

## The Effects of Input Modification on L2 Reading Comprehension and Incidental Vocabulary Learning

Weilong Huang<sup>a</sup> and Sarut Supasiraprapa<sup>b</sup>

<sup>a</sup> Graduate School of Language and Communication, National Institute of  
Development Administration, Thailand

<sup>b</sup> Graduate School of Language and Communication, National Institute of  
Development Administration, Thailand

\*Corresponding author: [sarut.sup@nida.ac.th](mailto:sarut.sup@nida.ac.th)

Article information	
<b>Abstract</b>	<p>Authentic reading texts intended for everyday communication, rather than those specifically designed for second language (L2) instruction, often provide rich, contextualized examples of natural language use that support L2 acquisition. However, such texts can be challenging for learners to comprehend because of unfamiliar vocabulary and complex sentence structures. To address this, Long (2020) identified two approaches instructors can use to enhance L2 comprehension and incidental vocabulary learning through reading: <i>elaboration</i> and <i>modified elaboration</i>. He further predicted that modified elaborated texts would be the most comprehensible and beneficial for incidental vocabulary learning, followed by elaborated texts and unmodified authentic texts, respectively. This study tested this prediction with 143 intermediate-level EFL university students in China, who were divided into three groups, each reading a different version of the same text: unmodified, elaborated, and modified elaborated. All groups completed the same comprehension test, and their incidental word learning was assessed through pre- and posttests. The results showed that both elaborated and modified elaborated texts significantly improved comprehension and vocabulary learning compared to the unmodified text, with no significant difference between the two elaboration methods. Theoretical and pedagogical implications are discussed.</p>

<b>Keywords</b>	Input modification, Elaborated input, Modified elaborated input, L2 reading comprehension, Incidental L2 vocabulary learning
<b>APA citation:</b>	Huang, W., & Supasiraprapa, S. (2025). The effects of input modification on L2 reading comprehension and incidental vocabulary learning. <i>PASAA Journal</i> , 71, 52–83.

## 1. Introduction

Authentic materials are records of real-life spoken or written texts between native speakers (and potentially non-native speakers) that are not created specifically for second language (L2) teaching purposes (Long, 2020). These materials (e.g., movies and magazine articles) offer meaningful real-world, contextualized examples of language use, making them invaluable for L2 learning (e.g., Gilmore, 2011). Contemporary L2 teaching approaches, including communicative and task-based language teaching, and Content and Language Integrated Learning (CLIL), encourage teachers to incorporate authentic L2 materials (Richards & Rodgers, 2014). These materials are also widely used in English as a medium of instruction (EMI) programs (Bi & Liu, 2024), in which students may not share the same first language. However, commercial L2 textbooks often include limited authentic materials, and CLIL or EMI instructors may need to find additional context-specific authentic reading texts (Richards & Rodgers, 2014). Therefore, instructors may often need to source their own materials to meet instructional goals and enhance student learning.

One challenge for teachers in adopting authentic materials is that these materials can be too challenging for students due to difficult vocabulary or complex grammar, often exceeding the optimal  $i+1$  level for L2 learning (Krashen, 1985). Moreover, because learners cannot interact with the original authors or speakers, they are deprived of interactionally modified input crucial for L2 development (Long, 1996). As Long (2020) argued, while L2 teachers can support learners' comprehension through classroom activities like schema building, vocabulary pre-teaching, or grammar explanation, these activities may reduce the communicative nature of L2 classrooms and fail to mirror real-life situations. More importantly, such activities may not promote incidental L2 learning, which Long (2020) defines as learning language features as a by-product of meaning-focused activities, such as reading or watching television. According to Long (2020), this type of learning

is crucial for developing learners' implicit L2 knowledge, which underlies fluent, spontaneous communication and differs from explicit knowledge.

Building on this view, Long (2020) further emphasized that comprehensible input is an important source of incidental L2 vocabulary learning because limited classroom time makes it impossible to teach all words explicitly. This perspective aligns with other scholars' emphasis on the importance of incidental word learning in successful L2 development (Nation, 2022; Teng, 2025). Empirical evidence also shows that L2 learners can acquire new words incidentally through comprehensible input, such as when unknown words occur in contexts containing familiar ones (e.g., Elgort et al., 2018; Webb, 2008). However, as Long (2020) observed, the comprehensibility of L2 texts depends not only on the familiarity of surrounding words but also on the complexity of contextual sentences.

To make authentic texts more comprehensible and supportive of incidental L2 word learning, Long (2020) identified three approaches: *simplification*, *elaboration*, and *modified elaboration*. First, simplification replaces difficult words and structures with known words and simpler structural alternatives. However, as Long (2020) and others (e.g., Yano et al., 1994) have observed, simplification is less ideal for L2 learning because it produces insufficient input made up mostly of familiar words and structures. It also often results in unnatural, choppy sentences and the omission of useful, authentic language features. Additionally, it may limit learners' ability to develop deep comprehension skills, such as making inferences. In contrast, elaboration retains challenging language but adds rephrasing, definitions, or extra details to support understanding. Finally, modified elaboration goes further by breaking up long sentences in elaborated input and adding connectors to clarify relationships between them. Long (2020) predicted that both elaborated and modified elaborated input should be more beneficial for L2 learning than overly difficult unmodified authentic input. He additionally predicted that modified elaborated input, which is easier to process, would be more comprehensible and more effective than elaborated input in promoting incidental L2 learning, including vocabulary learning.

To date, not all previous studies based on L2 reading texts have supported Long's (2020) prediction. These mixed results may have stemmed from several methodological issues, such as the lack of control of reading topic familiarity (e.g.,

Chung, 1995; Yano et al., 1994) or pretesting of target words for learning (e.g., Chung, 1995; Silva, 2000). In addition, to date, Hillman's (2020) study on listening input is the only study comparing modified elaboration with both elaboration and non-modification, but it also did not support the prediction and had methodological limitations which may have influenced the results. Because listening and reading involve distinct cognitive processes and affect incidental word learning differently (e.g., Vidal, 2011), further research on modified elaboration in reading contexts is needed.

Addressing these research gaps, the current study compared three types of reading texts—unmodified, elaborated, and modified elaborated—on L2 reading comprehension and incidental word learning. Following previous studies (e.g., Kim, 2006<sup>1</sup>), we use the term “unmodified texts” to refer to original authentic materials or those slightly adapted for research purposes without added elaboration or typographical enhancements. According to available information, our study is the first empirical test of Long's (2020) prediction on modified elaboration based on reading texts. Pedagogically, the findings can inform L2 and content instructors and material developers in designing reading texts that are more comprehensible and effective in promoting incidental L2 vocabulary learning.

## 2. Literature Review

### 2.1 Elaborated Input

Long (2020) described text elaboration as functionally similar to input modification in native–non-native interactions (i.e., foreigner talk), which occurs during communication breakdowns or meaning negotiation. The elaboration can be lexical or structural. Lexical elaboration involves providing definitions, synonyms, antonyms, paraphrases, or examples for unfamiliar words and can be either implicit or explicit. As noted by Chaudron (1985) and Kim (2006), implicit lexical elaboration does not explicitly link the target word to its elaboration, as shown in example (1), where the bolded word is elaborated without clear signals. In contrast, explicit lexical elaboration uses clear markers—such as questions (e.g., *What does...mean?*) or verbs of defining (e.g., *is*)—to indicate the connection, as in example (2).

---

<sup>1</sup> Kim's (2006) publication is based on the author's 2003 PhD dissertation. In this paper, when referring to Kim's (2006) study, we include some details of the author's research method drawn from the original dissertation (Kim, 2003).

- (1) ...*infants die in **cribs**, babies' beds, for no apparent reason* (Watanabe, 1997)<sup>2</sup>
- (2) **CrossFit** *is a type of exercise that focuses on high-intensity functional movement*.

