Linguistic evaluation of translation errors in Chinese–English machine translations of patent titles

Yvonne Tsai
National Taiwan University

The title of an invention allows the reader to understand the significance of a patent claim, and the wording of the title recurs throughout the subsequent patent documentation. If the translation of the title is erroneous, the quality of the translation in other parts of the patent documentation also suffers. This research involved using linguistic evaluation to identify common translation errors in Chinese–English machine translations of patent titles and examine the quality of machine-translated patent titles. Special focus was placed on orthographic, morphological, lexical, semantic, and syntactic errors found in patent titles. We sought to answer the following questions: (1) What are the trends in the application of machine translation in the Taiwan Intellectual Property Office (TIPO)? (2) How is the quality of machine translation controlled at TIPO? (3) What are common errors in machine-translated patent titles? Through analysis of our findings, it is possible to estimate the level of effort required from a posteditor following translation, and to suggest methods of improving machine translations of patent titles. This study also generates information applicable to the training of patent translators and posteditors.

Le titre d’une invention permet au lecteur de comprendre l’importance d’une revendication de brevet et le libellé du titre revient dans la documentation de brevet subséquente. Si la traduction du titre est erronée, la qualité de la traduction dans d’autres parties de la documentation de brevet souffre également. Cette recherche a consisté à utiliser l’évaluation linguistique pour identifier les erreurs de traduction courantes dans les traductions automatiques en chinois et en anglais des titres de brevets et examiner la qualité des titres de brevet traduits par machine. Un accent particulier a été mis sur les erreurs orthographiques, morphologiques, lexicales, sémantiques et syntaxiques des titres de brevets. Nous avons cherché à répondre aux questions suivantes: (1) Quelles sont les tendances dans l’application de la traduction automatique à l’Office taiwanaïs de la propriété intellectuelle (TIPO)? (2) Comment la qualité de
la traduction automatique est-elle contrôlée chez TIPO? (3) Quelles sont les erreurs courantes dans les brevets d’invention traduits par machine? Grâce à l’analyse de nos résultats, il est possible d’estimer le niveau d’effort requis d’un éditeur après la traduction et de proposer des méthodes d’amélioration de la traduction automatique de titres de brevets. Cette étude génère également des informations applicables à la formation des traducteurs et des expéditeurs de brevets.

**Keywords:** patent translation, machine translation, linguistic evaluation, error classification, postediting

**Mots-Clés:** traduction de brevet, traduction automatique, évaluation linguistique, classification d’erreur, post-édition

### 1. Introduction

Over the past decade, the growing demand for translations from Asian languages into English has resulted in an increase of translated titles and abstracts for patent documentation. For translating all titles and abstracts of patent applications along with all preliminary search and examination reports, the translation services at International Bureau (IB) of World Intellectual Property Organization (WIPO), have outsourced a large volume of translation work to contracted freelancers. In addition to outsourcing translation, IB has launched a project to ascertain the viability of using computer-assisted translation and workflow tools, thereby rendering the translation process more efficient, enabling the reuse of past translations, and simplifying the process of distributing vast numbers of translations. IB continues to develop terminology databases aimed at improving the quality of internally and externally produced translations, and plans to partner with external institutions to assist with the validation of technical terminology.

Without the rich resources that WIPO enjoys, small-scale regional patent offices, such as the Taiwan Intellectual Property Office (TIPO), receive fewer applications, and therefore, the demand for translation is lower. In Taiwan, the official filing language is Chinese, and the language of translation is English. Like IB, TIPO provides translation services to meet the requirements for providing accessible technological information. TIPO has also compiled a machine translation (MT) corpus for specific areas of technology, and has developed a terminology database. However, disregard and devaluation of professional expertise in the field of patent translation have resulted in both cost cuts for outsourcing contracts, and the development of MT, with a view of translating all titles and abstracts by using automated technology.
2. Development of MT at TIPO

In 2010, TIPO outsourced the ‘Corpora Optimizing Project for Machine Translation of Patent Specifications’ to WebGenie, a company that specialises in text mining (TIPO 2010). The goals of this outsourcing contract were to train MT systems by using corpora of patent specifications; to develop the Chinese–English Aligned Corpora and Sentence Analysis Inspection System; to construct a database containing specialised terminology; and to establish an automatic Chinese–English MT prototype for patent specifications (TIPO 2010, 2).

