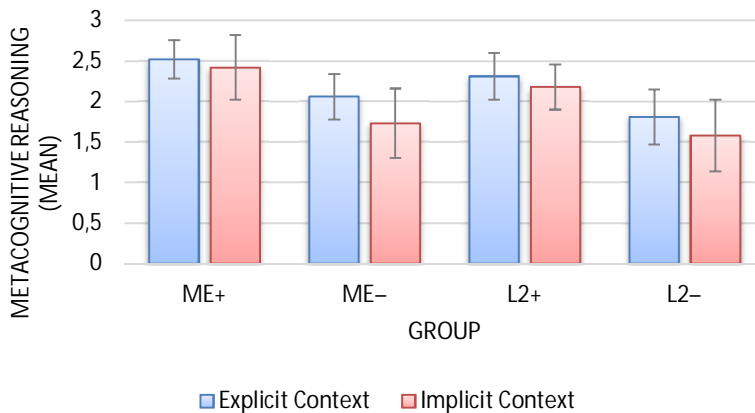


Figure 2 Metacognitive reasoning by context (means and SDs; the black lines indicate SDs for each group).

<sup>11</sup> The results of other interactions were as follows: *context* and *language* ( $F(1, 56) = .14, p = .72, \eta_p^2 = .002$ ); *context* and *reading* ( $F(1, 56) = 3.10, p = .08, \eta_p^2 = .05$ ); and *context, language, and reading* ( $F(1, 56) = .43, p = .52, \eta_p^2 = .008$ ).

<sup>12</sup> The results of others were as follows: *language* ( $F(1, 55) = .28, p = .60, \eta_p^2 = .005$ ); *context* ( $F(1, 55) = 1.77, p = .14, \eta_p^2 = .03$ ); *context* and *language* ( $F(1, 55) = .12, p = .73, \eta_p^2 = .002$ ); *context* and *reading* ( $F(1, 55) = 1.66, p = .20, \eta_p^2 = .03$ ); and *context, language, and reading* ( $F(1, 55) = .41, p = .53, \eta_p^2 = .007$ ).



In sum, after controlling for students' vocabulary size, students across groups defined words and used metacognitive reasoning better in explicit contexts than in implicit contexts, and stronger readers outperformed emergent readers irrespective of ME and L2 status.

### 5.3. Knowledge and strategies for inferring word meaning by context

Students' source of knowledge and strategies are shown in Figures 3 (explicit context) and 4 (implicit context). The figures indicate the average frequencies that the students used for each category and the standard deviations. In order to get a general picture of their use of knowledge and strategies, a series of two-way ANOVAs with repeated measures were conducted while *context* was included as the between-subject factor. Because multiple comparisons were made, *p*-value was adjusted to .008. The phonological category was excluded from the ANOVA analyses because the distributions of the residuals were heavily skewed. It turned out the overwhelmingly majority of the students did not use the phonological strategy; only a handful of students repeatedly used it. For the rest of the categories, the results indicated that there were significant differences in the average mean frequencies between the explicit and implicit contexts in *external information* ( $F(1, 56) = 10.56, p = .002, \eta_p^2 = .16$ ), *partial memory/knowledge* ( $F(1, 56) = 8.77, p = .004, \eta_p^2 = .14$ ), and *unspecified/no-response categories* ( $F(1, 56) = 21.37, p < .001, \eta_p^2 = .28$ ).<sup>13</sup> For *lexical information*, a significant difference was found in *reading* ( $F(1, 56) = 11.96, p < .001, \eta_p^2 = .18$ ). For *partial memory/knowledge*, in addition to *context*, main effects were found both in *reading* ( $F(1, 56) = 11.30, p < .001, \eta_p^2 = .17$ ) and *language* ( $F(1, 56) = 10.80, p = .002, \eta_p^2 = .16$ ). Namely, strong readers on average used the internal lexical information more than emergent readers. Strong readers and ME readers made more use of partial memory and knowledge. Although it did not reach the adjusted significant level, there was an approaching level in the main effect of *reading* for *unspecified/non-response* (emergent readers being higher) ( $F(1, 56) = 6.30, p = .01, \eta_p^2 = .10$ ). No interaction effect was found in any of the categories. Note, however, that the frequencies for each category were small and the standard deviations were relatively large, suggesting that there were substantial individual differences. Thus, we interpreted the statistical results only for a gross tendency, and then examined the data qualitatively for more detail.

---

<sup>13</sup> Statistical results for other strategies were:  $F(1, 56) = 5.33, p = .025, \eta_p^2 = .09$  for *lexical knowledge*; and  $F(1, 56) = 5.42, p = .024, \eta_p^2 = .09$  for *world knowledge*.