Additionally, according to Chung (1995) and Long (2007), structural elaboration involves adding structural redundancy, maintaining structural regularity, and incorporating structural signaling. Structural redundancy means adding words or phrases to clarify meaning without adding new information, such as paraphrasing with *in other words* and using full noun phrases instead of pronouns or concrete verbs instead of *do*-substitution. Structural regularity includes using canonical word order (e.g., *I have never been to Sicily* instead of *Never have I ever been to Sicily*), matching mention order to event order (e.g., *They arrived before we went out* instead of *We went out after they arrived*), and using anaphoric reference (e.g., placing a pronoun after its antecedent, as in *Before Mike left, he checked his emails*) rather than cataphoric reference (e.g., placing a pronoun before its antecedent, as in *Before he left, Mike checked his emails*). Lastly, incorporating structural signaling involves clarifying logical relationships by adding connectors or phrases to make idea links clearer.

## 2.2 Modified Elaborated Input

While text elaboration can improve comprehensibility and support incidental L2 learning, Long (2020) noted that it also lengthens the original text, requiring more reading or listening time. Elaboration may also increase sentence complexity, as measured by readability formulas like Flesch-Kincaid, which consider word syllables and sentence length. To address this, Long (2007, 2020) proposed *modified elaboration*, which breaks long sentences in elaborated texts into shorter ones and adds necessary linking words to clarify meaning. For example, sentence (3) from Long (2020) includes bolded authentic L2 examples, while sentence (4), the modified elaborated version, splits the sentence and adds *as a result* for clearer logical connection. Long also predicted that modified elaborated input would be the most comprehensible and effective for incidental L2 learning.

---

<sup>2</sup> The source of an example sentence is indicated in parentheses following the sentence. An absence of a source indicates that we constructed the sentence for illustrative purposes.

- (3) ...he **fled the scene**, driving away fast without stopping, so she could only provide the police with **a rough description** of him, not an accurate one.
- (4) ...he **fled the scene**, driving away fast without stopping. As a result, she could only provide the police with **a rough description** of him, not an accurate one.

### 2.3 Previous Work on the Effect of Input Modification on L2 Reading Comprehension

A few studies have reported that elaborated texts improve L2 reading comprehension compared to unmodified texts (Brown, 1985<sup>3</sup>; Keshavarz & Mobarra, 2003; Kong, 2017). In these studies, unmodified texts exceeded the learners' L2 proficiency levels, as indicated by readability indexes (i.e., above their school grade level) and/or low reading comprehension scores (43–55%). Another study by Oh (2001) also reported similar results, although it is unclear whether the unmodified texts, which were from an L2 textbook, were truly authentic. Finally, King (1987) reported mixed results in a study involving students across five levels (equivalent to U.S. grades 9–13) in a Hong Kong English immersion school. The findings suggest elaboration may not enhance comprehension if the unmodified text is not sufficiently challenging; however, it was unclear whether participants across text types were matched in L2 proficiency.

Three other studies did not support Long's (2020) prediction and attributed this result to methodological issues. First, Parker and Chaudron (1987) suggested that their lexically but not structurally modified magazine article may have remained too difficult. The content recall test they used as a measure of reading comprehension also seemed more challenging than multiple-choice tests used in other studies. Second, Yano et al. (1994) noted that their elaborated texts were 50% longer than the unmodified texts. However, all participants had the same amount of reading time, thus putting those reading elaborated texts at a disadvantage. Finally, Chung (1995) reported low comprehension scores (below 50%) for both text types and attributed this result to text difficulty and the low

---

<sup>3</sup> Brown's (1985) concept of *modified interactional structure* shares features with what Parker and Chaudron (1987) later termed *elaborative modification*, as well as with Long's (2020) concept of *elaboration*.

proficiency among 9th-grade Korean students. It was also unclear whether the participants were matched in L2 proficiency.

Previous research comparing unmodified and elaborated texts also suggests that text type may affect L2 learners' ability to answer different types of comprehension questions (Oh, 2001; Yano et al., 1994). Yano et al. (1994) identified three question types: *replication*, *synthesis*, and *inference*. Replication questions require readers to locate the same information from the question in the passage. Typically, both the question and the correct answer appear in the same sentence in the text. In contrast, synthesis questions involve connecting explicitly stated details across sentences or paragraphs, such as combining facts about an animal's diet to infer that it eats various foods. Finally, inference questions require drawing conclusions from implied information. Based on Davey's (1988) classification, these include *bridging inferences* (linking information across sentences), *gist inferences* (understanding the overall meaning), and *reader-based inferences* (applying readers' prior knowledge).

Yano et al. (1994) found no significant difference in overall comprehension scores between students reading elaborated and unmodified texts. However, by question type, those reading elaborated texts performed significantly better on replication and inference questions, but not on synthesis questions. The authors suggested that elaboration helped learners focus on key details and draw inferences, while the increased text length may have increased the difficulty of information integration. In contrast, Oh (2001) found that learners performed significantly better on all three question types with elaborated texts. This discrepancy may be due to differences in text length and test demands: Yano et al.'s (1994) participants read much longer texts and answered more questions, which may have limited the benefits of elaboration for synthesis questions.

In sum, previous studies suggest elaboration enhances L2 learners' comprehension of authentic reading texts when (1) the unmodified texts are challenging, (2) both lexical and structural elaborations are included, and (3) the elaborated texts are not too difficult or long for the reading time. Elaboration may also help learners better identify explicit details and make inferences. However, most studies, except for Oh's (2001), lacked clear criteria for selecting words or sentences to be elaborated. The selections were typically based on researchers'

subjective judgments, which may have missed difficult words or sentences or, as Chung (1995) noted, led to unnecessary elaboration which increases processing load. Additionally, studies demonstrating no benefits of elaboration did not explicitly address the influence of topic familiarity or assess students' prior topic knowledge (e.g., Chung, 1995; Yano et al., 1994), which may have affected reading comprehension.

## **2.4 Previous Work on Reading Text Elaboration and Incidental L2 Word Learning**

A few studies have directly compared the effectiveness of elaborated versus unmodified reading texts on incidental L2 word learning and found mixed results. Watanabe (1997) and Kim (2006) found that target word elaboration significantly improved word learning. Both studies controlled for participants' prior target word knowledge, used between-subjects designs, and ensured comparable L2 proficiency across groups. In contrast, Silva (2000) and Chung (1995) found no significant advantage of elaboration, but did not pretest participants' prior target word knowledge. Notably, only Chung (1995) elaborated on contextual sentences, whereas the other studies (Kim, 2006; Silva, 2000; Watanabe, 1997) focused solely on target word elaboration. A study by O'Donnell (2009) suggested that combining target word elaboration with elaboration of contextual sentences may better support incidental word learning than target word elaboration alone. However, the study did not include a pretest of participants' prior knowledge of the target words. In addition, these studies did not clarify whether the unmodified texts contained existing contextual clues which aided word learning and potentially reduced the benefits of elaboration. Finally, except for in Watanabe's (1997) study, it is unclear whether reading topic familiarity, which can influence incidental L2 word learning (e.g., Pulido, 2007), was adequately controlled.

In addition, to our knowledge, Hillman's (2020) study is the only study that has tested Long's (2020) prediction that modified elaborated input should be superior to elaborated and unmodified input in promoting L2 comprehension and incidental word learning. Focusing on advanced L2 Japanese language learners' comprehension of listening input (i.e., academic talks), the study found no support for this prediction: after controlling for participants' working memory, the three input types were similarly comprehensible and yielded comparable target word learning gains. This may be because the unmodified input was not sufficiently



challenging; participants' mean comprehension scores across the three text conditions were similar (71%–74%). Additionally, Hillman's (2020) study contained some of the previously discussed methodological limitations, such as unclear criteria for structural elaboration and no pretest of target word knowledge.

