Pre-Editing and Machine Translation: Developing Marine English Reading Materials

Chung-Ling Shih

This paper proposes the application of an online neural machine translation (MT) system and controlled language (CL) to develop marine English reading materials for senior high school students. Scant marine English reading materials have been issued by local publishers in Taiwan, making it challenging to integrate marine ecology and English education. The author of this paper proposes a technology-enhanced approach to self-develop marine English reading materials by adapting Chinese source texts with CL and then translating them into English by using the MT system of Google Translate. Textual analysis revealed that English MT output of CL texts, when compared with uncontrolled texts, used more words from levels one to six of the word list issued by Taiwan's College Entrance Examination Center (CEEC). The MT output of CL texts contained few complex, run-on sentences, so readability and comprehensibility were relatively high. Due to lower lexical and syntactic complexity, English MTs of CL texts received higher scores than those of uncontrolled texts when assessed by student respondents in a questionnaire-based survey. In response to a question regarding the linguistic factors that affect comprehension, 86.6% (71/82) of students reading uncontrolled texts complained about difficult words, whereas only 54.9% (45/82) of students reading CL texts encountered this problem. Regarding the syntactic factor, 48.8% (40/82) of students reading uncontrolled texts noticed English grammatical errors, compared with only 28% (23/82) of students reading CL texts. These findings highlight the effectiveness of boosting the comprehensibility of the MTs of marine English reading texts through pre-editing. This paper concludes by describing the technology-enhanced approach and calling for the use of CL and MT to produce comprehensible marine English reading texts for senior high school students.

Keywords: controlled language, machine translation, marine English reading materials, pre-editing, technology-enhanced approach

Chung-Ling Shih, Professor, Department of English, National Kaohsiung University of Science and Technology, E-mail: clshih@nkust.edu.tw

前編輯與機器翻譯：海洋英文閱讀教材之研製用途

史宗玲

　　臺灣本地出版社發行的海洋英文教材匱乏，導致很難推動海洋與英文整合之教育。基於此故，作者提議應用科技輔助方法自我研製海洋英文閱讀文本。該方法乃是使用控制性語言前編改寫海洋議題的中文文本，然後使用線上神經機器翻譯系統，譯成英文文本。本研究透過文本分析發現，如果與非控制性文本之英語機器文本相比較，控制性前編文本的英文機器譯文含有較多單字，可符合大學入學考試中心所頒布的一至六級高中詞彙表。此外，也較少使用複雜句與連寫句，所以具有較高的可讀性及可理解性。正因為控制性文本之英文機器譯文的詞彙與句構複雜程度較低，學生給予其可理解性之評分亦較非控制性文本之英語機器文本高。另一方面，針對學生閱讀機器譯文感到困難的阻礙要素，閱讀非控制性英語機器譯文時，高達 86.6%（71／82）的學生選擇困難的詞彙，而閱讀控制性英語機器譯文時，僅有 54.9%（45／82）贊同此要素。於句構方面，閱讀非控制性英語機器譯文時，48.8%（40／82）的學生選擇英文文法錯誤，但另一群組中，僅有 28%（23／82）的學生贊同此原因。由此可知，前編輯後的海洋英文機器譯本，可提升學生的可理解度，故本論文強調科技輔助教材研製的效益，也呼籲高中教師可使用控制性中文與機器翻譯為自己的學生研製可理解的海洋英文閱讀教材。

史宗玲，國立高雄科技大學應用英語系教授，E-mail: clshih@nkust.edu.tw。

Introduction

Over the past several years, the government of Taiwan has promoted the island-wide implementation of marine education, running from kindergartens, primary schools, junior high schools to senior high schools (Lee et al., 2019). To push for the enforcement of the policy, the government released "White Paper on Marine Education Policy" (The Ministry of Education, 2007), and "Marine Education Implementation Measures" (The Ministry of Education, 2008, as cited in Lee et al., 2019). In recent years, the government takes aim at making Taiwan transformed into a bilingual country by 2030. The government has increased its financial investment in bilingual education, reaching NT$2 billion (US$6.8 million) in 2021, and encouraged instructors to use English as a medium of instruction at school (Huang, 2020). Mulling how to achieve the two objectives with a killing-two-birds-with-one-stone approach at senior high schools, the author recommends the integration of marine and English education by using English reading texts for students to acquire marine knowledge. However, the goal can be a challenging mission because there is no adequate marine English reading text that is suitable for senior high school students to study.

To date, English textbooks that are used in Taiwan's senior high schools have been issued by local publishers, among whom the most famous ones are *Lungteng* 龍騰 and *Sanmin* 三民 . The *Lungteng* version has only one lesson dealing with ocean pollution, and the *Sanmin* version does not contain any lesson on marine topics. Although some English web-based texts address marine topics, the topics rarely discuss Taiwan's marine problems. To cope with the shortage of marine English reading materials for senior high school students, this paper proposes a tentative resolution by self-developing customized marine English reading texts using a technology-enhanced approach that requires the use of online neural

machine translation (MT) system and controlled language (CL) together. The approach starts with the adaptation of local marine texts using CL, followed by translating the adapted texts into English using an online MT system. The bold measure is proposed because current neural MT systems, e.g., DeepL and Google Translate, have improved their automatic translation quality, and pre-editing source texts contributes to a huge improvement in the accuracy and comprehensibility of MT outputs (Cardey et al., 2004; Lee, 1994; Lo, 2015; O'Brien, 2003; O'Brien & Roturier, 2007; Pym, 1990; Roturier, 2004).

Pre-editing with CL might be regarded as a common practice that teachers should do since the MT output of an uncontrolled source text will be more incomprehensible or contain more linguistic errors than that of a controlled text. However, not many teachers in Taiwan are familiar with or clearly aware of the CL strategies when they want to pre-edit a text. Furthermore, no research has addressed the effectiveness of MT and CL using marine texts as examples. Thus, the present research provides some clear guidance of CL strategies that can be used to pre-edit the Chinese web-based marine texts to create comprehensible and readable English MT texts. The research is also expected to offer empirical evidence that the MT outputs of controlled texts have a higher level of readability and comprehensibility than those of uncontrolled texts, and thus students can comprehend the MT outputs of controlled texts better than those of uncontrolled texts.

To conduct this experimental study, the author adapted ten marine texts with CL and translated them into English with online Google Translate. To gauge how far the English MT outputs can be understood by the senior high school students, textual analysis was conducted. All content words and sentence structures were measured to calculate how many words in the English MT outputs of controlled and uncontrolled texts met levels one to six of the word list issued by Taiwan's College Entrance Examination Center (CEEC), and how many difficult syntactic

structures fell short of the syntactic norms recommended by the editors of senior high school English textbooks. Additionally, to understand if students agree that the reading comprehensibility of the MT outputs of controlled texts is better than that of uncontrolled texts, a questionnaire-based survey was conducted. The implications derived from the findings are to be discussed in terms of the effectiveness of the technology-enhanced approach that is used to develop English marine reading materials for senior high school students.

