Adaptations in Controlled Cultural Writing for Effective Machine Translation: A Register-specific Probe

Chung-ling Shih
English Department, National Kaohsiung First University of Science and Technology, Taiwan

Abstract—Machine translation (MT) serves as a fast way of transmitting web information across countries and MT performance can be highly improved by adapting the source text using a controlled language (CL). Drawing on Halliday’s register-specific accounts as the theoretical framework, this paper investigates multi-faceted adaptations that undergo in the controlled cultural writing (CCW) considering the three dimensions of a language activity, such as field, tenor and mode. By analyzing 296 controlled sentences adapted from 22 folk cultural texts of online Encyclopedia of Taiwan, along with their corresponding good/excellent-level English machine translations, as a case study, this paper identifies some register-specific rules for adaptations. The findings show that variation in grammatical adaptations in CCW, related to the linguistic differences between English and Chinese, supports the mode of online MT application. The paraphrasing of all cultural references for clear transmission of cultural information helps achieve the function of intercultural communication (field), and the conversion from a heavy into lighter context meets western audiences’ pragmatic expectations (tenor). Above all, the MT-driven adaptations in CCW are performed within a context where the interplay of complex factors, such as the function of intercultural communication, cultural inadequacy of western audiences, and MT operational constraints, are taken into account for the creation of readable, scannable cultural machine translations to serve western audiences.

Index Terms—controlled cultural writing, machine translation, adaptation, tenor, mode, field

I. INTRODUCTION

Effective intercultural communication across countries depends on the adequate understanding of the cultures of foreign countries. Thanks to modern computer technology, web cultural texts provide an easy and convenient access to the information of foreign cultures. Web cultural texts inform diverse religious beliefs and attitudes towards family, fame, wealth, power, education, marriage, death and others. Knowing the thinking of different people and how they live helps reduce bias, misunderstanding and conflicts between countries. However, the information of foreign cultures cannot be easily understood without breaking language barriers through translation. Translation per se involves not only code-switching between source language (SL) and target language (TL) but also mediation between source and target cultural values. Thus, different degrees of adaptations must be involved in cultural translation. Diverse degrees of adaptations are exhibited by using diverse strategies for achieving different purposes in the translation process.

In investigating appropriate strategies for rendering cultural texts, some scholars (e.g., Zhang & Huang, 2007; Mo, 2003; Pang & Zhu, 2008) agree on using the foreignizing strategy for translating cultural references, but others (Peng & Wang, 2008; Faiq, 2008, Li, 2005) propose the use of the domesticating translation or manipulated intercultural representations for effective cultural communication. With an alternative focus, this paper conducts a register-specific investigation into multifaceted adaptations that undergo in the process of controlling cultural texts to create a satisfying machine translation (MT). Set within the framework of Halliday’s (1964, 1978) register theory, this paper investigates how web textual adaptations are governed by some factors such as the field in which cultural information is used or scanned, the global audiences who access the online translation, and the mode that conveys the translation. The interactive relations among tenor, field and mode govern the way of adapting web cultural texts for MT application. Figure 1 shows the interaction of the relevant factors.
The way of transmitting translation has changed on the web with the help of the MT tool, and so does the way of presenting the web source text. To allow Indo-European international audiences to get web information in their languages through MT application, the source cultural text must be revised using the controlled language (CL). In discussing the way of adapting the source text or, specifically described, the controlled cultural writing (CCW), we have to consider the linguistic difference between SL (e.g., Chinese) and TL (e.g., English). However, the constraints of some extra-linguistic factors, such as field, tenor and mode, must be considered as well. By analyzing 296 controlled sentences collected from 22 Chinese folk cultural texts of online Encyclopedia of Taiwan, along with their corresponding good/excellent-level English machine translations, as a case study, this paper aims to identify some rules for adaptation at grammatical, lexical and contextual levels within Halliday’s register theoretical framework. Additionally, the implications of the MT-enabled intercultural communication on the web will be explored. To achieve these two objectives, three research questions (RQ) are raised for investigation.

RQ1: How are grammar/lexicon-specific adaptations in CCW governed by the media of cultural transmission, i.e., MT application (mode)?
RQ2: How are contextual adaptations in CCW affected by western audiences (tenor)?
RQ3: How are overall adaptations in CCW related to the translation function of efficient intercultural communication on the web (field)?

RQ1 handles the relevance of grammar/lexicon-specific adaptations to the way of transmitting cultural information on the web through MT application; RQ2 addresses the contextual adaptations in relation to western audiences’ cultural inadequacy and stylistic preferences. RQ3 emphasizes the relevance of overall textual adaptation to the purpose of intercultural communication on the web. In short, this paper explores how adaptations in MT-driven CCW are undertaken considering the linguistic differences between SL (e.g., Chinese) and TL (e.g., English), and extra-linguistic, contextual factors such as online transmission, western audiences, web reading styles, MT operational constraints and relevant others.