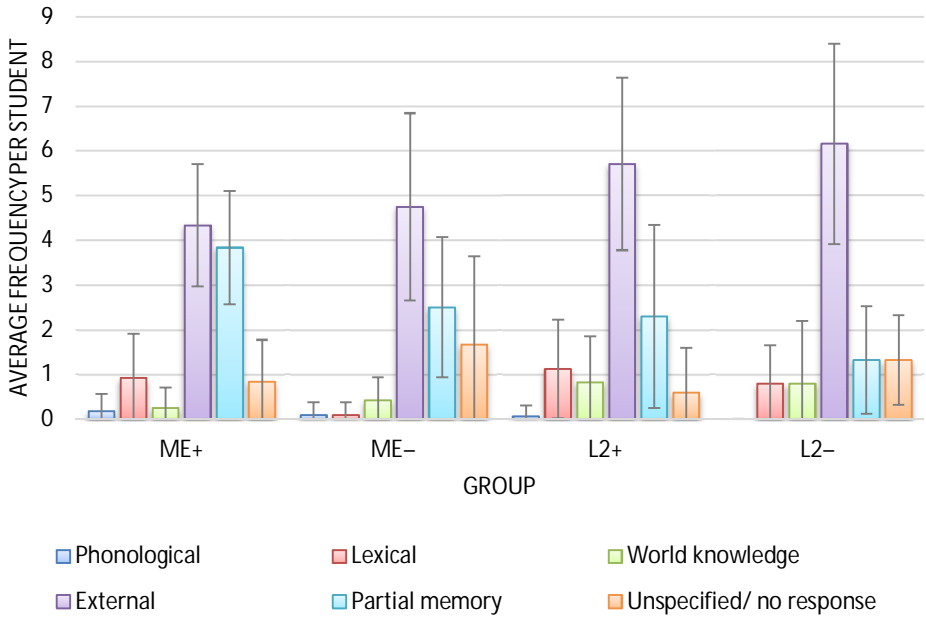


Figure 3 Knowledge and strategies to infer word meaning by group in the explicit context

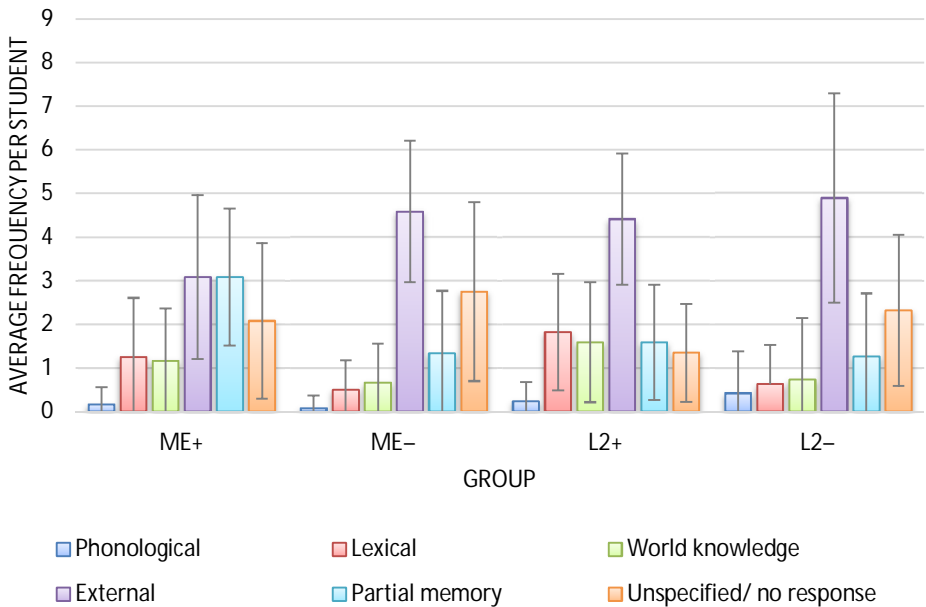


Figure 4 Knowledge and strategies to infer word meaning by group in the implicit context

In the explicit context, as expected, students frequently relied on external contextual information as a cue, but the sources used often appeared to differ between students who could and could not successfully construct meaning. A close look at the transcription showed that the students who succeeded in constructing meaning in context were usually better at using multiple sources of contextual information available in the entire paragraph. They also tended to monitor and clearly articulate their meaning-construction process, as exemplified by the following case of a strong reader:

Excerpt 1

*Student: (Reading aloud)*

*I haven't eaten anything for three days. I am famished.<sup>14</sup> I'm very hungry. I think I will go to the store and buy some food.*