To address the gaps above, the study focused on two issues: the lack of research on the effects of modified elaborated reading texts on comprehension and incidental word learning, and the methodological limitations of previous input elaboration studies. To address these gaps, the study posed the following questions:

1. What is the relative comprehensibility of unmodified, elaborated, and modified elaborated L2 texts?
2. What are the relative effects of unmodified, elaborated, and modified elaborated texts on incidental L2 vocabulary learning?

### **3. Methodology**

#### **3.1 Participants**

We adopted a pretest-posttest quasi-experimental design with 143 Chinese EFL learners from three intact second-year English-major classes at a southwestern Chinese university. Each class was randomly assigned to one of three reading groups: unmodified, elaborated, or modified elaborated texts. A between-subjects design was used to compare the groups because having participants read all three versions (a within-subjects design) would have made it difficult to isolate the effect of each text type.

Regarding their English proficiency, upon entering university, these students had achieved at least Level 4 on China's Standards of English Language Ability test, roughly equivalent to B1 on the CEFR (Peng et al., 2022). As administering a standardized English proficiency test at the time of the study was not feasible, we followed Brown and Grüter's (2022) recommendation and used Brown's (1980) cloze test scores as a proxy. This test has demonstrated concurrent validity through its correlation with a large-scale standardized English proficiency test (Brown, 1980). Although no specific criterion links scores to particular proficiency levels, the test has been widely used in empirical L2 studies (e.g., Chrabaszcz & Jiang, 2014; Kim & Rah, 2019) as a practical measure of general language proficiency for participant group assignment. As Table 1 shows, the three groups

had similar English learning backgrounds and scored an average of 51%-52% on the cloze test, indicating comparable proficiency.

**Table 1**

*Characteristics of the Participants (SD in Parentheses)*

<b>Group</b>	<b><i>N</i></b>	<b>Mean age</b>	<b>Gender (M/F)</b>	<b>Mean length of English education (years)</b>	<b>Mean cloze test scores</b>
Unmodified text	48	18.46 (0.74)	6/42	9.23 (0.59)	25.63 (5.39)
Elaborated text	50	18.66 (0.98)	7/43	9.52 (0.95)	26.54 (4.91)
Modified elaborated text	45	18.58 (0.75)	8/37	9.22 (0.56)	25.98 (4.84)

Note: The maximum cloze test score is 50.

To develop the materials for the three main reading groups, we recruited two additional intact English-major classes from the same university. They were randomly assigned to a material development group ( $N = 57$ , Mean age = 18.51,  $SD = 0.87$ ) and a pilot study group ( $N = 43$ , Mean age = 18.51,  $SD = 0.12$ ). As detailed in the following sections, the material development group assisted in selecting the reading text topic, target words for incidental vocabulary learning, and challenging sentences. The pilot study group helped determine the timing of the reading tests and activities and ensured that the materials were clear and understandable. The two groups had similar characteristics and English learning backgrounds to those of the three main reading groups. Their mean cloze test scores were 26.54 ( $SD = 5.52$ ) and 27.72 ( $SD = 5.94$ ), respectively. A one-way ANOVA of cloze test scores showed no significant proficiency differences across all five groups,  $F(4, 238) = 0.99$ ,  $p = .412$ ,  $\omega^2 < .001$ .

## **3.2 Materials**

### **3.2.1 Authentic Reading Text**

To select a suitably challenging authentic text and control for topic familiarity, we considered three articles (1,000–1,200 words each) from a 2018 issue of *The Economist* because their topics were likely unfamiliar to English majors. All texts scored at the U.S. 10th-grade level on the Flesch-Kincaid formula,

above the participants' final-year high school textbook level of 8.2 (Wu, 2023). Eight students from the material development group were asked to identify unknown words in these texts and rate text difficulty on a scale of 1–4 (1 = easiest, 4 = hardest). Average difficulty scores ranged from 3 to 3.25, with the number of unknown words ranging between 58 and 74. We selected a 994-word authentic reading text on global climate change ("The World Is Losing the War," 2018) because it had the lowest percentage (5.8%) of unknown words with a total of 58.

To ensure the three main reading groups were unfamiliar with the text topic, before collecting data from these groups, the material development and pilot study groups completed a survey rating their prior familiarity on a scale of 1 (not familiar at all) to 5 (extremely familiar). They reported low topic familiarity (mean = 1.10–1.13). As described below in the Procedure section, the three main reading groups completed the same survey after the reading comprehension test, topic sentence matching activity, and vocabulary posttest. These groups also reported low familiarity, and a Kruskal–Wallis test showed no significant differences among the five groups,  $H(4) = 0.92$ ,  $p = .94$ ,  $\eta^2 = 0$ .

### **3.2.2 Target Words for Learning and the Unmodified Baseline Reading Text**

We systematically selected unfamiliar words for incidental learning. First, the material development group took the Vocabulary Levels Test (VLT; Schmitt et al., 2001), which covers the General Service List (GSL) and Academic Word List (AWL). Approximately 91% of the group scored at least 26 out of 30 on the 2000-word frequency level, indicating mastery of these words based on Schmitt et al.'s (2001) guidelines. Therefore, words at or below this level were excluded as target words. We also excluded 3,500 words from the Chinese high school English curricula (Niu, 2022), which were likely known to the participants. Using *AntWordProfiler*, we identified words from the GSL, AWL, and the high school word list, and words not appearing in any of these lists were selected as candidate target words. We further excluded proper nouns, derivatives of listed words (e.g., *optimist* from *optimistic*), and compound words formed from listed words (e.g., *heatwave*), resulting in a final list of 75 candidate target words.

Second, to ensure the three reading groups were unfamiliar with the 75 words and unable to infer their meanings from context, we asked the material

development group to complete two activities, using the original article with the 75 words circled. First, they indicated whether they knew each circled word and, if so, provided all possible Chinese meanings. Next, they attempted to infer the meaning of each circled word from context and wrote it in Chinese. Based on the results, we selected 46 candidate target words that were unknown and not inferable for at least 80% of the group. The remaining 29 words were excluded as candidate targets but retained in the text: 21 were known by at least 80% of the group, and eight were unknown to at least 80% but could be correctly inferred by at least 20%. These eight were excluded to avoid confounding because their context alone might have promoted incidental learning independently of elaboration.

Finally, from the 46 candidate words, we selected 18 target words, listed in Table 2, which had not been lexically elaborated in the text. For ecological validity, these words varied in part of speech. We limited the number to 18 for two reasons. First, participants would encounter each target word only twice during classroom activities (once in a reading comprehension test and once in a topic sentence matching activity). Thus, expecting them to incidentally learn 46 words seemed unrealistic. Previous studies have shown that even more frequent exposure to 10 words in shorter texts yielded modest word learning gains (e.g., Hulstijn & Laufer, 2001; Webb, 2008). Next, we aimed to follow previous studies on incidental word learning that used reading texts with an unknown-to-known word ratio of about 1–2% (e.g., Chen, 2021; Hulstijn & Laufer, 2001); for a 994-word text, 18 target words would need to fall within this range. Because only 18 of the 46 candidates were selected as targets, the first author collaborated with an experienced native English-speaking EFL instructor to replace the remaining 28 non-target words with simpler synonyms or brief definitions, using words from the 2,000-word frequency level or the high school word list, which were expected to be known by the three main reading groups. The final unmodified baseline text consisted of 1,083 words. Since the text contained the 18 target words and the eight non-target words that were unknown to at least 80% of the material development group, the proportion of potentially unknown words for the three main reading groups was 26 out of 1,083, or approximately 2%. Thus, known words accounted for about 98%—a level considered optimal for L2 reading comprehension (Nation, 2006).