Two objectives of the paper are: (a) to identify lexical and syntactic differences between the English MT outputs of uncontrolled and controlled texts, and (b) to investigate whether students have a better reading comprehension of the English MT outputs of controlled texts than those of uncontrolled texts. To achieve the objectives, two research questions are answered through an investigation: (a) what are lexical and syntactic differences between the English MT outputs of uncontrolled and controlled texts? and (b) what's the respondents' reception of English MT outputs of uncontrolled and controlled texts? The practical implications of the present study will be also explored.

Theoretical Review

This section introduces the basic concepts of MT, CL, CEEC's words list, the sentence structures recommended by editors of senior high schools' English textbooks, and some factors to affect student English reading comprehension.

Machine Translation and Controlled Language

The first MT system was developed by the research team of Georgetown University in the 1950s under the sponsorship of IBM (Bellis, 2019). The US government used it to translate the information of military intelligence. However, the

rosy dream was soon diminished due to an infamous report released by the Automatic Language Processing Advisory Committee (ALPAC), which claimed that the MT system could not produce high quality fully automatic translation (Shih, 2020). The example of erroneous MT output ruined everyone's high expectations for the useful MT technology. In subsequent years, MT history went through twists and turns for more than 50 years. In the late 2010s, advanced, sophisticated MT systems were launched due to the application of AI technology. Baidu Translate (China), DeepL (Germany) and Google Translate (US), to cite some, are famous and popular systems. However, they still have their Achilles heels. To overcome their weakness, pre-editing source texts using a CL becomes a feasible solution.

The early form of CL, which was called simple English, was used for immigrants to understand the contents of English official documents from 1930s to 1940s in the US. The American official documents written in simple English aimed to help new residents understand the messages easily. From the 1950s to the 1970s, technical documents written in simple English were used to help engineers understand the technical texts easily. In recent years, CL has been used to pre-edit texts and help produce comprehensible multilingual MT outputs (Kamprath et al., 1998; Lee, 2016; Lehtola et al., 1998; Pym, 1990). Much evidence is given about the improvement of MT comprehensibility due to pre-editing with CL. Mercader-Alarcón and Sánchez-Martínez (2016) pointed out that the error rate in the Spanish MT output of a pre-edited English news text was 10% lower than a non-edited text. Babych et al. (2009) claimed that pre-edited phrases led to 40% error reduction in their MT outputs from French and Russian into English. In Lo's (2015) thesis, Thai respondents' understanding of Thai MT outputs of uncontrolled Chinese texts was 42% on average, but it rose up to 86.2% about those of controlled texts.

Pre-editing using CL varies with the source language and the requirements of accuracy and comprehensibility of MT outputs. The author of the present paper

recommends adapting marine Chinese texts by consulting some CL strategies that Shih (2013) has proposed in her book. In the lexical area, the Chinese CL strategies can be (a) the replacement of rarely-used, difficult words with often-used, easy words; (b) the removal of specialized terms; (c) the reduced use of proper nouns, and (d) the paraphrasing of fixed phrases, *Chengyu*s 成語 , idiomatic and metaphorical expressions (Shih, 2013). The strategies in the syntactic area can be (e) the division of a long sentence into shorter ones with one idea within one clause, and (f) the avoidance of using unclear and run-on sentences by omitting, adding words and changing word sequence (Shih, 2013). Furthermore, the author of the present paper suggests that to shorten the length of each MT text, redundant and less important information can be eliminated or rephrased.

CEEC's Word List and Editing Norms for Senior High School's English Textbooks

To provide teaching guidelines for senior high school's English instructions, CEEC invited university professors to design an English word list with a total of 6,480 words at six levels. The 4,320 words from levels one to four target the students attending the Scholastic Aptitude English Test (SAET), and all words from levels one to six, the students attending the more difficult Department Required English Test (DRET). The 4,320 words were extracted from 1,000 Most Common Words in English, nine types of high school readers used in Taiwan, the English word list released by The Ministry of Education and some cultural words used in U.S. and Taiwan (Zheng, 2002). The 2,160 words for levels five to six are collected from *Collins Cobuild English Dictionary for Advanced Learners* (HarperCollins Publishers, 2001), Mogilner's (1992) 2,000 words used in seven types of English readers for senior high school students, frequently-used American and Taiwanese cultural items (e.g., "blues," "calligraphy") and some words relevant to teenagers' daily life

(e.g., "acne," "dandruff") (Zheng, 2002, p. vi).

Zheng (2002) summarized some rules used to design the word list. The exclusion norms include that (a) all rarely used negative adjectives, which start with "in-," "im-," "ir-," "il-," "un-," "non-" and "re-," need to be excluded unless they are frequently used, such as "independence," "infinite," "innumerable" and relevant others; (b) the words, which end with "-ment," "-ing," "-ed," and "-ly," are also excluded unless they are frequently used, such as "argument," "commitment" and relevant others; (c) the words, which have special meanings, are not excluded, such as "promising" and "learned," and (d) many proper nouns should be excluded as they would affect reading comprehension, but "English," "Mandarin," "Confucius," "Christmas" and "Bible" can be included (Zheng, 2002, p. xi). Since the nouns that end with "-ion," "-ation," "-cation," "-ition," "-ison" or "-sion" can be included, "indication," "composition," "conclusion," and "comparison" are accepted and senior high school students need to learn them (Zheng, 2002, p. xii).

In the syntactic area, current English textbooks emphasize the use of easier syntactic structures for first graders, but increase complexity for second and third graders (Zhou, 2011). The sentence patterns range from the basic English structure S+V+O (e.g., "She is my sister"), S+V+Wh-to+VR (e.g., "I don't know what to eat for dinner") and What+S+V-is+NP (e.g., "What she worries most is her mother's health") to No matter "wh-"+S+V1, S+V2 (e.g., "No matter what happens, she would face it with courage"). The *Lungteng* version of English textbooks introduces some skills of English writing and advises students not to use run-on sentences and sentence fragments. "A run-on sentence occurs when two or more independent clauses are put together without a conjunction or punctuation" (Zhou, 2011, p. 216). A sentence fragment means that "a group of words dose not express a complete thought, but is punctuated like a sentence" (Zhou, 2011, p. 194). These two types of sentences are sometimes found in the English MT outputs of non-

edited or uncontrolled Chinese texts. To reduce them, the source texts need to be adapted using CL.

Reading Materials With Good Comprehension

It is better for instructors to abide by some principles when self-developing English reading materials. According to some scholars (Littlejohn, 2012; McDonough et. al., 2013; Solak & Cakır, 2015; Tomlinson, 2012), reading and teaching materials should be focused on boosting student interest and enhancing their learning motivation. In addition, scaffolding students to comprehend the texts with pictorial aids remains important. However, in the present paper, the author only focuses on the linguistic issue, so the visual aids are not discussed.

