

博 士 論 文

**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)

穴 沢 良 子

Table of Contents

LIST OF APPENDICES

ABBREVIATIONS

PUBLICATIONS RELATED TO THE DISSERTATION

ACKNOWLEDGMENT

ABSTRACT1

Chapter 1. Introduction

1.1 Necessity of reading English literature for nursing professionals4

1.2 Translation and MT 3

1.3 Overview and evaluation of MT: Literature review5

1.3.1 Development of methods of MT4

1.3.2 The era of the Internet: Online MT- free, easy-access, convenient language tool
.....7

1.3.3 Evaluation of MT quality

1.3.3.1 Evaluation methods- manual and automatic8

1.3.3.2 Evaluation of online MT systems.....9

1.3.3.3 MT Evaluation in health-related materials.....11

1.4 MT for nursing purposes.....13

1.5 Scope.....14

Chapter 2 Study 1: Testing Evaluation Method

2.1 Purpose16

2.2 Methods

2.2.1 Measurements.....16

2.2.2 Online MT systems17

2.2.3 Preparation of study materials.....18

2.2.4 Procedural steps.....20

2.2.5 Data analysis21

2.3 Result

2.3.1 IRR of the evaluation method.....21

2.3.2 MT translation evaluation

2.3.2.1 EJ translations	22
2.3.2.2 KJ translations	25
2.4 Discussion	
2.4.1 Reliability of the evaluation criterion of online MT.....	26
2.4.2 Which is possibly the best online MT system?.....	27
2.5 Limitations.....	30
2.6 Conclusion	30

Chapter 3 Study 2: Evaluation by and Current Status of Online MT Use in Nursing

Users

3.1 Purpose	31
3.2 Methods	
3.2.1 Preparation of study material	33
3.2.2 Study participants and collection of data.....	35
3.2.3 MT evaluation measurements.....	35
3.2.4 Questionnaire items for online MT use	36
3.2.5 Data analysis	37
3.3 Ethical Considerations.....	38
3.4 Preliminary Study.....	39
3.5 Results of Main Survey	
3.5.1 Participants' characteristics	42
3.5.2 GT evaluation results of EJ and KJ translations	
3.5.2.1 Evaluation of EJ translations	45
3.5.2.2 Evaluation of KJ translations	46
3.5.2.3 Comparison of the performance between EJ and KJ	47
3.5.3 Characteristics of source sentences and evaluation results.....	47
3.5.4 Use of online MT systems among nursing population	
3.5.4.1 Previous and current use of online MTs and perceived usefulness.....	50
3.5.4.2 Differences in the number of English articles read and perceived language barrier according to participant characteristics	52
3.5.4.3 Differences in the frequency of use and the perceived usefulness of online MTs according to participant characteristics.....	54
3. 6 Discussion	

3.6.1 How are GT translations rated by Japanese nurses?	57
3.6.2 Influence of the characteristics of source sentences on the evaluation results.....	58
3.6.3 Participants' online MT use and its perceived usefulness.....	59
3.6.4 Factors affecting the frequency of online MT use and its perceived usefulness.....	60
3.6.5 Practical and educational implications.....	61
3.7 Limitations.....	64
3.8 Conclusion.....	66
 Chapter 4 Conclusions	
4.1 General Summary	68
4.2 Future Works.....	70
4.3 Concluding Remark	72
 Contributions	73
 References	73

LIST OF APPENDICES

1. Questionnaire form for user evaluation of GT translations by nursing participants (one of 23 kinds of forms)
2. Questionnaire form for participants' demographic background, frequency of reading nursing literature in foreign languages, frequency of experiencing a language barrier, and online MT use and perceived usefulness

ABBREVIATIONS

ALPAC	Automatic Language Processing Advisory Committee
APRA	Advanced Research Project Agency
BL	BizLingo
BT	Bing Translator
CL	Cross Language
EBP	Evidence-based Practice
EJ	English to Japanese
GT	Google Translate
HSD	Tukey's honestly significant difference
IRR	Inter-rater reliability
KD	Kodensha
KJ	Korean to Japanese
MT	Machine Translation
PTR	Proper Translation Rate

AUTHOR'S PUBLICATIONS RELATED TO THE DISSERTATION

Anazawa R, Ishikawa H, Kiuchi T. Evaluation of Online Machine Translation by Nursing Users. *Computers, Informatics, Nursing*. 31(8) 2013.

Anazawa R, Ishikawa H, Park MJ, Kiuchi T. Online Machine Translation Use with Nursing Literature: Evaluation Method and Usability. *Computers, Informatics, Nursing*, 31 (2), 59–65, 2013.

Anazawa R, Ishikawa H, Takahiro K. Use of Online Machine Translation for Nursing Literature: A Questionnaire-Based Survey. *The Open Nursing Journal*, 7, 22-28, 2013.

Anazawa R, Ishikawa H, Park MJ, Kiuchi T. Preliminary study of online machine translation use of nursing literature: quality evaluation and perceived usability, *BMC Research Notes*, 5:635, 2012.

Anazawa R, Ishikawa H, Kiuchi T. The perceived need for Japanese nursing faculty members to learn English: issues related to career development. *Journal of Continuing Education in Nursing*, 43(4), 188–191, 2012.

ACKNOWLEDGMENTS

I would like to gratefully acknowledge the instruction and assistance given by Dr. Takahiro Kiuchi and Dr. Hirono Ishikawa. I also would like to thank Dr. Park MJ, Dr. Fujisawa Asami, and Ms. Kurihata Toshie for their great cooperation.

Furthermore, I am grateful to all the participants of the study survey, finally, to my family who have been very patient with their unfailing support for me.

ABSTRACT

Background: Nurses in Japan are increasingly required to read international journals to stay abreast of developments in their fields; however, much of this material is in English, and so a language barrier exists. Freely accessible online machine translation (MT) offers a possible solution to this problem. However, the quality of online MT for nursing-related material has not been reported, and the current prevalence of use of MT among nurses is also unknown.

Methods: Translations of nursing abstracts using four online MT systems were evaluated by two expert raters using an existing manual evaluation method. The evaluation criteria were intelligibility, structural accuracy, and the proper translation rate for specific terms. In addition, 250 nursing faculty members across Japan were asked to evaluate the intelligibility of online MT using a 5-point Likert scale. The current status of MT use by nurses was also explored.

Results: Inter-rater reliability was secured for the evaluation method for both English–Japanese and Korean–Japanese language pairs. Google Translate (GT) scored highest in terms of both intelligibility and PTR. Among nursing respondents, the quality of GT English–Japanese translations was minimally acceptable ($M = 2.39$). The intelligibility score was lower when source texts had more word/letter counts and unidentifiable words. A total of 73.8% of respondents had used online MT, and nurses who used MT more often experienced less of a language barrier ($F = 5.195, P = .007$).

Conclusion: The word count of source texts and the existence of unidentifiable words are indicative of translation quality. The perceived language barrier is related to the frequency of online MT use. Language proficiency is an important factor affecting the utilization of MT. There is a requirement for education in English, reading scientific papers, and online MT training. Cooperation with developers and providers of MT to improve these systems is expected to be beneficial.

Chapter 1

Introduction

1.1 Necessity of reading English literature for nursing professionals

The use of research to justify nursing interventions has become a prevalent issue among nurses in evidence based practice (EBP) worldwide and also Japanese nurses are increasingly required to have knowledge of the research literature published internationally that is relevant to their clinical practices, education, and research activities [1,2]. In addition to requiring literacy in terms of using computer technology, nurses need to have the ability to understand the information in nursing-related literature [3,4]. The language barrier can be a significant obstacle for nurses in non-English-speaking countries such as Japan, because most major research journals are published internationally in English [5,6].

In Japan, nurses are required to improve their clinical competency by maintaining an up-to-date knowledge of evidence-based nursing (EBN) practices. EBN is defined as “utilizing scientific evidence that is available and the best possible at the present time rather than conventional nursing care based on the experiences and knowledge of experts” [7]. Many Japanese hospitals require nurses to be involved in clinical research activities as a part of their continuing education program. Clinical nurses typically face the need to read nursing and medical articles written in English. They should read current research and make use of this work in their clinical care delivery. Nursing faculty members at nursing colleges (i.e.,

educators and researchers) should read a substantial number of research papers in international journals to support their research and educational activities.

At the same time, however, Japanese nurses do not often use English on a daily basis, providing little incentive to learn English, even though they understand the need to review research published in English. In fact, it is more important for Japanese nurses to be able to obtain nursing information quickly to provide quality care than it is to spend time mastering the English language [6]. Hence, along with English education for nurses, it is also important to provide Japanese nurses with convenient access to the information provided in international nursing literature in a way that is easy for them to access and to understand. To this end, the use of machine translation (MT) tool, which generates automatic translations and is readily accessible, may be an important step. This could help these nurses to read English research even if the translations are not perfect. Especially for nursing faculty members, who are in need to read English literatures but lacking English language proficiency, desperately would want to find any tool to lessen the burden. Indeed, an increase in the nursing population that uses MT could potentially create a better environment for accessing and retrieving high-quality nursing information [6].

1.2 Translation and MT

Translation is converting a semantic content (message) being expressed in one language into an expression in another language retaining the substantial meaning of the original content as much as possible [8]. It is also “reproducing in the receptor language the closest natural equivalent of the source-language message, first in terms of meaning and secondly in terms of style” [9]. This process can be divided into the following levels: translations depending on the syntax (replacing the words and structure of given sentences with those of the corresponding language), translations based on semantics (generating sentences so that the uttered sentences can be equivalent to the target language in meaning based on communication theory), and translation based on pragmatics (generating translations of uttered sentences which can render the same understanding for both two different language societies) [10].

MT is mostly used for assimilation, that is for comprehension [11]. For MT systems, generating translations with understanding is an ongoing difficult challenge [8]. It is impossible for MT systems to achieve the generation of the translation which is most appropriate interpreting the utterances in relation to the situation and social knowledge producing, which is just as a human approach [10]. However, experts of the technical fields may better understand the translations from MT systems inferring the contents of the original utterances. Sentences include only facts can be most tractable by MT system, and technical

documents which are not related social and cultural aspects as far as possible are better to be translated [10].

In the next section, the major points in MT technology will be described.

1.3 Overview and evaluation of MT: Literature review

1.3.1 Development and methods of MT

MT has long been used as an aid to multilingual communication worldwide. It is the system in which text in one language is automatically translated into a different language [12].

The original form of MT began in the United States in 1933, and with the emergence of computers, the idea of MT truly materialized during the 1950s [13, 14]. In Japan, a national MT project for Japanese/English science journals, called the Mu Project, was conducted as early as 1982 to 1986 [13-16]. Different kinds of MT approaches have been studied and used.

Originally, transfer-based MT in which it analyzes the input text for morphology and syntax to generate a translation was typically used [17]. Statistical and example-based methods based on parallel translations and bilingual corpus data emerged, and they have been the major trends in MT [18]. Statistical-based MT is used in popular online MT systems such as Google.

MT systems have been developed for various domains of documents for users who need to read documents in foreign languages. Before the era of the Internet, MT systems were in limited use and were typically expensive [13].

Methods of MT are generally divided into three types: rule-based, example-based, and statistical MT. The followings are the basic characteristics of these MT methods [19-22].

1) Rule-based MT: conducts the translation by language generation in the target language after the conversion between two languages, analyzing grammatical components of the source language. Transfer method is included in rule-based method. It conducts morphological analysis, syntactic parsing, and semantic analysis for the source language to make the structure which depends on the source language, then transfers it into the structure which depends on the target language generating sentences in the target language. The term, rule refers to mainly grammatical rules (structure of the source sentences) and converting rules (differences between the source languages and translated sentences).

2) Example- based MT: conducts translation based on stored translation examples before-hand, retrieving translation examples similar to the words entered into the system. Entered sentences are broken down into several parts, and the resembled examples of translations are combined to be translated.

3) Statistical-based MT: uses a dataset of a huge number of parallel translations. This method has been greatly improved in recent years. Statistical MT makes translations probabilistically based on statistics obtained from the statistical calculation of words and phrases with the highest probability which come after one word or phrase.

Parallel translation data are used in both rule-based and statistical MT methods.

Statistical MT method disassembles the original sentence into words making maximum efforts to generate a translation with preferred matches of words which occur frequently.

Contrarily, rule-based MT prioritizes larger matches of examples to the extent possible.

1.3.2 The era of the Internet: Online MT- free, easy-access, convenient language tool

The emergence of the Internet has boosted the number of people who want to read documents in foreign languages around the world, and MT systems offered via the Internet have become popular and increasingly convenient. Online MT systems available via the Internet are freely accessible and may enable Japanese nurses to access technical literature published in English. Many such MT systems are now available [23-25], and although many of them have problems with translations and have much room for improvement, they are generally valuable, practical, free of charge, and commonly used.[13, 26, 27] Some Japanese hospital nurses have a positive view of using an MT system as an aid for reading nursing information written in English[28]. Hence, it is important to mention that online MT systems could help Japanese nurses access and use research to improve their clinical abilities in EBP. The current situation of the use of online MT systems by Japanese nursing professionals is not known [29].

1.3.3 Evaluation of MT quality

1.3.3.1. Evaluation methods- manual and automatic

Traditionally, MT quality has been evaluated both manually and automatically. In manual evaluation, raters evaluate translation quality according to the criteria. ALPAC (Automatic Language Processing Advisory Committee) approach is a representative human evaluation method which measures fidelity (or accuracy) and intelligibility of MT outputs: fidelity (or accuracy) measures how much information the translated sentences retained compared to the original sentences (on a scale of 0 – 9) and intelligibility was a measure of how understandable the automatic translation was (on a scale of 0 – 9) [30]. ARPA (the Advanced Research Project Agency) started the evaluation system which has been typically used to now this date [31, 32], and it consists of criteria of adequacy and fluency. These two criteria usually use a 5-point Likert scale as shown in Table 1 [30-32]. Questions for the two criteria include “How much of the meaning expressed in the gold-standard translation is also expressed in the target translation?” and “How do you judge the fluency of this translation?” for fluency. In Japan, many previous studies evaluated translations quality from MT with fluency and adequacy as well as structural “fidelity” and “intelligibility” of meanings translation outputs [33-39]¹. An early national project of MT evaluation called “Mu Project”¹

¹ Mu Project was conducted by the science and technology agency of Japan from 1982 to 1985 aiming at developing the MT system for translating paper abstracts of science and technology fields in English to Japanese and Japanese to English.

used the evaluation method of using intelligibility (on a scale of 1 – 5) and fidelity (on a scale of 0 – 6).