II. THEORETICAL REVIEW

In this section, Halliday’s register-specific concepts will be introduced because this paper draws on them as the theoretical framework to support some arguments for multifaceted adaptations in CCW. Additionally, some concepts in relation to adaptation, CL, and MT are necessarily discussed at some length.

The Register Theory

This paper investigates how the adaptation of web cultural texts using the CL for online machine translation is governed by some register-specific factors. The concept of register is used to distinguish the variety in language use in different situations within the user-use framework developed by Halliday et al. (1964). When a language activity takes place in various contexts, there are some differences between “the types of languages selected as appropriate to different types of situation” (Hatim & Mason, 1990, p. 46). From Halliday’s (1978, p. 111) point of view, register refers to “the configuration of semantic resources that the member of the culture associates with a situational type” (qtd. in Martin, 1992, p. 498). It also means “the meaning potential that is accessible in a given social context” (qtd. in Martin, 1992, p. 498). This argument suggests that the meaning of an utterance is created through the use of linguistic codes that are specific to a given social situation. Simply put, the concept of register is used to decide what people do to whom in what linguistic format or in what style of writing that is specific to a given situation and a specific function.

Register is often identified as the three dimensions of language use: field, mode and tenor. Field means what is going on; mode, the way of presenting the message; and tenor, the receptors of the textual message or the style appropriate to specific target audiences. Viewed in an alternative way, field means “the purposive role” (Gregory & Carroll, 1978; qtd. in Hatim & Mason, 1990, p. 48) of the language use or social function of the text. The content of the text reflects some happenings in a society for a specific purpose. Mode refers to the medium in which a language activity is performed and it is “the manifestation of the nature of the language code being used” (Hatim & Mason, 1990, p. 49). For examples, there is a distinction between writing and speaking when the same message is presented. The linguistic features between writing and speaking are overtly different. Another factor, tenor, denotes “the relationship between the addressee and the addressee” (Hatim & Mason, 1990, p. 50) and it can be identified by distinguishing between the formal and informal ways of presentation.

In Halliday’s (1989) description, field refers to “social action”; namely, “what is actually taking place” and “what is happening” (qtd. in Martin, 1992, p. 499). It also means what the participants are engaged in. Tenor refers to “who is taking part, to the nature of the participants, their statuses and roles” (Halliday, 1989; qtd. in Martin, 1992, p. 499). The relationship among the participants is considered. Mode refers to “what part language is playing, what is it that the participants are expecting the language to do for them in the situation” (Halliday, 1989; qtd. in Martin, 1992, p. 500). In this regard, the channel of communication, either speaking or writing or a mixture of them, plays a crucial role to achieve an intended purpose, say, persuasive, informative, affective, expository or didactic function and the like. Above all, the points at issue are content/information (what), the audience (who) and the style/form of presentation (how), which constitute the three aspects of discourse in the register theory. Normally, an interplay of the complex factors such as field, mode and tenor outlines how a language activity is performed or how a language behavior is enacted in a given
Adaptation, Controlled Language and Machine Translation

Adaptation, an inevitable strategy in the translation process, is the main issue in this research and needs to be introduced in brief. According to Bastin (1998), adaptation is often performed at the local level due to some linguistic differences between SL and TL, and also at the global level to achieve a specific function. The former is viewed as obligatory modification and the latter, optional modification. Normally, human translation can easily achieve the tasks of local or global adaptation by directly modifying the SL text in the translation, but in the machine-enabled translation, the source text must be adapted before it is submitted to the machine to improve the MT performance. For example, human translator can directly adapt and translate the Taiwanese slang, 吃飯配菜脯 (lit: eat meal accompanied by pickles), as “eat like a beggar”, but for MT application, the phrase must be adapted as 吃像—乞丐 (lit: eat like a beggar) so that its English machine translation [eat like a beggar] can be semantically clear to western audiences. In this respect, we notice that adaptation is necessary for effective MT application, and this adaptation refers to rewriting a source cultural text in a CL.