Many factors, including the text, context and the reader, can affect student reading comprehension (Snow, 2002), but vocabulary, among many others, is a vital factor (Iqbal et al., 2015). From Laufer's (1989) viewpoint, 95% of words of the text should be familiar to learners. For Nation (2001), only when 98% of its words are understood can a text show a good reading comprehensibility. In addition to vocabulary, some scholars (Barry & Lazarte, 1995; Johnson, 1981; Yorio, 1971) called our attention to the problem of syntactic complexity, and insisted that the syntactic factor also hindered student reading comprehension. Barry and Lazarte (1995) declared that syntactic complexity would result in reading interference when the number of embedded clauses per sentence increased. To boost students' reading comprehension, the instructors should consider whether the reading content meets students' English lexical and syntactic proficiencies, and whether the topics are familiar to students when self-developing English reading materials for their students.

Methodology

This section introduces local marine texts collected for adaptation using CL and illustrates the mixed-method approach adopted to conduct the present research.

Materials

Ten web-based Chinese texts were collected for adaptation from websites, including The Ministry of Internal Affairs of the Republic of China (2011), The National Museum of Marine Biology and Aquarium (n.d.), Taiwan Fertilizer Co., Ltd. (2017), and others. The chosen texts address Taiwan's Ocean day, marine pollution, marine debris, marine ecological crisis, incorrect fish-eating and incorrect fish-catching methods, and marine culture. These topics are related to ocean environmental pollution, natural ecology and water resources (including fish) protection, which are important subjects for Taiwan's marine education. In the comprehensive senior high school's curriculum implemented in the 1999 school year in Taiwan, the government of Taiwan recommended a series of ocean-related subjects, including "biology and environment" in the basic-level Biology course, "climate and hydrology" in the Geography course, and "natural ecology," "energy and environment," "water resources and their utilization and protection" in the course of Introduction to Environmental Science (The Ministry of Education, 2007, pp. 10-11). Thus, the chosen marine topics for the present research are what Taiwanese students of senior high school should be concerned about and pay attention to.

The chosen marine texts contain many proper nouns and professional terms, so they need to be eliminated or adapted. There are also many long sentences, and need to be shortened and rewritten. Lexical and syntactic adaptation is expected to improve the accuracy and comprehensibility of the English MT outputs. Two

examples of controlled and uncontrolled passages with their English MT outputs are shown in Tables 1 and 2 respectively. The author of the present paper wrote the controlled texts by revising/pre-editing the uncontrolled texts. She has worked on pre-editing study for years and has taught pre-editing in her translation class. Since she is a native Chinese speaker, she can control pre-editing with her intuition and Chinese literacy skills. The English translations were produced by online Google Translate in 2021.

Table 1

Controlled and Uncontrolled Texts With Their English MT Outputs

Types	Contents
Uncontrolled Text	(a)「旬魚」的旬，就是當季的意思，讓人去食用季節對的食物。(b) 以螃蟹為例，雖然說一年四季都能吃得到，秋季卻是螃蟹盛產也是最肥美的時候。(c) 當消費者選擇「不旬」的海鮮，這些食品有可能是冷凍了很久、或是從別的國家運送過來的（Hsieh, 2018）。
English MT	The ten days of "Xunyu" means the season, and people eat the right food in the season. Take crabs as an example. Although they can be eaten all year round, autumn is the time when crabs are abundant and plump. When consumers choose "non-decade" seafood, these foods may have been frozen for a long time or shipped from other countries.
Controlled Text	(a) 我們應該吃當季的食物。(b) 秋天螃蟹很多，所以我們應該在秋天的時候吃螃蟹。(c1) 過季的食物通常是冷凍的食物。(c2) 這些食物從別的國家運送過來時，它們可能是冷凍了很久（adapted by the author）。
English MT2	We should eat seasonal food. There are many crabs in autumn, so we should eat crabs in autumn. Out-of-season food is usually frozen food. When these foods are shipped from other countries, they may have been frozen for a long time.

In Table 1, in accordance with Shih's (2013) Chinese CL strategy (6), sentences (a) and (b) in the uncontrolled text were paraphrased by omitting, adding words and changing word sequence. Sentence (c) in the uncontrolled text was divided into two shorter ones (c1) and (c2) in the controlled text by using Shih's (2013) CL strategy (5). *Xun* 旬 (in season) and *bu xun* 不旬 (off season), were adapted as *dangji de*

當季的 (seasonal) and *guoji de* 過季的 (non-seasonable) by conforming to Shih's (2013) CL strategy (1) of using common words to replace rarely-used words. After the two archaic words were replaced, they were translated accurately as "seasonal" and "out-of-season," not as "ten days" and "non-decade" in the MT output.

Table 2

Controlled and Uncontrolled Texts With Their English MT Outputs

Types	Contents
Uncontrolled Text	(a) 乘著洋流來去的海洋廢棄物，在人類驚覺嚴重前，已快速地全球化。(b) 海洋廢棄物不分國界，對生物的殺傷力同樣不分領土。(c) 人類垃圾是現今海洋生物生存最大的敵人。(d) 數不清因誤食、纏繞而死的生命，正在海底看不見的幽暗角落上演（Chen & Zheng, 2017）。
English MT	The marine debris that comes and goes by ocean currents has rapidly globalized before mankind becomes seriously alarmed. Marine debris knows no borders, and its lethality to organisms also does not distinguish territories. Human trash is the biggest enemy of marine life today. Countless lives that died due to ingestion and entanglement are unfolding in an invisible dark corner of the ocean floor.
Controlled Text	(a) 在人類瞭解嚴重情形之前，海洋垃圾搭乘著洋流漂流至全世界。(b) 海洋垃圾不分國界，對所有生物造成的傷害是一樣的。(c) 人類垃圾是現今海洋生物的最大敵人。(d) 在海洋底層，許多海洋動物吞食垃圾袋，因此就生病死亡（adapted by the author）。
English MT2	Before mankind understood the serious situation, marine trash drifted to the world on ocean currents. The ocean trash knows no borders and causes the same damage to all living things. Human trash is the greatest enemy of marine life today. At the bottom of the ocean, many marine animals swallow garbage bags and become sick and die.

In table 2, sentence (a) in the uncontrolled text was revised by adopting Shih's (2013) CL strategy (6), so it changed its original word sequence. Meanwhile, the metaphorical expression *kuaisu de quanqiuhua* 快速地全球化 (rapidly globalized) was adapted as *piaoliu zhi quanshijie* 漂流至全世界 (drift to the world) following Shih's (2013) CL strategy (4), so the senior-high school students can have a better understanding of the message. In sentence (b), the specialized term *haiyang feiqiwu*

海洋廢棄物 (marine debris) was omitted and revised as *haiyang lese* 海洋垃圾 (marine trash) in accordance with Shih's CL strategy (2). Abiding by Shih's (2013) CL strategy (1), *shashangli* 殺傷力 (harmful force) was adapted as a common word, *shanghai* 傷害 (harm). In sentence (d), *wushi* 誤食 (eat by mistake) and *chanrao er si* 纏繞而死 (entangled to death) were adapted as *tunshi lesedai* 吞食垃圾袋 (swallow trash bags) and *shengbing siwang* 生病死亡 (get sick and die) by adopting Shih's (2013) CL strategy (1)–to replace rarely-used, difficult words with often-used, easy words. The long metaphorical expression *zhengzai haidi kanbujian de youan jiaoluo shangyan* 正在海底看不見的幽暗角落上演 (performing in the dark, invisible corner at the bottom of the ocean) and the short metaphor *shubuqing de shengming* 數不清的生命 (countless lives) were paraphrased as *zai haiyang diceng* 在海洋底層 (at the bottom of the ocean) and *xuduo haiyang dongwu* 許多海洋動物 (many ocean animals), adopting Shih's (2013) CL strategy (4). After the revision, the MT output of the controlled text can be understood by senior-high school students more easily and more clearly.