**Table 2***Target Words*

Part of speech	Target words
Noun	swathe, hemisphere, fumes, inertia, lobby, panjandrum, oomph
Verb	divest, entrench, underpin, countenance, cripple, avert
Adjective	freakish, ubiquitous, sweltering, sturdy, vulnerable

**3.2.3 Elaborated Text**

The same native English-speaking instructor assisted in creating both the elaborated and modified elaborated texts from the unmodified baseline. In the elaborated text, half of the 18 target words were implicitly elaborated and half were explicitly elaborated. Although previous research suggests that explicit lexical elaboration, rather than implicit lexical elaboration, is more effective than no elaboration in promoting incidental vocabulary learning (Kim, 2006; Watanabe, 1997), we included implicit lexical elaboration to avoid repetitive and unnatural phrasing from overusing explicit cues (e.g., *which means*). We then incorporated structural elaboration using three devices: structural regularity, structural signaling, and structural redundancy. After applying structural regularity and signaling, the text consisted of 57 sentences, which students in the material development group rated on a four-point difficulty scale (1 = least difficult, 4 = most difficult). Because 11 of these sentences had mean difficulty scores above 2.60, we added structural redundancy to these sentences to further support comprehension, using paraphrases (e.g., *That is to say, in China and India, the pollution from coal is so bad...*) or summaries (e.g., *To conclude, the weather is becoming more dangerous...*). The resulting elaborated text consisted of 1,518 words.

**3.2.4 Modified Elaborated Text**

As noted, modified elaboration involves breaking long sentences in elaborated texts into shorter ones and adding linking words where necessary to clarify meaning. However, Long (2007, 2020) did not specify a criterion for identifying excessively long sentences in elaborated texts, and this likely varies with learners' L2 proficiency. In this study, we based the criterion on the average sentence length from two EFL textbooks used by participants in their previous academic year. According to Song (2016), the average number of words per

sentence in the reading texts in these books was 16.22 and 15.24, respectively. We used 16.22 words per sentence, the longer of the two, as the benchmark for identifying long sentences. In the elaborated text, 43 sentences exceeded this length. We split 12 of these sentences into short ones and added appropriate connectors, resulting in a 1,539-word modified elaborated text. In consultation with the teacher who assisted in text development, we chose not to split 31 sentences to avoid overly choppy and unnatural sentences that may impede reading comprehension. For example, splitting sentence (5) into short ones, as shown in sentence (6), would result in three short sentences that carry separate emphasis and confuse readers about what is most important.

(5) *The second reason is economic and political inertia—the lack of economic and political willingness and interest, as many governments don’t want to stop using coal.*

(6) *The second reason is economic and political inertia. This refers to a lack of economic and political willingness. This is because many governments are not interested in stopping the use of coal.*

We ensured that the final three texts (unmodified, elaborated, and modified elaborated) were cohesive and sounded natural. Table 3 summarizes the linguistic features of the three texts (see excerpts of the texts in Appendix 1). The modified elaborated and elaborated versions were approximately 42% and 40% longer than the unmodified text, respectively. Consistent with Long (2020) and prior studies (e.g., Yano et al., 1994), the elaborated text was more complex than the unmodified one, based on average sentence length and Flesch-Kincaid Grade Level. Moreover, as in Hillman’s (2020) study, the modified elaborated text was less complex than the elaborated text.

**Table 3**

*Linguistic Characteristics of the Reading Texts*

	Text version		
	Unmodified	Elaborated	Modified elaborated
Number of words	1,083	1,518	1,539
Number of sentences	57	68	80
Number of words per sentence	19	22.32	19.24
Flesch-Kincaid Grade Level	11.1	12.3	11.5

### **3.2.5 Reading Comprehension Test**

We created a 15-item multiple-choice test for all three reading groups, consisting of five replication, five synthesis, and five inference questions based on Yano et al.'s (1994) classification, excluding inference questions requiring prior knowledge. The first author collaborated with the text creator and another native English-speaking EFL instructor to ensure appropriateness and create an answer key. Three experienced Chinese EFL instructors at the participants' university rated whether each item reflected its intended category (e.g., synthesis requiring integration of explicit facts across sentences), using a scale of 1 (matches), 0 (unsure), and -1 (does not match). Item-objective congruency scores ranged from 0.67 to 1, exceeding the 0.5 threshold recommended by Rovinelli and Hambleton (1977).

### **3.2.6 Topic Sentence Matching Activity Material**

Based on each of the three reading texts, we also created a topic sentence matching activity, which the three reading groups completed immediately after the reading comprehension test. The purpose was to support incidental learning of the target words, which appeared only once per text, by providing participants with an additional exposure. Since each text consisted of 12 paragraphs with the same content, we developed the same 12 topic sentences for all three texts, each expressing a paragraph's main idea. The same three EFL instructors rated whether each topic sentence matched the intended paragraph on a scale of 1 (matches), 0 (unsure), or -1 (does not match). All sentences received a mean rating of 1.

In this activity, each of the three reading groups reread the same text version they had encountered in the comprehension test. Paragraphs were numbered, and the 12 topic sentences were presented in random order. Participants were asked to identify the paragraph corresponding to each sentence. No feedback or correct answers were provided because these might influence incidental word learning. This activity was not used to assess reading comprehension, which was measured by the multiple-choice test, but was used solely to provide the three groups with an additional encounter with the target words. Consequently, each group encountered the target words in the same text twice: once during the comprehension test and once during the topic sentence matching activity.

### 3.2.7 Vocabulary Pretest and Posttest

We used a meaning-recall pretest and posttest to assess incidental learning of the 18 target words (see Appendix 2). The same 18 target words appeared in both tests but in a different order. In each test, participants provided the Chinese meaning of each English word in a blank or selected *I don't know*. Only the pretest included nine additional distractor words, which they had studied previously, to prevent their discouragement from encountering only unknown words and to reduce the chance that they could remember the target words simply from taking the pretest. The answer key was created using Chinese meanings based on the reading text context and the *Oxford Advanced Learner's English-Chinese Dictionary* (Hornby, 2009).

### 3.3 Procedure

After developing the materials with assistance from the material development group, we asked the pilot group to complete a series of activities over three weeks. In Week 1, they took Brown's (1980) cloze test. In Week 2, they completed a vocabulary pretest without being informed of a posttest. In Week 3, they were randomly assigned to read one of the three texts (16, 13, and 14 students for the unmodified, elaborated, and modified elaborated texts, respectively) and answer 15 comprehension questions within 40 minutes (with the text available). They then completed a 20-minute topic sentence matching activity, followed by an immediate vocabulary posttest and a survey on familiarity with the reading content. The survey was administered last to prevent prior exposure to the text content and target words. During Week 3, students in the pilot group were not allowed to talk to each other, consult the classroom teacher, or use electronic devices. The main study followed the same 3-week schedule, with each class randomly assigned to one of the three reading conditions. Based on the pilot, the three groups had 40, 20, and 8 minutes for the comprehension test, the topic sentence matching activity, and the vocabulary tests, respectively. The three groups were informed about reading activities, but not the vocabulary posttest. As noted earlier, after the posttest, they also completed the content familiarity survey. All participants were compensated for their participation.

### 3.4 Test Scoring

The reading comprehension test, as well as the vocabulary pretest and posttest, was independently scored by the first author and another experienced



EFL instructor who was involved in the material development. Responses were marked correct (1) or incorrect (0). The comprehension test included 15 questions (maximum score = 15), and each vocabulary test included 18 target words (maximum score = 18). In both the pilot and main studies, inter-rater agreement for the comprehension test was 100%, as scoring followed the same answer key. In the main study, inter-rater reliability for the vocabulary pre- and posttests was high ( $\kappa = 0.91$  and  $0.92$ , respectively), and scoring disagreements were resolved through discussion. The vocabulary pretest showed internal reliability (Cronbach's  $\alpha = 0.35$ ), likely due to the participants' limited prior knowledge of the target words, which resulted in low score variance (mean score =  $0.31$ ; over 90% scored 0 on each item). This outcome was expected because the target words should be unfamiliar to them. In contrast, the posttest demonstrated acceptable internal consistency (Cronbach's  $\alpha = 0.70$ ).