A CL is often identified as a sub-language or an artificial language. As Arnold et al. (1994, p. 211) have put, the CL means “a specially simplified version of a language”. This language is “a subset of natural language with an artificially restricted vocabulary, grammar and style” (Kaji, 1999, p. 37). For example, controlled English, one type of CL, is characterized by “the high frequency of specialized terminology and often also by a restricted set of grammatical patterns” (Arnold, et al., 1994, p. 216). Shuttleworth and Cowie (1997) claim that a CL is essentially a “variant of SL in which texts are composed according to a set of rules designed to enhance the clarity and readability of what is said” (29; qtd. in Shih, 2006, p. 129). Some renowned corporations, such as Boeing, Caterpillar and others in Western countries, have authored technical texts, user’s manuals, and product instructions in controlled English (CE) for effective multilingual machine translations. Some evaluations on the MT performance (Roturier, 2004; Cardey, Greenfield & Wu, 2004) have confirmed the cost-effective benefits of using CE to save time and efforts for post-MT editing.

When using controlled English, a set of rules can be followed, including “avoidance of the gerund and noun clusters of more than three nouns, the use of an article (the, a, an) or a demonstrative adjective (this, these) before a noun, the use of specific instructions, use of approved words from the dictionary and use of relative pronouns (“who”, “which” or “that”)” (O’Brien, 2003, p. 110). In the present research, web cultural adaptation using controlled Chinese (CC) has greatly improved its English machine translation. To create satisfying MTs in Indo-European languages, the CC can imitate the grammar of English and use plain words, common expressions and short, simple sentences. The idiomatic or local attributes of lexicons and sentence structures must be adapted so that machine translations can be semantically clear and grammatically correct. The following gives a general introduction to MT.

MT is here referred to as a translated text that is automatically created by a machine/computer from a source text within seconds. Early rule-based MT systems create a translation through analysis, transfer and generation with an unsatisfying result. In contrast, contemporary statistics/corpus-based MT systems greatly improve the performance by stringing up the statistically most likely segments selected from various candidate translations. Google Translate is one of free online statistics/corpus-based MT systems, and is here used to test the English machine translations of controlled web cultural texts. Google Translate allows users to correct suggested translations and submit them back, so that its corpus size can be enlarged as time goes by. A bigger corpus is definitely helpful to improve the MT performance. Yet, at the present time, as a tentative solution, we can control the input data by adapting them using the CL and their machine translations have highly promoted the target audience’s comprehension.

III. METHODOLOGY

This paper investigates how we adapt web cultural texts using the CL by following some rules in the areas of lexicon, grammar, context and style from the register-related perspective. The rules are subject to the constraints of tenor, field and mode. The method of analyzing the controlled cultural texts and their corresponding comprehensible English MTs is used to complete the investigation. This section introduces the analyzed web cultural texts, some criteria for assessing readable English MT and some adaptation rules.

Web Cultural Texts

There are controlled 296 sentences, adapted from 22 web cultural texts of online Encyclopedia of Taiwan, for analysis. These texts describe some cultural practices in the five stages of human life, including birth, adulthood, marriage, pregnancy/childbirth and death. Table 1 shows the textual titles and Appendix 1 provides detailed information.
By reading the cultural texts as illustrated in Table 1, foreigners can understand well some concepts about Taiwanese life and culture, including: 1) continuing the ancestral line, 2) outweighing men over women; 3) remembering ancestors; 4) fulfilling one’s marital duty; 5) asking for God’s protection; 6) respecting the deceased; 7) emphasizing filial duty, and 8) choosing spouses based on similar family backgrounds.

Criteria for MT Assessment
For effective intercultural communication, MT performance must meet some expectations or standards. The four-level criteria proposed by O’Brien and Roturier (2007) are used to assess the English machine translations of 296 sentences. The excellent level means that the machine translation does not need to be edited, and the target audience can easily understand it without having to cross-reference to the ST text. The good level means that the machine translation is acceptable and readable even though it has some grammatical or lexical errors. Table 2 gives one example to illustrate the four-level machine translations created by Google Translate. ST=source text. The word-for-word glossing of the ST is given for reference.

<table>
<thead>
<tr>
<th>Cultural Themes</th>
<th>Text Titles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adulthood</td>
<td>1. Coming-of-age ceremony; 2. A thank-heavenly-lord rite</td>
</tr>
<tr>
<td>Pregnancy/childbirth</td>
<td>1. Praying for a child; 2. Women’s postnatal care; 3. Rite of belly-change,</td>
</tr>
<tr>
<td>Death</td>
<td>1. Guarding the corpse; 2. Guiding the dead soul; 3. Worshipping memorial tablets of ancestors</td>
</tr>
</tbody>
</table>

