論文の内容の要旨

論文題名　Evaluation and Use of Online Machine Translation of Nursing Literature in Japanese Nursing

Population

（看護学文献を用いたオンライン機械翻訳：看護ユーザの評価と利用状況に関する研究）

氏名　穴　沢　良　子

The use of research to justify nursing interventions has become a prevalent issue among nurses in evidence-based practice, and Japanese nurses are increasingly required to have a good knowledge of the research literature that is relevant to their clinical practice, education, and research activities. However, most of this literature is in English, and thus nurses whose native language is not English may face difficulties. Among nursing professionals, the requirement to read this literature varies between clinical nurses and nursing researchers/educators. Clinical nurses should read research papers and make use of them in their clinical care delivery, whereas nurses at educational institutions should read a substantial number of research papers to support their research and educational activities; these nurses have a greater requirement for English language skills. Therefore, a freely accessible online machine translation (MT) tool is desirable to help nurses access this research material.

MT is a system whereby text in one language is translated into another. Such systems have long been used to aid to multilingual communication worldwide. The emergence of the Internet has boosted the number of people who wish to read documents in foreign languages, and MT systems offered via the Internet (online MT) have become increasingly popular and convenient. The quality of MT systems has been evaluated using both manual and automatic methods. In manual evaluation, people typically evaluate the translation quality according to the criteria of fidelity (or accuracy) and intelligibility of the MT output. Automatic evaluations are typically carried out using statistical algorithms. The major advantage of automatic evaluation is the reduced workload, including time required for the evaluation. However, automatic evaluations do not consider the meaning of the translated passages, which is only possible to assess through human evaluation. Many existing approaches to the evaluation of MT have been conducted in different fields. For MT of documents related to the medical domain and concerning the Japanese language, it has been reported that the evaluation criteria should include structural accuracy and intelligibility of the translated sentences, as well as the proper translation rate (PTR) for specific

terms. However, very few additional studies have addressed MT evaluation for medical documents in Japan, and no attention has been paid to the use of MT for nursing literature. The appropriateness of evaluation criteria for documents in the nursing field remains unknown, despite the requirement for Japanese nurses to access research reports that are in English. In this study, we examined and discuss the feasibility of online MT technology used by nursing professionals to obtain technical information from articles written in a foreign language.

There were two phases to this study. In Study 1, an existing method was used to evaluate the quality of online MT through manual evaluation. The structural accuracy, intelligibility, and the usability of the translations were discussed. Two language pairs were considered: Korean–Japanese (KJ) and English–Japanese (EJ). In total, 297 English and 290 Korean sentences from the abstracts of nursing journal articles were evaluated for accuracy and intelligibility. Then 75 words were identified as specific terms in English sentences. The inter-rater reliability (IRR) of the evaluation method was measured by two raters. The results showed a sufficient level of reliability of the existing evaluation method. The EJ accuracy was $r = 0.778$ with $P = < .001$, and the EJ intelligibility was $r = .752$ with $P = < .001$. The KJ accuracy was $r = .729$ with $P = < .001$, and the KJ intelligibility was $r = .620$, $P = < .001$. A readily accessible online MT system offered in Japan was identified as Google Translate (GT), which performed the best in terms of intelligibility with both language pairs (EJ: 2.93 and KJ: 4.04) and PTR (BizLingo-GT, $P < .001$; Cross Language-GT, $P < .001$; Bing Translator -GT, $P < .015$). Specific terms appeared to influence the extent of the intelligibility of the translations.

Then we included more samples of the nursing population to examine the usefulness of online MT systems. The limitations of this part of the study include the relatively small number of data sources, limited availability of evaluators, and possible bias in the process of testing the IRR. The raters knew what each online MT system was when evaluating the translations. Furthermore, the order in which the raters evaluated the translations of each online MT system was not considered. A single-blind procedure should be used to eliminate possible influences on scoring due to raters' assumptions and bias.

In Study 2, we investigated the intelligibility of EJ and KJ translations using GT in a larger nursing population. The current status of MT use and its perception by nurses was explored. Translations of nursing article abstracts using GT were evaluated by 250 randomly selected nursing faculty members at nursing colleges across Japan, using a previously verified method to evaluate intelligibility. Abstracts were evaluated according the title, background, aim, method, result, discussion, conclusion, and relevance to clinical practice. We also asked for participants' impressions of the perceived usefulness of the

translated abstract as a whole. The results indicate that GT had a minimum acceptable quality in EJ translations, and the KJ translations were considered to be significantly more intelligible than English–Japanese (for intelligibility, EJ, M = 2.39, KJ, M = 3.38, t = 28.349, P = <.000; for usefulness EJ, M = 2.74, KJ, M = 3.75, t = 10.354, P = <.000). Word and character counts, as well as the existence of unidentifiable words, were considered factors that were indicative of the quality of the translation. Participants answered a questionnaire examining the current use of online MT and the perceived usefulness among Japanese nurses, along with the number of articles read in English and the perceived language barrier. The items were rated on 5-point Likert scale. Of the participants, 73.8% had used online MT, and of these, 63.7% felt it was useful. The language barrier was felt to be strong, and academic degrees and English proficiency levels were associated factors. The perceived language barrier was related to the frequency of online MT use (F = 5.195, P = .007). No associated factor was found for the perceived usefulness of online MT.

Language proficiency is an important factor for optimum utilization of MT. Nurses may benefit from GT when reading literature in a foreign language; however, the results of this work suggest that an effort should be made to improve the performance of online MT, and we emphasize the importance of providing training and education for nursing professionals to utilize this technology. We may assume that those with less of a language barrier have less difficulty using MT when reading nursing literature in English. As previous studies suggest, language knowledge is required for effective use of MT. English language training for nurses is actually even more desirable for post-graduate and continuing education students.

The limitations of Study 2 include using only one online MT system, having participants evaluate translations without reading the English source texts, and limiting participants to two types of nursing faculty member. Nursing professionals from a broader range of settings and levels may be necessary for a deeper understanding of general trends. In addition, the participants' reading and writing skills should be assessed, and an analysis based on such criteria would be desirable. Recall bias may have been introduced when the participants were asked about the number of articles read within the previous 3 months. A separate analysis of source sentences for reports using qualitative and quantitative research methods is expected to be valuable, because researchers tend to use different linguistic styles and terminologies in the two approaches.

Future work should be conducted to analyze the translations according to the types of nursing research to assess how online MT can be useful for understanding nursing literature. Evaluating MT in an

environment in which raters compare translations from the system with the sentences of the source language is also expected to be beneficial, as well as exploring better ways of using online MT systems according to a user's level of language proficiency.

Issues related to reading articles published in foreign languages, especially in English, should be considered in many other non-English-speaking countries. The work described in this thesis will provide a starting point for further exploration of possible linguistic support for nurses via advanced Internet-based technologies, leading to more effective use of MT tools, and eventually more opportunities for nurses to access internationally published material and obtain up-to-date knowledge for better care delivery.