*Um, he's amazed? I don't know, he's [?]*

*Teacher: WHAT DO YOU THINK?*

*Student: I am hungry. He's really starving.*

*Teacher: OKAY. WHY DO YOU THINK IT MEANS HUNGRY OR STARVING?*

*Student: Because it goes, I haven't eaten anything for three days. That's a clue that he's really hungry...*

*And it says right here, I'm very hungry. And he said, I think I'll go to the store and buy some food, like starving.*

*Teacher: GOOD JOB, AND YOU'RE RIGHT. FAMISHED MEANS STARVING OR HUNGRY. GOOD.*

By contrast, when the students could not infer the word meaning successfully, they tended to pay attention only to the immediate context of the word in question and guess the meaning by relying on their own knowledge associated with the limited information that they focused on. They also often missed syntactic or cohesion cues such as conjunctions, determiners, and pronouns that would provide readers with relational information among propositions and ultimately help them make suitable inferences in context. Below is an example from an emergent reader:

Excerpt 2

*Student: (Reading aloud)*

*Anthony had been working hard for several months, and he was feeling very tired. He decided to take a trip to Hawaii. The getaway was just what he needed to relax.*

*Teacher: GOOD. WHAT DO YOU THINK GETAWAY MEANS?*

*Student: A jacuzzi?*

*Teacher: A JACUZZI? WHY DO YOU THINK IT MEANS A JACUZZI?*

*Student: 'Cause he wants to relax.*

---

<sup>14</sup> In the excerpts, the underlined words are the target words. The teacher's utterances are indicated in capital letters.

"A jacuzzi" was a creative guess, and one can imagine why this student associated this idea with "relaxing." However, the definite article "the" in front of "getaway" indicates that "the getaway" is supposed to be mutual knowledge between the reader and the writer. Thus, introducing a new piece of information (in this case, "a jacuzzi") is not the most natural reading in this particular context.

Another notable characteristic of the students' inference knowledge/strategies in the explicit context condition was relatively high frequencies of the *partial memory/knowledge* category. The contextual information provided in the explicit condition seemed to help students improve or modify their partially acquired knowledge of the word in question. Some students retrieved incidences of encountering the target word somewhere else and used contextual information associated with the previous encounters as well as contextual information in the task at hand to construct the word meaning. For example, when asked what "pending" means after reading "the school board decided to build a new gym at the school last year. But this plan is pending because they don't have enough money. We don't have a gym yet," a student remembered her mother, who was trying to sell a house, saying to her neighbor that "the sale is pending." Another student, after being asked the meaning of "aviator," realized that he had heard of "aviation" on TV recently and guessed the meaning of "aviator" to be a pilot. Students with larger vocabularies seemed to have more advantage in using partial word knowledge/memory (the frequency was the highest among the ME+ group, as shown in Figure 3), and extra contextual information given in the paragraph helped them refine the meaning of the word in context. This mechanism may in turn assist them in further developing their vocabulary.

In the implicit context condition, we can expect that students need to rely more on internal word knowledge and their own world knowledge, given that there is relatively less contextual information available in the paragraphs themselves. This assumption is not necessarily supported in our data. But strong readers, regardless of ME and L2 backgrounds, used lexical knowledge more frequently than emergent readers, irrespective of the contexts. For example:

Excerpt 3

*Student: (Reading aloud)*

*Ken really wanted to wear a new suit to the party, but he couldn't because his new penn-i-less.*

*Teacher: BECAUSE HE WAS.*

*Student: . . . he was penni...*

*Teacher: PENNILESS.*

*Student: . . . penniless.*

*Teacher: CAN YOU GUESS WHAT THE WORD PENNILESS MEANS?*

*Student: Broke.*

Teacher: AND WHAT MAKES YOU THINK THAT?

Student: Because, uh, I know that penny means and what less means, so I thought he probably doesn't have any money.

Teacher: OKAY. GOOD. PENNILESS MEANS BEING POOR OR HAVING NO MONEY.

Other students seemed to use internal word knowledge to ensure that the meaning they constructed based on (limited) external contextual information was appropriate.

Excerpt 4

Student: (Reading aloud)

At first, the doctor worried about John. Then, the new medicine prolonged John's life.

Teacher: SO, WHAT DO YOU THINK PROLONGED MEANS?

Student: I think it made it longer, it made his life longer.

Teacher: OKAY, WHY?

Student: Well, if, at first the doctor worried about John. Then the new medicine . . . John's life. John, that, the last sentence actually gives me a hint.

Teacher: HMM, OKAY. AND IS THERE ANYTHING IN THE WORD THAT ALSO GAVE YOU A HINT OR...

Student: Yes, long!

Teacher: THERE'S THE WORD LONG IN THERE; RIGHT.