Table 1. The titles of collected cultural texts

<table>
<thead>
<tr>
<th>Levels</th>
<th>Assessment Criteria</th>
<th>English Machine Translations and ST</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excellent</td>
<td>The MT output is perfect and does not need to be edited. The end-user does not have to cross-reference to the ST text and could understand the MT output easily.</td>
<td>In Taiwan, before the wedding, a feast is often held for the bride and her sisters. This feast means separation from her family and her sisters’ congratulations. [ST: 在臺灣，結婚當天，新娘會在她的姐妹們的祝福中，與她的姐妹們離別。] [Lit: In Taiwan, before marriage, a feast is often held for the bride and her sisters. This feast means separation from family and her sisters’ congratulations.]</td>
</tr>
<tr>
<td>Good</td>
<td>The MT output is acceptable and readable though it has some minor grammatical or lexical mistakes. The end-user can still understand the MT output without consulting the SL text.</td>
<td>According to Taiwan’s traditional practices, before marriage, a woman will eat together with her sisters. This feast means the bride and her sisters will say goodbye to each other, and this is the blessing of the sisters. [ST: 根據台灣的傳統習俗，在結婚當天，新娘將與她的姐妹們聚餐，這就是姐妹們的祝福。] [Lit: According to Taiwan’s conventional custom, before marriage, a girl would eat a meal together with her sisters. This banquet means the bride and her sister would say good-bye to each other, and this is her sisters’ blessing.]</td>
</tr>
<tr>
<td>Mediocre</td>
<td>The MT output has significant errors and the end-user can simply get the gist of the MT output without referring to the SL text.</td>
<td>Wedding customs in Taiwan, when the girls get married the same day, she and the sisters dinner, which is the bride with her sisters farewell feast, but also a feast for sisters congratulations bride. [ST: 臺灣的結婚習俗，當女生出嫁當天，她會與她的姐妹們同吃飯，這是新人與姊妹們離別的筵席，同時也是姊妹們祝賀新人的筵席。] [Lit: In Taiwan’s wedding custom, on the day when a girl marries, she would eat a meal with her sisters together; this is the farewell banquet for the bride and her sisters, simultaneously is also a feast for her sisters to congratulate the bride.]</td>
</tr>
<tr>
<td>Poor</td>
<td>The MT output is unreadable and incomprehensible. It contains serious errors. The end-user is unable to catch any message from the MT output without reading the SL text.</td>
<td>Taiwan’s traditional practices, the girls married the same day, out the door to dinner with her sisters, commonly known as the Sisters table, which is the bride with her sisters farewell feast, the bride’s sisters congratulated feast. [ST: 臺灣的傳統習俗，女生出嫁當天，出門前要與她的姐妹們一同吃飯，這是新嫁娘與姐妹們離別的筵席，也是姐妹們祝賀新人的筵席。] [Lit: In Taiwan’s conventional custom, on the day when a girl marries, she must eat a meal with her sisters before leaving home, conventionally called “sister table”; this is the farewell banquet for the bride and her sisters, also is the feast for sisters to congratulate the bride.]</td>
</tr>
</tbody>
</table>

The example above shows that the shorter the source sentence is, the better performance the machine creates.
Additionally, some supplementary and specific rules can be followed to create good-level and excellent-level machine translations in terms of lexical, grammatical, syntactic accuracy and pragmatic appropriateness. All the collected English MTs in this research are identified as reaching the excellent level or good level.

### Rules for Multi-faceted Adaptations in MT-driven CCW

Some grammatical, lexical and contextual adaptations must be performed to create the excellent-level and good-level English machine translations. Once when the English MT has a good quality, its transfer into other Indo-European languages has a higher percentage of good-level or excellent-level performance than the machine renditions from unacceptable, mediocre or/and poor-level English MT. In the following section, some rules for adaptation in the areas of lexicon, grammar and context are introduced with supportive examples.

#### Rules for Grammatical Adaptation

A host of rules for grammatical adaptation in CCW, to cite most important ones, include: 1) use of a subject (or/and an object) for the SVO structure; 2) use of an article such as –yi (lit: a) or a demonstrative pronoun, such as –zhe/gai/na (lit: this/the/that); 3) use of a connective to join two clauses; 4) use of a preposition marker -zai; 5) use of a possessive adjective such as -t de (lit: to link a noun; 6) use of a post-noun modifier, not a pre-noun modifier; 7) use of a passive voice, not an active voice, and 8) use of a time/place-specific prepositional phrase at the initial or final position of a sentence.