Table 1. Two criteria for human evaluation

Score	Adequacy*	Fluency**
5	All	Flawless
4	Most	Good
3	Much	Non-native
2	Little	Disfluent
1	None	Incomprehensive

Automatic evaluations are typically conducted using statistical algorithms. One such method is the Bilingual Evaluation Understudy, which examines the degree of similarity between the translation outputs of an MT and one prepared by human translators [40, 41]. Reduced human cost, including time required for the evaluation, is the major advantage of automatic evaluations. However, automatic evaluations that gauge translation quality in this way do not consider the actual meaning of the translated sentences, which is possible only by human evaluation [42].

1.3.3.2 Evaluation of online MT systems

Many studies have attempted to evaluate MT quality for various languages and domains, and such studies have frequently used online MT systems such as Babelfish, Google Translate (GT), Systran, and Freetranslation [43,44]. Although the methods of evaluation (ie,

human or automatic) have varied depending on the study, the online MT system of Google has been reported to outperform other systems [29, 34, 44-46].

Previous studies on MT users have indicated various points of view regarding the use of online MT and wide-ranging perceptions of usability by general users [24, 47-50]. Some user feedback on MT translations *via* the Internet has demonstrated a wide range of opinions, varying from gratitude to harsh criticism of the output [24]. Surveys on MT use by general Internet users (which includes a large portion of online MT users) have been conducted in Japan and many different ways of using this tool have been suggested [48]. Also, Japanese English learners have reported the perceived usefulness of online MT in a large number of subjects [49]; moreover, students whose level of English proficiency was intermediate or higher were able to assess the quality of the MT systems on websites more critically [50].

According to a user survey conducted by Japanese research companies, GT was rated as best among the online MT systems offered in Japan, followed by the systems in the search engines of Yahoo and Excite, although most users responded that they were not satisfied with any system currently available [51].

In previous studies, factors of the source sentences that may have influences on users' comprehension of MT translations have been indicated. Some studies have addressed the factors such as sentence length, complexity of sentences, and unknown words (i.e., those that

could not be translated by the MT systems) [53, 55]. These factors should be considered for translations of nursing literature.

1.3.3.3 MT Evaluation in health-related materials

Regarding medical- and healthcare-related documents, one study addressed the usability of online MT systems for multilingual health information (medical records) [27]. This study, which evaluated online MT translations using criteria of comprehensibility and accuracy, suggested some potentially useful features of MT systems and mentioned that future improvement of MTs may help address language barriers [15, 27]. A study to examine the usefulness of GT translations for drug information was conducted by using the evaluation criteria of adequacy and inadequacy [52]. Also, another study examined the quality of GT for patient educational materials with the language pair of English-Spanish [53]. The study results suggest that the quality of GT is enough only when the original sentences are short and that GT's performance is equivalent to professional translators in terms of retaining information and meanings of the source message [53].

Regarding MT evaluation with documents related to the medical domain concerning the Japanese language, one study reported in the English-Japanese pharmaceutical literature for medical doctors considered the evaluation criteria of structural accuracy and intelligibility of the translated sentences as well as the proper translation rate (PTR) for specific terms[54,

55]. Table 2 shows the evaluation items for the two criteria of structural accuracy and intelligibility used in the previous study [55]. Since then, very few other studies have addressed MT evaluations for medical document translation in Japan, and no attention has been given to MT use for nursing literature. The appropriateness of evaluation criteria for documents in the nursing field remains unknown despite the need for Japanese nurses to access research information in English. In order to know the quality of online MT outputs for nursing literature written in the English language, reliability of the evaluation criteria needs to be verified. Also, nurses may be benefitted from knowing the most usable online MT system among currently available ones, since several systems are being offered via the Internet retrieval engines. Furthermore, in terms of the evaluation of online MT systems by users, ones by general users have been reported as state above, but there has not been by health professionals, or even by nurses. To accelerate online MT use in nursing practice, it is essential to evaluate the use of this tool by nursing populations. However, such investigations have rarely been conducted.

Table 2. Evaluation scale items for the two criteria

Accuracy	Intelligibility
5. Accurate	5. Clear and natural
4. Accurate, but some problems with translation of general use terms or others	4. Fully understandable, if not natural expression
3. Only one mistake in terms of basic sentence structure sentence is properly understandable.	3. Easily surmised. The basic thread of the whole
2. Two mistakes in terms of basic sentences structure, or mistake in recognition of part of speech of word unit	2. Not easily surmised. Only some parts of the sentence are understandable
1. More inaccurate than the aforementioned standards	1. Not intelligible

1. 4 MT for nursing purposes

As stated earlier, there is an imminent need to be able to read English professional literatures among nursing population regardless of the areas of clinical or educational contexts. MT technology, then, might be proposed as an aid which makes this situation improved. Also mentioned above, MT tools are widely used globally in various fields and populations, and they have much potential to contribute to the nursing community. Even for nurses with some proficiency in English, MT may improve the speed of assimilating written material, as well as improve their English language skills. The rate of improvement in language proficiency varies significantly between individuals.

To date, online MT evaluation studies have only been conducted on healthcare consumer users. However, it is expected to be beneficial to conduct a study on Japanese

nurses, who need to read journal articles that are published in English. In addition, there are many nurses whose native language is not English worldwide, who face the same problem.

1.5 Scope

The framework of this thesis is based on nursing informatics in health communication in preference to the perspectives of engineering science or natural language processing. In health communication, the transmission of information among health professionals is a significant factor in improving health care delivery. Nursing informatics applies information technology to the healthcare skills and tasks of nurses. It deals with the literacy of nurses in computer science and informatics for communicating data, information, and knowledge of nursing practice [56]. If nurses want to read nursing articles in an unfamiliar foreign language, they are hindered from utilizing relevant research and are thereby deprived of an opportunity to extend their knowledge. MT technology may contribute to solving this problem of language barrier by providing nurses in non-English speaking countries the additional motivation of being able to read technical literature with fewer burdens, improving their clinical, educational, and research practices and eventually leading to greater patient satisfaction. This study attempts to examine and discuss the feasibility of online MT technology use by nursing professionals to obtain technical information from articles written in an unfamiliar foreign language.

In addition, multiple language pairs were needed to be considered in the view of more potential use of MT systems. The Korean language was one option because nursing exchanges between Japan and its neighboring country South Korea have increased markedly in recent years in all clinical, education, and research fields. Accessing nursing information written in Korean might be an important option for Japanese nurses in the future. In terms of translation quality of MT outputs between Japanese and Korean, due to the similarity of these two languages, their usability has been said to be better than other pairs of languages [57-60].

In this dissertation paper, two separate studies are included. In Study 1 (Chapter 2), the investigation of method to evaluate online MT systems with nursing literature currently provided within Japan is tested. Then in Study 2 (Chapter 3), the quality evaluation of online MT translations by nursing users as well as the status of its usage are investigated and discussed. Investigations were conducted based on the research questions guided by the literature review. Specifically, the research questions that ask the reliability of existing evaluation criteria of online MT systems and that confirm possibly the best online MT system for nursing literature were investigated in Study 1; and those ask the evaluation scores of online MT translations by Japanese nursing users, the association between the characteristics of source sentences and intelligibility of the online MT outputs, and how Japanese nurses use online MT and perceive the usefulness were inquired in Study 2. In Chapter 4, an overall discussion of the studies is described.

Chapter 2

Study 1: Testing Evaluation Method

2.1 Purpose

This part of the study aims to report the performance quality of freely accessible online MT systems for nursing literature, examining the reliability of the evaluation method for translations from English to and Korean to Japanese (EJ; KJ). Specifically, the following research questions were investigated:

- (1) Are evaluation criteria of online MT systems for nursing literature reliable?
- (2) Which online MT system(s) is best for enabling Japanese nurses to read nursing literature published in English and Korean?

2.2 Methods

2.2.1 Measurements

As mentioned above, translation quality depends strongly on meaning, and this is qualitative in nature. Thus, human evaluation, which has traditionally been conducted, may be suitable for the evaluation of translations from MT systems in view of the need to consider sentence structure and comprehensiveness of the meaning of translations [42]. For this reason, it was decided to use the previously developed evaluation method for MT systems in the medical domain [54,55], in which the evaluation criteria included the accuracy of sentences

and the intelligibility of content (Table 1). Also, specific terms, by the PTR for EJ translations (PTR is not applicable to KJ translations due to the nature of the Korean language), were examined to evaluate this element contributing to MT usability. Specific terms are operationally defined as terms that depend on the context of the sentences for their meaning [54, 55]. For example, the process “administration of medications/drugs” means “giving (someone) medications/drugs” in the medical context. The noun “complication” means making something complicated, as a general term, but it is a secondary disease or effect in the medical context. The PTR thus represents the translation quality of semantic aspects. Specific terms usually do not affect sentence structure, but they may influence contextual comprehensibility, as the examples above show. The extent to which PTR is achieved was found to be associated with the level of intelligibility in a previous study [55]. In this study, the PTR was determined by judging whether the specific term was translated appropriately or not based on references in major encyclopedias and dictionaries, such as the Nursing Science Dictionary (version 5), Steadman’s Electronic Medical Dictionary (version 5.0), and the online Life Science Dictionary.

2.2.2 Online MT systems

Freely accessible online MT systems offered via Internet search engines in Japan were used for the analysis. Table 3 lists the online MT systems with the abbreviations of their

names used in this study. The four systems for each language pair shown are offered by major Internet search engines that have a high share of users in Japan, and they are provided by different manufacturers. At this point of time, Excite and Nifty offer different systems for EJ and KJ: BizLingo for EJ and Kodensha for KJ (Table 3).

Table 3. Online MT systems used in this study and corresponding Internet search engines (as of July, 2011)

Online MT Systems	Abbreviations	Internet Search Engines in Japan
English to Japanese		
Cross Language	CL	Yahoo, Infoseek, So-net, Livedoor
Google Translate	GT	Google
Bing Translator	BT	MSN
BizLingo	BL	Excite, Nifty
Korean to Japanese		
Cross Language	CL	Yahoo, Infoseek, So-net, Livedoor
Google Translate	GT	Google
Bing Translator	BT	MSN
Kodensha	KD	Excite, Nifty

2.2.3 Preparation of study materials

In total, 297 English and 290 Korean sentences from texts of nursing research were translated via MT for analysis. The number of sentences translated was decided based on previous studies in which approximately 100 to 400 sentence units of translation were used as well as on feasibility with respect to manual evaluation [15, 25, 53, 55].

The sentences used were from 23 nursing research abstracts from prominent nursing journals (eg, *Journal of Nursing Research*, *Journal of Advanced Nursing*, *Journal of Clinical Nursing*), which were freely accessible via the Internet. The subject areas (cancer, psychiatric, maternal/child, geriatric, and chronic illnesses) were selected according to the presented papers at the annual convention of the Japan Academy of Nursing Science [61], presuming that these areas were of major interest among the Japanese nursing population. The nursing abstracts were randomly extracted from the Pubmed queries in reverse chronological order by date of publication.

The source nursing abstracts needed to be consistent in the English and Korean languages for common understanding of the contexts in both languages; therefore, the Korean abstracts used in this study were translations from English. The English abstracts were translated into natural Korean expressions by a speaker who is a nurse researcher, and is also trilingual in Korean, Japanese, and English. These Korean translations were then further validated by a Japanese professional medical translator with skill in English-Korean-Japanese translation. Thus, the naturalness and authenticity of the Korean versions of the abstract texts have been secured.

2.2.4 Procedural Steps

The study, conducted between July and October 2011, followed the procedure outlined below.

(1) All sentences in English and Korean were translated into Japanese by each of four MT systems and entered into MS Excel sheets for evaluation. Each sentence was to be evaluated by the existing method using criteria of both structural accuracy and intelligibility for all four MT systems.

(2) Specific terms for EJ translations were identified from all sentences and counted numerically and using the PTR. The percentage of properly translated specific terms according to reference dictionaries was calculated as the PTR.

(3) Two researcher evaluators (raters A and B) for each language pair were involved in this study. Rater A was a professional medical translator and an RN being in charge of the evaluation for both language pairs; rater Bs were an RN with high English proficiency (language interpreter level) and graduate level education and a professional translator of KJ with various experiences in translating technical documents including medical and healthcare related ones. These personnel evaluated randomly extracted sentences (88 sentences for each language pair: 352 sentences in total by the four online MT systems) to examine the reliability of the existing evaluation method. These were evaluated for both structural accuracy and intelligibility; then the inter-rater reliability (IRR) between the raters was calculated.

(4) All sentences were then evaluated by one researcher evaluator (Rater A).

2.2.5 Data analysis

Descriptive statistics were used to calculate the mean values of evaluation scores for outputs from the online MT systems. Inter-rater reliability between two raters for each language pair was measured and determined using Spearman rank correlation coefficient. Differences in the evaluation scores of the systems were compared using analysis of variance (ANOVA). The PTR for EJ translations was examined with Cochran Q test. Analyses were conducted using the SPSS (Statistical Package for the Social Sciences) software (version 20; IBM Corp., Armonk, NY, USA).

2.3 Results

2.3.1 IRR of the Evaluation Method

The IRR of the evaluation method was examined for structural accuracy and intelligibility. In total, 88 sentence units of MT output were evaluated, and mean values for each criterion were determined. Table 4 shows the Spearman correlation coefficients of the IRR between raters A and B. Fair correlation was found between the raters for both structural accuracy and intelligibility in the two language pairs.