The MT tool was developed as part of a larger project, entitled the ‘Automatic Translation System for English Translations of Chinese Patent Specifications’, initiated by TIPO in 2008. TIPO published the Chinese–English Technical Patent Glossary in 2008 for public access, and the Automatic Translation Prototype for English Translations of Chinese Patent Specifications in 2009 for internal use. The automatic translation system was published and released to the public in 2014. TIPO anticipates that in the long run, the development of an automatic translation system will reduce the cost of translating patent titles, abstracts, and patent claims into English.


![Screenshot of Google Translate](image-url)
Google Translate into their website to enable users to automatically translate patent application specifications on the main Taiwan patent search website (http://twpat.tipo.gov.tw/tipotwoc/tipotwekm). As shown in Figure 1, users can select a target language from the drop-down menu on the left, and quickly translate text from Chinese into the target language.

MT provides multilingual access to patent documents. The marked increase in the number of patent applications from Asia in recent years has attracted international attention to the development of MT. Examples include the Patent Machine Translation Task at NTCIR-9 (http://ntcir.nii.ac.jp/PatentMT), the European project Pluto (Tinsley, Way, and Sheridan 2010), the collaboration between the EPO and Google Translate (Tager 2011), and the training of statistical MT systems from parallel texts (Koehn 2010).

Although MT seldom produces perfect results, it is still necessary for global patent documentation. Systematic manual translation is no longer a viable option because of the increase in the number of patent applications. MT systems can be deployed as tools to assist human translators in working more efficiently, or to provide sufficient solutions when human translators are unavailable (Dillinger and Lommel 2004, 4). With the emergence of translation technology, translators must now learn how to use and optimise such technology to benefit in business environments. Postediting is an effective approach for optimising MT efficiency.

2.1 Translation evaluation method at TIPO

In most patent offices, the most difficult challenges faced when providing translation services are those relating to quality and timeliness (USPTO 2012). In this study, the term ‘postediting’ refers to the part of the translation process that involves editing, modifying, and correcting machine-translated text. The required skill set of a posteditor incorporates proficiency in both the source and target languages; domain knowledge, including the involved terminology and the type of text in question; confidence in his or her own translation ability and technical expertise; and a positive attitude towards MT (O’Brien 2002, 102–103). Other valuable abilities include adequate knowledge of MT systems, terminology management, and controlled language writing skills.

At TIPO, qualified language providers are contractually obliged to provide services for the translation and reviewing of abstracts using specified evaluation criteria, as well as the postediting of machine-translated titles and abstracts (TIPO 2005). Manually translated abstracts are peer-reviewed by third-party translators. Because all third-party translators are contracted based on predetermined qualifications, the peer-review quality control process is cost effective, and provides a
Yvonne Tsai

systematised and streamlined evaluation process for all translations, while eliminating the need to recruit additional reviewers.

Translation quality is reviewed based on three criteria: technical terminology, grammar, and meaning. On a scale ranging from A (highest quality) to D (lowest quality), language service providers evaluate the specialised terminology, grammar, and quality of delivery presented in a translation to determine its overall quality (TIPO 2007). The evaluation of technical terminology is focused on domain-specific vocabulary. The second part of the evaluation involves investigating grammatical accuracy, and determining whether the translation complies with the rules that govern the structure and composition of clauses, phrases, and words. The final evaluation criterion is concerned with ensuring that the meaning of the source text is clearly and accurately expressed in the translation.