The lack of a subject or object or an article or a possessive adjective tends to create a grammatically incorrect English machine translation. Thus, the uncontrolled Chinese sentence, **離開前，會與姊妹們吃飯。** [lit: before leaving, will eat meal with sisters], must be adapted as **離開前，該新娘將會吃飯與她的姊妹們。** and then its English MT is correct as **Before leaving, the bride will eat a meal with her sisters.** It is also noted that a Chinese sentence tends to use a comma, not a conjunction, to connect two clauses, and therefore its English MT is grammatically incorrect. Thus, the sentence **新娘不哭泣，她會被視為不夠孝順** [lit: Bride does not cry, she will be regarded as inadequately filial] must be modified as **若新娘不哭泣，她會被視為不夠孝順** and its English MT is grammatically correct as **If the bride does not cry, she will be deemed not filial.**

#### Rules for Lexical Adaptation

In addition to grammar-specific rules, some lexicon-specific rules for adaptation are identified, including: 1) modifying fixed phrases into explanatory expressions, 2) replacing idiomatic terms with common words, and 3) directly using well-established English names for proper nouns. Some examples are given as follows. 嫁香煙 [lit: marry cigarette] is adapted as **探門風** [lit: explore door style] as 未婚前調查 [lit: premarital investigation]; 三界神明 [lit: three god] as 三個神明, 以掌管陸地、海洋及天空 [lit: three gods, to control land, ocean and sky], and 四方金 [lit: four square gold] as **紙錢** [lit: paper money]. Western audiences do not know source cultural information, so local cultural references or proper nouns should be paraphrased to explicating the implicit meanings in the controlled texts. Thus, the MT outputs of paraphrased cultural references are semantically clear and the target audiences would easily understand them.

#### Rules for Contextual Adaptation

More importantly, a majority of western audiences favor the low context so that specific details must be condensed, and overlapping, redundant information can be deleted. These context-specific adaptation rules are followed to reduce contextual density and help the target audiences easily read and clearly understand the message of the MT output.

### IV. FINDINGS & DISCUSSIONS

After some specific rules are identified, Chinese-to-English machine translations of 296 sentences are analyzed to identify which rule takes up a higher or lower frequency, and the implications of the findings will be discussed within the register-specific theoretical framework.

#### Grammar/Lexicon-specific Adaptations in CCW

In response to RQ1, the findings show that the mode of MT application has governed the adaptation of web cultural texts in the lexical and grammatical areas. All the English machine translations evaluated as good and excellent levels have followed eight grammar-specific rules. Since the statistics/corpus-based Google Translate cannot automatically change the original wording and word sequence, grammatical and lexical adaptations are needed to make the source text similar to the grammatical and lexical forms of English. In this respect, if a linguistic feature in Chinese shows a greater difference from its counterpart in English, a higher degree of adaptation is required. This phenomenon supports that variation in the use of eight grammar-specific rules is highly related to the linguistic differences between Chinese and English. The top three grammar-specific rules are: 1) adding a subject or an object or a verb (69 occurrences; 69/296=23.31%), 2) adding an article or a demonstrative pronoun (68 occurrences; 68/296=22.97%), and 3) adding a connective (52 occurrences; 52/296=17.56%). The bottom three grammatical rules are 1) adding a prepositional marker (8 occurrences; 8/296=0.27%), 2) adding a possessive adjective (12 occurrences; 12/296=4.05%) and 3) moving a preposition-led temporal or locative phrase to the final position of a clause or a sentence (18 occurrences; 18/296=6.08%). The other two rules that lie in between are changing active voice into passive voice (48 occurrences; 48/296=16.21%) and changing a pre-noun modifier into a post-noun modifier (21 occurrences; 21/296=7.09%). Figure 2 shows varied frequencies of using eight grammar-specific adaptation rules.
Adding a subject ranks first and adding an article/demonstrative pronoun ranks second. One inference is that these two grammatical features are distinctively different between English and Chinese, so controlled cultural writing must undergo adaptation when these differences show up. Normally, Chinese sentences use the topic-comment structure so the clauses often have no subjects, but English sentences use the SVO structure and are grammatically incorrect without subjects. In addition, Chinese sentences rarely use an article before a noun, but native Chinese readers can understand the message. However, an English sentence without an article or a demonstrative pronoun is grammatically incorrect. Adding a connective ranks third because Chinese sentences tend to use commas to connect two clauses but English uses connectives to show diverse intra-sentential relations. By imitating English sentence structure, connectives must be added in controlled Chinese cultural writing.

Adding a preposition marker shows the lowest frequency and this can be attributed to the fact that the prepositional marker such as -yu or -zai is often used by native Chinese in describing a date or a location. Only few sentences do not use it. The second lowest frequency of adaptation is found in adding a possessive adjective, because machine translations of both nouns with and without possessive adjectives are equally comprehensible. For example, two machine translations, such as Ghost/Spiritual marriage can reduce the anxiety of her family and Ghost/Spiritual marriage can reduce the anxiety of the family, are equally understandable to western audiences.

On the other hand, the MT application cannot handle well cultural references so that lexical adaptations are needed. The statistical result shows that eight out of fifteen fixed phrases must be modified with explanatory accounts, taking up 53.33% (8/15). Among twenty-one idioms, fifteen must be represented with common expressions, holding 71.42% (15/21), and of twenty proper nouns, twelve must be paraphrased, holding 60% (12/20). These findings suggest that more than half of cultural references have to be adapted or modified using the CC to create comprehensible machine-enabled translations. Figure 3 shows the frequencies of three types of lexical adaptation.

