Shaoyi He  
University of North Carolina at Chapel Hill

Concept-based vs. Word-based Measures of Medical Information Transfer via English-Chinese and Chinese-English Translation of Medical Titles

Abstract: This comparative informational analysis is conducted on 200 article titles and their 200 translations from two English medical journals and two Chinese medical journals. The informativeness of the original and translated medical titles are calculated and compared for concept-based and word-based measures of medical information transfer via English-to-Chinese and Chinese-to-English translation. The results have shown convincing evidences that concept-based measure provides a more accurate and reliable evaluation on medical information transfer via English-to-Chinese and Chinese-to-English translation of medical journal article titles.

1. Introduction

Many scholars have studied translation as information transfer between languages (Anderson, 1978; Carpenter, 1988; Ganeshsundardm, 1980; Godman and Vetman, 1990; Ibuki et al., 1983; Kaliyan and Rao, 1993; Large, 1983; Olshtain, 1986; Riley, 1992), and some of them have studied the information transfer through English-Chinese and Chinese-English translation (Kovacs, 1986; Liu, 1983; Zhang et al., 1994). But, relatively fewer scholars have investigated translation of journal article titles as information transfer between English and Chinese, and vice versa (Loh and Kong, 1977; Wang et al., 1991). Since medical journal article titles function as condensed summaries of medical information transferred in the medical journal articles in English and Chinese, studies on translation of medical journal article titles in English and Chinese offer opportunities to examine medical information transfer between English and Chinese. This paper concentrates on the comparison between the concept-based and word-based measures for medical information transfer via English-to-Chinese and Chinese-to-English translations of medical journal article titles.

Words and concepts both play very important roles in information representation and transfer. Generally speaking, a concept could be represented by a word (Barnwell, 1980, 141; Rowley, 1992, 253), two words (Haas and Losee, 1994, 626; Ku, 1987, 146), or a phrase (Addison, 1991, 8; Yin and Felley, 1990, 134). Also, a concept represented by a word in one language could be represented by a word, or two words, or a phrase in another language, depending on different situations, since "a concept is a recognizable unit of meaning in any given language (Larson, 1984:55)". Take English and Chinese for example. The English word *patient* has its Chinese counterpart *rongrende* when it means "having ability to endure without complaining". But, when it is used to refer to "an individual receiving medical treatment", its Chinese counterpart *bing ren* means "sick person". Another English word *memorialize* could be a Chinese phrase *chengdi qingyuan-shu* which means "present a written statement to authorities for or against something". Since the concepts represented by such single words as *patient* and *memorialize* in English are not necessarily represented by single words in Chinese, how to measure informativeness of journal article titles has become an important and interesting issue for information representation and transfer via article title translation between English and
There are two types of informativeness measure: one is based on word-counting and the other is based on concept-counting. The word-based measure for document titles was first proposed by Richard Diener in 1984. The title informativeness, also known as *Informational Value (IV)*, is defined as "counting the total number of words ($Y_t$) in a title, counting the number of substantive words ($Y_s$), or keywords, in a title, and computing the proportion of substantive words ($Y_p = Y_s/Y_t$)" (Diener, 1984:222). While it was useful for calculating informativeness of English document titles, the word-based measure is unable to account for conceptual representation in medical information transfer between English and Chinese via medical title translation. Thus, such an incapability of the word-based measure calls for another better measure. In this paper, the author proposes a concept-based measure that is called *Conceptual Information Value (CIV)*. It is defined as "counting the total number of words in a title ($W$), counting the total number of concepts in the same title ($C$), and computing the proportion of concepts ($C/V = W/C$)". and used to evaluate medical information transfer via medical title translation between English and Chinese, and vice versa.

The goal of this paper is to conduct a comparative informational analysis between the concept-based and word-based measures of medical title informativeness, via English-Chinese and Chinese-English translation of medical titles, to see which measure provides greater accuracy and is more reliable in information representation and transfer.