### 3.5 Analysis

Regarding the first research question (RQ 1), we excluded four outliers, who scored more than 2 standard deviations below their group means. Because the remaining data from 138 participants met normality and homogeneity of variance assumptions, we conducted a one-way ANOVA on comprehension scores and subsequently Bonferroni-adjusted post hoc tests. This adjustment was applied to control the overall Type I error rate (Field, 2024). We then analyzed group differences by question type. Given that the data by group violated ANOVA assumptions, we used a Kruskal-Wallis test, followed by Bonferroni-adjusted Mann-Whitney  $U$  tests for pairwise comparisons. For the second research question (RQ 2), we additionally excluded three outliers based on vocabulary posttest scores. Because the remaining data violated parametric assumptions, we used a Kruskal-Wallis test to examine the effect of text type on word learning gain scores, followed by Bonferroni-adjusted Mann-Whitney  $U$  tests for pairwise comparisons.

## 4. Results

### 4.1 RQ 1: What is the relative comprehensibility of unmodified, elaborated, and modified elaborated L2 texts?

Table 4 shows that the modified elaborated text group achieved the highest mean score of 10.78 out of 15 points, or approximately 72%, followed by the elaborated text group (9.57, or 64%) and the unmodified text group (6.88, or 46%). A one-way ANOVA revealed a significant effect of text type on overall reading

comprehension,  $F(2, 135) = 30.44$ ,  $p < .001$ , with a large effect size ( $\omega^2 = 0.30$ ), based on Kirk's (1996) interpretation guidelines. The post hoc Bonferroni-adjusted pairwise comparisons (Table 5) indicated that both the elaborated and modified elaborated texts were significantly more comprehensible than the unmodified text, with large effect sizes ( $d = 1.10$  and  $d = 1.71$ , respectively) according to Plonsky and Oswald's (2014) guidelines. However, the difference between elaborated and modified elaborated texts was not significant, with a small-to-moderate effect size ( $d = 0.47$ ).

**Table 4**

*Descriptive Statistics of Reading Comprehension Test Scores by Group*

Group	Mean	SD	Range
Unmodified text ( $N = 48$ )	6.88	2.16	3-11
Elaborated text ( $N = 49$ )	9.57	2.72	5-15
Modified elaborated text ( $N = 41$ )	10.78	2.39	6-15

Note: Total score = 15

**Table 5**

*Results of Post Hoc Pairwise Comparisons of Overall Reading Comprehension Scores by Group*

Group	Mean difference	SE	P	95% CI	d
Unmodified vs. elaborated texts	-2.70	.50	< .001**	[-3.90, -1.49]	1.01
Unmodified vs. modified elaborated texts	-3.91	.52	< .001**	[-5.16, -2.65]	1.71
Elaborated vs. modified elaborated texts	-1.21	.52	.062	[-2.46, .04]	0.47

Note: SE refers to standard error; CI denotes confidence interval; \*\* $p < .00$

Table 6 presents mean reading comprehension test scores by question type and group. Kruskal-Wallis tests revealed significant effects of text type on all question types: replication,  $H(2) = 20.04$ ,  $p < .001$ , synthesis,  $H(2) = 22.25$ ,  $p < .001$ , and inference,  $H(2) = 35.43$ ,  $p < .001$ . Post hoc pairwise comparisons using Mann-Whitney  $U$  tests with Bonferroni adjustments and effect sizes—or Wendt's (1972) rank-biserial correlation coefficients ( $r_v$ )—are reported in Table 7. For all question types, the elaborated and modified elaborated text groups significantly

outperformed the unmodified text group. Effect sizes ranged from medium to large ( $r_u = 0.41$ ), except for a small-to-medium effect size for replication questions when comparing elaborated and unmodified texts ( $r_u = 0.29$ ). No significant differences were found between the elaborated and modified elaborated text groups for any question type, and effect sizes were consistently small.

**Table 6**

*Descriptive Statistics of Reading Comprehension Scores by Group and Question Type*

Group	Question Type					
	Replication		Synthesis		Inference	
	Mean	SD	Mean	SD	Mean	SD
Unmodified text	2.10	0.88	2.27	1.11	2.50	1.29
Elaborated text	2.65	1.18	3.14	1.16	3.78	1.26
Modified elaborated text	3.17	1.12	3.46	1.12	4.15	0.96

Note: Total score for each question type is 5.

**Table 7**

*Pairwise Post Hoc Comparisons of Reading Comprehension Scores by Question Type*

Question type	Comparisons	Z	U	p	Adjusted p	$r_u$
Replication	Unmodified vs. elaborated texts	-2.54	840.00	.013*	.038*	0.29
	Unmodified vs. modified elaborated texts	-4.43	464.00	< .001**	< .001**	0.53
	Elaborated vs. modified elaborated texts	-2.10	753.50	.037*	.111	0.25
Synthesis	Unmodified vs. elaborated texts	-3.54	699.50	< .001**	.001**	0.41
	Unmodified vs. modified elaborated texts	-4.41	462.00	< .001**	< .001**	0.53

	Elaborated vs. modified elaborated texts	-1.11	872.50	.278	.835	0.13
Inference	Unmodified vs. elaborated texts	-4.41	577.50	< .001**	< .001**	0.51
	Unmodified vs. modified elaborated texts	-5.63	316.00	< .001**	< .001**	0.68
	Elaborated vs. modified elaborated texts	-1.21	862.50	.205	.615	0.14

Note: \* $p < .05$ . \*\* $p < .001$

#### 4.2 RQ 2: What are the relative effects of unmodified, elaborated, and modified elaborated texts on incidental L2 vocabulary learning?

Table 8 presents the vocabulary test pretest, posttest, and gain scores for each group. All groups had mean pretest scores near 0, indicating minimal prior knowledge of the target words. The elaborated and modified elaborated text groups showed similar mean gain scores of about 4 out of 18 words, while the unmodified text group had the lowest mean gain of 1 word. A Kruskal-Wallis test revealed a significant effect of text type on gain scores,  $H(2) = 36.45$ ,  $p < .001$ . As Table 9 shows, post hoc Mann-Whitney  $U$  tests with Bonferroni correction showed that both elaborated and modified elaborated text groups learned significantly more words than the unmodified text group, with large effect sizes ( $r_u = 0.60$  and  $0.66$ , respectively). No significant difference was found between the elaborated and modified elaborated text groups, with a small effect size ( $r_u = 0.08$ ).

**Table 8**

*Mean Vocabulary Pretest, Posttest, and Gain Scores by Group (SD in Parentheses)*

Group	Pretest	Posttest	Gain score	Gain score range
Unmodified text ( $N = 46$ )	0.35 (0.64)	1.43 (1.41)	1.09 (1.15)	0-4
Elaborated text ( $N = 48$ )	0.44 (0.77)	3.98 (2.78)	3.54 (2.48)	0-8
Modified elaborated text	0.12	4.05	3.93	0-8

( <i>N</i> = 41)	(0.40)	(2.63)	(2.60)
------------------	--------	--------	--------

Note: The maximum score in the pretest and posttest is 18.