Table 4. IRR for the MT Evaluation method

Evaluation Criteria		Rater	Mean (SD)	IRR	
				<i>r</i>	<i>P</i>
English to Japanese	Accuracy	A	3.52 (1.30)	.778	<.001
		B	3.57 (1.26)		
	Intelligibility	A	2.78 (1.03)	.752	<.001
		B	2.78 (1.03)		
Korean to Japanese	Accuracy	B	3.33 (1.17)	.729	<.001
		A	4.55 (.65)		
	Intelligibility	B	4.60 (.67)	.620	< .001
		A	3.78 (.94)		
		B	4.29 (.81)		

2.3.2 MT translation evaluation

2.3.2.1 EJ translations

For EJ translations, the prepared 297 sentences were analyzed with each MT system, for a total of 1,188 MT sentences. As shown in Table 5, in terms of EJ, the mean score of the evaluation was highest with CL for structural accuracy. The next highest was BL, and the third highest was GT. Regarding intelligibility, GT had the best mean score, and CL and BL were second. BT had the lowest mean scores for both criteria. In terms of the total evaluation mean score in which accuracy and intelligibility were combined (the full score is 10), CL had the highest value. Analysis of variance showed significant differences among the four online MT systems regarding both structural accuracy and intelligibility. Results of a multiple-comparisons procedure (Tukey's honestly significant difference [HSD]) among the four MT systems in quality differences of structural accuracy and intelligibility are shown in

Table 6. Regarding structural accuracy, the quality of BT was significantly lower than that of the other three MT systems. There was a significant difference between CL and GT. For intelligibility, there was a significant difference between BT and the other MT systems, whereas the difference between GT and CL/BL was not significant.

Table 5. Quality differences of structural accuracy and intelligibility among four MT systems for EJ and KJ

Evaluation Criteria		Mean (SD)				<i>F</i>	<i>P</i>
		CL	GT	BT	BL -EJ KD -KJ		
EJ <i>n</i> = 297	Accuracy	3.98 (1.12)	3.60 (1.21)	2.60 (1.14)	3.80(1.20)	83.416	< .001
	Intelligibility	2.88 (.90)	2.93 (1.11)	2.25 (.99)	2.88 (.98)	31.345	< .001
	Total	6.86	6.53	4.85	6.68		
KJ <i>n</i> = 290*	Accuracy	4.72 (.03)	4.67 (.04)	4.04 (.05)	4.84 (.03)	97.257	< .001
	Intelligibility	3.35(.86)	4.04 (.95)	3.32 (1.03)	3.71 (.96)	37.694	< .001
	Total	8.07	8.71	7.36	8.19		

* During the preparation process of translating the source English sentences into Korean, *n* of the source sentences in Korean changed to 290 from 297 to secure the authenticity of Korean grammars and expression.

Table 6. Multiple Comparisons (Tukey's HSD) among four MT systems on quality differences of structural accuracy and intelligibility: for EJ

MT Systems	Structural accuracy				Intelligibility			
	Mean difference	P	95% CI		Mean difference	P	95% CI	
			LL	UL			LL	UL
GT	.380	<.001*	.13	.63	-.051	.927	-.26	.16
CL BT	1.387	<.001*	1.14	1.63	.630	<.001*	.42	.84
BL	.182	.231	-.07	.43	.000	1.000	-.21	.21
CL	-.380	<.001*	-.63	-.13	.051	.927	-.16	.26
GT BT	1.007	<.001*	.76	1.25	.680	<.001*	.47	.89
BL	-.199	.163	-.45	.05	.051	.927	-.16	.26
CL	-1.387	<.001*	-1.63	-1.14	-.630	<.001*	-.84	-.42
BT GT	-1.007	<.001*	-1.25	-.76	-.680	<.001*	-.89	-.47
BL	-1.205	<.001*	-1.45	-.96	-.630	<.001*	-.84	-.42
CL	-.182	.231	-.43	.07	.000	1.000	-.21	.21
BL GT	.199	.163	-.05	.45	-.051	.927	-.26	.16
BT	1.205	<.001*	.96	1.45	.630	<.001*	.42	.84

*significance level for mean difference is .05, SE for the structural accuracy is .096

SE for the intelligibility is .082

In total, 75 word items were identified as specific terms, and these words were used for calculating the PTRs. The PTRs for each online MT system are shown in Table 7. As the table shows, the percentage was best for the GT MT system; 74.70% of the specific terms were properly translated by this system. The second best in terms of PTR was BT, followed by CL and BL. The average PTR was 53.3%. There was a significant difference between GT and the other three systems (BL-GT, $P < .001$; CL-GT, $P < .001$; BT-GT, $P < .015$).

Table 7. Proper translation rate (PTR) for specific terms among four MT systems.

MT system	# of proper translation (N=75)	PTR (%)
CL	34	45.3
GT	56	74.7
BT	40	53.3
BL	30	40.0

* significance level is .05, SE is .071

Table 8. Comparison of PTR in each pair of MT systems

Pair of MT system	<i>Q</i>	<i>P</i>	Adjusted <i>P</i>
BL-CL	.053	.450	1.000
BL-BT	.133	.059	.353
BL-GT	.347	.000	<.001*
CL-BT	-.080	.257	1.000
CL-GT	-.293	.000	<.001*
BT-GT	.213	.002	<.015*

2.3.2.2 KJ translations

The analysis was also conducted for 290 KJ translated sentences with each online MT system currently available, for a total of 1160 sentences. As with the evaluation of EJ, although KD marked best score in structural accuracy, GT was best rated among the four systems in terms of intelligibility, and there was a significant difference between GT and other three systems (Table 5 and 9). The combined mean score of structural accuracy and intelligibility was highest for GT, followed by KD, CL and BT. In KJ translations, BT also marked lowest performance.

Table9. Multiple Comparisons (Tukey's HSD) among four MT systems on quality differences of structural accuracy and intelligibility: for KJ

		Structural accuracy				Intelligibility			
MT systems		Mean Difference	P	95% CI		Mean Difference	P	95% CI	
				LL	UL			LL	UL
CL	GT	.045	.820	-.09	.18	-.697*	<.001	-.90	-.49
	BT	.679*	<.001	.55	.81	.031	.979	-.17	.23
	KD	-.121	.089	-.25	.01	-.359*	<.001	-.56	-.16
GT	CL	-.045	.820	-.18	.09	.697*	<.001	.49	.90
	BT	.634*	<.001	.50	.77	.728*	<.001	.52	.93
	KD	-.166*	.007	-.30	-.03	.338*	<.001	.13	.54
BT	CL	-.679*	<.001	-.81	-.55	-.031	.979	-.23	.17
	GT	-.634*	<.001	-.77	-.50	-.728*	<.001	-.93	-.52
	KD	-.800*	<.001	-.93	-.67	-.390*	<.001	-.59	-.19
KD	CL	.121	.089	-.01	.25	.359*	<.001	.16	.56
	GT	.166*	.007	.03	.30	-.338*	<.001	-.54	-.13
	BT	.800*	<.001	.67	.93	.390*	<.001	.19	.59

*significance level for mean difference is .05, SE for the structural accuracy is .096
SE for the intelligibility is .079

2.4 Discussion

2.4.1 Reliability of the evaluation criteria of online MT

The findings indicate that the IRR of the existing evaluation method of structural accuracy and intelligibility criteria was statistically reliable for evaluating EJ/KJ translations of nursing documents from MT systems. This evaluation method can be used to further evaluate the translation quality of different MT systems for nursing literature. Regarding intelligibility, the range of evaluation scores between the two raters was larger than that of

structural accuracy. This may be due to the subjective nature of the intelligibility criterion.

Although the accuracy of grammatical structures can be relatively objectively identified if the raters have appropriate knowledge of the languages to be evaluated and the evaluation is suitably guided by criteria, the determination of intelligibility inevitably involves subjective judgments, depending on each rater. However, consistency between the raters was obtained with this evaluation method.

2.4.2 Which is possibly the best online MT system?

As the results show, CL performed best for structural accuracy among the four online MT systems, and GT had the highest mean value in intelligibility. Simple total values, which combined accuracy with intelligibility, indicated the best score for CL, and there was a significant difference between CL and GT in structural accuracy. However, it is most important for Japanese nurses to understand the meaning of translations from English to Japanese, so the level of intelligibility helps to account for whether the MT system is actually usable. According to the results of the PTRs, GT was significantly highest among the online MT systems used. The findings are consistent with a previous study that reported a positive relationship between PTR and intelligibility [55]. As mentioned above, specific terms play an important role in whether readers can understand meaning in a specific context; thus, the appropriateness of translations of the specific terms could have influenced the degree of

intelligibility. From this point of view, based on the results of the present study, GT may be the most usable online MT system for nursing literature.

Potential usability of MT systems for Japanese nurses has been indicated in previous studies [6, 28], and the current study identified the online MT system that can be most usable. This is perhaps a good news for Japanese nurses because GT is offered by Google, the world's largest search engine, and can be accessed conveniently by users. The evaluation score for intelligibility tended to be proportional to that of structural accuracy. Nevertheless, none of the online MT systems evaluated scored a 4 or 5 on the 5-point scale of quality for EJ in the current study, and we are not suggesting that current online MT systems are ideal or even practical enough for wide and extensive use when reading English professional articles. Although online MT systems are basically for general purposes, and there is much room for improvement, the findings imply the potential efficacy of using these systems, which could show promise not only for Japanese nurses but also for nurses in other non-English-speaking countries who are in the same situation. As for KJ translations evaluation, the rated scores were relatively favorable confirming the feasible usability which has been previously stated in some literatures [57-60]. This fact might contribute to activating educational and research exchanges in nursing between Japan and Korea.

At the same time, the study results indicated the necessity to further explore the perceptions of the usability of online MT systems in reading English and Korean nursing literature among Japanese nursing population.

This study also has implications for the providers of MT systems. We confirmed the importance of the influence of specific terms. If an MT system focuses on the narrower domain of the nursing field, the vocabularies and grammar might be better controlled, and it would be beneficial for nursing populations [24]. Furthermore, their motivation to use such technology may be increased. Also, new phrases are used to describe original concepts in nursing literature from day to day, and it usually takes some time to find valid translations in another language to be disseminated. Failing to immediately and accurately translate new and highly technical words is a downside of MT technology, but it may be a limitation of current MT services. Efforts to improve the corpus data and dictionary in health and nursing domains will be needed.

Because translation quality can depend on the context of the sentences, the type of the document may influence the translation evaluation, so differences based on each nursing abstract may exist. Nevertheless, unlike literary works, such as novels, this study dealt with nursing literature, which features scientific writing that follows certain rules; thus, it seems less likely that there will be variation in the quality of translation depending on the extracted document, but this could be a subject of further investigation.

2.5 Limitations

The limitations of this study include the relatively small number of data sources. More source data could produce more reliable results, and especially for the PTR, the findings could be better confirmed by using more word items and increasing the number of data sentence units. Adding more evaluators could increase the generalizability of the evaluation results. Bias may exist in the perception of the raters. During testing of the IRR, two raters evaluated the MT translations, but they both knew which online MT system produced the translations they were evaluating. Furthermore, the order in which the raters evaluated the translations of each online MT system was not preliminarily considered. A single-blind procedure should have been followed to eliminate possible influences of any existing assumptions and forward bias. Finally, it should be mentioned that the study result shows the status of online MT performance only at a specific point in the time. This is an inherent limitation for study of ever-improving technology of online MT systems, therefore, periodic quality evaluation of the systems is needed.

2.6 Conclusion

The reliability of an existing method for evaluating the MT quality of both EJ and KJ using the criteria of accuracy and intelligibility was verified for nursing literature. We identified the most useful online MT system among those offered via major Internet search

engines in Japan. GT was considered to be the best for the Japanese nursing population, based on its better scores for intelligibility but especially its high PTR, which can significantly impact intelligibility.

This study, the first to discuss the value of using online MT systems for nursing literature from English into Japanese as well as Korean to Japanese, provides helpful information for Japanese nurses who need to read technical, international research. The next research step is the examination of the general usability of online MT systems rated by more Japanese nurses.

Chapter 3

Study 2: Evaluation by and Current Status of Online MT Use in Nursing Users

3.1 Purpose

The purpose of Study 2 was to examine and describe the evaluated GT performance for nursing literature by nursing users including linguistic factors as well as to explore how they use online MT and perceive its usefulness in reading English articles and to discuss what should be considered for better utilization of online MT lessening the language barrier.

Specifically, we attempted to investigate the following research questions:

- (1) How would GT translations be rated by Japanese nurses?
- (2) Would word count, rate of unidentifiable words, and complexity of source text be related to the intelligibility of GT translations?
- (3) How do Japanese nurses use online MT and perceive its usefulness?
- (4) What factors are associated with the online MT use and perceived usefulness?

The inter-rater reliability of the method was verified in Study 1, and GT was identified as a possibly useful online MT system among the four major systems offered in Japan [62]. The performance of GT had been similarly tested in several previous studies, and GT was reported to outperform other systems [43, 44-46]. Translations from GT were thus used in the current study, in which Japanese nurses evaluated the quality of translations of

professional nursing articles. The nurses manually evaluated the outputs from GT in terms of their intelligibility and perceived usefulness.

3.2 Methods

3.2.1 Preparation of study materials

For both language pairs, sentences translated *via* GT from 23 randomly selected nursing abstracts in Study 1, which was considered to be most useful among tested four online MT systems offered within Japan, were provided. In this study, in order to obtain more realistic evaluation results, the study participants read and rated each section in a nursing abstract, not each sentence. A total of 142 sections of nursing abstracts were prepared to be evaluated for each language pair of EJ and KJ. The section types included title, aim/background, method, results, conclusions/discussion, and implications for clinical practice, in accordance with the structured abstract form generally employed in nursing journals. The number of sections in an abstract varied from five to seven, depending on the study and the journal rules (e.g., some abstracts did not include the discussion and/or implications for clinical practice sections). Each section contained one to three sentences or phrases.

3.2.2 Study participants and collection of data

We aimed to select a study population of native Japanese speakers with some nursing experience who would understand the clinical contexts of technical papers. Exclusion criteria were not defined. Assistants and research associates at nursing universities were assumed to be appropriate.² Commonly, 3–5 years of nursing experience is required when applying for these positions. There were 200 4-year nursing programs in Japan as of February 2012, and the number of assistants and research associates across the nation was assumed to be over 2,000.