2. Materials and Methodology

The corpus for this research include two types of translations: one is English-to-Chinese Translations of medical journal article titles and the other is Chinese-to-English translations of medical journal article titles. The 400 original and translated medical article titles in English and Chinese are as follows:

One hundred original and 100 translated article titles are obtained from two medical journals in English and their translated counterparts in Chinese: 1) 50 article titles in English are from *JAMA* and their counterparts of 50 Chinese translations are from *Meiguo Yixuehui Zazhi Zhongwen Ban* (Chinese Edition of *JAMA*); and 2) 50 article titles in English are from *Archives of Ophthalmology* and their counterparts of 50 Chinese translations are from *Meiguo Yixuehui Yanke Zazhi Zhongwen Ban* (Chinese Edition of *Archives of Ophthalmology*).

One hundred original and 100 translated article titles are obtained from two Chinese medical journals and their translation in English: 1) 50 article titles in Chinese are from *Zhonghua Yixue Zazhi* (National Medical Journal of China) and their counterparts of 50 English translations are from the translated *table of contents* in English; 2) 50 article titles in Chinese are from *Zhonghua Yanke Zazhi* (Chinese Journal of Ophthalmology) and their counterparts of 50 English translations are from the translated *table of contents* in English.

The measurement of title informativeness of the original and translated medical journal article titles are compared and analyzed as follows:

1) The average *Informational Value (IV)* per title of the original medical journal article titles is first calculated and then compared with that of the translated medical journal article titles: the less difference between them, the more accurate the measure, and vice versa.

2) The average *Conceptual Information Value (CIV)* per title of the original and translated medical journal article titles is first calculated and then compared: the less difference between them, the more accurate the measure, and vice versa.
3. Results

First, the average word-based measure is used to calculate the information value (IV) of the original and translated article titles from the two English medical journals and their Chinese counterparts. The results are as follows:

<table>
<thead>
<tr>
<th>Journal Title</th>
<th>Average IV of Original Titles</th>
<th>Average IV of Translated Titles</th>
<th>Difference of IV in Ori. &amp; Tran. Titles</th>
</tr>
</thead>
<tbody>
<tr>
<td>JAMA</td>
<td>0.742</td>
<td>0.828</td>
<td>0.086</td>
</tr>
<tr>
<td>Archives of Ophthalmology</td>
<td>0.753</td>
<td>0.829</td>
<td>0.076</td>
</tr>
<tr>
<td>JAMA &amp; Archives of Ophthalmology</td>
<td>0.748</td>
<td>0.829</td>
<td>0.081</td>
</tr>
</tbody>
</table>

Table 1: IV in English-Chinese translation of medical journal article titles

Second, the average information value (IV) of the original and translated article titles from the two Chinese medical journals is calculated with the word-based measure. The following are the results:

<table>
<thead>
<tr>
<th>Journal Title</th>
<th>Average IV in Original Titles</th>
<th>Average IV in Translated Titles</th>
<th>Difference of IV in Ori. &amp; Tran. Titles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhonghua Yixue Zazhi</td>
<td>0.807</td>
<td>0.732</td>
<td>0.075</td>
</tr>
<tr>
<td>Zhonghua Yanke Zazhi</td>
<td>0.871</td>
<td>0.734</td>
<td>0.137</td>
</tr>
<tr>
<td>Zhonghua Yixue Zazhi &amp; Zhonghua Yanke Zazhi</td>
<td>0.839</td>
<td>0.733</td>
<td>0.106</td>
</tr>
</tbody>
</table>

Table 2: IV in Chinese-English translation of medical journal article titles

Third, the average conceptual information value (CIV) of the original and translated article titles in the two medical journals in English and their counterparts in Chinese is calculated with the concept-based measure. The results are shown as follows:

<table>
<thead>
<tr>
<th>Journal Title</th>
<th>Average CIV in Original Titles</th>
<th>Average CIV in Translated Titles</th>
<th>Difference of CIV in Ori. &amp; Tran. Titles</th>
</tr>
</thead>
<tbody>
<tr>
<td>JAMA</td>
<td>0.434</td>
<td>0.432</td>
<td>0.002</td>
</tr>
<tr>
<td>Archives of Ophthalmology</td>
<td>0.455</td>
<td>0.451</td>
<td>0.004</td>
</tr>
<tr>
<td>JAMA &amp; Archives of Ophthalmology</td>
<td>0.445</td>
<td>0.442</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Table 3: CIV in English-Chinese translation of medical journal article titles