Student: Yes.

Teacher: GOOD, AND YOU'RE RIGHT. PROLONG MEANS TO MAKE LONGER, TO STRETCH, ANYTHING THAT HAPPENS TO...

Student: Even bigger.

Teacher: UM-HMM, TO EXTEND, RIGHT, GOOD.

As regards lexical knowledge, students, strong readers in particular, could use knowledge of suffixes and prefixes appropriately in context, when such information was available. In contrast, students' efforts to use compound word knowledge did not work well on occasion despite their rather creative attempts. Students' difficulty constructing the meaning of compound words reflects the complexity of compound word meanings in English. For instance, one student reasoned that "outgoing girl" means "she goes out a lot," while another reasoned that "to oversee" means "to double-check" because "over means to do it over, to do it again." The next example, from an emergent reader, illustrates that segmenting a word is not always easy for students.

Excerpt 5

Student: (Reading aloud)

The school board decided to build a new gym. But the plan is pending.

Teacher: THE PLAN IS PENDING. WHAT DOES PENDING MEAN?

*Student: The plan is pending. Um, pending means, like, the plan is working?*

*Teacher: UM-HMM. WHAT MAKES YOU THINK IT MEANS WORKING?*

*Student: Because they're using a pen.*

Across different groups, students did not frequently rely on phonological sources of knowledge, but when students used this source, they usually arrived at inappropriate or confusing meanings. Those students tended to simply guess the meaning of words based on phonologically similar words, which were completely irrelevant to the context. Examples include “bending” and “depending” for the meaning of “pending,” “nice” for the meaning of “durable” (based on the association that this was related to the word “adorable”), and “to show” for the meaning of “apparel” because of its phonological similarity to “appear.” Relying on phonologically similar words often resulted in the misidentification of lexical categories as well (e.g., “apparel” is a noun but “appear” is a verb).

Finally, it is important to note that the frequency of “unspecifiable or no response” was higher among participants in the implicit context (more authentic contexts) than in the explicit contexts (more pedagogically oriented contexts). This finding can serve as a basis for implications for the use of explicit pedagogical texts in vocabulary learning for young readers, emergent readers in particular.

## 6. Discussion

This study investigated fourth-grade students’ abilities to infer the meaning of words in context as well as their abilities to employ metacognitive knowledge/strategies to do so. It examined how such abilities differ between strong and emergent readers and between ME and L2 readers. It also explored how the explicitness of contextual information influenced the ability to construct meaning among students with different reading proficiency and language backgrounds. How students employ metacognitive sources in two contextual conditions was also investigated. Quantitative analyses indicate significant differences between strong and emergent readers in performance in overall fluency, reading aloud, defining words, identifying word categories, and metacognitive reasoning. Although performance in defining words and metacognitive reasoning was initially significantly different between ME and L2 readers, such differences disappeared after controlling for their receptive vocabulary size. As discussed above, reporting on L2 learners’ vocabulary in the target language only can seriously mislead their “true” lexical knowledge, which is assumed to be spread across languages (e.g., De Houwer, 2009). In the present study, after controlling for students’ vocabulary size in English, in addition to other potentially confounding variables (e.g., basic oral proficiency levels in English, SES backgrounds, and

the amount and types of instruction previously received at school), the performances of ME and L2 students were found to be compatible. In contrast, differences between strong and emergent readers were unaffected even after controlling for their vocabulary size. The challenges that emergent readers faced, regardless of ME and L2 backgrounds, went beyond the need to increase their vocabulary size, including abilities for allocating sufficient mental resources (e.g., memory), making a good use of external information, and enhancing various types of lexical knowledge (i.e., vocabulary depth) in order to process texts for comprehension and articulate responses based on higher order reasoning.

All students, irrespective of their backgrounds, were better at defining words and using metacognitive reasoning to construct word meanings in explicit contexts (more pedagogically oriented texts) than in implicit contexts (more authentic texts). Providing more explicit contextual information, such as restating the meaning of the target words and offering synonyms and concrete examples, certainly helped all participants construct word meaning; the pedagogical benefits of having explicit information were evident. Although all students relied heavily on such external contextual information, readers who successfully could infer the word meaning seemed to be better at fully taking advantage of multiple sources of contextual information in the entire paragraph. By contrast, unsuccessful inference often came from paying insufficient attention to cohesive devices (e.g., conjunctions and determiners) across sentences and focusing on a certain word or phrase that appeared immediately before or after the target word within the same sentence. This finding is consistent with the study conducted by Can (2016) among L1 secondary school students which showed that their understanding of cohesive relations, conjunctions in particular, was associated with their reading comprehension levels. The finding in the present study suggests that young readers would also benefit from receiving instructional assistance on how to effectively use more global contextual cues and cohesive devices in meaning construction when they read.