A Mix-Methods Approach

A mix-methods approach is used to conduct the research, including (a) textual analysis to identify lexical and syntactic differences between two types of English MT outputs, (b) a questionnaire-based survey, and (c) a probe into research implications.

Analysis of Lexical and Syntactic Complexity Across Two Types of Machine Translation Outputs

To distinguish lexical and syntactic complexity of MT outputs of uncontrolled and controlled texts, content words and sentence structures were analyzed. In the lexical area, all content words were measured based on three groups of words (A),

(B) and (C). Group (A) means the words within CEEC's levels one to four; group (B), the words within levels five to six, and group (C), the words beyond the sixth level. The words, e.g., "distinguish," "submarine," and "individuals," belong to group (A); "organism," "unfolding," and "transparent," to site some examples, belong to group (B), and "picky," "debris," and "entanglement" belong to group (C). If the MT output contains many group (C) words, it must increase students' interpretation burden.

In the syntactic area, complex sentence structures and run-on, fragmentary sentences in the MT outputs were calculated. A complex sentence, for example, is composed of more than two relative-pronoun-led clauses, more than two conjunctions, more than three prepositional phrases, more than three noun phrases, or more than two present-participle-led clauses. One MT example is:

> Perhaps you <u>who are reading the paper</u> are not like me <u>who eat seafood every day</u>, but from <u>the seafood buffet in restaurants</u>, seafood shops on <u>the coast of fishing ports</u>, <u>seafood stalls in streets and lanes, and seafood snacks at night markets</u>, seafood is the main focus.

The MT is rendered from a sentence:

> 也許正閱讀文章的您也許並不像我每天都會食用海鮮，但從飯店的海鮮吃到飽 Buffet、漁港海岸的海產店、街道巷弄的海產攤和夜市的海鮮小吃，從各類以海鮮為主打的飲食。（Bai, 2017）

Additionally, run-on sentences are identified when a noun or a noun segment is inserted in a sentence without using a conjunction or a relative pronoun to link its preceding clause. One MT example is "About one-third of the plastic produced is disposable, <u>single-use plastic products, which usually turn into waste within a year,</u> and there is a high chance of wandering to the beach." This is the MT output of a sentence: "這些被生產出的塑膠約有 3 分之 1 是拋棄式、單次使用的塑膠產品，通常一年內就會變成廢棄物，並有很高的機會流浪到海邊" (Chen & Zheng, 2017). The current MT output also produces a single noun as an independent

sentence that is viewed as a run-on sentence. One MT example is "Because of this, he captured images and images of many creatures threatening their lives due to marine debris. photo." The underlined word "photo" should not be an independent sentence and needs to be removed. The MT is rendered from a sentence: "也因為這樣，他捕捉到許多生物因為海洋廢棄物而遭受生命威脅的影像與照片" (Chen & Zheng, 2017).

The Questionnaire-Based Survey

The Respondents. The current research proposes the use of comprehensible English MT outputs of controlled marine texts as reading materials for senior high school students, so the experiences of student users are crucial for consultation. If student users are not satisfied with MT texts/products, the proposal in the present paper would be ineffective and infeasible. Thus, a questionnaire-based survey was conducted to understand whether many student users comprehended the MT outputs of controlled texts better than those of uncontrolled texts. Only after many student users accept the comprehensible and readable English MT outputs of controlled marine texts can we further investigate whether many high school teachers agree to use CL and the MT system to develop their English marine texts.

A total of 82 third-grade students (42 males and 40 females) from a private senior high school in Taichung, Taiwan, were invited to participate in the questionnaire. To enhance the validity of findings, students from the classes of natural science and social science were mixed. They were randomly invited, not limited to very high or very low English achievers. Seventy-seven students (97.5%) had the experience of using MT systems, and all of them were willing to help complete the survey. They were aware of the purpose of the questionnaire and signed the letter of consent. Noticeably, high-school students are still underaged, so the author asked their English instructor and their school for the approval of

administering the questionnaire. After the questionnaire, each student participant and the instructor got a small bag as a reward.

The questionnaire copies were delivered to students in classes. All participants were asked to read the English MT outputs translated from four uncontrolled and four controlled texts. After reading the eight texts, the participants were asked to (a) rate MT outputs based on their overall comprehension, and (b) choose the possible factors to cause their difficult reading comprehension. To ensure grading consistency among the respondents, the English instructor on the site gave some instructions on the method of scoring the MT texts. The respondents from two classes were given the same length of time, namely 50 minutes, to complete the questionnaire.

Questionnaire Design. The questionnaire written in Chinese consists of four parts. Part 1 asks participants to provide demographic information, including their gender and experiences of using MT systems. Part 2 asks them to grade the English MT outputs of uncontrolled and controlled texts based on their overall comprehensibility. The entire MT text–not individual sentences–is used as the unit of measurement. If the MT outputs can be used for students to acquire marine knowledge and learn English at the same time, they must be readable and comprehensible. This means that students must be able to comprehend the messages easily and clearly, so they can acquire marine information. To measure whether the MT outputs can be used as reading materials for students to learn marine knowledge and English, students were asked to assess the overall MT readability based on their comprehension of MT outputs.

The students are informed that the measurement uses a 1-100 scale. A score of 90-100 means that the MT output can be easily and clearly understood. A score of 80-89 indicates that the general message of the MT output can be understood, but it contains some difficult words and/or sentences. A moderate score of 70-79 means

that the general ideas of the MT output are not easily grasped because of serious lexical and syntactic errors. The less-than-70 score means that the MT output cannot be understood at all.

Part 3 has six questions to ask respondents how six linguistic factors have affected their reading comprehension of MT outputs of uncontrolled and controlled texts. The six factors include (a) some difficult words, (b) some words with unclear meanings, (c) some grammatical errors, (d) complex sentence structures, (e) incomplete sentences, and (f) odd, awkward expressions. Since reading comprehension is the process of simultaneously constructing and extracting meanings from words, sentences (including syntax) and working memory (Tánczikné, 2017), questionnaire design with a focus on the areas of words and sentences can help us to know whether the respondents have difficulty of comprehending the marine English texts. Words with unclear meanings and awkward expressions cannot be understood easily and clearly, and thus affect the reader's understanding of a text. The unclear meanings are also difficult to be guessed indirectly from the context. Meanwhile, grammar and sentence structure play a key role since readers often use these two linguistic features to decode the meaning of a sentence. If a sentence is incomplete and lacks a complete thought, it certainly hinders the reader's understanding of a text. Due to their close correlations, semantically ambiguous words, awkward expressions, incomplete sentences, and relevant others are chosen as the benchmarks to measure the student respondents' comprehension of marine English texts translated from the controlled and uncontrolled texts.