All the nursing colleges across Japan accessible through the Internet at that point were the subject of the study participants. Two to four individuals were randomly selected from the Website of each university, for a total of 500 candidates. The actual participants were recruited by mail. Each candidate was sent a cover letter asking for his/her participation and the questionnaire, and 498 of them received it. Two candidates could not receive the questionnaire due to address unknown. The questionnaire included the GT translations of EJ and KJ of one nursing article abstract for each language pair, as well as the question items for participants' background information and the use of online MT. Each participant completed the questionnaire and returned it by mail. Each individual was assigned five to seven sections

² In Japan, faculty members with the title of “assistant” are assigned the duties of “assisting teaching and research,” and those with the title of “research associate” are charged with “conducting teaching and research. Nevertheless, the necessary qualifications and actual duties of academic nursing staff members are not always consistent at nursing universities across Japan.

of an article abstract which was one of the randomly selected 23 abstracts. A total of 250 individuals responded to the survey, a response rate of 50.2%. The data were collected from March to April of 2012.

3.2.3 MT evaluation measurements

In Study 1, which employed a manual method of MT quality evaluation using the criteria of structural accuracy and intelligibility, the reliability of the evaluation method was verified. In the current study, each section of the source abstract was evaluated with the criterion of intelligibility, as in Study 1, while a criterion of perceived usefulness, as in a previous study of general users [63], was used to evaluate the translation quality of the whole abstract. The wording for the criterion of intelligibility was modified for general users. A five-point Likert scale was used for criteria, as shown in Table 10. The nursing abstracts in their source language were not presented to the participants because of potential bias, depending on individuals' degree of knowledge of the source language.

Table 10. Evaluation criteria of GT translations used by nursing participants

Intelligibility*	Usefulness
1. Not intelligible at all	1. Not useful at all
2. Only partially intelligible	2. Not so useful
3. Somewhat intelligible, but not sure	3. Neither useful nor not useful
4. Almost intelligible	4. Useful to some degree
5. Very intelligible	5. Very useful

*The wording of describing the criteria was modified for general users.

3.2.4 Questionnaire items for online MT use

The questionnaire included background questions, previous and current online MT usage, and perceived usefulness of online MT. The question items for participants' background information include sex, job title, age group, academic level, English language proficiency level, years of professional experiences. In addition, the number of technical articles read within approximately the previous 3 months was asked and scored on a five-point Likert scale (1: 0, 2: 1–5, 3: 6–10, 4: 11–20, and 5: 21 or more), and the frequency of language barrier perception was asked about and scored also on a five-point Likert scale (1: never, 2: rarely, 3: occasionally, 4: often, and 5: always). Table 11 shows the questions and scoring regarding previous online MT use. Respondents who answered that they had never used online MT were asked to select a reason from among the following: “I do not need it”, “I do not think it is useful”, “I didn’t know there were online MT services”, and “any other reason(s)” (to be specified). In addition, those who had used online MT, but stopped using it could answer “I do not use it anymore”, optionally selecting a reason from “I do not need it”, “I do not think it is useful”, or “any other reason(s)” (to be specified).

Table 11. Questions and scoring used to examine the status of online MT use and perceived usefulness

	Domains	Questions	Scoring
Online MT use	Previous use of online MT	Have you ever used online MT system?	Nominal: Yes or No
	Frequency of using online MT	How often do you use online MT system?	Liker scale 1= rarely 2= occasionally 3= always, almost always Do not use anymore
	Online MT systems which have been used	Which search engines' online MT system(s) have you used before?	Multiple choice 1) Yahoo, 2) Google, 3) Bing, 4) Excite, 5) Nifty, 6) Livedoor, 7) OCN, 8) So-net, 9) other, 10) do not remember the name of search engines
Perceived usefulness of online MT	How did you feel about the online MT system(s)?	Likert scale 1= not useful at all 2= not useful very much 3=neither useful nor unuseful 4= useful in some degree 5= very useful	

3.2.5 Data analysis

Quantitative data were analyzed using the SPSS (Statistical Package for the Social Sciences) software (version 20; IBM Corp., Armonk, NY, USA). Descriptive statistics were used to summarize the baseline data. Raw data were analyzed after calculating the means, standard deviations (SDs), distributions, and percentages for questionnaire items. The *t*-test was used to examine the relationships between the translation evaluation values and the

characteristics of source texts and analysis of variance (ANOVA) - to assess the differences among evaluation values. The chi-square test was used to identify the differences in the number of English articles read and perceived language barrier according to the respondents' demographic data, where the number of English articles read was measured based on the binary responses "0 - 5" and "6 or more", and perceived language barrier was measured based on the binary responses "less" and "more". Also, the chi-square test was used to determine the differences in perceived usefulness, measured based on the binary responses "not useful" and "useful", according to the participants' characteristics. The *t*-test and ANOVA were also used to detect the differences in the frequency of online MT system use according to participant characteristics. A post hoc examination was performed, using a multiple comparison procedure (Tukey's HSD test).-The relationships between professional experience and the outcome variables were examined using a Spearman's correlation coefficient.

3.3 Ethical Considerations

The Ethical Review Board of the University of Tokyo approved this study. Participants were informed that all information would be treated confidentially, and that they could contact the researcher if they had any questions. The approval number: 3675.

3.4 Preliminary Study

Preliminary study was conducted to predict the response rate for the survey and the tendency of evaluation scores by nursing users, as well as their opinions and impressions on online MT systems.

In the preliminary study, 50 candidates from nursing universities throughout Japan were randomly chosen and the survey was distributed. A total of 28 individuals responded to the survey, a response rate of 56%. As there were 23 kinds of nursing abstract, each was distributed to two or three individuals for translation evaluation. The data were collected during February and March, 2012.

Respondents to this survey included 25 (89%) women and 3 (11%) men. Respondents' age groups were as follows: 20s ($n = 5$, 18%), 30s ($n = 16$, 57%), 40s ($n = 6$, 21%), 50s ($n = 1$, 4%). Concerning job title, 19 (68%) of the respondents were research associates, and 9 (32%) were assistants. Academic degrees held by the respondents included a master's degree ($n = 16$, 57%), bachelor's degree ($n = 10$, 36%), and associate degree ($n = 2$, 7%). The average number of years of clinical and teaching professional experience was 7.93 and 3.96, respectively. In terms of the evaluation of online MT quality, Table 12 presents the evaluation results. A total of 169 and 174 sections were evaluated for intelligibility of the translations in the language pairs of EJ and KJ, respectively. Usefulness was evaluated for the translation of the entire abstract as a whole, and 27 abstracts were evaluated for each

language pair. When the mean scores were simply compared between English-Japanese and Korean-Japanese, translation performance was significantly better in Korean-Japanese with respect to both intelligibility and usefulness. Also, as shown in Table 13, the mean scores of intelligibility for each abstract section was lowest for the translations of “results” in both language pairs (EJ: 2.14; KJ: 2.96).

Table 12. Evaluation of translation of EJ and KJ in the preliminary study

	EJ	KJ	<i>t</i>	<i>P</i>
Intelligibility Mean (SD)	<i>n</i> * = 169 2.46 (0.99)	<i>n</i> * = 174 3.38 (1.14)	8.046	< .001
Usefulness Mean (SD)	<i>n</i> ** = 27 2.59 (1.19)	<i>n</i> ** = 27 3.59 (1.12)	3.189	.002

*n** the total number of unit sections for translation evaluation, according to structured abstracts, e.g., “Title” “Aim” “Background” “Method” “Results” Discussion” and “Conclusions”.

*n*** the total number of whole abstracts evaluated for usefulness of the translation.

For current online MT use and perceived usefulness by participants, of 28 participants, 17 (61%) had experience using an online MT system. Of the major Internet retrieval engines that provide online MT systems used by the participants included, Excite was the most popular system among the study participants (*n* = 13). Yahoo (*n* = 10), Google (*n* = 9), MSN (*n* = 2), Livedoor (*n* = 1), and OCN (*n* = 1) followed. The only language pair that the participants had experienced with online MT systems was English-Japanese. In answer to the question asking with what frequency they had used online MT systems, one (6%) respondent

answered “always,” six (35%) answered “sometimes,” six (35%) answered “often,” and four (24%) answered “do not use any more.” When asked about the usefulness of online MT systems in general, 10 (59%) respondents answered that the system is “not very useful” and 7 (41%) answered that it is “useful to some degree.”

The participants were able to comment on their perceptions of usefulness and their impressions of online MT systems. The contents of written comments were obtained from 17 respondents. These comments majorly included the perspectives regarding online MT use and usability, participants’ own English proficiency, and participants’ expectations for online MT systems. Specifically, comments such as “improper translations of technical terms by online MT system,” “difficulty in translating contexts by online MT systems,” “no use due to improper translations,” “use only for grasping the meaning of source sentences,” and “use only for checking single words” were listed for online MT use and usability. Concerning the comments on participants’ own English proficiency, “limited English proficiency” and “need to improve English ability” were mentioned. For participants’ expectations for online MT systems, the comment “enriched technical terminologies in online MT systems” was received.

The result of this preliminary study implied that KJ translations could be read through an online MT system of GT at an acceptable level of comprehension but the MT was not sufficient for EJ translations. Respondents with experience using MT use this tool largely to grasp the overall meaning of the original text. Enrichment in technical terms in the online MT

systems appeared to be the key to better usefulness.

The main survey with larger sample was immediately conducted, as the results shown in the following section.

Table 13. Evaluation scores according to abstract sections in the preliminary survey
Average evaluation score of translations (SD) for abstract sections

	Sections of an abstract*							
	1	2	3	4	5	6	7	8
EJ	2.41 (0.80)	2.44 (0.96)	2.36 (0.99)	2.82 (1.16)	2.14 (0.89)	2.22 (0.83)	2.67 (1.19)	2.67 (0.82)
<i>n</i>	27	25	28	28	28	9	18	6
KJ	3.63 (1.28)	3.57 (1.00)	3.33 (1.14)	3.46 (1.14)	2.96 (1.07)	3.80 (1.10)	3.17 (1.17)	3.43 (0.98)
<i>n</i>	27	28	27	28	28	5	24	7

1: Title, 2: Background, 3: Purpose, 4: Method, 5: Results, 6: Discussion
7: Conclusion, 8: Relevance to clinical practice

3.5 Results of Main Survey

3.5.1 Participants' characteristics

The respondents were predominantly women (90.4%), with a most frequently reported age group of 30s (54%) (Table 14). One hundred seventy-one (68.4%) respondents had a master's degree and with regard to job titles, 75.6% of them were research associates. In average, the participants' had longer professional experiences in clinical settings than in teaching.

Forty-five point five percent of the respondents reported their English language proficiencies, based on their scores or grades from the following tests: the Test of English for International Communication (TOEIC), the Test of English as a Foreign Language (TOEFL), the Society for Testing English Proficiency (STEP) test, and the United Nations Association's Test of English (UNATE). The English proficiency levels were operationally defined and divided into two categories: above and below intermediate (Table 14). Above intermediate refers to a score of 600 or above on the TOEIC, 2nd grade on the STEP test, or a grade of B or above on the UNATE, while scores and grades below these thresholds were referred to as below intermediate [64]. The mean number of years of clinical experience was more than twice the mean number of years of teaching (Table 14).

Fifty-seven point nine percent of the participants responded that they had read one to five technical articles in English language within approximately the previous 3 months (Table 14), and 23.1% reported having read no article in English. Although they amounted to less than 4%, a few nurses had read more than 21 articles in English. More than 80% said they often or always feel a language barrier when reading nursing articles in English, and, overall, as many as 98.7% of the respondents perceived a language barrier (Table 14).

Table 14. Participants' background information (nr = no response)

Answers		n	%	
Sex	Men	22	8.8	
	Women	226	90.4	
	nr	2	.8	
Title	Assistant	59	23.6	
	Research			
	Associate	189	75.6	
	nr	2	.8	
Age group	20s	19	7.6	
	30s	135	54	
	40s	81	32.4	
	50s	13	5.2	
	60s	2	.4	
Academic degree	Diploma	4	1.6	
	Associate	6	2.4	
	Bachelor	47	18.8	
	Master	171	68.4	
	Doctoral	20	8.0	
	nr	2	.8	
English language proficiency (N = 113)	> intermediate	54	47.79	
	≤ intermediate	59	52.21	
		N	Mean	SD
Professional experiences (years)	Clinical	246	8.84	5.26
	Teaching	248	4.32	3.67
		Answers	Value	%
Number of articles read in English within approximately three months (N=247)	0	57	23.1	
	1 - 5	143	57.9	
	6 - 10	28	11.3	
	11 - 20	10	4.0	
	21 ≤	9	3.6	
Frequency of perceived language barrier (N=245)	Never	1	.4	
	Rarely	2	.8	
	Occasionally	42	17.1	
	Often	74	30.2	
	Always	126	51.4	

3.5.2 GT evaluation results of EJ and KJ translations

3.5.2.1 Evaluation of EJ translations

A total of 1,527 different sections were evaluated by the nursing participants. Each participant also evaluated one whole abstract for the overall perceived usefulness of the translation, resulting in a total of 246 evaluation values. The mean value for intelligibility was 2.39 on the five-point Likert scale (SD = .910), and that for perceived usefulness was 2.74 (SD = 1.124). For intelligibility, the largest number of participants indicated that the translations were “only partially intelligible” (43.4%), and “somewhat intelligible, but not sure” (29.5%) was the second most common answer (Figure 1). For overall usefulness, “not so useful at all” and “not useful so much” together accounted for 50.4% of responses (Figure 1).