![Figure 2 Frequencies of using eight grammar-specific adaptation rules](image)

![Figure 3 V aried frequencies of three types of lexical adaptation](image)

Adaptation of idioms shows a higher frequency. One inference is that the idioms describe cultural rites and practices, and their genuine meanings cannot be clearly presented based on the combination of surface meanings of the words together. For example, the term 变肚 [lit: change belly] suggests that a married woman is eager to be pregnant, not really converting her belly. Its literal translation will be misleading to western audiences, so it is paraphrased as 去改变一情侣的肚子，使她懷孕，且會生出一男孩子 [lit: to change a woman’s belly, make her pregnant and will produce a boy] and its English machine translation “to change a women’s belly, so that she is pregnant and will give birth to a boy” is semantically clear and grammatically correct. Furthermore, the idiomatic expressions, such as 坐月子 [lit: rest for one month] cannot be appropriately translated by the machine, so it must be adapted as 產後照顧 [lit: care after giving birth], and its machine translation “post-natal care” is semantically accurate.

Adapting proper nouns holds the second high because the names of some Gods, such as 玉皇大帝、觀世音、七娘媽、媽祖, can be correctly translated by the machine as “Jade Emperor”, “Guanyin”, “Qinianma” and “Matsu”, and thereby they do not need any adaptation. In contrast, the names of some gods such as 註生娘娘 [lit: registering birth Goddess] · 床母 [lit: bed mother] and 龍水夫人 [lit: close-to-water Madame] cannot be appropriately translated, so they have to be paraphrased, and their English machine translations, such as “fertility Goddess”, “bed Goddess” and “Lady Linshui”, are pragmatically appropriate. Furthermore, the names of some plants, such as banana flowers and pomegranate, and the name of the festival, such as the Lantern Festival, do not need to be adapted because their English machine translations are correct. However, some fixed phrases, such as 養生送死 [lit: raise living persons, send away...
dead persons].

In general, cultural references have distinctive linguistic features and special implications, so that they must be adapted into general expressions. The modified expressions not only provide a commonly shared cultural background to foreign audiences, but also improve the semantic clarity and pragmatic appropriateness of their machine translations. Today, the data of the corpus of Google Translate remain insufficient, particularly the translations of idioms and proper nouns. The idioms and proper nouns in cultural texts must be adapted for effective MT application. Above all, lexical and grammatical adaptation concurs with the register-specific argument that language activity must be adjusted to meet the need of using a special way of presentation in a new situation, and meet western audiences’ expectations for successful cultural acquisition by scanning the understandable machine translations of cultural texts on the web.

**Contextual Adaptations in CCW**

In response to RQ2, the finding shows that contextual adaptations through omission, condensation and division are needed, so the MT output can meet western audiences’ preferences for the light context and concise writing style. However, in the collected data, no many contextual adaptations are identified in controlled cultural writing, and therefore the statistical result is not significantly meaningful. For this reason, only some examples are given here to support the three types of contextual adaptations. With regard to omission, all the redundant, overlapping and unnecessary segments are omitted. One example can be found with the sentence, “...死不瞑目...” and “...床女神不能吃太久，床神會允許小孩早些...” which can reduce the audience’s interpreting efforts and optimize the communication effect.

**Division** is another strategy for adapting texts. By dividing a long sentence into two or three shorter ones, its machine translation can reduce the audience’s interpreting efforts and optimize the communication effect. For example, there is a long sentence, “...民间信仰望拜床神的时间不能太久，主要是怕床神会寵孩子，导致小孩懒床；另一说是要床神吃饭快点，才不会宠坏小孩。“ Traditionally, people believe that the time for worshipping the Bed Goddess cannot be too long, the main reason is that they fear that the Bed Goddess would spoil children, and children do not want to get up; another reason is that if the Bed Goddess eats faster, she can take good care of children. ] We can divide it into two shorter ones, such as “...This ritual cannot proceed too long, because Bed Goddess will allow the child to stay in bed too long.” and “...Bed Goddess must eat faster, so she can take care of the child earlier.” In this way, the machine translations of the two shorter sentences are correct as “This ceremony is not too long, because Bed Goddess [bed goddess] will allow children to stay in bed too long” and “In addition, Bed Goddess must eat quickly, so she can take care of the children earlier.” At this point, we know that the strategy of division not only improves the grammatical accuracy of the machine translation, but also highlights the implicit and logical relationships for semantic clarity.