In the questionnaire design, the MTs of four reading passages that are extracted from four chosen texts address ocean pollution, beach clean-up, hand fishing and *Mazu* 媽祖 (sea goddess). The author chose the four topics due to high-school students' higher familiarity with them than other topics, such as world ocean

day, *Yamei* 雅美 flying fish festival and slow food. The MTs of the controlled four passages are averagely shorter than those of uncontrolled passages by ten words. No huge difference in length between the two types of MTs is expected to avoid the respondents' prejudicial favor with shorter MT texts. Respondents were asked to read all MTs of the four uncontrolled passages on four topics, and then read the MTs of the four controlled passages on the same topics. A time lag between the same topics for reading is expected to reduce the repetition (familiarity) effect that would cause respondents to favor the second MT texts on the same topics.

However, student subjectivity and attitudes toward the questionnaire might affect the reliability of the findings though their English instructor asked them to answer the questions of the questionnaire carefully. To ensure that the respondents answered the questions with a careful manner, their English instructor informed them of the importance of the questionnaire. Students knew that the survey was conducted to get the findings to determine if marine reading materials could be developed using a technology-aided method in the future. Their cooperation to do the questionnaire was helpful and crucial. Additionally, the author's research assistant helped monitor the student respondents doing the questionnaire. According to his observation, the respondents worked hard on the questionnaire. None of them finished the questionnaire less than 40 minutes; most of them took about 50 minutes. The questionnaire content was presented in Chinese except the English MT texts, so none would misunderstand what the questionnaire asked them to do.

Rating is based on a four-level scale, composed of "strongly agree" (four points), "agree" (three points), "disagree" (two points) and "strongly disagree" (one point). "Neither agree nor disagree" is not included because it is easily chosen by the irresponsible respondents who avoid making a clear decision. This is also one way the questionnaire design has used to call on student respondents to answer the questions properly.

Findings and Discussion

This section reports the findings in two areas: (a) lexical and syntactic variation between the English MT outputs of uncontrolled and controlled texts, and (b) respondents' assessment and reading comprehension. Theoretical and practical implications of the findings would be also discussed.

Lexical and Syntactic Variation Between Two Types of English Machine Translation Outputs

After the English MT outputs were analyzed, the findings showed that the MTs of ten uncontrolled texts contained a total of 148 words within CEEC's levels one to four, 263 words within levels five to six, and 264 words beyond the sixth level. As opposed to it, the English MT outputs of ten controlled texts have 373 words within CEEC's levels one to four, 84 words within levels five to six, and 44 words beyond the sixth level. The number of difficult words within levels five to six drops from 263 to 84 in the MTs of controlled texts, suggesting that adapting the source texts using CL has reduced the lexical difficulty of the English MT outputs. Table 3 shows the amounts of three groups of words in the English MT outputs of uncontrolled and controlled texts. Group (A) means CEEC's levels one to four; group (B), levels five to six; group (C), beyond the sixth level.

Table 3

Lexical Variation Between English MTs of Uncontrolled and Controlled Texts

Text Titles	Source Texts	Source Text Wds	MT Wds	(A) Wds	(B) Wds	(C) Wds
(1) Uncontrollable marine debris	Uncontrolled	3,356	2,560	13	46	51
	Controlled	632	355	37	7	2
(2) Marine resources cannot be exhausted in our generation	Uncontrolled	2,963	1,916	2	44	32
	Controlled	680	411	38	6	1
(3) Problems of the coastal environment	Uncontrolled	3,232	1,926	5	50	70
	Controlled	720	406	36	14	3
(4) World Ocean Day: Nine ocean crises	Uncontrolled	2,366	1,410	14	56	59
	Controlled	695	437	22	17	16
(5) You cast, he picked. The plastic bottles gained from beach clean-up are higher than Jade Mountain	Uncontrolled	1,483	1,025	40	37	22
	Controlled	623	428	62	7	1
(6) Sustainable management of slow food & slow fish	Uncontrolled	2,100	1,404	33	44	41
	Controlled	655	425	24	11	3
(7) How can we have seafood culture without the solid education of eating fish?	Uncontrolled	1,837	1,186	7	29	27
	Controlled	586	355	16	7	3
(8) Why can Mazu become the biggest sea goddess in China?	Uncontrolled	2,396	1,761	16	36	35
	Controlled	565	415	21	5	5
(9) Yamei (Dawu) flying fish festival	Uncontrolled	2,210	1,566	10	40	43
	Controlled	630	444	56	4	3
(10) What is Taiwan's marine culture?	Uncontrolled	1,884	1,197	8	23	30
	Controlled	640	376	61	6	7
Total	Uncontrolled	23,827	15,951	148	405	410
	Controlled	6,426	4,052	373	84	44

The statistical results indicate that words from group (B) and words from group (C) together take up 5.1% (815/15951) in the MT outputs of uncontrolled texts, but they account for only 3.1% (128/4052) in the MT outputs of controlled texts. The finding suggests that the MT outputs of uncontrolled texts must be difficult for student reading because it does not meet Laufer's (1989) claim that 95% of words in a text need to be familiar to learners, and so the text can be easily understood. In contrast, the English MT outputs of controlled texts comply with Laufer's (1989) requirement, so they could be more easily understood. The advantage of lower lexical complexity to enhance student reading comprehension has been supported by the scholars (Hsu, 2004; Jiang, 2003; Lin, 2008; Lu, 2005).

The syntactic aspect also shows a huge difference. The average sentence length of English MT outputs of ten uncontrolled texts is 22.1 words, but it reduces to 16.6 words in the MTs of ten controlled texts. It is calculated by dividing the total word count of each English MT text by the total count of periods, exclamation marks, and question marks, which are shown in each text. On the other hand, there are 28 complex sentences in the MTs of ten uncontrolled texts, but they drop to one in the MTs of ten controlled texts. The total number of run-on sentences falls from 26 to 0 in the MTs of ten controlled texts. Table 4 shows syntactic variation between the English MT outputs of controlled and uncontrolled texts.

MT outputs of controlled texts contain only one complex sentence and zero run-on sentences. The finding suggests that English MT outputs of controlled texts have reduced their syntactic complexity after their source texts are adapted using CL. The strategies of CL include the avoidance of using run-on sentences and sentence fragments. Reduction of syntactic complexity in the English MT outputs of controlled texts helps students understand the content easily, supported by the research of Barry and Lazarte (1995), Johnson (1981) and Yorio (1971).