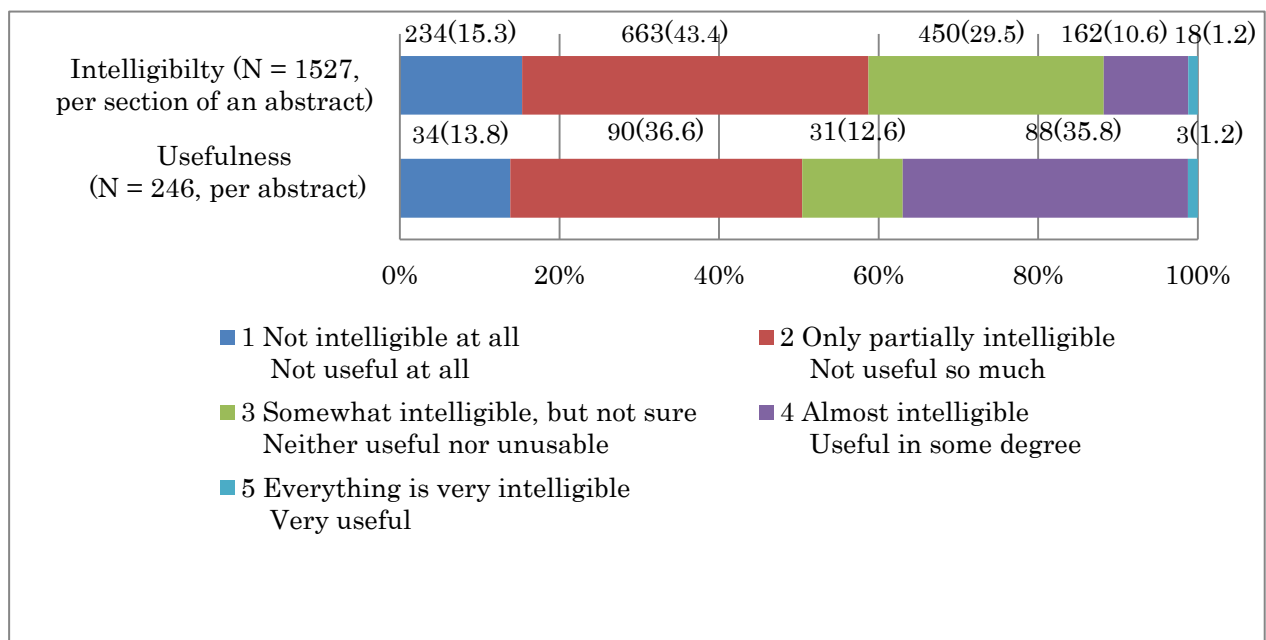


Figure 1. Distribution of evaluation scores by the respondents for intelligibility and usefulness of EJ translations (%)

3.5.2.2 Evaluation of KJ translations

A total of 1,525 different sections were evaluated by the nursing participants. As with rating the EJ translations, each participant also evaluated one whole abstract for the overall perceived usefulness of the translation, resulting in a total of 244 evaluation values. The mean value for intelligibility was 2.38 on the five-point Likert scale (SD = 1.023), and that for perceived usefulness was 3.75 (SD =1.045). For intelligibility, the largest number of participants indicated that the translations were “almost intelligible” (34.2%) followed by “somewhat intelligible, but not sure” (32.6%) (Figure 2). For overall usefulness, “useful to some degree” accounted for 45.9% of responses (Figure 2).

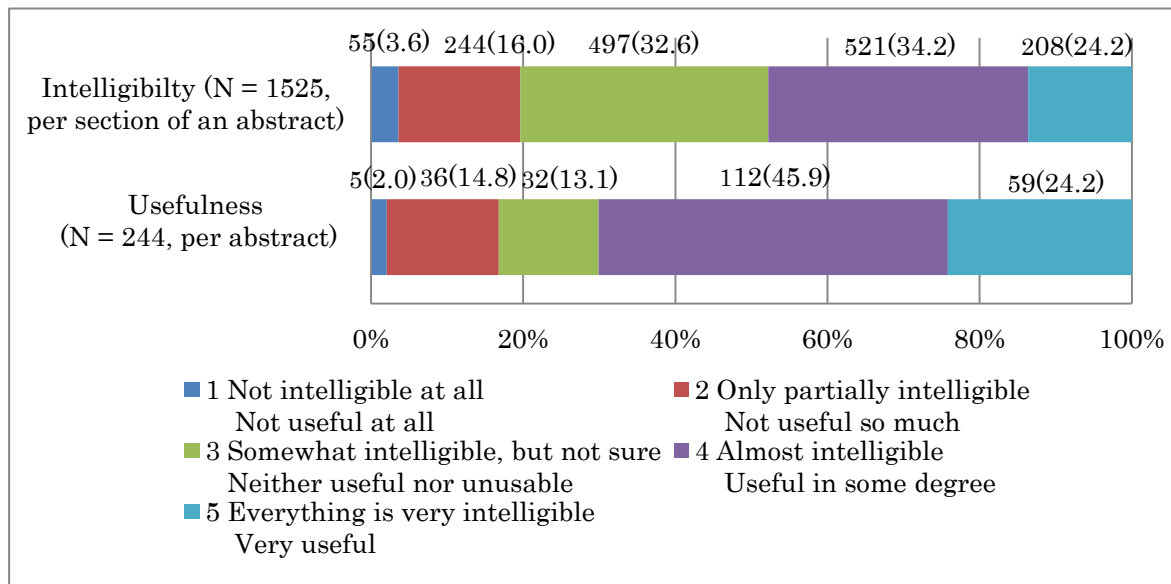


Figure 2. Distribution of evaluation scores by the respondents for intelligibility and usefulness of KJ translations (%)

3.5.2.3 Comparison of the performance between EJ and KJ

The GT performance was compared between EJ and KJ assuming that KJ translations would be better than those of EJ based on the original differences in the language structure in which Japanese and Korean are similar. As it was expected the scores of KJ translations significantly outperformed EJ in both intelligibility and usefulness (Table 15).

Table 15. Comparison of mean scores between EJ and KJ for each criteria

	Language pair	N	Mean	SD	SEM	<i>t</i>	<i>P</i>
Intelligibility	EJ	1527	2.39	.91	.023	28.349	<.000
	KJ	1525	3.38	1.02	.026		
Usefulness	EJ	246	2.74	1.12	.072	10.354	<.000
	KJ	244	3.75	1.05	.067		

Significance level: < .05

3.5.3 Characteristics of source sentences and evaluation results

EJ intelligibility and the characteristics of the source texts per section type are shown in Table 16 and 17. The mean values of the word counts and complexity of the English source texts per article section were 44.49 (SD = 27.64) and 1.40 (SD = .52). We found 15 unidentifiable words in total in 11 out of 142 article sections (7.8%). Broken down by abstract section type, the results section had the lowest evaluation score with the highest average word count (Table 16). The relationship between each characteristic of the source text and GT intelligibility was examined using a *t*-test. In terms of intelligibility, significant differences

were observed according to word count and with or without unidentifiable words in source texts (Table 17).

Table 16. EJ intelligibility and the characteristics of the source texts per section type

	1	2	3	4	5	6	7	8
<i>N</i> of section (total = 142)	23	22	23	23	23	6	16	6
<i>N</i> of respondents (total = 1527)	245	237	247	249	249	90	142	68
Mean of intelligibility (SD)	2.46 (.87)	2.36 (.92)	2.42 (.89)	2.48 (.94)	2.16 (.84)	2.33 (.98)	2.44 (.98)	2.62 (.88)
Average word count	12.68	36.03	33.78	55.48	79.36	52.29	43.55	55.18
Unidentifiable words (mean per section) (total = 15)	0(0)	2(.09)	1(.04)	3(.13)	8(.34)	0(0)	1(.06)	0(0)
Average complexity	1.06	1.62	1.44	1.25	1.33	1.62	1.58	1.91

1: Title, 2: Background, 3: Purpose, 4: Method, 5: Results, 6: Discussion
7: Conclusion, 8: Relevance to clinical practice

Table 17. EJ Differences in intelligibility according to the characteristics of original texts.

Variables	Intelligibility (per abstract section)				
		<i>N</i> (1527)	Mean (SD)	<i>t</i>	<i>P</i>
Word count*	< 39	741	2.49 (.93)	4.231	< .001
	≥ 39	786	2.29 (.88)		
Unidentifiable words **	Without	1401	2.41 (.91)	3.180	.002
	With	126	2.14 (.85)		
Complexity ***	= 1	788	2.38 (.89)	.311	.756
	≥1	739	2.40 (.93)		

* 39 = median, ** sentences without or with unidentifiable words, *** = 1: there was no noun clause, adverb clause, or adjective clause in a source text sentence
≥1: there was one or more clause in a source text sentence

KJ intelligibility and the characteristics of the source texts per section type are shown in Table 18. The mean value of the letter counts was 99.7 (SD = 65.32). We found 16 unidentifiable words in total in 14 out of 142 article sections (9.9%). Broken down by abstract section type, the results section had the lowest evaluation score with also the highest average letter count (Table 18). The relationship between each characteristic of the source text and GT intelligibility was examined using a *t*-test. In terms of intelligibility, significant differences were observed according to letter count and with or without unidentifiable words in source texts (Table 18).

Table 18. KJ intelligibility according to letter count and unidentifiable words

	1	2	3	4	5	6	7	8
<i>N</i> of section (total = 142)	23	22	23	23	23	6	16	6
<i>N</i> of respondents (total = 1525)	250	249	236	250	250	59	176	55
Mean of intelligibility (SD)	3.43 (1.12)	3.57 (1.01)	3.46 (1.00)	3.36 (.97)	3.13 (.99)	2.86 (1.07)	3.45 (.97)	3.58 (.88)
Letter count (mean per section)	25.35	64.48	66.78	140.43	174.17	124	100.63	105.67
Unidentifiable words (mean per section) (total = 16)	1(.04)	2(.1)	2(.1)	6(.26)	1(.04)	0(0)	2(.13)	2(.33)

1: Title, 2: Background, 3: Purpose, 4: Method, 5: Results, 6: Discussion
7: Conclusion, 8: Relevance to clinical practice

Table 19. Intelligibility assessment of K-J translations according the characteristics of the source language

Variables	Intelligibility (per abstract section)				
	<i>n</i> (1525)	Mean (SD)	<i>t</i>	<i>P</i>	
Letter count*	< 91	754	3.57 (1.95)	7.059	< .001
	≥ 91	771	3.20 (.97)		
Unidentifiable words **	Without	1386	3.43 (1.02)	5.815	< .001
	With	139	2.91 (.93)		

3.5.4 Use of online MT systems among nursing population

3.5.4.1 Previous and current use of online MTs and perceived usefulness

Seventy-three point eight percent of respondents indicated that they had indeed used MTs (in terms of the language pair of JE) (Table 20). Among the 65 individuals who said they had never used an online MT, 9 (13.8%) indicated “I do not need it”, 36 (55.4%) reported “I do not think it is useful”, and 12 (18.5%) responded “I never knew there were online MT services”. Eight (12.3%) respondents specified other reasons (e.g., “I cannot learn to read English with MT”, “Reading English by myself is somewhat faster and more efficient for me”, “I would rather use a commercially available MT software package”, “I never read articles written in any foreign languages”, “I do not have any opportunities to be involved in research activities”, and “I am lazy about using online MT”). As for the frequency of using online MT, 43.4% of respondents were occasional users (Table 20). Additionally, 31 (18%) respondents who had previously used online MTs said that they did not use them anymore. Among the

reasons why they stopped using MTs, 3 answered that they “do not need it”, and 28 answered that online MT was “not useful”.

Table 20. Use of online MT and perceived usefulness

Variables	Answers	<i>n</i>	%
Previous online MT use (<i>n</i> = 248)	Yes	183	73.8
	No	65	26.2
Frequency of online MT use (<i>n</i> = 143)	Rarely	49	34.3
	Occasionally	62	43.4
	Always/almost always	32	22.4
Perceived usefulness (<i>n</i> = 182)	not useful at all	3	1.6
	not useful very much	51	28.0
	neither useful nor unuseful	25	13.7
	useful in some degree	91	50.0
	very useful	12	6.6

Online MT systems which have been used by the nursing professionals were shown in this study. The respondents chose Internet search engines that offer online translation tools within Japan, and Yahoo (*n* = 106), Google (*n* = 100), and Excite (*n* = 102) were remarkably popular compared with Nifty (*n* = 14), Bing (*n* = 17), Livedoor (*n* = 16), OCN (*n* = 16), So-net (*n* = 13), and others (*n* = 19). Eleven respondents answered that they did not remember the names of the Internet engine(s) they had used.

Fifty percent of the respondents evaluated online MT as “useful to some degree”, followed by “not very useful”, “neither usable nor unuseful”, and “very useful”. The percentage that answered “not useful at all” was the lowest (Table 20).

As for the language pairs other than Japanese – English, the respondents have experienced using online MT systems with those of Japanese – Spanish (n = 2), – Chinese (n = 2), – Finnish (n = 1), – Korean (n = 1), – German (n = 4), – Dutch (n = 1), and – French (n = 1).

3.5.4.2 Differences in the number of English articles read and perceived language barrier

according to participant characteristics

Significant differences in the number of English articles read were observed according to job title and the level of English proficiency (Table 21). For each of these, participants at higher levels tended to read more articles in English. Also, the amount of teaching experience was weakly correlated with the number of articles read. There were significant differences in the perceived language barrier according to academic degree and English proficiency level (Table 22). Participants with lower academic degrees and lower English proficiencies were more likely to feel the language barrier when reading articles written in English (Table 22).