**Table 9**

*Results from Post Hoc Pairwise Comparisons Based on Vocabulary Gain Scores*

Group comparison	<i>Z</i>	<i>U</i>	<i>P</i>	Adjusted <i>p</i>	<i>r<sub>v</sub></i>
Unmodified vs. elaborated texts	-5.09	840.00	<.001**	< .001**	0.60
Unmodified vs. modified elaborated texts	-5.34	464.00	<.001**	< .001**	0.66
Elaborated vs. modified elaborated texts	-.652	753.50	.555	1.000	0.08

## 5. Discussion

### 5.1 L2 Reading Comprehension Across Text Types

Regarding RQ 1, the mean comprehension score for the unmodified text (46%) indicates that it was appropriately challenging. In the current study, we ensured sufficient reading time for all participants through pilot testing, and our findings support Long's (2020) prediction that elaborated texts are significantly more comprehensible than unmodified authentic texts, due to added rephrasing, definitions, and details that aid understanding of unknown vocabulary and complex sentences. These results align with prior studies using between-subjects designs, participant groups with comparable L2 proficiency, appropriately challenging authentic texts, and controlling for topic familiarity (Brown, 1985; Keshavarz & Mobarra, 2003). However, the present research increased methodological rigor by using a pilot study and students' own judgments to identify difficult words and structures for elaboration rather than relying solely on the researchers' intuition. Moreover, it is the first to empirically confirm Long's (2020) prediction that modified elaborated texts are significantly more comprehensible than unmodified texts. Our results suggest that the mixed or non-significant findings in earlier reading studies may stem from the previously discussed methodological issues.

Our results across the three question types (replication, synthesis, and inference) mirrored the overall comprehension findings: both the elaborated and modified elaborated texts were significantly more comprehensible than the unmodified text. This aligns with the findings of Oh (2001) but contrasts with those

of Yano et al. (1994), who found no benefit of elaboration over unmodified texts on synthesis reading questions. As noted, the longer elaborated texts and greater number of questions in Yano et al.'s (1994) study may have contributed some cognitive load. In contrast, although our elaborated text was longer than the modified version, we used shorter texts overall, fewer comprehension questions, and ensured adequate reading time through pilot testing. This may have reduced cognitive burden, thus allowing participants to better benefit from elaboration when answering synthesis questions.

However, our findings for overall comprehension and question types do not support Long's (2020) prediction that modified elaborated texts are more comprehensible than elaborated texts. The only difference between the two texts was that the modified version split 12 long sentences and added cohesive connectors. Since the literature lacks a clear standard for excessive sentence length, we used 16.22 words—the higher average from participants' textbooks in the previous academic year—as a benchmark. The results suggest that elaboration was effective and that participants, who were actively enrolled in English classes, may have had reading skills surpassing this threshold. This may have reduced the added benefits of sentence splitting and extra connectors, leading to no significant difference in comprehensibility between the two texts.

## **5.2 Incidental L2 Vocabulary Learning Across Text Types**

In regard to RQ 2, we used a target word meaning recall test to measure incidental L2 word learning. Our results support Long's (2020) prediction that more comprehensible elaborated texts significantly enhance incidental L2 vocabulary learning compared to unmodified authentic texts. These findings align with previous studies which controlled for prior word knowledge, adopted a between-subjects design, and ensured L2 proficiency homogeneity (Kim, 2006; Watanabe, 1997). Compared to these studies, our study adopted more rigorous methods by (1) ensuring the unmodified text lacked contextual clues that could reduce the benefit of target word elaboration and (2) controlling for reading topic familiarity. Additionally, unlike most previous studies that elaborated only target words, we also elaborated difficult contextual sentences. Our results further suggest that the insignificant difference between elaborated and unmodified texts in some previous studies likely resulted from a lack of a pretest to assess prior target knowledge,

the exclusive focus on target word elaboration, and the overall difficulty of both text types.

This study is also the first to empirically support Long's (2020) prediction that modified elaborated texts are significantly more facilitative of incidental L2 word learning than elaborated texts. However, like Hillman (2020), we found no advantage of modified elaborated texts over elaborated texts. This suggests that our elaborated text, which included structural elaboration and approximately 2% potentially unknown vocabulary, was already sufficiently comprehensible for the participants, so the extra modification was not beneficial.

### **5.3 Pedagogical Implications**

As Long (2020) argued, elaborated and modified elaborated input supports incidental L2 learning, crucial for developing implicit knowledge. Our results suggest instructors can improve comprehension of authentic L2 texts by applying lexical and structural elaboration while retaining challenging vocabulary and grammar to promote incidental learning. In addition, the results of this and previous research (e.g., Parker & Chaudron, 1987) suggest that the benefits of text modification for reading comprehension and incidental word learning depend on both text difficulty and learners' English proficiency. Therefore, English instructors should know their students well or pilot modified materials with target learners to identify words or sentences that may require elaboration and to ensure the texts are appropriate for their English levels.

In the current study, the authentic reading text was elaborated manually. This may raise a question of whether instructors should use technological tools to support text elaboration and modified elaboration. According to Odo (2023), while software tools can simplify authentic texts, these tools may have limitations, such as placing words in unnatural contexts. Drawing on previous research and his own findings, Odo (2023) also noted that texts that are carefully and manually simplified are more likely to be comprehensible than texts that are authentic and unmodified, a benefit not always observed from software-simplified texts. Further research is needed to determine whether this finding extends to elaborated and modified elaborated texts. In addition, although advanced artificial intelligence tools may assist instructors in text modification, it remains important for instructors to understand and be able to apply the elaboration techniques discussed in this

study. This knowledge will help ensure that their elaborated and modified elaborated texts are both comprehensible to learners and incorporate target vocabulary and structures, rather than relying solely on AI-generated content.

The results also suggest that while elaborated and modified elaborated texts better promote incidental L2 learning than unmodified texts, overall vocabulary gains were modest. This result aligned with previous reading research (e.g., Brown et al., 2008). In this study, each target word appeared only once in the text, and participants encountered them twice—during the comprehension test and the topic sentence matching activity. Given the importance of repeated exposure for incidental word learning (e.g., Hulme et al., 2019; Supasiraprapa, 2019), instructors should provide more frequent encounters with target words. Additionally, meaning-focused activities which require learners to use target words should further enhance their learning (e.g., Laufer & Hulstijn, 2001).

It should also be noted that promoting incidental L2 learning through comprehensible input does not imply that explicit L2 instruction has no place in the classroom (see Long, 2020, for a discussion). In particular, in vocabulary learning, studies have shown that intentional learning can be more effective than incidental learning (e.g., Stratton, 2022). Thus, combining explicit and incidental approaches is likely most effective. However, given the vast scope of learning a second language, teaching all words and structures explicitly is impractical, and, as Long (2020) argued, incidental learning through comprehensible input remains essential for learners' development of implicit L2 knowledge needed for fluent communication. We also believe that in EFL contexts like China, instructors can further support incidental L2 learning by providing elaborated and modified elaborated reading texts for out-of-class study.

## 6. Limitations and Future Research

Like other studies, the current study has limitations, which in turn lead to several recommendations for future research. First, future studies could use alternative comprehension measures which do not allow guessing and explore other aspects of word knowledge. In line with Long's (2020) recommendations, future research may also investigate *enhanced incidental learning* through elaborated and modified elaborated input, which incorporates teaching techniques such as bimodal input (e.g., students reading an elaborated text while listening to



its audio). Although some studies support the benefits of bimodal input for vocabulary learning (e.g., Malone, 2018), its effects on input elaboration remain underexplored. Finally, future studies could include a delayed posttest to assess vocabulary retention over time.

## 7. Conclusion

This study supports Long's (2020) prediction that elaborated and modified elaborated texts enhance L2 comprehension and incidental vocabulary learning, providing the first empirical evidence for the effectiveness of modified elaborated input. The findings suggest that instructors in L2 task-based and communicative classrooms, as well as in CLIL or EMI contexts, can use lexical and structural elaboration to support learner comprehension and facilitate their incidental L2 vocabulary learning.

## 8. About the Authors

Weilong Huang is a Ph.D. student at the Graduate School of Language and Communication, National Institute of Development Administration, Thailand. He also serves as an associate professor at Neijiang Normal University, China, where he teaches undergraduate English courses. His research interests include English vocabulary instruction and English writing pedagogy.