To maximise objectivity, the evaluation process is completed anonymously. The reviewer is required to offer suggestions for improvements relating to the three criteria, in addition to submitting an overall rating. As shown in Table 1, ratings are determined based on the error rate of a translation, in accordance with the following scale: less than 5% (A); 6%–15% (B); 16%–25% (C); and more than 26% (D).

Table 1. TIPO Translation Evaluation Table (資料服務組第一科 2011b)

<table>
<thead>
<tr>
<th>No.</th>
<th>Volume</th>
<th>Patent Number</th>
<th>Overall rating</th>
<th>Technical terminology</th>
<th>Grammar</th>
<th>Meaning</th>
<th>None</th>
<th>Reasons for retranslation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The information in the column ‘Overall rating’ is provided by the reviewer, who selects the appropriate rating (i.e., A, B, C, or D) from a drop-down menu. The column ‘Suggestions for improvement’ is mandatory and can be completed by marking the appropriate items. The final column, ‘Reasons for retranslation’, is an open-ended column, which requires completion only when a translation must be returned for retranslation.
Manually translated abstracts are peer-reviewed by external translators, who also postedit machine-translated titles and abstracts. For machine-translated text, the same evaluation table is applied by TIPO; however, three broad error categories may not be sufficient to classify errors from MT output. Common errors, such as misplaced punctuation marks in front of prepositions, or word order errors resulting from literal translation, may fall under the same category. This limits the MT system’s ability to identify and classify errors. To improve the performance of the MT system and decrease the workload of human evaluators, the classification of errors must be more intricate.

2.2 Using linguistic error classification to evaluate MT output

This study employed Hsu’s (2014) linguistic error classification model targeted at Chinese–English translation. This model was developed as an adjustment of the model proposed by Farrús et al. (2010). In the original error classification model, there were five basic categories of linguistic errors: orthographic, morphological, lexical, semantic, and syntactic. However, the original classification model was designed for Catalan–Spanish MT. Hsu (2014) reexamined the criteria and made adjustments to include errors applicable to Chinese–English translation (see Table 2).

<table>
<thead>
<tr>
<th>Linguistic levels</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orthographic</td>
<td>Punctuation, capitalization, spelling</td>
</tr>
<tr>
<td>Morphological</td>
<td>Verb, noun, other</td>
</tr>
<tr>
<td>Lexical</td>
<td>Extra words, missing words</td>
</tr>
<tr>
<td>Semantic</td>
<td>Polysemy, homonymy, and incorrect use of expressions</td>
</tr>
<tr>
<td>Syntactic</td>
<td>Conjunction, prepositions, articles, and syntactic reordering</td>
</tr>
</tbody>
</table>

According to Hsu’s adjusted model, orthographic errors are capitalisation, punctuation, and spelling errors; morphological errors are the misuse of noun and verb forms; lexical errors occur when there are either unnecessary words or missing words in the translation; semantic errors comprise incorrect uses of expressions in a broad sense; and syntactic errors are the misuse of conjunctions, prepositions, or articles and incorrect syntax.

Postediting of MT output is a common practice; however, a posteditor’s revisions are generally restricted to the words provided by the MT system, many of which are incorrect or misleading. Certain machine-translated segments take longer to postedit than others. Occasionally, translating problematic sentence segments manually is more efficient than using an MT system. If MT systems were
more efficiently trained, translation quality would be more effectively controlled, and the pay structure provided for posteditors would be fairer.

This discussion demonstrates the importance of conducting research on translation errors in Chinese–English MTs of patent titles. While most studies related to patent translation have focused on the development of MTs, more attention should be paid to the classification of errors and the human evaluation of Chinese–English MT output. The present study identified common MT errors in patent translation by observing machine-translated patent titles from Chinese to English through linguistic evaluation. The study analysed the quality of machine-translated patent titles by identifying linguistic errors in Chinese–English MT, and the results provide suggestions for how to improve the quality of MTs. We anticipate that our findings will serve as a reference for the training of patent translators and posteditors.