Table 21. Relationship numbers of articles read in English and participant characteristics

Variables		Numbers of articles read in English				
		<i>n</i>	0-5(%)	≥ 6 (%)	χ^2	<i>P</i>
Title	Assistant	59	55 (93.2)	4 (6.8)	7.713	.005
	Research Associate	186	143 (76.9)	43 (23.1)		
Age group	≤ 30s	152	127 (83.6)	25 (16.4)	1.709	.191
	≥ 40s	95	73 (76.8)	22 (23.1)		
Academic degree	≤ bachelors	56	50 (89.3)	6 (10.7)	3.359	.067
	≥ masters	189	148 (78.3)	41 (21.7)		
English level	< intermediate	59	53 (89.9)	6 (10.1)	6.026	.014
	≥ intermediate	53	38 (71.7)	15 (28.3)		
		<i>n</i>	<i>r</i> *	<i>P</i> **		
Professional experience	Clinical	243	-.045	.485		
	Teaching	245	.186	.004		

*Spearman's correlation coefficients

** Significance level is 1%

Table 22. Relationship between perceived language barrier, according to participant characteristics

Variables		Perceived language barrier			χ^2	<i>P</i>
		<i>n</i>	less (%)	more (%)		
Title	Assistant	55	6 (10.9)	49 (89.1)	2.484	.115
	Research Associate	188	38 (20.2)	150 (79.8)		
Age group	≤ 30s	149	24 (16.1)	125 (83.9)	1.295	.255
	≥ 40s	96	21 (21.9)	75 (78.1)		
Academic degree	≤ bachelors	58	4 (7.5)	49 (92.5)	5.097	.024
	≥ masters	190	40 (21.1)	150 (78.9)		
English level	< intermediate	56	8 (14.3)	48 (85.7)	4.629	.031
	≥ intermediate	53	17 (31.5)	37 (68.5)		
		<i>n</i>	<i>r</i> *	<i>P</i> **		
Professional experience	Clinical	241	.140	.029		
	Teaching	243	-.112	.018		

*Spearman's correlation coefficients

** Significance level is 1%

3.5.4.3 Differences in the frequency of use and the perceived usefulness of online MTs

according to participant characteristics

There was a significant difference in the frequency of online MT use according to the perceived language barrier (Table 23). The results of the post-hoc multi-comparison procedure (Tukey HSD test) revealed that the difference was significant between the “always” and “never–occasionally” categories of the frequency of online MT use, as shown in Table 25.

There were no statistically significant differences in perceived usefulness according to respondent characteristics.

Table 23. Relationship between frequency of online MT use and participant characteristics

Variables		Frequency of online MT use				
		<i>N</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>P</i>
Title	Assistant	28	2.18	.82	.471	.638
	Research Associate	115	2.1	.73		
Age group	≤ 30s	94	2.2	.74	1.87	.064
	≥ 40s	49	1.96	.74		
Academic degree	≤ bachelors	26	2.23	.82	.827	.414
	≥ masters	115	2.09	.73		
English level	< intermediate	20	2.20	.70	1.016	.314
	≥ intermediate	39	2.00	.73		
		<i>N</i>	<i>M</i>	<i>SD</i>	<i>F</i>	<i>P*</i>
Number of articles ^a	zero	26	1.81	.75	.462	.631
	1-5	87	1.87	.73		
	≥ 6	27	2.00	.83		
Language barrier ^b	Never-Occasionally	18	2.28	.75	5.195	.007
	Often	40	2.00	.72		
	Always	84	1.73	.72		
		<i>N</i>	<i>r</i>	<i>P**</i>		
Professional experiences	Clinical	142	.048	.57		
	Teaching	42	.052	.54		

*ANOVA p-value

**Spearman's correlation coefficient p-value

^a Number of professional articles read in English within approximately the previous 3 months

^b Frequency of language barrier perception when reading nursing articles in English

Table 24. Relationship between perceived usefulness of online MT use and participant characteristics

Variables		Perceived usefulness				
		<i>n</i>	Unuseful (%)	Useful (%)	χ^2	<i>P</i>
Title	Assistant	38	17 (44.7)	21 (55.3)	.053	.818
	Research	143	61 (42.6)	82 (57.3)		
	Associate					
Age group	≤ 30s	113	47 (41.6)	66 (58.4)	.399	.528
	≥ 40s	69	32 (46.4)	37 (53.6)		
Academic degree	≤ bachelors	37	19 (51.4)	18 (48.6)	1.219	.269
	≥ masters	143	59 (41.3)	84 (58.7)		
English level	< intermediate	33	16 (48.5)	17 (51.5)	.402	.526
	≥ intermediate	46	19 (41.3)	27 (58.7)		
		<i>n</i>	Unuseful (%)	Useful (%)	χ^2	<i>P*</i>
Number of articles ^a	zero	33	15 (45.5)	18 (54.5)	.852	.653
	1-5	113	51 (45.1)	62 (54.9)		
	≥ 6	33	12 (36.4)	21 (63.6)		
Language barrier ^b	Never-Occasionally	25	13 (52.0)	12 (48.0)	.879	.644
	Often	55	24 (43.6)	31 (56.4)		
	Always	101	42 (41.6)	59 (58.4)		
		<i>n</i>	<i>r</i>	<i>P****</i>		
Professional experiences	Clinical	175	.128	.093		
	Teaching	176	-.072	.341		

*Kruskal-Wallis test p-value

**Spearman's correlation coefficient p-value

^a Number of professional articles read in English within approximately the previous 3 months

^b Frequency of language barrier perception when reading nursing articles in English

Table 25. Multiple comparison procedure: post-doc test for the difference of frequency of MT use according to the levels of perceived language barrier (Tukey's HSD)

Dependent variable: Frequency of online MT use

Frequency of language barrier perception when reading nursing articles in English	Mean Difference	SE	P	95% CI		
				LL	UL	
Never-Occasionally	Often	-.278	.205	.366	-.76	.21
	Always	-.552*	.187	.011	-1.00	-.11
Often	Never-Occasionally	.278	.205	.366	-.21	.76
	Always	-.274	.139	.122	-.60	.05
Always	Never-Occasionally	.552*	.187	.011	.11	1.00
	Often	.274	.139	.122	-.05	.60

* significance level for mean difference is .05

3. 6. Discussion

3.6.1 How are GT translations rated by Japanese nurses?

The results suggest that Japanese nurses do not perceive GT as having adequate quality especially for EJ. Factors that may have influenced this perception could include the fact that the participants evaluated the translations without reading the English source texts. Comparing MT outputs to source texts for comprehension is a natural behavior for MT users regardless of language proficiency. However, by not allowing participants to view source texts, they were forced to base their evaluations on the true quality of the translations themselves, providing a more substantive assessment. For KJ translations, evaluation scores were significantly better than EJ implicating its actual usefulness.

The perceived usefulness of GT translations was given relatively high scores compared with intelligibility in both EJ and KJ. Many respondents gave usefulness a rating of

4 out of 5. It is difficult to explain why perceived usefulness could be rated better than intelligibility. This outcome might have been influenced by the option to select “neither useful nor not useful” as an answer in the questionnaire. It was considered that this answer option was not appropriate to measure the nature of the extent of perceived usefulness, therefore, the means by which perceived usefulness is assessed needs to be improved.

In nursing studies, there are many qualitative research papers. When reporting the methods of such research, authors tend to use different linguistic style and terminology than those reporting quantitative studies. In this work, however, this distinction was not considered, and is an interesting area for future exploration that may reveal more precise trends in the effectiveness of MT.

3.6.2 Influence of the characteristics of source sentences on the evaluation results

For both EJ and KJ, translation quality was found to depend in part on the word/letter count and with or without unidentifiable words in source texts. The fact that longer sentences in original text could result in poor translation quality has been demonstrated by the findings of previous studies [36, 55]. Additionally, the results indicate that the readability of translated sentences may be affected when the translation system is unable to translate certain words. Thus, more efficient use of online MT could be achieved by solving this problem. Such words often included those with British spellings, with prefixes (e.g., preidentified, nonadherence),

and so forth for EJ translations. For KJ, unidentifiable words included unfamiliar terms, an exotic word, incorrect spaces between words. This problem could be addressed by pre-editing [65, 66], a process for simplifying source sentences. However, it remains unknown how well general users are able to edit source text before entering it into an MT system; it seems self-evident that it would depend entirely on one's level of understanding of the source language. MT developers' attempt to minimize the number of untranslatable words in professional healthcare articles is required.

3.6.3 Participants' online MT use and its perceived usefulness

A large number of participants indicated that they had previously used online MTs when reading technical papers. It is perhaps only natural that people would enjoy the convenience of automatic translations on websites in this era of the Internet. However, this result may imply that many nurses require help when reading technical English language papers in their own fields. Accordingly, it indicates that online MT is much in demand among nursing professionals, highlighting the necessity of language support for them. Also, the comments written on the questionnaire indicated that they tended to use online MT services to grasp the basic meaning of the original text, similar to the results of a previous study [65]. In addition, some participants used multiple online MT systems for the same text, to identify the most useful translation (which is an advantage of the free accessibility of online translators).

Some reported that they use online MT systems in combination with a dictionary of technical terminology, indicating that the respondents do not accept translations from online MT systems without question. Also, from the fact that 43% of the respondents used online MT “occasionally” when reading technical literature, it is assumed that they did not regard it as adequate for full-time use, although their perception was that this tool was somewhat helpful.

More than half of the respondents perceived that online MT systems are useful, and hence the general tendency was toward relatively positive impressions. In a previous study, university students rated online MT performance, and more than 30% of the respondents judged it favourably [19]. More of study subjects reported positive perceptions, and in one previous study, online MT received a far better evaluation [46]. The reasons for these differences among study populations are not known but could involve the factors used in the study samples, respondent needs, and the kinds of source texts.

3.6.4 Factors affecting the frequency of online MT use and its perceived usefulness

The perceived language barrier was associated with frequency of use. Respondents who felt less of a language barrier were likely to use online MT systems more often. This may indicate that MT is relatively advantageous for users with higher language proficiency. Because Japanese users with extremely poor English language skills cannot judge MT performance by looking at a source text written in English, it is difficult for them to utilise MT,

and thus they will use it less often. Hence, this result appears to be consistent with previous studies that have indicated the importance of English language ability for the effective use of MT [47, 65]. The perceived language barrier was associated with the participants' academic degree (Table 23). In Japan, nurses have more opportunities to read English in graduate programs. In such programs, nurses might become accustomed to reading research articles or presenting their own research projects, eventually improving their English skills. In fact, some graduate nursing programs in Japan offer a course in reading English [69].

In contrast, no significant correlations were found between perceived usefulness and participant background information. Factors that could be associated with perceived usefulness include motivation for use, type of use, and the expected effectiveness of online MT systems, in addition to the demographic attributes of participants. To obtain a beneficial outcome for users, factors related to usefulness, such as willingness to use, and specific ways and expected benefits of using online MT services, should be investigated in future studies.

3.6.5 Practical and educational implications

The comments written on the questionnaire by the participants included their expectations for and suggestions to developers and providers to improve the systems for the benefit of nursing and healthcare professionals. Specifically, they wanted better translations of technical terms as with the comments in the preliminary study. Their awareness of language

difficulties when reading and writing English was also often mentioned in connection with desired improvements in MT performance. Their feedback should be helpful to MT providers for improving translations in nursing fields, as well as beneficial to nursing professionals. Furthermore, the study respondents pointed out the important issue of transmitting nursing information from Japan in English. They were concerned with producing internationally interactive nursing personnel resources. Thus, online MT systems are expected to produce better translations not only from English to Japanese, but also from Japanese to English.

The results highlight that MT systems are currently ancillary tools, considering their low level of performance, and changes must be made in educating nurses to achieve optimal use of this technology. English education using MT systems may be helpful. However, a negative aspect of MT in foreign language education is that learners may accept an output without verifying its correctness, and stop trying to improve their own language skills [46]. But previous studies have indicated some effectiveness in the use of MTs in English language education. It has been claimed that the use of MT lessens students' aversion to reading and writing English sentences, and helps to train them in logical thinking [64].

The results of the current study suggest that nurses who perceive a smaller language barrier are more likely to use online MT systems. Therefore, we may assume that those with less of a language barrier experience fewer difficulties in using this technology. Good knowledge of languages is required to compare the translation output from an MT with the

source material [50, 67, 68]. However, even a relatively small level of language proficiency may aid the utilization of MT systems. Therefore, basic language education is important for effective use of MT systems.

Basic English is taught at nursing colleges; however, English education for nurses should continue in post-graduate level and continuing education. These educational opportunities should be provided continuously both in clinical settings and research and educational institutions. Concurrent training in utilizing online MT systems and improving basic knowledge of English language skills is expected to enhance English reading comprehension, and aid the use of MT systems. As a result, nurses may read more papers written in English, helping with EBN and education and research. Therefore, English education for nurses may be expected to bring about improvements in the nursing community and to patients.

Also, Japanese nurses need to have training to learn how to specifically read scientific papers in English because this type of document has its own rules for readers. Knowing these rules can save time and reduce inconvenience. This education should be provided in the form of integrated and organisational programs that have yet to be established in the nursing workplaces of Japan. The understanding and collaboration of managerial level nurses will be required to make this continuing education a reality, and we will need to investigate their opinions and attitudes regarding education in the English language and reading scientific

papers as part of the occupational development of nursing professionals.

3.7 Limitations

In study 2 we only considered GT, although there are other major systems on the market. We chose to use GT because it would reduce the workload for the respondents, and a lower response rate would be expected when using multiple MT tools. However, including more MT systems in the quality evaluation may provide useful data on which systems are more effective for nursing research papers.

As mentioned above, the participants were from the general users of nursing population, and they evaluated translations without reading the English source texts. This is different from how users behave under real circumstances; users typically compare MT output to source text to comprehend the meaning of a translation. However, by excluding source texts, we were able to obtain evaluations based solely on the quality of translations. Future studies should further assess the relationship between MT utilization and nurses' English language skills with a more practical evaluation method in which nursing users can compare translations and source sentences.

The sample was limited to two types of nursing faculty members, therefore, their academic degrees may have been higher than those typically held by nurses in clinical settings. Nursing professionals from a broader range of settings and levels will be necessary for a

deeper understanding of general tendencies. In addition, the sample might have contained a disproportionate number of respondents that were already interested in MT services and/or users of online MT systems.

The English language test scores and grades reported by our respondents did not appear to accurately represent their English reading skills. The participants' English language proficiencies were judged solely in terms of scores and grades on general language tests. Japanese English language examinations usually measure overall skills, and half of the score is based on listening and/or speaking skills. Therefore, the total score on one of these examinations could be lower if the examinee's listening or speaking skills are poor, even if his/her reading proficiency is higher. In the present study, the participants' reading and writing skills should have been directly assessed, and analyses based on these assessments are needed. Also, recall bias could have been introduced when the participants were asked about the number of articles read within approximately the previous three months.

A separate evaluation of MT for qualitative and quantitative research papers may be beneficial because of the differences in linguistic style and terminology between the two areas of research. In future work, a suitable grouping of material into qualitative and quantitative methods should be considered.