Explicit contextual information also helped students evoke their partially acquired word knowledge and refine the word meaning in context. As noted previously, word knowledge is multifaceted and gradual (Schmitt, 2014). As children are exposed to new words, they refine and modify the meanings every time they encounter them in different contexts. Offering explicit contextual information seems to facilitate such a process. It is well known that over time, children with larger vocabulary size tend to develop their vocabulary size more rapidly than their peers with smaller vocabulary size. This phenomenon of cumulative advantage is called *the Matthew effect* (Duff, Tomblin, & Catts, 2015). Indeed, students with larger vocabularies in this study appeared to benefit more from the above-mentioned mechanism of refining word meaning, and this mechanism may partially explain why gaps in vocabulary size among students increase over time.

With respect to internal word information, strong readers, regardless of ME and L2 backgrounds, used such sources more frequently than emergent readers, irrespective of the textual context conditions. Students used knowledge of prefixes and suffixes strategically, when they used such knowledge at all. A study among L1-learning students in Grades 4 and 7, as well as high school undertaken by Nagy, Diakidoy, and Anderson (1993) indicated that the participants' knowledge of suffixes was so influential over their reading comprehension that it can be used as a diagnostic tool. The researchers also found that knowledge of suffixes developed substantially between the fourth- and seventh-grade levels. Moreover, individual differences in their morphological knowledge grew during the same time period. Considering that fourth graders who were strong readers in our study already seemed to benefit from using morphological knowledge to construct meaning in context, it may be a good idea to introduce explicit instruction on morphology to emergent readers at this grade level, if not earlier, before gaps in knowledge between strong and emergent readers become even more substantial.

Unlike participants' experiences with morphology, as exemplified in the "outgoing" and "oversee" cases reported above, the students appeared to find it challenging to identify the meaning of compound words regardless of their background. Indeed, syntactic and semantic relations within compounds in English are not straightforward. For example, "a magnifying glass" is a glass that magnifies, but "a looking glass" is not a glass that looks. Moreover, the meaning of a compound is not necessarily a combination of the meanings of each root. Complicated internal structures of English compounds are language specific. In English, there is no limit to the number of words that are allowed to be put together. Various types of combinations of lexical categories are possible, and determining the lexical category of compounds can be confusing for young learners. The lexical category of compounds usually follows the lexical category of the final root. For example, *headstrong* (noun + adjective) is an adjective and *cellphone* (noun + noun) is a noun. There are exceptions, however. *Hands-on* (noun + preposition) is not a preposition but an adjective, whereas *higher-up* (adjective + preposition) is a noun. Children need to understand such complexity associated with compounds. Again, it may be beneficial to have occasional explicit instruction on word formation in English as part of reading instruction in order to raise students' awareness of complex internal structures of compounds, as addressed above.

Curiously, there was hardly any instance of L2 students relying on cognates or their L1 knowledge in this study. This may be related to the fact that the L2 students in this study did not receive formal academic instruction in their L1 at school; they were not taught how to make use of their L1 resources at school. The lack of explicit reliance on L1 knowledge may also be due to the fact that Vietnamese was the L1 of the majority of the L2 students in this study, and Vietnamese



is not related to English and thus has no cognates. Small sample sizes and an unbalanced number of students with Vietnamese and Spanish backgrounds did not allow us to conduct any systematic analyses of the influence of their L1 over their performance in the inference assessment. For a future study, it would be worthwhile to systematically investigate the role of L1 in students' meaning-making processes and strategies during their L2 reading (and ideally in their L1 reading as well).

In implicit contexts, namely more authentic texts, young readers had greater frequencies of *unspecified and no responses* than in the explicit context (a more pedagogical condition). It seems that less authentic texts more effectively support young readers (emergent readers in particular) in constructing word meaning in context with their more explicit cues. Until young readers have developed a certain level of ability to use both external and internal information to make sense of word meaning in context, providing them with authentic texts without any assistance (independent reading) may have a limited effect on improving their reading comprehension.