3. Research method and data collection

This study focused exclusively on patent titles. The linguistic composition of patent titles are mostly nominal phrases and incomplete clauses. It would be interesting to see how effective MT can deal with these linguistic characteristics. The title of an invention repeatedly appears throughout its respective patent document, and therefore errors found in the translation of patent titles closely resemble many of the errors found in the translation of the subsequent documentation. Since the title of an invention should be informative, meaningful, concise, and present the claimed rights, error analysis of patent title translation can reveal opportunities for improvement for patent translators.

Prior to the publication of translations online, TIPO assigned translated titles and abstracts to contracted translators for peer-reviews. Upon completion, translators sent back the results to TIPO. There were around 25 contracted translators who translate and peer-review patent translations. To identify MT errors in English translations of Chinese patent titles, a total of 473 machine-translated titles in Section H (electricity) of the International Patent Classification were collected from TIPO. The selected translations were generated by the statistical MT system developed by TIPO. This system is trained using English data from the European Patent Office.

Machine-translated titles collected for this study were evaluated manually, and errors detected in each title were classified into five main categories: orthographic, morphological, lexical, semantic, and syntactic. One title may contain errors that fall in different categories, and some error categories may have common grounds, especially frequent errors such as lexical errors and semantic errors. Lexical errors refer to words that are redundant or have been omitted in the translation.
and semantic errors are words and expressions that are used incorrectly. Overuse of words and intentional omission of words may result in incorrect expression. However, semantic errors are words that are completely different from the meaning of the source text and lexical errors are words that may seem similar, but are incorrect in one or two of the words used. These are further exemplified in the examples in Section 4.

4. Analysing patent translation quality from patent title translation errors

Analysis of the 473 machine-translated titles revealed a total of 692 errors; an average of approximately 1.5 errors per entry. The overall results indicated that on average, syntactic errors were the most common type, while morphological errors were the least common (see Figure 2). This has two implications: (1) concerning language, translations rated A were not necessarily error free; and (2) capitalisation and article errors could be considered minor.

Figure 2. Proportions of errors found in machine-translated patent titles

Based on the analysis of peer-review results, the errors identified in the title translations were further categorised into ‘word order’, ‘preposition’, ‘capitalisation’, ‘article’, and ‘word choice’, for the purpose of this study. The categories ‘preposition’ and ‘article’ refer to linguistic categories of words. Prepositions are used to show spatial relations between words, or serve to mark various syntactic functions and semantic roles, while articles convey the grammatical definiteness of nouns. The
capitalisation category follows the accepted rules of capitalisation in English, and according to TIPO, the first word, abbreviations, and specialised terminology should all be capitalised in a title (資料服務組第一科 2011a).

4.1 Orthographic errors

Based on Hsu’s model (2014), orthographic errors are errors in punctuation, capitalisation, and spelling. ‘Electrical connector Metal Shell structure’ is an example of a capitalisation error taken from our research. ‘Metal shell’ is neither an abbreviation nor a specialised term, and therefore, should not be capitalised. Another such example is ‘Side CONDUCTIVE Pins Improved Power converter’. In this instance, the first letter of almost every word has been capitalised, with the exception of the last word. Such inconsistency in capitalisation signifies errors, and the ungrammatical structure impedes a reader’s ability to understand the title. The word ‘conductive’ is also unnecessarily fully capitalised.

Misspelling is not often found in the translation of patent titles; one of the few cases to be identified was ‘Wireing plate of the entire devices’ (配線盤之整線裝置), where the word ‘wireing’ should be ‘wiring’. However, according to the source text, this translation should actually be ‘Cable box of wire plate’.

4.2 Morphological errors

Morphological errors are those relating to verb and noun usage. The first example is a title starting with ‘may’, where the subject is unknown. In the title ‘May reduce the activated by the vibration motor structure’ (可降低作動時所產生震動之馬達構造), how ‘the activated’ ‘may’ be ‘reduced’ does not make sense to the reader, and the relationship with the ‘vibration motor structure’ is even more confusing. The main object of this title is ‘motor structure’, and the technical solution it wishes to deliver is a reduction in vibration when the motor is activated. Therefore, this translation could be revised as ‘Motor structure with minimal vibration upon activation’.