3.8 Conclusion

In the present study, we attempted to investigate how Japanese nursing users rated the quality of online MT output based on their needs of obtaining updated information published internationally. We found that GT a representative online MT, has only minimally acceptable performance for EJ and potentially useful quality for KJ as assessed by nursing professionals. The word/letter count and the existence of unidentifiable (untranslatable) words of source text both affected GT translation quality.

Online MT had been used by a sizeable percentage of the study participants as an aid to reading international nursing journals. The level of the perceived language barrier was related to the likelihood of using online MTs.

The current study highlighted how nursing professionals evaluate online translations of nursing literature as well as its use and perceived usefulness, and the results indicate the need for developers to improve MT systems and for users to improve their ability to use them. For optimal use of online MT, it is necessary to improve users' language skills. We should consider the English language education of nurses, as well as the possibility of using online MTs in English education, and providing training in reading scientific papers and the use of MT tools. It will also be a role of the work to promote user cooperation with the developers and providers of MT systems to improve MT dictionaries with regard to nursing-specific terms.

Future work will include the investigation with more practical way of evaluation method in which the users rate the translations comparing with the source sentences, along with more precise measurement for general impression of online MT translations held by nursing users. Also, further studies should include considering the possible variations in translations evaluation according to the type of research (qualitative/quantitative), involving a wider range of nurses, assessing English proficiency precisely, investigating the usefulness of online MT considering possible influencing factors (such as willingness to use the systems and specific ways and expected benefits of using this technology), examining the need for research education on using international journals, and promoting postgraduate English education according to the perceptions of nurses at the managerial level.

Chapter 4

Conclusions

4.1 General Summary

In Chapter 1, the needs and importance of utilizing research literatures written in foreign languages, especially in the English language in Japanese nurses for nursing practice, education, and research activities were firstly discussed. Then the overview of MT technology including its history and its use as a possible solution option to tackle the explicit barriers such as language difficulties were discussed. The technology of MT has been improved in recent years, especially with the emergence of the Internet, and more and more people are using this innovation worldwide *via* online. Moreover, evaluation methods for quality of MT outputs was reviewed and described. Up to now, MT evaluation methods include those by human being and by automatic manner. In this thesis manual evaluation method was focused based on the importance of essential understanding of translations by human being. Evaluation method of online MT systems regarding nursing related documents has not been specified so far, and online MT use by nursing users was also unknown. The author also stated the necessity of the investigation of quality evaluation for another language pair, Japanese and Korean, which is a neighbor country's language, was newly to be attempted in this dissertation study.

In Chapter 2, the study which attempted to examine and discuss the evaluation method of currently available online MT systems and feasibility of online MT technology use by nursing professionals was described. Online MT outputs of another language pair, Korean into Japanese, were also assessed, to better determine the potential use of online MT for the Japanese nursing population. The reliability of the evaluation method based on the existing criteria at structural accuracy and intelligibility was verified, and GT was identified as possibly most useful online MT at this time.

In Chapter 3, the next step was investigated for the intelligibility of translations of English- and Korean-Japanese (EJ, KJ) from the online system of GT in larger nursing population. At the same time, the current status of MT use and its impressions perceived by nurses, which has not known, was attempted to be explored. The next study explored the level of understanding of online MT outputs from GT in general nursing population, mainly using the criteria of intelligibility. The study participants of nursing users rated the translations performance according to the typical structural sections of an abstract of science papers. In addition, they also rated overall impression of usefulness of an abstract itself. The results showed minimum intelligibility for EJ translations, and KJ translations were significantly more intelligible than that of EJ. Many of the nursing participants have experienced using online MT systems, and it seemed to be limited to an ancillary use. As for the relationship between the participants' background information and use of online MT systems, the

perceived language barrier was associated with the frequency of online MT use. As for KJ translations via online MT, raters thought they are understandable to a certain extent, and it has been revealed that the usefulness of GT translation of KJ for nursing users was reasonably ensured.

The quality of EJ translations of GT is problematic, however, we need to consider the ways to optimize the use of these outputs from online MT systems for nursing users. English education is needed even more in their post-graduate and continuing course, both in clinical settings and research and educational institutions. Basic knowledge of English grammars and vocabularies should be taught. Improved language knowledge in nurses would enable them to utilize online MT in an effective way. Providing nurses with the opportunity to learn fundamental science writing to read necessary papers for clinical practice, teaching, and research and their voluntary involvement in learning will constitute the basic literacy skills. As a result, the goal of reading professional literatures in English for expertise in nursing fields will be achieved, eventually leading to bringing the benefits to nursing community and patients.

4.2 Future works

Future works to be done have been highlighted from this dissertation study. As for the importance of analyzing the translations according to the types of nursing

research, more precise analyses with the same number of abstracts for each type. Then it is necessary to know how online MT can be useful for nursing literatures and to further explore its utilization.

In real situation, most users of MTs try to understand the translation outputs from the system comparing them with source sentences, especially for the translations of English, which is relatively a familiar language for Japanese people. In the present study, the questionnaire was distributed to nursing participants without giving them source sentences in English and Korean to purely know the quality of GT translations, eliminating the influence of individual English language proficiency on their rating the EJ translations. For Korean source sentences, it was assumed that most of the participants were not able to read texts written in this language. However, more realistic method of rating the quality of translation outputs may require providing the raters with both the original sentences in the source language (specifically, English) and online MT outputs in the target language (Japanese). Therefore, the next work should consider this aspect of the evaluation method, and, in doing so, the difference of English proficiency among raters should be focused exploring better ways of using online MT systems according to the raters' level of the language proficiency. Such an investigation would provide implications for English education for nursing population.

Also, the evaluation criterion for impression of usefulness needs to be further considered for a refinement. The answer option for this criterion, “Neither useful nor unuseful,” was not appropriated to measure the degree of the usefulness felt by the respondents. Furthermore, the current study which used nursing population of faculty members asked the subjects the frequency to read professional papers in a foreign language, and the result might have varied depending on different nursing fields: that is to say there might be differences in the frequency between clinical nurses and nursing teachers, hence the status of using and expectations for online MT systems may also differ. Therefore, nurses in broader fields will need to be investigated for this limitation.

4.3 Concluding Remark

Issues related to reading articles published in foreign languages, especially in English, should be seriously considered, to be resolved not only in Japan, but also in many other non-English-speaking countries. This study thesis will provide incentive for further exploration of possible linguistic support for nurses via advanced Internet-based technology, leading to more effective use of MT tools, and eventually more opportunities for nurses to access internationally published nursing information and obtain updated knowledge for better care delivery.

Contributions

Contributions to this thesis are as follows:

Study design: Ryoko Anazawa, Hirono Ishikawa, and Takahiro Kiuchi

Preparation of the source data: Ryoko Anazawa and Park MJ

Evaluation Rating: Ryoko Anazawa, Asami Fujishima, and Toshie Kuriahata

Data analysis: Ryoko Anazawa, Hirono Ishikawa, and Takahiro Kiuchi

Writing the manuscript: Ryoko Anazawa

References

- [1] Strickland RJ, O’Leary-Kelly C. Clinical nurse educator’s perceptions of research utilization: barriers and facilitators to change. *Journal of Nursing Staff Development*, 25(4),164–171, 2009
- [2] Pravikoff DS, Tanner AB, Pierce ST. Readiness of U.S. nurses for evidence-based practice. *American Journal of Nursing*, 105(9), 40–51, 2005
- [3] Jacobs SK, Rosenfeld P, Haber J. Information literacy as the foundation for evidence-based practice in graduate nursing education: a curriculum-integrated approach. *Journal of Professional Nursing*, 19(5), 320–328, 2003
- [4] 穴沢良子. 看護師の英語文献利用と英語学習. *小児看護*, 34(4), 508–509, 2011.
- [5] Anazawa R, Ishikawa H, Kiuchi T. The perceived need for Japanese nursing faculty

- members to learn English: issues related to career development. *Journal of Continuing Education in Nursing*, 43(4), 188–191, 2012
- [6] 小島通代, 秋山照男. 機械検索と機械翻訳による看護職の文献利用の効果. *看護と情報*, 1, 41-47, 1994.
- [7] 草間朋子. EBN (Evidence-Based Nursings) を考える. *大分看護科学研究*, 4 (1), 12-15, 2003.
- [8] 新田義彦. 機械翻訳の原理と活用法-古典的機械翻訳再評価の試み. pp. 18, 2012.
- [9] Nida E, Taber C. *The Theory and Practice of Translation*. pp. 12, 1969.
- [10] 長尾真. 機械翻訳はどこまで可能か. pp. 43, 1986.
- [11] Munday J. *Introducing Translation Studies: Theories and Applications*. pp. 281, 2012.
- [12] Kirchhoff K, Turner AM, Axelrod A, Saavedra F. Application of statistical machine translation to public health information: a feasibility study. *Journal of American Medical Informatics Association*, 18(4), 473–478, 2011.
- [13] 宮平知博, 渡辺日出雄, 田添英一, 神山淑朗, 武田浩一. インターネット機械翻訳の世界. pp. 12-13, 2000.
- [14] 田口尚三. 機械翻訳の現状と課題. *オフィス・オートメーション*, 6 (3), 4-8, 1985.
- [15] 佐藤理史. アナロジーによる機械翻訳. pp. 6-7. 1997.
- [16] Kurohashi S, Aramaki E. Status and issues of machine translation. *Journal of Information Science Technology Association*, 55(8), 326-330, 2005.

- [17]佐藤理史. アナロジーによる機械翻訳. pp. 21-35. 1997.
- [18]宮平知博, 渡辺日出雄, 田添英一, 神山淑朗, 武田浩一. インターネット機械翻訳の世界. pp. 34-41.
- [19]宮平知博, 渡辺日出雄, 田添英一, 神山淑朗, 武田浩一. インターネット機械翻訳の世界. pp. 16-44.
- [20]後藤正浩. 機械翻訳における前処理の適用とその効果. 京都大学修士論文, 2004.
http://www.ai.soc.i.kyoto-u.ac.jp/publications/thesis/M_H23_goto-masahiro.pdf
(2013年2月14日閲覧)
- [21]林良彦. 機械翻訳概論 (J-to-E を中心に) & 統計的機械翻訳, 言語情報科学論 A, 2007.
<http://ocw.osaka-u.ac.jp/language-and-culture-jp/language-and-information-science-jp/lecturenoteno-11.pdf> (2013年2月14日閲覧)
- [22]株式会社アイタス. 初心者にもわかる機械翻訳入門 (2010.10.07).
<http://www.itas.co.jp/jp/2010/10/post-d079.html> (2013年2月14日閲覧)
- [23]Garcia-Santiago L, Olvera-Lobo MD. Analysis of automatic translation of questions for question-answering systems. Information Research, 15 (4), 2010.
<http://informationr.net/ir/15-4/paper450.html>.(2011年11月23日閲覧)
- [24]Gaspari F. The Role of Online MT in Webpage Translation [dissertation]. University of Manchester, 2007.

- [25]熊谷哲孝. インターネット翻訳サイトを用いた英日翻訳についての予備的考察.
富士大学紀要, 41(1), 115-124, 2008.
- [26]Chunyu K, Tak MW. Comparative evaluation of online machine translations systems with
legal texts. Law Library Journal, 100(2), 299-322, 2008
- [27]Zeng-Treitler Q, Kim H, Roseblat G, Keselman A. Can multilingual machine
translation help make medical record content more comprehensible to patients? Stud
Health Technol Inform, 160 (Pt 1), 73-77, 2010.
- [28]和田 佳代子. 看護職への支援--情報の信頼性と評価をふまえて (特集 日本病院
ライブラリー協会 2008 年度第 1 回研修会). ほすびたるらいぶらりあん, 33(3),
178-184, 2008.
- [29]Anazawa R, Ishikawa H, Takahiro K. Use of online machine translation for nursing
literature: a questionnaire-based survey. The Open Nursing Journal, 7, 22-28, 2013.
- [30]Gim'enez J. Empirical machine translation and its evaluation [doctoral thesis]. Barcelona,
Maig de Universitat Polit`ecnica de Catalunya, 2008.
- [31]White JS. Approaches to black box MT evaluation. In: Proceedings of Machine
Translation Summit V, Machine Translation Conference, Exhibition & Tutorial
Luxembourg, July 10-13, 1995 Hemicycle of the European Parliament, 10.
- [32]White JS, O'Connell T, O'Mara F. The ARPA MT evaluation methodologies: evolution,
lessons, and future approaches. Proceedings of the 1st Conference of the Association for

Machine Translation in the Americas (AMTA) (pp. 193-205), 1994.

- [33]Miyashita K, Yasuda K, Yamamoto S, Yanagida M. Quality evaluation method of machine translated sentences by comparing text retrieved from WWW. IPSJ SIG Tech Rep, 171(15), 89-94, 2006
- [34]Yasuda K, Sumida E. Automatic evaluation for machine translation R & D and its applications. Jpn Soc Artif Intell, 23(1), 2-9, 2008.
- [35]成田一. 翻訳ソフトの実力評価. 自然言語処理, pp. 125-14, pp. 123-130, 1998.
- [36]開原成允, 荻野敏, 木内貴弘, 長瀬淑子, 小島基之. 英語医学文献の機械翻訳の試み. 第8回医療情報学連合大会論文集, 593-596, 1988.
- [37]隅田英一郎, 佐々木 裕, 山本 誠一. 機械翻訳システム評価法の最前線,情報処理 46(5), 552-557, 2005.
- [38]黒田 由加, 鈴木 博和. 機械翻訳文の自然さに影響を与える要因の分析(テーマセッション)思考と言語, 111 (98), 31-36, 2011.
- [39]宮部真衣, 吉野孝. ユーザの翻訳精度判定に対する既判定精度提示の影響, 情報処理学会論文誌, 53 (1), 39-48, 2012.
- [40]Papineni K, Roukos S, Ward T, Zhu WJ. BLEU: a method for automatic evaluation of machine translation. In: 2002 Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL). 2002:311-318. Association for Computational Linguistics Stroudsburg, PA, USA.

[41]内山将夫. 自動評価尺度 BLEU. <http://mastarpj.nict.go.jp/~mutiyama/corpm/4.pdf>.