Fourth, the average conceptual information value (CIV) of the original and translated article titles from the two Chinese medical journals is calculated with the concept-based
measure. The results are given below:

<table>
<thead>
<tr>
<th>Journal Title</th>
<th>Average CIV in Original Titles</th>
<th>Average CIV in Translated Titles</th>
<th>Difference of CIV in Ori. &amp; Tran. Titles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhonghua Yixue Zazhi</td>
<td>0.425</td>
<td>0.412</td>
<td>0.013</td>
</tr>
<tr>
<td>Zhonghua Yanke Zazhi</td>
<td>0.403</td>
<td>0.410</td>
<td>0.007</td>
</tr>
<tr>
<td>Zhonghua Yixue Zazhi &amp;</td>
<td>0.414</td>
<td>0.411</td>
<td>0.003</td>
</tr>
</tbody>
</table>

Table 4: CIV in Chinese-English translation of medical journal article titles

4. Discussion

From the tables 1 to 4 which are shown above, we can see clearly the differences between the word-based measure of information value (IV) and the concept-based measure of conceptual information value (CIV). Quite obviously, the concept-based measure of conceptual information value (CIV) has yielded much better results for both English-Chinese and Chinese-English translation of medical journal article titles.

Let us first look at the word-based measure of information value (IV) in both English-Chinese and Chinese-English translation of medical journal article titles. On the one hand, for the word-based measure of the information value (IV) in the English-Chinese translation, the information value (IV) of the translated article titles in Chinese is about 8.1% more than that of the original article titles in English. That is to say, the translated article titles in Chinese have more substantive words than their English counterparts. On the other hand, for the word-based measure of the information value (IV) in the Chinese-English translation, the information value (IV) of the translated article titles in English is about 10.6% less than that of the original article titles in Chinese. In other words, the translated article titles in English have fewer substantive words than their Chinese counterparts. With such differences between the original and translated information value (IV), it is not reliable to measure the informativeness of the original and translated article titles just by counting their substantive words in English and Chinese.

As for the concept-based measure of conceptual information value (CIV) in both original and translated medical journal article titles, the results are much better in both English-Chinese and Chinese-English translation. In the translation from English into Chinese, the concept-based measure of conceptual information value (CIV) in the translated article titles in Chinese is only about 0.3% less than that of original titles in English. This means that there is hardly a change in the conceptual information value (CIV) with the concept-based measure for English-Chinese translation of medical journal article titles. Such is the same case with the translation from Chinese into English. The concept-based measure of conceptual information value (CIV) in the translated article titles in English is also about 0.3% less than that of original titles in Chinese. This again means that there is almost no change at all in the conceptual information value (CIV) with the concept-based measure for Chinese-English translation of medical journal article titles. With almost no difference between the original and translated conceptual information value (IV), it is thus reliable to measure the informativeness of the original and translated article titles by
counting the concepts in English and Chinese.

The above results also indicate the advantage of concept-based measure of conceptual information value (CIV) in translation as information transfer between English and Chinese because the informativeness of article titles in both English and Chinese is solely represented by concepts rather than words. In addition, the differences between word-counting and concept-counting for title informativeness have shown the latter is more practical in measuring conceptual information transfer between English and Chinese, and vice versa.

5. Conclusions

Based on the comparative informational analysis between the concept-based measure and the word-based measure, this paper shows convincing evidences that concept-based measure is more accurate and thus more reliable for evaluating medical information transfer via English-to-Chinese and Chinese-to-English translation of medical journal article titles. Further studies with a larger research corpus of medical journal article titles in both English and Chinese are needed to validate the results of this exploratory comparative analysis. The methodology used in this study to compare word-based and concept-based measures of medical article titles in English and Chinese should then be applied to article titles in different fields other than medicine and different languages other than English and Chinese.

References


Unschuld, P. U. (ed.), *Approaches to Traditional Chinese Medical Literature*, 85-96.