One more strategy, condensation, aims to create a concise text by condensing overlapping messages. Table 3 shows two examples in a contrastive form of pre-condensation and post-condensation with their English translations. CW=controlled writing; UCW=uncontrolled writing.
The two examples above show that information condensation in the CCW results in the concise, clear English translation created by Google Translate, whereas the uncontrolled text leads to the erroneous and vague English machine translation. In short, contextual adaptation, along with lexical and grammatical adaptation, supports the register-specific argument that textual re/presentation must be appropriate to the given situation, the textual function, the target audience and the mode of transmitting the translation.

**MT-driven Adaptations for Intercultural Communication**

In response to RQ3, the findings denote that re-contextualizing and re-presenting web cultural texts through multifaceted adaptations helps achieve the function of intercultural communication on the web, not for rhetorical/linguistic appreciation. Today, for international audiences, acquiring web cultural information is more important than appreciating the distinct linguistic traits of the source text. Most of western audiences want to know certain rituals and cultural practices in each folk culture, so these data have to be appropriately and accurately translated into English or other Indo-European languages. Actually, western audiences would not practice the source cultural performances, but also makes western audiences easily and quickly catch the cultural message within a short timeframe.

Since web cultural texts serve as an easily accessible means of providing cultural information to a global audience, they are a valuable avenue for intercultural communication and local cultural promotion. Understanding foreign cultural practices helps one accept diverse ways of handling the same issue and tolerate different cultural concepts. However, to improve MT performance, we have to adapt the source text lexically, grammatically and contextually. Textual adaptation must clearly convey core messages to optimize the communication effect, and emphasize the factual reality by deleting un-rooted rumors or/and groundless information. After source cultural texts are properly adapted, the improved multilingual machine translations will help achieve the goal of effective intercultural communication.

V. CONCLUSION

In this paper, the findings of investigation support that adaptation for MT-enabled communication must consider field (web cultural translation for intercultural communication and information acquisition on the web), tenor (international web audiences) and mode (MT application for online real-time information transmission). Being constrained by the operational function of corpus/statistic-based MT system, adaptation has to notice the linguistic differences between English and Chinese, and follow some grammar-specific rules to close up the differences. In addition, all idiomatic expressions or proper nouns have to be adapted to make the machine translations culturally relevant to the target audiences’ situation. Still importantly, to comply with the pragmatic and socio-cultural conventions of western audiences, contextual condensation, omission and syntactic division are adopted. Contextual restructuring also facilitates the ease of web scanning. Overall, the register-specific probe helps us identify how web cultural texts must be adapted following a set of rules under the governance of MT operation, intercultural communication purpose and western audiences’ stylistic preferences, all of which dialectically interact with each other.
Web textual adaptation for MT-enabled intercultural communication has shed light on the significance of multiculturalism. Textual representation and contextual reconstruction using the CL helps transmit local cultural information to foreign countries through multilingual machine translations. This research is irrelevant to power control and ideological manipulation. Instead, it emphasizes how the local cultural identity, whatever it represents, can be equally recognized on the Internet through MT-enabled intercultural communication. Above all, this research supports how linguistic behavior in the technological world can be adjusted to facilitate multilingual and multicultural communication by re-presenting and reconstructing web cultural texts using the CL for successful MT application.

ACKNOWLEDGEMENT

The author thanks Taiwan’s NSC for the funding grant, so she can complete the research and write the findings into the paper.

APPENDIX: TWENTY CULTURAL TEXTS FOR ADAPTATION (EXTRACTED FROM ONLINE ENCYCLOPEDIA OF TAIWAN)