Another example is in the title ‘Separable riding recording assembly’ (分離式行車紀錄組件), where the recording device can be applied to the use of any vehicle, including a bicycle, motorcycle, or car. Therefore, the word ‘riding’ would restrict the use of the recorder by excluding vehicles that cannot be ridden. The verb should be replaced with ‘drive’, or alternatively, the title could be altered to ‘vehicle recording assembly’.
4.3 Lexical errors

Lexical errors consist of unnecessary words and missing words. In this study, lexical errors accounted for 18.64% of all errors, and were therefore the second most frequently occurring type of error. Sometimes, a title may combine several types of errors, such as ‘INFRARED-DETECTOR solder Water filling device’ (紅外線偵測器防水裝置). Here, we can see examples of orthographic, lexical, and syntactic errors. The first two capitalised words are unnecessarily combined with a hyphen, and the full capitalisation is erroneous, since no emphasis is required. The word ‘Water’ is also arbitrarily capitalised, which is not only confusing for the reader, but also affects the quality of the disclosure. Lexically, the ‘waterproof’ or ‘water-resistant’ feature of this invention was mistranslated as ‘water filling’, which is the opposite of the purpose of the invention. Syntactically, the translated title follows the structure of the source text, but the main object concerned is the ‘water-resistant device’ of the ‘infrared detector’, and therefore should be reordered as follows: ‘Water-resistant device of infrared detector’.

4.4 Semantic errors

Distinct from lexical errors, semantic errors comprise polysemy, homonymy, and the incorrect use of expressions. These errors accounted for 14.6% of all errors found in the present study. One such example is the title ‘CIGS twilight follow control device’ (C I G S 微光追蹤控制裝置), where ‘twilight’ (微光) refers to the light from the sky when the day is coming to an end and night time is beginning, and is remotely related to semiconductor materials, or CIGS. The term ‘CIGS’ is found to collocate with ‘optical’, and therefore, ‘twilight’ should be replaced with ‘micro-optic’. The verb ‘follow’ (追蹤) is also incorrect. The feature of the invention is to ‘track’ (追蹤) CIGS micro-optics, not to ‘follow’ (追蹤) it. Here, the source text is problematic because ‘track’ and ‘follow’ are expressed by the same word in Chinese.

4.5 Syntactic errors

Most of the linguistic errors detected in this study are syntactic errors, which include those relating to conjunctions, prepositions, articles, and word order.

One example of an article error is ‘The mobile power supply structure’ (行動電源供應器結構), where the article ‘the’ is redundant. Since the wording of the title presents the specificities of one single invention, definitiveness is not necessary and therefore the article should be omitted.
With regard to the categories of ‘word choice’ and ‘word order’, one example is the title ‘Green can be provided with a turnable output portable charging device’ (具綠能之可調變輸出可攜帶充電裝置). The source text exhibits a typical structure of a Chinese patent title, where the main subject, ‘charging device’ (充電裝置), is at the end. Placed in front of the subject are its capabilities, ‘adjustable output’ (可調變輸出) and ‘portable’ (可攜帶), and the specific feature of the invention, ‘contains green energy’ (具綠能).

The translation is grammatically incorrect and does not make any sense. The translation of 可調變輸出 (adjustable output) as ‘turnable output’ is an example of word choice error. Judging from the literal sense of 調變, it is a combined term generated from 調整 (adjustment) and 變化 (change), neither of which have any relevance to the word ‘turn’ or the idea of causing something to move around.

In addition to the mistranslation of 具綠能 as the adjective ‘green’, the translation contains word order errors; the colour ‘green’ is misused as a noun and acts as the subject, where the subject should be ‘charging device’. The need to readjust word order usually indicates grammatical errors.