Sarut Supasiraprapa is an assistant professor at the Graduate School of Language and Communication, National Institute of Development Administration, Thailand. His research focuses on second language acquisition, and he has published in leading journals such as *TESOL Quarterly*, *Bilingualism: Language and Cognition*, and *Studies in Second Language Acquisition*.

## 9. Acknowledgement

We would like to express our sincere thanks to the EFL instructors who assisted with data collection.

## 10. References

- Bi, P., & Liu, Y. (2024). Effects of EMI on learners' linguistic development. In S. M. Curle & J. Pun (Eds.), *Researching English medium instruction: Quantitative methods for students and researchers* (pp. 150–167). Cambridge University Press. <https://doi.org/10.1017/9781009425407.014>

- Brown, J. D. (1980). Relative merits of four methods for scoring cloze tests. *Modern Language Journal*, 64(3), 311–317. <https://doi.org/10.1111/j.1540-4781.1980.tb05198.x>
- Brown, J. D., & Grüter, T. (2022). The same cloze for all occasions? Using the Brown (1980) cloze test for measuring proficiency in SLA research. *International Review of Applied Linguistics in Language Teaching*, 60(3), 599–624. <https://doi.org/10.1515/iral-2019-0026>
- Brown, R. L. (1985). *A comparison of the comprehensibility of modified and unmodified ESL reading materials* (Publication No. 28238693) [Master's thesis, University of Hawai'i at Mānoa]. ProQuest Dissertations & Theses Global. <http://hdl.handle.net/10125/37660>
- Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Reading in a Foreign Language*, 20(2), 136–163. <https://doi.org/10.10125/66816>
- Chaudron, C. (1985). Intake: On models and methods for discovering learners' processing of input. *Studies in Second Language Acquisition*, 7(1), 1–14. <https://doi.org/10.1017/S027226310000512X>
- Chen, Y. (2021). Comparing incidental vocabulary learning from reading-only and reading-while-listening. *System*, 97(3), 1–13. <https://doi.org/10.1016/j.system.2020.102442>
- Chrabaszcz, A., & Jiang, N. (2014). The role of the native language in the use of the English nongeneric definite article by L2 learners: A cross-linguistic comparison. *Second Language Research*, 30(3), 351–379. <https://doi.org/10.1177/0267658313493432>
- Chung, H. J. (1995). *Effects of elaborative modification on second language reading comprehension and incidental vocabulary learning* (Publication No. 28276109) [Master's thesis, University of Hawaii at Manoa]. ProQuest Dissertations & Theses Global. <http://hdl.handle.net/10125/40765>
- Davey, B. (1988). Factors affecting the difficulty of reading comprehension items for successful and unsuccessful readers. *The Journal of Experimental Education*, 56(2), 67–76. <https://doi.org/10.1080/00220973.1988.10806468>
- Elgort, I., Brysbaert, M., Stevens, M., & Van Assche, E. (2018). Contextual word learning during reading in a second language: An eye-movement study. *Studies in Second Language Acquisition*, 40(2), 341–366. <https://doi.org/10.1017/S0272263117000109>
- Field, A. (2024). *Discovering statistics using IBM SPSS statistics* (6th ed.). Sage.

- Gilmore, A. (2011). "I prefer not text": Developing Japanese learners' communicative competence with authentic materials. *Language Learning*, 61(3), 786–819. <https://doi.org/10.1111/j.1467-9922.2011.00634.x>
- Hillman, K. K. (2020). *Effects of different types of auditory input on incidental vocabulary learning by L2 Japanese speakers* (Publication No. 28155402). [Doctoral dissertation, University of Maryland, College Park]. ProQuest Dissertations & Theses Global. <https://drum.lib.umd.edu/bitstreams/2df6041b-3025-4f92-8215-6387914f2140/download>
- Hornby, S. A. (2009). *Oxford advanced learners' English-Chinese dictionary* (7<sup>th</sup> ed., Y. Wang, C. Zhao, & X. Zou, Trans.) The Commercial Press/Oxford University Press. (Original work published 2005)
- Hulme, R. C., Barsky, D., & Rodd, J. M. (2019). Incidental learning and long - term retention of new word meanings from stories: The effect of number of exposures. *Language Learning*, 69(1), 18–43. <https://doi.org/10.1111/lang.12313>
- Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*, 51(3), 539–558. <https://doi.org/10.1111/0023-8333.00164>
- Keshavarz, H., & Mobarra, M. K. (2003). The effects of simplification and elaboration on reading comprehension of Iranian EFL learners. *Iranian Journal of Applied Linguistics*, 6(1), 101–117. <https://www.sid.ir/paper/55209/en>
- Kim, H., & Rah, Y. (2019). Constructional processing in a second language: The role of constructional knowledge in verb-construction integration. *Language Learning*, 69(4), 1022–1056. <https://doi.org/10.1111/lang.12366>
- Kim, Y. (2003). *Effects of input elaboration and enhancement on second language vocabulary acquisition through reading by Korean Learners of English* (Publication No. 3099065). [Doctoral dissertation, University of Hawaii at Manoa]. ProQuest Dissertations & Theses Global. <http://hdl.handle.net/10125/6863>
- Kim, Y. (2006). Effects of input elaboration on vocabulary acquisition through reading by Korean learners of English as a foreign language. *TESOL Quarterly*, 40(2), 341–373. <https://doi.org/10.2307/40264526>
- King, T. W. (1987). Text modifications in ESL reading comprehension. *REL C Journal*, 18(2), 31–44. <https://doi.org/10.1177/003368828701800203>

- Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement*, 56(5), 746–759.  
<https://doi.org/10.1177/0013164496056005002>
- Kong, D.-K. (2017). Effects of text elaboration on Korean reading comprehension. *The Korean Language in America*, 21(1), 53–88.  
<https://doi.org/10.5325/korelangamer.21.1.0053>
- Krashen, S. (1985). *The input hypothesis: Issues and implications*. Longman.
- Laufer, B., & Hulstijn, J. H. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1–26. <https://doi.org/10.1093/applin/22.1.1>
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). Academic Press.
- Long, M. H. (2007). *Problems in SLA*. Lawrence Erlbaum Associates.
- Long, M. H. (2020). Optimal input for language learning: Genuine, simplified, elaborated, or modified elaborated? *Language Teaching*, 53(2), 169–182.  
<https://doi.org/10.1017/S0261444819000466>
- Malone, J. (2018). Incidental vocabulary learning in SLA: Effects of frequency, aural enhancement, and working memory. *Studies in Second Language Acquisition*, 40(3), 651–675. <https://doi.org/10.1017/S0272263117000341>
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63(1), 59–82.  
<https://doi.org/10.3138/cmlr.63.1.59>
- Nation, I. S. P. (2022). *Learning vocabulary in another language* (3rd ed.). Cambridge University Press.
- Niu, S. (2022). *Gaozhong Yingyu 3500 Ci [3500 words for high school English]*. Liaoning Jiaoyu Chubanshe [Liaoning Education Press].
- Odo, D. M. (2023). The effect of automatic text simplification on L2 readers' text comprehension. *Applied Linguistics*, 44(6), 1030–1046.  
<https://doi.org/10.1093/applin/amac057>
- O'Donnell, M. E. (2009). Finding middle ground in second language reading: Pedagogic modifications that increase comprehensibility and vocabulary acquisition while preserving authentic text features. *Modern Language Journal*, 93(4), 512–533. <https://doi.org/10.1111/j.1540-4781.2009.00928.x>