Table 4

Syntactic Variation Between Two Types of English MTs

	English MTs of	Average Sentence Length (Words)	Complex Sentences	Run-on Sentences
Text 1	Uncontrolled texts	20.8　(2560/123)	5	5
	Controlled texts	15.4　(355/ 23)	0	0
Text 2	Uncontrolled texts	20.8　(1916/ 92)	2	2
	Controlled texts	14.2　(411/ 29)	0	0
Text 3	Uncontrolled texts	17.8　(1296/ 73)	7	6
	Controlled texts	19.3　(406/ 21)	1	0
Text 4	Uncontrolled texts	21.3　(1410/ 66)	2	1
	Controlled texts	16.2　(437/ 27)	0	0
Text 5	Uncontrolled texts	20.5　(1025/ 50)	0	1
	Controlled texts	15.3　(428/ 28)	0	0
Text 6	Uncontrolled texts	21.2　(1404/ 66)	4	2
	Controlled texts	19.3　(425/ 22)	0	0
Text 7	Uncontrolled texts	19.1　(1186/ 62)	3	3
	Controlled texts	15.4　(355/ 23)	0	0
Text 8	Uncontrolled texts	22.9　(1761/ 77)	0	3
	Controlled texts	17.3　(415/ 24)	0	0
Text 9	Uncontrolled texts	23.7　(1566/ 66)	2	1
	Controlled texts	19.3　(444/ 23)	0	0
Text 10	Uncontrolled texts	25.5　(1197/ 47)	3	2
	Controlled texts	15.7　(376/ 24)	0	0
Total	Uncontrolled texts	22.1 (15951/722)	28	26
	Controlled texts	16.6　(4052/244)	1	0

MT outputs of controlled texts contain only one complex sentence and zero run-on sentences. The finding suggests that English MT outputs of controlled texts have reduced their syntactic complexity after their source texts are adapted using CL. The strategies of CL include the avoidance of using run-on sentences and sentence fragments. Reduction of syntactic complexity in the English MT outputs of controlled texts helps students understand the content easily, supported by the research of Barry and Lazarte (1995), Johnson (1981) and Yorio (1971).

Respondents' Reception

The Overall Comprehensibility of Machine Translation Outputs

The findings elicited from Part 2 of the questionnaire showed that the respondents grading 90-100 scores took up 10% (8/82) about the English MTs of four uncontrolled texts, but they accounted for 45% (37/82) about MTs of controlled texts. Only 4% (3/82) of students assessed the MT outputs of controlled texts as less than 70 scores but 5% (4/82) of students did so about the MT outputs of uncontrolled texts. Table 5 shows the respondents' assessments.

Table 5

Assessments on the Comprehensibility of Two Types of English MT Outputs

Scores	MTs of Uncontrolled Texts	MTs of Controlled Texts
90-100	10% (8/82)	45% (37/82)
80- 89	50% (41/82)	38% (31/82)
70-79	35% (29/82)	13% (11/82)
< 70	5% (4/82)	4% (3/82)

According to the assessment results, four students comprehended less than 70% of the uncontrolled texts, but one of them changed his mind and chose 70-79

score when reading the MT outputs of controlled texts. There was a huge gap of 35% between students who comprehended more than 90% of MTs of controlled texts and those who comprehended those of uncontrolled texts. Generally viewed, no students whose assessment scores decreased when reading the MT outputs of controlled texts.

Overall, many student respondents agree that MT outputs of controlled texts have increased their comprehensibility, so the advantage of pre-editing using CL is justified. The respondent assessment conforms to some research on the improvement in the overall comprehensibility of MT outputs through pre-editing (Babych et al., 2009; Lo, 2015; Mercader-Alarcón & Sánchez-Martínez, 2016).

Factors to Cause Difficult Reading Comprehension

In their reactions to the factors that cause their difficult reading comprehension, all respondents placed the highest focus on some difficult words. When "strongly agree" and "agree" were conflated, the findings showed that 86.6% (71/82) of respondents reading the MTs of four uncontrolled texts agreed to the factor of difficult words, but only 54.9% (45/82) of the respondents reading the MTs of four controlled texts did so. Furthermore, respondents reading the MTs of uncontrolled texts were more homogenous about the factor as their responses showed a lower SD (0.64) than that (0.75) from those reading the MTs of controlled texts. The second high factor is some words with unclear meanings. More respondents (69.5%; 57/82) reading the MTs of uncontrolled texts agreed to the factor with a higher Mean (2.8) and a lower SD (0.71) than those respondents (42.7%; 35/82) reading the MTs of controlled texts with a lower Mean (2.4) and a higher SD (0.73). Still, more respondents (31.7%; 26/82) reading the MTs of uncontrolled texts agreed to the factor of odd, awkward expressions with a higher Mean (2.27) than those (18%; 15/82) reading the MTs of controlled texts with a lower Mean (2.07).

Another consensus was reached about the syntactic factors. More respondents (48.8%; 40/82) reading MTs of uncontrolled texts agreed to the factor of grammatical errors with a higher Mean (2.51) than those respondents (28%; 23/82) reading the MTs of controlled texts. However, the former showed a higher SD (0.67) than the latter (0.6), suggesting that more respondents reading the MTs of controlled texts were homogenous about grammatical errors than those reading MTs of uncontrolled texts. In a similar manner, more respondents reading the MTs of uncontrolled texts agreed to the factors of complex sentence structures and incomplete sentences than those reading the MTs of controlled texts. The former's responses showed the higher Means (2.36, 2.23) and the higher SDs (0.68, 0.59) than the Means (2.1, 2.05) and the SDs (0.65, 0.57) from the latter's responses. The findings denote that although more respondents reading the MTs of uncontrolled texts agreed to the difficult syntactic factors, they stayed less homogenous than those reading the MTs of controlled texts. Table 6 shows the respondents' reactions to the six factors that cause their difficult reading comprehension.

Table 6

Assessment of the Factors to Cause Difficult Reading Comprehension

Items	MTs	*M*	*SD*	Percentage %			
				SA	A	D	SD
(1) Some difficult words	Uncontrolled	3.19	0.64	31	56	12	1
	Controlled	2.62	0.75	11	44	40	5
(2) Some words with unclear meanings	Uncontrolled	2.80	0.71	14	55	27	4
	Controlled	2.40	0.73	6	36	50	8
(3) Some English grammatical errors	Uncontrolled	2.51	0.67	6	42	48	4
	Controlled	2.21	0.60	2	26	64	8

(continued)

Table 6

Assessment of the Factors to Cause Difficult Reading Comprehension (continued)

Items	MTs	*M*	*SD*	Percentage %			
				SA	A	D	SD
(4) Complex sentence structures	Uncontrolled	2.36	0.68	7	27	61	5
	Controlled	2.10	0.65	4	13	70	13
(5) Incomplete sentences	Uncontrolled	2.23	0.59	2	26	65	7
	Controlled	2.05	0.57	1	15	72	12
(6) Odd, awkward expressions	Uncontrolled	2.27	0.68	5	26	60	9
	Controlled	2.07	0.60	2	16	69	13

SA = strongly agree; A = agree; D = disagree; SD = strongly disagree

The variation in the responses to lexical factors is due in part to more words of CEEC's levels five to six and beyond the sixth level in the MTs of uncontrolled texts than in MTs of controlled texts. The more difficult words a text contains, the more difficult it is understood. The finding can be supported by Yorio (1971) who argues that second language learners, either native or international students, consider vocabulary their biggest obstacle to reading comprehension. Iqbal et al. (2015) also claim that the vocabulary, including difficult words and unclear-meaning words, may cause the learner's difficult reading comprehension.