Errors in the preposition category were found mostly in B- and C-rated translations. The average number of prepositional errors in B- and C-rated translations demonstrates that prepositional errors occur frequently in translations with average ratings, and this finding is statistically significant. Errors involving ‘of’ and ‘with’ account for the majority of prepositional errors, most of which were identified in B-rated translations. Among the three most common types of prepositional errors in our data set, the use of the preposition ‘of’ proved to be the most problematic (see Table 4). This is in part because of the translation strategy applied by the translator, and in part because of the translation device adopted by the patent applicant.

According to Cross (2007), patent translations often involve one-to-one correspondence between the source text and the target text, yet employing such a strategy can easily generate difficulties, especially regarding the conversion of prepositions, which may result in misreading or mistranslation.

In the use of the preposition ‘of’, the most common type of error is the misplacement of nouns before and after the preposition. One example of misplaced nouns concerns the title ‘Backup power system of traffic light’ (交通號誌之備援電源系統), which was mistranslated as ‘Traffic light of backup power system.’ The Chinese function word 之 (of/possessive) presents the relationship between the two nouns before and after 之, and is one of the most frequently used words in patent texts (Tsai 2010). In patent documentation language, what lies before 之 is usually the main subject, and the invention, a specific feature, function, or device, is described after 之.

In this example, the main subject is the ‘traffic light’, which is implied to be one part of the ‘backup power system.’ To be valid for the purposes of patent law, the
disclosure of the invention must be sufficiently clear and complete for the invention to be produced by an expert in the field. This being said, any user who drives would know that the translated title is not only inapplicable, but also irrelevant, as it is not necessary to invent a traffic light for a backup power system. The translation also contains word order errors that are the result of word-for-word translation.

An example of the underused (or missing) ‘of’ is the title ‘Heat dissipation device of ignition coil’ (點火線圈散熱裝置), which was translated into ‘Ignition coil heat dissipation device’. In the Chinese source text, the relationship between the two noun phrases, ‘ignition coil’ (點火線圈) and ‘heat dissipation device’ (散熱裝置), is denoted by their positions in the title. Despite the omitted 之 in the Chinese title, the ignition coil can be acknowledged to be one part of the heat dissipation device. However, without specifying the relationship between the two, the translated title is abstruse and confusing. This is another example of word-for-word translation that disregards actual meaning.

Another example of a ‘one-to-one correspondence between source and target’ is the translation of ‘With feedback charging wheel system’ for 具有回饋充電的輪動系統. The first part of the source title, ‘with feedback charging capability’ (具有回饋充電), specifies the function of the subject ‘wheel system’ (輪動系統). The two parts are connected by the function word 的 (of/possessive), which is essentially the same as 之. Here, the result of direct translation is the entire title as a complement with an unidentifiable subject. In this instance, neither the grammatical structure nor the meaning was retained in the translation.

The error count results illustrate that syntactic and semantic errors were the two most common types. When comparing the evaluation results of the specific error categories, D-rated translations were found to contain mostly semantic errors and syntactic errors. From Cross’s definition, a patent is ‘a long, precisely worded legal definition of an invention. As such, the meaning conveyed by the words is important, but so is the wording itself’ (Cross 2007, 19). From these results we can advise that patent translators should work on improving their word choice and grammar.