<table>
<thead>
<tr>
<th>No.</th>
<th>Authors</th>
<th>Text Titles</th>
<th>Copyright Dates and URLs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Fongming Lee &amp; Liling Xu</td>
<td>Rite of the first birthday</td>
<td>2011 <a href="http://taiwanpedia.culture.tw/web/content?ID=1996">http://taiwanpedia.culture.tw/web/content?ID=1996</a></td>
</tr>
<tr>
<td>2</td>
<td>Ruuyi Tu</td>
<td>Rite of hair-cutting</td>
<td>2009 <a href="http://taiwanpedia.culture.tw/web/content?ID=11521">http://taiwanpedia.culture.tw/web/content?ID=11521</a></td>
</tr>
<tr>
<td>4</td>
<td>Ruuyi Tu</td>
<td>Long-life chain/lock</td>
<td>2009 <a href="http://taiwanpedia.culture.tw/web/content?ID=11520">http://taiwanpedia.culture.tw/web/content?ID=11520</a></td>
</tr>
<tr>
<td>6</td>
<td>Zheting Xu</td>
<td>Reporting the child-birth news</td>
<td>2009 <a href="http://taiwanpedia.culture.tw/web/content?ID=11537">http://taiwanpedia.culture.tw/web/content?ID=11537</a></td>
</tr>
<tr>
<td>8</td>
<td>Huuying Cai</td>
<td>A thank-heavenly-lord rite</td>
<td>2009 <a href="http://taiwanpedia.culture.tw/web/content?ID=11557">http://taiwanpedia.culture.tw/web/content?ID=11557</a></td>
</tr>
<tr>
<td>9</td>
<td>Yiping Xin</td>
<td>Crying for good life</td>
<td>2009 <a href="http://taiwanpedia.culture.tw/web/content?ID=11463">http://taiwanpedia.culture.tw/web/content?ID=11463</a></td>
</tr>
<tr>
<td>10</td>
<td>Wenxian Li</td>
<td>Spiritual marriage</td>
<td>2009 <a href="http://taiwanpedia.culture.tw/web/content?ID=2018">http://taiwanpedia.culture.tw/web/content?ID=2018</a></td>
</tr>
<tr>
<td>11</td>
<td>Juming Zhidang</td>
<td>No marriage between two persons carrying the same family name</td>
<td>2009 <a href="http://taiwanpedia.culture.tw/web/content?ID=11443">http://taiwanpedia.culture.tw/web/content?ID=11443</a></td>
</tr>
<tr>
<td>12</td>
<td>Fusheng Xu</td>
<td>Marrying a concubine</td>
<td>2009 <a href="http://taiwanpedia.culture.tw/web/content?ID=11470">http://taiwanpedia.culture.tw/web/content?ID=11470</a></td>
</tr>
<tr>
<td>13</td>
<td>Yujun Dai</td>
<td>Pre-marital spying</td>
<td>2009 <a href="http://taiwanpedia.culture.tw/web/content?ID=11473">http://taiwanpedia.culture.tw/web/content?ID=11473</a></td>
</tr>
<tr>
<td>14</td>
<td>Yucheng Zhu</td>
<td>Marriage within 100 days after parental death</td>
<td>2009 <a href="http://taiwanpedia.culture.tw/web/content?ID=11447">http://taiwanpedia.culture.tw/web/content?ID=11447</a></td>
</tr>
<tr>
<td>15</td>
<td>Wenxia Lin</td>
<td>Visiting the bride’s natal family</td>
<td>2009 <a href="http://taiwanpedia.culture.tw/web/content?ID=2011">http://taiwanpedia.culture.tw/web/content?ID=2011</a></td>
</tr>
<tr>
<td>16</td>
<td>Pinyao Wu</td>
<td>Child bride</td>
<td>2009 <a href="http://taiwanpedia.culture.tw/web/content?ID=11439">http://taiwanpedia.culture.tw/web/content?ID=11439</a></td>
</tr>
<tr>
<td>18</td>
<td>Yanrou Zhu</td>
<td>Women’s Postnatal Care</td>
<td>2012 <a href="http://taiwanpedia.culture.tw/web/content?ID=100144">http://taiwanpedia.culture.tw/web/content?ID=100144</a></td>
</tr>
<tr>
<td>19</td>
<td>Xiaoyuan Wei</td>
<td>Rite of belly-change</td>
<td>2009 <a href="http://taiwanpedia.culture.tw/web/content?ID=11538">http://taiwanpedia.culture.tw/web/content?ID=11538</a></td>
</tr>
<tr>
<td>20</td>
<td>Yanfang Zhuang</td>
<td>Guarding the corpse</td>
<td>2009 <a href="http://taiwanpedia.culture.tw/web/content?ID=11579">http://taiwanpedia.culture.tw/web/content?ID=11579</a></td>
</tr>
<tr>
<td>21</td>
<td>Fangyi Zeng</td>
<td>Guiding the dead soul</td>
<td>2009 <a href="http://taiwanpedia.culture.tw/web/content?ID=11569">http://taiwanpedia.culture.tw/web/content?ID=11569</a></td>
</tr>
<tr>
<td>22</td>
<td>Wenching Wu</td>
<td>Worshipping memorial tables of the ancestors</td>
<td>2009 <a href="http://taiwanpedia.culture.tw/web/content?ID=11606">http://taiwanpedia.culture.tw/web/content?ID=11606</a></td>
</tr>
</tbody>
</table>

REFERENCES


Chung-ling Shih, English Professor, is now teaching at the English Department of National Kaohsiung First University of Science and Technology. She received a Ph. D. in British and American Literature from National Taiwan Normal University in 1998. However, her major interests and current researches have shifted from literature to translation studies. Her study focuses encompass training and research in MT, TM, CTS, cultural studies, translation theories and relevant others.