- Oh, S.-Y. (2001). Two types of input modification and EFL reading comprehension: Simplification versus elaboration. *TESOL Quarterly*, 35(1), 69–96. <https://doi.org/10.2307/3587860>
- Parker, K., & Chaudron, C. (1987). The effects of linguistic simplifications and elaborative modifications on L2 comprehension. *University of Hawaii Working Papers in English as a Second Language*, 6(2), 107–133. <http://hdl.handle.net/10125/38614>
- Peng, C., Liu, J., & Cai, H. (2022). Aligning China’s standards of English language ability with the Common European Framework of Reference for Languages. *The Asia-Pacific Education Researcher*, 31(6), 667–677. <https://doi.org/10.1007/s40299-021-00617-2>
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878–912. <https://doi.org/10.1111/lang.12079>
- Pulido, D. (2007). The relationship between text comprehension and second language incidental vocabulary acquisition: A matter of topic familiarity? *Language Learning*, 57(1), 155–199. <https://doi.org/10.1111/j.1467-9922.2007.00415.x>
- Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching* (3rd ed.). Cambridge University Press.
- Rovinelli, R. J., & Hambleton, R. K. (1977). On the use of content specialists in the assessment of criterion-referenced test item validity. *Tijdschrift Voor Onderwijsresearch*, 2(2), 49–60. <https://files.eric.ed.gov/fulltext/ED121845.pdf>
- 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>
- Silva, A. D. (2000). *Text elaboration and vocabulary learning* (Publication No. 28442431) [Master’s thesis, University of Hawaii at Manoa]. ProQuest Dissertations & Theses Global. <https://www.proquest.com/openview/a6fe360bbe4edc2f1f9a4daa8a244e79/>
- Song, X. (2016). Jiyu yuliaoku de Yingyu zhuanye jiaocai cihui yanjiu [A corpus-based study on vocabulary in English major textbooks]. *Jiaoyu Pinglun [Education Review]*, 206(8), 123–126.

<https://faculty.fjut.edu.cn/songxiaozhou/en/lwcg/40045/content/8492.htm>

- Stratton, J. M. (2022). Intentional and incidental vocabulary learning: The role of historical linguistics in the second language classroom. *Modern Language Journal*, 106(4), 837–857. <https://doi.org/10.1111/modl.12805>
- Supasiraprapa, S. (2019). Frequency effects on first and second language compositional phrase comprehension and production. *Applied Psycholinguistics*, 40(4), 987–1017. <https://doi.org/10.1017/S0142716419000109>
- Teng, M. F. (2025). *Introduction to researching incidental vocabulary learning in a second language*. In M. F. Teng & B. L. Reynolds (Eds.), *Researching incidental vocabulary learning in a second language* (pp. 1–15). Routledge.
- The Economist. (2018, August 2). The world is losing the war against climate change. *The Economist*. <https://www.economist.com/leaders/2018/08/02/the-world-is-losing-the-war-against-climate-change>
- Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, 61(1), 219–258. <https://doi.org/10.1111/j.1467-9922.2010.00593.x>
- Watanabe, Y. (1997). Input, intake, and retention: Effects of increased processing on incidental learning of foreign language vocabulary. *Studies in Second Language Acquisition*, 19(3), 287–307. <https://doi.org/10.1017/S027226319700301X>
- Webb, S. (2008). The effects of context on incidental vocabulary learning. *Reading in a Foreign Language*, 20(2), 232–245. <https://doi.org/10.125/66826>
- Wendt, H. W. (1972). Dealing with a common problem in social science: A simplified rank-biserial coefficient of correlation based on the U statistic. *European Journal of Social Psychology*, 2(4), 463–465. <https://doi.org/10.1002/ejsp.2420020412>
- Wu, F. (2023). Keduxing ceshi zai gaozhong Yingyu zhong de yingyong—jiyu jin sannian (2020–2022) Gaokao (Quanguo Juan I) he xin jiaocai yupian ceshi fenxi [The application of readability tests in high school English: Analysis based on the reading comprehension texts of Gaokao (National Test I) and the national English curriculum textbooks in the past three years (2020–2022)]. *Chuxiong Shifan Xueyuan Xuebao [Journal of Chuxiong Normal*

*University*], 38(4), 154–160. <https://doi.org/10.3969/j.issn.1671-7406.2023.04.022>

Yano, Y., Long, M. H., & Ross, S. (1994). The effects of simplified and elaborated texts on foreign language reading comprehension. *Language Learning*, 44(2), 189–219. <https://doi.org/10.1111/j.1467-1770.1994.tb01100.x>

## 11. Appendices

### Appendix 1: Excerpts of the Three Reading Texts

As noted in the Materials section, the three text versions below were developed from an article in *The Economist* (2018). Target words for incidental learning are in bold. In the elaborated text, lexical elaboration is underlined with a dotted line. Structural signaling is marked with square boxes. Structural regularity is underlined with a wavy line, and structural redundancy is underlined with a solid line. In the modified elaborated text, connectors added after the splitting of long sentences are shown in italics.

#### ***Unmodified Text***

Reshaping existing ways of doing things can take years. In 2017 Britain enjoyed its first coal-free day since starting the Industrial Revolution in the 1800s. Coal generates not merely 80% of India's electricity, but also **underpins** the economies of some of its poorest states. **Panjandrums** in Delhi are not keen to **countenance** the end of coal, lest that **cripple** the banking system, which lent it too much money, and the railways, which depend on it.

#### ***Elaborated Text***

On the other hand, reshaping existing ways of doing things can take years. In 2017 Britain enjoyed its first coal-free day since starting the Industrial Revolution in the 1800s. Coal generates not merely 80% of India's electricity, but also **underpins** (i.e., provides vital support to) the economies of some of its poorest states. **Panjandrums**, or the so-called senior authorities, in Delhi are not keen to **countenance** or support the end of coal, lest the end of coal should **cripple** or seriously weaken the banking system, which lent it too much money, and the railways, which depend on it. In other words, senior authorities in India will not support moving away from coal because it might cause a financial crisis. This is because banks invested money in coal, and all the train systems still use it.

**Modified Elaborated Text**

On the other hand, reshaping existing ways of doing things can take years. In 2017 Britain enjoyed its first coal-free day since starting the Industrial Revolution in the 1800s. Coal generates not merely 80% of India's electricity. *Also*, it **underpins** (i.e., provides vital support to) the economies of some of its poorest states. **Panjandrums**, or the so-called senior authorities, in Delhi are not keen to **countenance** or support the end of coal. *Otherwise*, the end of coal should **cripple** or seriously weaken the banking system, which lent it too much money, and the railways, which depend on it. In other words, senior authorities in India will not support moving away from coal because it might cause a financial crisis. This is because banks invested money in coal, and all the train systems still use it.

**Appendix 2: Vocabulary Test**

Directions: Please provide the Chinese meaning of each word in the blank. If you don't know the meaning of a word, you can mark "I don't know" in the bracket.<sup>4</sup>

- |                 |       |                  |
|-----------------|-------|------------------|
| 1. swathe       | _____ | ( ) I don't know |
| 2. hemisphere   | _____ | ( ) I don't know |
| 3. freakish     | _____ | ( ) I don't know |
| 4. ubiquitous   | _____ | ( ) I don't know |
| 5. sweltering   | _____ | ( ) I don't know |
| 6. divest       | _____ | ( ) I don't know |
| 7. fumes        | _____ | ( ) I don't know |
| 8. inertia      | _____ | ( ) I don't know |
| 9. lobbies      | _____ | ( ) I don't know |
| 10. entrench    | _____ | ( ) I don't know |
| 11. underpin    | _____ | ( ) I don't know |
| 12. panjandrums | _____ | ( ) I don't know |
| 13. cripple     | _____ | ( ) I don't know |
| 14. countenance | _____ | ( ) I don't know |
| 15. oomph       | _____ | ( ) I don't know |
| 16. sturdy      | _____ | ( ) I don't know |
| 17. avert       | _____ | ( ) I don't know |
| 18. vulnerable  | _____ | ( ) I don't know |

<sup>4</sup> The original instructions the participants received were in Chinese, but the English translations are provided here.