There are a few limitations in this study. First, because paragraphs in the inference assessment were rather short, we were able to examine students' ability to use relatively local cohesion relationships but had limited capacity to thoroughly examine the role of global contextual information – information that goes beyond a single lexical or phrasal cue that connects propositions. Global coherence inference requires readers to construct relevant lexical and mental networks in order to fill gaps in concepts that are not directly stated in the text. It would be interesting to investigate the role of global contexts in children's meaning-making processes. Second, only the students' receptive vocabulary size in English was obtained in this study. For L2 readers, it would have been better if their vocabulary knowledge in their L1 had been taken into account (even though the participants in this study had academic instruction exclusively in English). Moreover, information on participants' expressive vocabulary would have been useful, even though all the participants were identified as orally proficient by a standardized measure (i.e., IPT). This is because one could argue that expressive vocabulary might be a better indicator of children's ability to define word meanings and express their metacognitive reasoning. Having information on students' expressive vocabulary size would be particularly helpful for L2 readers because larger gaps between receptive and productive vocabulary are often found among bilingual children (e.g., Gibson, Jarmulowicz, & Oller, 2018). Furthermore, in this study we focused on understanding how young students construct the meaning of unknown words but not on their actual learning of such words. Research examining how the meaning-making process relates to the learning of words would be of great value.

## 7. Conclusion

Although there is substantial research on young learners' vocabulary size and its relation to reading comprehension, students' ability to construct meaning in context is not well understood. This study investigated fourth-grade students' ability to infer and define meanings of unknown words in context and how they use metacognitive knowledge and strategies to arrive at meanings in two different contexts (i.e., explicit and implicit contexts). The study found that such abilities differed between strong and emergent readers as well as between ME and L2 readers. Importantly, however, after controlling for the students' English receptive vocabulary size, ME and L2 readers were compatible in their performance. The study also found that the students used more strategies in the explicit context and used different strategies depending on context.

Given these results, the present study can provide a few practical implications for pedagogy. The first suggestion is to recognize the potential merit of using texts with explicit contexts to provide instruction on strategies for inferring word meaning. Second, it would be pedagogically useful to provide emergent readers with explicit instruction on how to use external information. Metacognitive knowledge used by strong readers can be viewed as an important source of information when designing instruction to assist emergent readers. Useful strategies can include paying attention to syntactic or cohesive information and using wider contexts rather than focusing on the immediate local contexts of the words in question. Similarly, it would be helpful to give students explicit instruction on how to use internal lexical information by relying on morphological knowledge, compounds, and so forth. Fostering such knowledge in emergent readers would strengthen their reading skills.

References

- Anglin, J. M., Miller, G. A., & Wakefield, P. C. (1993). Vocabulary development: A morphological analysis. *Monographs of Society for Research in Child Development, 58*, 1-186. <http://doi.org/10.2307/1166112>
- August, D., & Shanahan, T. (Eds.). (2006). *Developing literacy in second-language learners: Report of the National Literacy Panel on language-minority children and youth*. Mahwah, NJ: Lawrence Erlbaum.
- Beck, I. L., McKeown, M. G., McCaslin, E. S. (1983). Vocabulary development: All contexts are not created equal. *The Elementary School Journal, 83*(3), 177-181. <http://doi.org/10.1086/461307>
- Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition, 13*(4), 525-531. <http://doi.org/10.1017/s1366728909990423>
- Butler, Y. G. (2019). Teaching vocabulary to young second- or foreign-language learners: What can we learn from the research? *Language Teaching for Young Learners, 1*(1), 4-33.
- Cain, K., & Oakhill, J. (2014). Reading comprehension and vocabulary: Is vocabulary more important for some aspects of comprehension? *L'Année Psychologique, 114*, 647-662. <http://doi.org/10.4074/s0003503314004035>
- Cain, K., Oakhill, J., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: The influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology, 96*(4), 671-681. <http://doi.org/10.1037/0022-0663.96.4.671>
- Can, R. (2016). The relationship between readers' comprehension skills and answers of cohesive relationships. *Anthropologist, 21*, 85-92. <http://doi.org/10.1080/09720073.2016.11891927>
- Carlo, M. S., August, D., McLaughlin, B., Dressler, C., Lippman, D. N., Lively, T. J., & White, C. E. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Reading Research Quarterly, 39*, 188-215. <http://doi.org/10.1598/rrq.39.2.3>
- Carton, A. S. (1971). Inferencing: A process in suing and learning language. In P. Pimsleur & T. Quinn (Eds.), *The psychology of second language learning* (pp. 45-58). Cambridge: Cambridge University Press.
- Chall, J. S. (1987). Two vocabularies for reading: Recognition and meaning. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 7-17). Hillsdale, NJ: Lawrence Erlbaum.
- Cronbach, L. J. (1942). Measuring knowledge of precise word meaning. *The Journal of Educational Research, 36*(7), 528-534. <http://doi.org/10.1080/00220671.1943.10881192>