Cross (2007, 22) described literal translation as ‘an exact and accurate reproduction of the entire content of the source text without embellishment or modification,’ and he stated that patent translators should limit themselves to reproducing precisely what is said in the original patent. However, such precision refers to adhering to both the words and grammatical structure to recreate the original content in the target text. From his perspective, literal translation of patents should ‘(1) reproduce the meaning; (2) reproduce the register; (3) respect sentence breaks and carriage returns; (4) be consistent in the use of vocabulary and phrasing; (5) maintain a one-to-one correspondence between source and target; and (6) provide appropriate annotation’ (ibid.).
Translation errors are the result of a complex interplay between carelessness, strategy, and competence. The title 低複雜度之反轉換計算方法 was translated into ‘Inverse transform calculation method for low complexity’. From previous discussions of 之 (of/possessive), we can safely identify the subject as ‘inverse transform calculation method’ (反轉換計算方法). The first part of the title, 'low complexity' (低複雜度), refers to the feature of the invention. The translation ‘Inverse transform calculation method for low complexity’ shows the result of using such a calculation method, which does not exactly correspond with the source text’s meaning. The preposition ‘for’ should be replaced with ‘with’ to emphasise the uniqueness of the invention.

The overall results demonstrate the average quality of machine-translated patent titles in TIPO. Most suggestions for improvement relate to ‘grammar’ and ‘meaning’. The influence of the linguistic characteristics of titles, including nominal phrases and incomplete clauses, on the errors identified in this study are self-explanatory. A further analysis revealed that syntactic errors (especially word order and preposition use) were the most frequently occurring error type. The linguistic errors found in this study are commonly found in translations, but are particularly frequent in the patent field. In patent title translations, word order and prepositional errors are often a result of literal translation, which is suggested by Cross (2007) as the translation strategy for patent documents. Nevertheless, in Cross’s definition of literal translation, meaning should be reproduced in the target text alongside word-for-word translation. Resulting from the pursuit of one-to-one correspondence, the errors found in this study disregard the meaning of the source text, and therefore can be rated only as mistranslations.

5. Conclusion

The title of an invention acts as an introduction to the text, and provides an insight into the context of the patent application, to allow the reader to understand the significance of the text. In Nelson’s words, ‘a text is more (not) less than a work; the part is greater than the whole’ (Nelson 1998, 14). Thus, the translation of titles creates an awareness that helps the reader to overcome assumptions suggested in the source text. Since the wording in the title is usually the subject of the invention, and often recurs in the patent claim, if the translation of the title is erroneous, then it is likely that the translation quality of other sections will be low.

Our results indicated that most translations contain at least one error. Some error types, such as orthographic errors, have less influence on the overall quality of the translation, and therefore do not significantly decrease the quality rating. Some
common error types, such as semantic and syntactic errors, have greater effects on grammatical structure and meaning, and thus engender lower ratings.

In the literal translation of patents, precise reproduction refers to both the words and the grammatical structure, and that the meaning should be considered along with the need for one-to-one correspondence between the source text and the target text. The examples of prepositional errors mentioned in this study were mostly the result of word-for-word translation without due consideration of meaning. The disregard for structural differences between the source text and the target text produced word order and grammatical errors. The misreading of the relationships between words appearing before and after the Chinese function word 之 further elicits a translation of an invention that is either inapplicable or nonexistent. Consequently, patent translators and trainees alike are advised to pay extra attention when translating prepositions, and to factor the study of prepositions into the early stages of translation pedagogy.

This study provides important new insights into patent language. As with any endeavour of error analysis, a large amount of work must be completed before the research results can be analysed. It is essential to include a substantial set of data along with elaborate processing (POS-tagging and error-tagging). Results should also be interpreted in the light of current translation theory, and incorporated into syllabus and material design. Evidently, this study has major limitations as regards sample size, yet it is a building block in the foundation of the discipline, and thus has the potential to be built on.

Acknowledgements

This work was supported by the College of Liberal Arts, National Taiwan University, under Grant 104R3101.

References


**Author’s address**

Yvonne Tsai

1, Section 4, Roosevelt Road

Taipei City

Taiwan

yvtsai@ntu.edu.tw

**Biographical notes**

Yvonne Tsai is an Associate Professor in the Department of Foreign Languages and Literatures at National Taiwan University. She received her MA in Translation and Interpreting from the University of Bath and her Ph.D. in Translation Studies from Newcastle University. Her major research interests include patent translation, translation technology, and translation pedagogy. Email: yvtsai@ntu.edu.tw.