The student reactions to the syntactic factors can be attributed to the greater amount of complex, and run-on sentences in the MTs of uncontrolled texts than in the MTs of controlled texts. As aforementioned, MTs of uncontrolled and controlled texts show a gap of 28:1 in complex sentences, and a gap of 26:0 in run-on sentences. More complex, run-on sentences must cause student difficult understanding of MTs of uncontrolled texts. The finding can be supported by Barry and Lazarte (1995) who maintain that syntactic complexity would cause reading interference and reduce the learner's reading comprehension.

Research Implications

The findings yielded from the questionnaire-based survey shed some light on the practical functions of CL use, an investigation of the linguistic factors to cause student difficult reading and the application of the technology-enhanced approach as follows.

The Effectiveness of Adapting Source Texts Using Controlled Language

The results of the questionnaire-based survey showed that the respondents graded the MT outputs of controlled texts with higher scores than those of uncontrolled texts. This finding justifies that pre-editing with CL can raise the comprehensibility of MT outputs and make respondents understand them more easily. The finding can be supported by previous researchers who claim that adapting source texts using CL contributes to the production of more accurate and more comprehensible MT outputs (Babych et. al. 2009; Lo, 2015). Without being pre-edited, the MT outputs of uncontrolled texts show more lexical and grammatical errors, and become less accepted by students.

Erroneous translations of uncontrolled texts add difficulty to student reading comprehension. For example, the professional terms of fishing methods, *dian yu* 電魚 (electrocute fish), *du yu* 毒魚 (poison fish) and *zha yu* 炸魚 (blast fish), are literally translated by Google Translate as "electric fish," "poisoned fish" and "fried fish." The proper noun, *jiukong* 九孔 (baby abalone), is translated by Google Translate as "nine holes." The correct translations should be "electroshock fishing," "poison fishing," "blast fishing" and "baby abalone." The example supports that an uncontrolled text runs a high risk of producing mistranslation from the MT system. Only adapting source texts using CL helps boost the readability and comprehensibility of the English MT output.

The Linguistic Factors to Cause Student's Difficult Reading Comprehension

The findings of a questionnaire-based survey denote that all respondents agree that their reading comprehension has been affected by the six linguistic factors. Among the three lexical factors, the respondents reading either MTs of controlled or controlled texts have identified difficult words as the most important factor. In the syntactic area, grammatical errors receive more attention than incomplete and run-on sentences. The findings remind us that lexical difficulty and grammatical errors play the top two roles in affecting student reading comprehension. Thus, customizing vocabulary and reducing grammatical errors are necessary if we want to develop our own reading texts. The understanding of the respondent reaction to the linguistic factors enables us to know what should be more emphasized when we adapt source texts and then use the neural MT system to produce accurate and comprehensible translations.

The Application of the Technology-Enhanced Approach

The findings through textual analysis have shown that after source texts are adapted using CL, their English MT outputs contain more words that meet CEEC's levels one to six and use some sentences that comply with the writing norms of current English textbooks-avoidance of complex and run-on sentences (Zhou, 2011). The finding underscores the effectiveness of the technology-enhanced approach, which calls for the application of MT and CL together by tailoring reading materials to student expectations. All instructors who want to develop marine English reading materials can try this approach. However, to what extent the source texts need to be controlled depend on the students' English proficiency level. Seeking relevance to student marine background knowledge also helps enhance student interest in learning, which can be viewed as one vital variable that is worth our consideration in designing reading materials for students (Littlejohn, 2012; McDonough et al., 2013; Solak & Cakır, 2015; Tomlinson, 2012).

Conclusion

Never have CL and MT been used to develop English reading materials for senior high school students. This experimental research proposes the technology-enhanced approach used to help produce marine English reading texts for the implementation of integrated marine and English education at senior high schools. The findings have indicated that the English MT outputs of controlled texts are more comprehensible due to lower lexical and syntactic complexity, and respondents have a higher level of satisfaction with MTs of controlled texts than MTs of uncontrolled texts. Customizing marine source texts with limited vocabulary size and simpler sentence structures helps produce more comprehensible English MT outputs. The findings reinforce the CL benefit for the application of MT systems, and justify the effectiveness of boosting student reading comprehension using the texts with lower lexical and syntactic complexity.

There are limitations about the present paper. The sample size of respondents is small, so the effectiveness of the proposed approach might not be comprehensively gauged. It raises a concern as to whether different groups of students and different MT texts in the questionnaire would change the results. Thus, the future study needs to involve more students from public and private senior high schools and use more MT texts on diverse topics. Additionally, the student respondents might provide subjective reaction toward the reading materials, so the future study can administer an interview with students, asking them how specific words and some sentences in the MT text help them comprehend the text easily and clearly. This method enables us to get more insightful and genuine findings.

Despite some limitations, the results yielded from the present findings have provided us some insights into the affordances of MT and CL to help produce comprehensible marine English reading materials for senior high school students.

To facilitate the future integration of marine and English education, the English instructors can start with self-developing marine English reading materials with the technology-enhanced approach.

References

Babych, B., Hartley, A., & Sharof, S. (2009). Evaluation-guided pre-editing of source text: Improving MT-tractability of light verb constructions. In L. Màrquez & H. Somers (Eds.), *Proceedings of the 13th annual conference of the European Association for Machine Translation* (pp. 36-43). European Association for Machine Translation.

Bai, S. R. (2017, September 26). *Meiyou jianquan de shiyu jiaoyu*, *he lai haixian wenhua*? [How can we have seafood culture without good fish-eating education?] Environmental Information Center. https://e-Info.org.tw/node/207511

Barry, S., & Lazarte, A. A. (1995). Embedded clause effects on recall: Does high prior knowledge of content domain overcome syntactic complexity in students of Spanish? *The Modern Language Journal*, *79*(4), 491-504. https://doi.org/10.1111/j.1540-4781.1995.tb05449.x

Bellis, M. (2019, February 5). *The history of international business machines and IBM computers*. https://www.thoughtco.com/the-ibm-701-1991406

Cardey, S., Greenfield, P., & Wu, X. H. (2004). Designing a controlled language for the machine translation of medical protocols: The case of English to Chinese. In R. E. Frederking & K. B. Taylor (Eds.), *Machine translation: From real users to research* (pp. 37-47). Springer Verlag. https://doi.org/10.1007/978-3-540-30194-3_5

Chen, Z. H., & Zheng, H. W. (2017, June 6). *Shikong de haiyang feiqiwu – Sujiao qicheng de haigui fenchang* [Uncontrollable marine debris – A turtle cemetery made of plastic]. The Reporter. https://www.twreporter.org/a/marine-debris-seaturtle-tomb

HarperCollins Publishers. (Ed.). (2001). *Collins COBUILD English dictionary for advanced learners* (3rd ed.).