- Dale, E. (1965). Vocabulary measurement: technique and major findings. *Elementary English*, 42, 895-901.
- de Bot, K., Paribakht, T. S., & Wesche, M. B. (1997). Towards a lexical processing model for the study of second language vocabulary acquisition: Evidence from ESL reading. *Studies in Second Language Acquisition*, 19, 309-329. <http://doi.org/10.1017/s0272263197003021>
- De Houwer, A. (2009). *Bilingual first language acquisition*. Bristol: Multilingual Matters.
- Diakidoy, I. N. (1998). The role of reading comprehension in word meaning acquisition during reading. *European Journal of Psychology of Education*, 13(2), 131-154. <http://doi.org/10.1007/bf03173086>
- Duff, D., Tomblin, J. B., & Catts, H. (2015). The influence of reading on vocabulary growth: A case for a Matthew effect. *Journal of Speech, Language, and Hearing Research*, 58, 853-864. [http://doi.org/10.1044/2015\\_jslhr-l-13-0310](http://doi.org/10.1044/2015_jslhr-l-13-0310)
- Duke, N. K., & Carlisle, J. (2011). The development of comprehension. In M. L. Kamil, P. D. Pearson, E. B. Moje, & P. P. Afflerback (Eds.), *Handbook of reading research*, (Vol. IV, pp. 199-228). New York: Routledge.
- Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test-Revised*. Circle Pines, MN: American Guidance Service.
- García, G. E. (1991). Factors influencing the English reading test performance of Spanish-speaking Hispanic children. *Reading Research Quarterly*, 26(4), 371-392. <http://doi.org/10.2307/747894>
- Gibson, T. D., Jarmulowicz, L., Oller, D. K. (2018). Difficulties using standardized tests to identify the receptive expressive gaps in bilingual children's vocabularies. *Bilingualism: Language and Cognition*, 21(2), 328-339. <http://doi.org/10.1017/s1366728917000074>
- Goodrich, J. M., & Lonigan, C. J. (2018). Development of first- and second-language vocabulary knowledge among language-minority children: Evidence from single language and conceptual scores. *Journal of Child Language*, 45(4), 1006-1017. <http://doi.org/10.1017/s0305000917000538>
- Goriot, C., van Hout, R., Broersma, M., Lobo, V., McQueen, J. M., & Unsworth, S. (2018). Using the picture vocabulary test in L2 children and adolescents: Effects of L1. *International Journal of Bilingual Education and Bilingualism*. Advance online publication. <http://doi.org/10.1080/13670050.2018.1494131>.
- Haynes, M. (1993). Patterns and perils of guessing in second language reading. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 46-64). Norwood, NJ: Ablex.
- Huckin, T., & Block, J. (1993). Strategies for inferring word meaning in context. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 153-178). Norwood, NJ: Ablex.

- Jimenez, R. T., Garcia, G. E., & Pearson, P. D. (1996). The reading strategies of bilingual Latina/o students who are successful English readers: Opportunities and obstacles. *Reading Research Quarterly, 31*(1), 90-112. <http://doi.org/10.1598/rrq.31.1.5>
- Ke, S., & Koda, K. (2019). Is vocabulary knowledge sufficient for word-meaning inference? An investigation of the role of morphological awareness in adult L2 learners of Chinese. *Applied Linguistics, 40*(3), 456-477. <http://doi.org/10.1093/applin/amx040>
- Kucan, L., & Palincsar, A. S. (2011). Locating struggling readers in a reconfigured landscape. In M. L. Kamil, P. D. Pearson, E. B. Moje, & P. P. Afflerback (Eds.) *Handbook of reading research*, (Vol. IV, pp. 341-358). New York: Routledge.
- Laufer, B. (1997). The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 20-34). Cambridge: Cambridge University Press.
- Mancilla-Martinez, J., & Lesaux, N. K. (2011). The gap between Spanish speakers' word reading and word knowledge: A longitudinal study. *Child Development, 82*(5), 1544-1560. <http://doi.org/10.1111/j.1467-8624.2011.01633.x>
- Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of Educational Research, 80*(3), 300-335. <http://doi.org/10.3102/0034654310377087>
- Masrai, A. (2019). Vocabulary and reading comprehension revisited: Evidence from high-, mid-, and low-frequency vocabulary knowledge. *SAGE Open*. <http://doi.org/10.1177/2158244019845182>
- Moghadam, S. H., Zainal, Z., & Ghaderpour, M. (2012). A review on the important role of vocabulary knowledge in reading comprehension performance. *Procedia – Social and Behavioral Sciences, 66*(7), 555-563. <http://doi.org/10.1016/j.sbspro.2012.11.300>
- Nagy, W. E. (1995). *On the role of context in first- and second-language vocabulary learning*. Technical Report No. 627, University of Illinois at Urbana-Champaign. <https://files.eric.ed.gov/fulltext/ED391152.pdf>
- Nagy, W. E., Diakidoy, I., & Anderson, R. C. (1993). The acquisition of morphology: Learning the contribution of suffixes to the meanings to derivatives. *Journal of Reading Behavior, 25*, 155-170. <http://doi.org/10.1080/10862969309547808>
- Nagy, W. E., & Herman, P. A. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 19-35). Hillsdale, NJ: Lawrence Erlbaum.