Hsieh, T. (2018, October 5). *Haiyang ziyuan buneng zai women zheyidairen zhongjie: Xunyu manshi, lihai yongxu* [Marine resources cannot be exhausted in our generation: Developing the right attitude of eating seasonable fish and maintaining the sustainable relationship between humans and sea]. NOM Magazine. https://nommagazine.com/?p=45313

Hsu, S. C. (2004). *Reading comprehension difficulties and reading strategies of junior high school EFL students in Taiwan* [Unpublished master's thesis]. National Kaohsiung Normal University.

Huang, T. T. (2020, August 26). *Taiwan's goal of going bilingual by 2030 unattainable: Education expert*. Taiwan News. https://www.taiwannews.com.tw/en/news/3995094

Iqbal, M., Noor, M., Muhabat, F., & Kazemian, B. (2015). Factors responsible for poor English reading comprehension at secondary level. *Communication and Linguistics Studies*, *1*(1), 1-6. http://dx.doi.org/10.2139/ssrn.2602630

Jiang, P. S. (2003). *The relationship between EFL learners' vocabulary size and reading comprehension* [Unpublished master's thesis]. National Kaohsiung Normal University.

Johnson, P. (1981). Effects on reading comprehension of language complexity and cultural background of a text. *TESOL Quarterly*, *15*(2), 169-181. https://doi.org/10.2307/3586408

Kamprath, C., Adolphson, E., Mitamura, T., & Nyberg, E. (1998). *Controlled language for multilingual document production: Experience with Caterpillar Technical English*. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.44.6405

Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In C. Lauren & M. Nordman (Eds.), *Special language: From humans thinking to thinking machines* (pp. 316-323). Multilingual Matters.

Lee, A. (1994). Controlled English with and without machine translation. *Aslib Proceedings*, *46*(5), 131-133. https://doi.org/10.1108/eb051356

Lee, C. J. (2016, April 19). *Fanyi yu rengong zhihui jiehe de muqian qingkuang* [Current situation of translation integrated with artificial intelligence]. KKNEWS. https://kknews.cc/zh-tw/tech/6pl68mp.html

Lee, H. S., Liu, S. Y., & Yeh, T. K. (2019). Marine education through cooperation: A case study of opportunity in a remote school in Taiwan. In G. Fauville, D. L. Payne, M. E. Marrero, A. Lantz-Andersson, & F. Crouch (Eds.), *Exemplary practices in marine science education* (pp. 191-205). Springer.

Lehtola, A., Bounsaythip, C., & Tenni, J. (1998). *Controlled language technology in multilingual user interfaces*. http://ui4all.ics.forth.gr/UI4ALL-98/lehtola.pdf

Lin, S. Y. (2008). *The effects of memorizing word lists on vocabulary learning and reading* [Unpublished master's thesis]. National Taiwan Normal University.

Littlejohn, A. (2012). Language teaching materials and the (very) big picture. *Electronic Journal of Foreign Language Teaching*, *9*(1), 283-297. http://e-flt.nus.edu.sg/v9s12012/littlejohn.pdf

Lo, P. H. (2015). *A study of normative controlled language writing for Thai MT application* [Unpublished master's thesis]. National Kaohsiung First University of Science and Technology.

Lu, C. F. (2005). *The effects of reading articles and memorizing word lists on EFL high school students' reading and vocabulary abilities* [Unpublished master's thesis]. National Taiwan Normal University.

McDonough, J., Shaw, C., & Masuhara, H. (2013). *Materials and methods in ELT: A teacher's guide* (3rd ed.). Wiley-Blackwell.

Mercader-Alarcón, J., & Sánchez-Martínez, F. (2016). *Analysis of translation errors and evaluation of pre-editing rules for the translation of English news texts into Spanish with Lucy LT*. https://www.dlsi.ua.es/~fsanchez/pub/pdf/mercader-alarcon16.pdf

Mogilner, A. (1992). *Children's writer's word book*. Writer's Digest.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

O'Brien, S. (2003). Controlling controlled English: An analysis of several controlled language rule sets. In CLAW & EAMT (Eds.), *Proceedings of EAMT-CLAW-03, controlled language translation* (pp. 105-114). Dublin City University.

O'Brien, S., & Roturier, J. (2007). How portable are controlled language rules? A comparison of two empirical MT studies. In B. Maegaard (Ed.), *Proceedings of machine translation summit XI* (pp. 345-352). Centre for Language Technology.

Pym, P. J. (1990). Pre-editing and the use of simplified writing for MT: An engineer's experience of operating an MT system. In P. Mayorcas (Ed.), *Translating and the computer 10* (pp. 80-96). Aslib.

Roturier, J. (2004). Assessing a set of controlled language rules: Can they improve the performance of commercial machine translation systems? In Aslib (Ed.), *Proceedings of the international conference translating and the computer 26* (pp. 18-19). Aslib.

Shih, C. L. (2013). *Alternative web textual writing: Effective communication across borders*. Bai-xiang.

Shih, C. L. (2020). *Translation technology: Development and application*. Bookman.

Snow, C. E. (2002). *Reading for understanding: Toward an R&D program in reading comprehension.* RAND Corporation.

Solak, E., & Cakır, R. (2015). Exploring the effect of materials designed with augmented reality on language learners' vocabulary learning. *Journal of Educators Online*, *13*(2), 50-72. https://doi.org/10.9743/JEO.2015.2.5

Taiwan Fertilizer Co., Ltd. (2017, October 1). *Manshi manyu de yongxu jingying - Rang xiayidai ye chidedao xinxian de yu!* [Sustainable management of slow food & slow fish!] Taifei Kanwu. https://www.taifer.com.tw/PublicationArticle DetailC004000.aspx?Cond=5e90e1e3-9977-468f-9b7c-7f4a61f5acbd

Tánczikné, S. V. (2017). Factors affecting reading comprehension. *Gradus*, *4*(2), 41-47.

The Ministry of Education. (2007). *White paper on marine education policy*. https://ws.moe.edu.tw/001/Upload/3/relfile/6315/55805/40f900df-a70b-4e1f-847f-e53ed94b4c77.pdf

The Ministry of Internal Affairs of the Republic of China. (2011, December 13). *Yami (Dawu) zu feiyu ji* [Yami (Dawu) flying fish festival]. Cultural Heritage in Taiwan. https://www.taiwangods.com/html/cultural/3_0011 aspx?i=198

The National Museum of Marine Biology and Aquarium. (n.d.). *Haian shui yu shengtai huanjing* [Problems of the coastal environment]. https://www.nmmba.gov.tw/cp.aspx?n=5efaf2d5eb3360fd&s=04D50237E37C9779

Tomlinson, B. (2012). Materials development for language learning and teaching. *Language Teaching*, *45*(2), 143-179. https://doi.org/10.1017/S0261444811000528

Yorio, C. A. (1971). Some sources of reading problems for foreign language learners. *Language Learning*, *21*(1), 107-115. https://doi.org/10.1111/j.1467-1770.1971.tb00494.x

Zheng, H. X. (2002). Preface. In College Entrance Examination Center (Ed.), *High school reference English glossary developed for college entrance examination center: A report of the editing research project* (pp. ii-xiii). College Entrance Examination Center.

Zhou, Z. T. (Ed.). (2011). *English reader 2*. Lungteng.