- 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). Mahwah, NJ: Lawrence Erlbaum.
- Nassaji, H. (2003). Vocabulary learning from context: Strategies, knowledge sources, and their relationship with success in L2 lexical inferencing. *TESOL Quarterly*, 37(4), 645-670. <http://doi.org/10.2307/3588216>
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge: Cambridge University Press.
- National Institute of Health. (2000). *Report of the National Reading Panel: Teaching children to read* (NIH Publication No. 00-4769). Washington, DC: US Department of Health and Human Services.
- Paribakht, T. S., & Wesche, M. (1999). Reading and "incidental" L2 vocabulary acquisition: An introspective study of lexical inferring. *Studies in Second Language Acquisition*, 21, 195-224. <http://doi.org/10.1017/s027226319900203x>
- Peña, E. D., Bedore, L. M., & Zlatic-Giunta, R. (2002). Category-generation performance of bilingual children: The influence of cognition, category, and language. *Journal of Speech, Language, and Hearing Research*, 45, 938-947. [http://doi.org/10.1044/1092-4388\(2002/076\)](http://doi.org/10.1044/1092-4388(2002/076))
- Peregoy, S. F., & Boyle, O. F. (2000). English learners reading English: What we know, what we need to know. *Theory into Practice*, 39(4), 237-247. [http://doi.org/10.1207/s15430421tip3904\\_7](http://doi.org/10.1207/s15430421tip3904_7)
- Pressley, M., Levin, J. R., & McDaniel, M. A. (1987). Remembering versus inferring what a word means: Mnemonic and contextual approaches. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 107-127). Hillsdale, NJ: Lawrence Erlbaum.
- Prior, A., Goldina, A., Shany, M., Geva, E., & Katzir, T. (2014). Lexical inference in L2: Predictive roles of vocabulary knowledge and reading skill beyond reading comprehension. *Reading and Writing*, 27(8), 1467-1484. <http://doi.org/10.1007/s11145-014-9501-8>
- Proctor, C. P., Silverman, R. D., Harring, J. R., & Montecillo, C. (2012). The role of vocabulary depth in predicting reading comprehension among English monolingual and Spanish-English bilingual children in elementary school. *Reading and Writing*, 25, 1635-1664. <http://doi.org/10.1007/s11145-011-9336-5>
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 64(4), 913-951. <http://doi.org/10.1111/lang.12077>
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *Modern Language Journal*, 95(1), 26-43. <http://doi.org/10.1111/j.1540-4781.2011.01146.x>

- Sénéchal, N. M. (2006). Testing the home literacy model: Parent involvement in kindergarten is differentially related to grade 4 reading comprehension, fluency, spelling, and reading for pleasure. *Scientific Studies of Reading, 10*, 59-87. [http://doi.org/10.1207/s1532799xssr1001\\_4](http://doi.org/10.1207/s1532799xssr1001_4)
- Takanishi, R., & Menestrel, S. L. (2017). *Promoting educational success of children and youth learning English: Promising futures*. Washington, DC: The National Academies Press.
- Tannenbaum, K. R., Torgesen, J. K., & Wagner, R. K. (2006). Relationships between word knowledge and reading comprehension in third-grade children. *Scientific Studies of Reading, 10*, 381-398. [http://doi.org/10.1207/s1532799xssr1004\\_3](http://doi.org/10.1207/s1532799xssr1004_3)
- Verhoeven, L. (2011). Second language reading acquisition. In M. L. Kamil, P. D. Pearson, E. B. Moje, & P. P. Afflerback (Eds.), *Handbook of reading research*, (Vol. IV, pp. 661-683). New York: Routledge.
- Walters, J. (2004). Teaching the use of context to infer meaning: A longitudinal survey of L1 and L2 vocabulary research. *Language Teaching, 37*, 243-252. <http://doi.org/10.1017/s0261444805002491>
- Westermann, G., & Mani, N. (Eds.). (2018). *Early word learning*. London: Routledge.

APPENDIX

Target words (in alphabetical order)

apparel  
aviator  
confirm  
disadvantage  
dishonesty  
durable  
envy  
famish  
getaway  
loiter  
malicious  
masterpiece  
navigate  
outgoing  
oversee  
pending  
penniless  
prolong  
trustworthy  
unicycle