(2011年12月29日閲覧) .

[42]辻井潤一. 機械翻訳の評価：客観と主観. Japio 2008 Year Book, 250-251, 2008.

[43]NIST. NIST 2005 Machine translation evaluation official results.

http://www.itl.nist.gov/iad/mig/tests/mt/2005/doc/mt05eval_official_results_release_20050801_v3.html. Accessed January 21, 2012.

[44]Toral A, Gaspari F, Naskar SK, Way A. Comparative evaluation of research vs. online

MTsystems. In: Proceedings of the 15th Conference of the European Association for Machine Translation. 2011:13-20. European Association for Machine Translation, Allschwil,Switzerland.

[45]Kit C, Tak MW. Comparative evaluation of online machine translation systems with legal texts. Law Libr J, 100 (2), 299-321, 2008.

[46]Naskar SK, Toral, A, Gaspari F, Way A. A framework for diagnostic evaluation of MT based on linguistic checkpoints. 2011. <http://www.mt-archive.info/MTS-2011-Naskar.pdf>. Accessed January 21, 2012.

[47]Yang J, Lange ED. SYSTRAN on AltaVista: a user study on realtime machine translation on the Internet. In Farwell D, Gerber L, Hovy E, Eds. Proceedings of the 3rd conference of the Association for Machine Translation in the Americas: Machine translation and the information soup. Langhorne PA, USA. Berlin: Springer 1998; pp.

275-85.

[48]AAMTワーキンググループ2. 英日・日英機械翻訳に関するアンケート結果報告.

AAMTジャーナル, 50, 52-60, 2011.

[49]Matsui J, Magnusson D. Six Pre-edit techniques for enhancing Japanese to English machine translations. *Interpret Transl Stud*, 11, 173-87, 2011.

[50]Osuka N, Mano C. The negative effects of using translation sites on college students' language learning. *J Facult Int Stud Bunkyo Uni.*, 13, 133-44, 2003.

[51]Daily Research (2011).インターネットコム

<http://japan.internet.com/research/20110621/1>. (html 2013年2月14日閲覧)

[52]Bedrick DS, Mauro A. A multi-lingual web service for drug side-effect data. *AMIA Annu Symp Proc*. 2009 Nov 14, 34-38, 2009.

[53]Khanna RR, Karliner LS, Eck M, Vittinghoff E, Koenig CJ, Fang MC. Performance of an online translation tool when applied to patient educational material. *J Hosp Med*, 6 (9), 519-525, 2011.

[54]Kiuchi T, Kaihara S. Practical experience of English-Japanese machine translation of medical literatures: quantitative evaluation of the effect of the dictionaries. *Informatics: Symposium Proceedings 1990*. 59-71, 1990, Information Processing Society of Japan, Tokyo, Japan.

[55]Kiuchi T, Kaihara S. Quantitative evaluation of English-Japanese machine translation of

medical literature. *Methods Inf Med*, 30 (3), 199-205, 1991.

[56]太田勝正, 猫田泰敏 (編) . 看護情報学. pp. 11- 14. 2008.

[57]長瀬久明, 橘守亮, 機械翻訳に適応した日本語文の書き方.

<http://repository.hyogo-u.ac.jp/dspace/bitstream/10132/3569/1/1003241806.pdf> (2013年2

月19日閲覧)

[58]金政仁, 大駒誠一. 日韓機械翻訳における動詞の多訳性処理. 全国大会講演論文集第45回平成4年後期 (3), 97-98, 1992.

[59]成田一. 機械翻訳をどう使うの? - 言語対を踏まえた実務的翻訳 - . *Japio YEAR BOOK 2012*, 238-249, 2012.

[60]Akamine S, Okumura A, Muraki KA. Multi-lingual translation system based on an English-Japanese translation system. In: *Proceedings of the 45 the Annual Meeting of Information Processing Society of Japan*, 11-14 October, 1992, Tokushima, Japan, pp. 119-120, 1992.

[61]日本看護科学学会. <http://jans.umin.ac.jp/index.html>. (2011年12月26日閲覧)

[62]Anazawa R, Ishikawa H, Park MJ, Kiuchi T. Online machine translation use with nursing literature: evaluation method and usability. *Computer, Informatics, Nursing*, 31 (2), 59-65, 2013.

[63]Matsui J, Magnusson D. Six Pre-edit techniques for enhancing Japanese to English machine translations. *Interpreting and Translation Studies*, 11, 173-187, 2011.

- [64]Kobe City University of Foreign Studies. TOEIC/TOEFL/IELTS/STEP Score Converting System [internet]. Available from:
http://www.kobe-cufs.ac.jp/about/contribution/license/files/toEIC_toEFL_score.pdf (2012年7月28日閲覧)
- [65]Yoshimi T, Sata I. Improvement of translation quality of English newspaper headlines by automatic preediting. *Journal of Natural Language Processing*, 7(2), 27-43, 2000.
- [66]Yoshimi T, Sata I. Automatic preediting of English sentences for a robust English-to-Japanese MT system. *Journal of Natural Language Processing*, 7 (2), 99-117, 2000.
- [67]安藤進. 機械翻訳(MT)とサーチエンジンを英語学習に活用する.情報処理学会研究報告.97 (65), 9-15, 1997.
- [68]荒木健治. 自然言語こと初め一言葉を覚え会話のできるコンピュータ. pp. 76, 2004.
- [69]Kawagoe K. ESP education for graduate students of nursing science. *Japan Assoc College English Teachers (JACET)*, 42,154-5, 2003.

APPENDICES

翻訳された日本語文を以下の「理解度の評価基準」に沿って評価してください

- 訳出された日本語訳
1. 全体的にまったく理解できない
 2. 部分的に理解可能な箇所はあるが、全体としては理解できない
 3. 全体的になんとか理解はできるが、理解に自信がもてない
 4. 全体がほぼ、良く理解できる
 5. まったく問題なく、すべて良く理解できる

翻訳文 1 (看護学論文の抄録 1 件分)



日本語へ翻訳した文	理解度の評価基準				
	1	2	3	4	5
1 (タイトル) ジェンダーと癌性疼痛の経験では民族の違い：米国における多民族の調査。					
2 (背景) 癌性疼痛の経験では民族の違いで一貫性のない調査結果は、適切ながん性疼痛管理のためのこのトピックに関するさらなる研究の必要性を示唆している。					
3 (目的) 米国の 4 つの民族グループの癌性疼痛の経験では民族の違いを判別するには。					
4 (方法) フェミニストの視点は、理論的な基礎として使われました。これは、社会人口統計学的特性と健康や病気の状態、3 次元癌性疼痛スケール、2 次元癌性疼痛スケール、メモリアルの症状評価尺度、および癌治療のスケールの機能評価の質問を 480 癌患者の多民族のサンプルの調査である。データは ANOVA と階層的重回帰分析を含めて記述し、推測統計学を用いて分析した。					
5 (結果) 結果は、痛みと患者が経験した症状の種類の特定の民族の違いを示した。また、結果は癌性疼痛および機能状態の重要な民族の違いを示した。					
6 (考察) 調査結果は、深い文化的な価値とこのトピックの少数民族の大きい数と癌の各民族の痛みと全国の研究に関係する信念に関する定性的探査さらに示唆している。					

次ページへ

「翻訳文1」全体について、あてはまる数字に○をつけて下さい。

この翻訳文全体は、原文の意味を把握するのに役立ったと感じますか

- 1) まったく役に立たない
- 2) あまり役に立たない
- 3) どちらともいえない
- 4) 少しは役に立つ
- 5) かなり役に立つ

翻訳された日本語文を以下の「理解度の評価基準」に沿って評価してください

- 訳出された日本語訳を
1. 全体的にまったく理解できない
 2. 部分的に理解可能な箇所はあるが、全体としては理解できない
 3. 全体的になんとなく理解はできるが、理解に自信がもてない
 4. 全体がほぼ、良く理解できる
 5. まったく問題なく、すべて良く理解できる

翻訳文2（看護学論文の抄録1件分）




日本語へ翻訳した文	理解度の評価基準				
	1	2	3	4	5
1（タイトル） 韓国人移民のための非公的支援サービス提供の経験と困難	1	2	3	4	5
2（目的） この論文は、韓国系アメリカ人の支援サービス提供の経験を移動するための研究報告である。	1	2	3	4	5
3（背景） アジアの人口の増加している高齢者のことにより、長期的、アジアのケア提供者たちに求められている長期の面倒重要視提起されるが、これらのケア提供者たちの問題としては、ほとんど注目されなかった。本研究では、韓国系アメリカ人のケア提供者の視点から独自のケア提供の経験の格差を把握しようとする。	1	2	3	4	5
4（方法） フォーカスグループアクセスを利用した質的研究のデザインを利用して、文化的文脈におけるケア提供者の経験を考察した。データは2005年8カ月間、収集された。様々な面のケア提供状況の24人の情報提供（主に女性）が、フォーカスグループインタビューに参加した。インタビューの時間は、それぞれのフォーカスグループ当たり約1（1/2）時間から2時間だった。2人のバイリンガルの研究者がテーマの分析を行った。	1	2	3	4	5
5（結果） ケア提供者の役割として、まず成果の信念、ケア提供の範囲と影響、そして、教育や文化への適応のための支援システムの必要性という3つのキーテーマが確認された。10のサブテーマが同定された：（1）二重の困難に直面していること、（2）の子として果たす虫にとっては、態度の変化（Hyo）、（3）ケアを提供すること、（4）憤むことができないと感ずること、（5）変化する家族の力関係を経験すること、（6）に接続されていること vs. 接続を提供すること、（7）恵みを戻してくれること、（8）自分で学ぶこと、（9）の違いを認識すること、（10）高齢者ケアのシステムを再考すること。	1	2	3	4	5

日本語へ翻訳した文	理解度の評価基準				
<p>6（結論） 韓国系アメリカ人の家族が記述したケア提供者の経験から、この急速に増加する人口に対して、もっと文化的に適切なサポートサービスだけでなく、より焦点を絞ったアウトリーチプログラムを確認して証明し、開発する必要性が指摘される。</p>	1	2	3	4	5

「翻訳文 2」全体について、あてはまる数字に○をつけて下さい。

この翻訳文全体は、原文の意味を把握するのに役立ったと感じますか

- 1) まったく役に立たない
- 2) あまり役に立たない
- 3) どちらともいえない
- 4) 少しは役に立つ
- 5) かなり役に立つ

次ページへ 

属性とオンライン機械翻訳の利用に関する以下の質問にお答えください。

質問1. 各項目のあてはまるものに○をつけるかご記入ください

- ・性別 1) 女性 2) 男性
- ・年齢 1) 20代 2) 30代 3) 40代 4) 50代 5) 60代
- ・職位 1) 助手 2) 助教
- ・学位 1) 専門士 2) 準学士 3) 学士 4) 修士 5) 博士
- ・臨床経験年数 []年 ・教員経験年数 []年
- ・語学力に関する資格試験のスコアなどがお分かりでしたらご記入ください

[英語]

TOEIC []点 TOEFL []点 英検 []級

IELTS []点 その他 []点、級、レベル

[英語以外の外国語]

()語、試験名():[]点、級、レベル

()語、試験名():[]点、級、レベル

質問2. 最近の3ヶ月間位で、ご自身の専門領域に関連する外国語の文献を何本程度読みましたか。以下の本数の中からあてはまるものを選び、番号に○をつけてください。

言語		本数
英語文献		① 0本 ② 1~5本 ③ 6~10本 ④ 11~20本 ⑤ 21本以上
英語以外の外国語の文献を読んだ場合	()語)	① 1~5本 ② 6~10本 ③ 11~20本 ④ 21本以上
	()語)	① 1~5本 ② 6~10本 ③ 11~20本 ④ 21本以上

次ページへ

質問3. 専門領域の外国語文献を読む際、どの程度の頻度で言葉の壁を感じますか。英語文献と、英語以外の文献を読まれる場合はその言語について、以下の数字からあてはまるものを選び、番号に○をつけてください。

言語		言葉の壁を感じる頻度
英語文献		① 感じたことは一切ない ② めったに感じない ③ ときどき感じる ④ 頻繁に感じる ⑤ 常に感じる
英語以外の外国語の論文を読む場合	(言語)	① 感じたことは一切ない ② めったに感じない ③ ときどき感じる ④ 頻繁に感じる ⑤ 常に感じる
	(言語)	① 感じたことは一切ない ② めったに感じない ③ ときどき感じる ④ 頻繁に感じる ⑤ 常に感じる

質問4. 専門領域の外国語文献を読む際に、オンライン機械翻訳を利用したことがありますか。あてはまる番号に○をつけてください。

1. 利用したことは無い（下記の1. の質問にお答えください）
2. 利用したことがある（下記の2. の質問にお答えください）

1. 「利用したことは無い」とお答えの場合：その理由について、あてはまるものに○をつけてください

- ① 必要ないから
② 役に立たないと思うから
③ オンライン機械翻訳の存在を知らなかった
④ その他の理由

[]

2. 「利用したことがある」とお答えの場合：以下1)～4)の質問に対して、あてはまるものに○をつけるかご記入ください。

1) ご利用の頻度はどの程度ですか

- ① 常に/ほぼ常に利用する ② ときどき利用する ③ ごくたまに利用する
④ 現在は利用していない → 理由：i 必要ないから ii 役に立たないと思うから
iii その他の理由 []

2) どの検索エンジンの機械翻訳を利用されたことがありますか（複数回答可）。

- ① ヤフー ② グーグル ③ MSN (Bing) ④ エキサイト ⑤ ニフティ

⑥ ライブドア ⑦ OCN ⑧ So-net ⑨ その他 []

⑩ どの検索エンジンか覚えていない

3) どの言語の組み合わせで利用されましたか（複数回答可）

①英語から日本語

②英語以外の外国語から日本語：[]語、[]語

4) 機械翻訳をご利用になった感想はいかがでしたか。英語と、英語以外の外国語（ご使用経験がある場合）に関して、以下の感想の中から、ご自身が経験した言語に関して選び、番号に○をつけてください。

言語		感想
英語→日本語		①かなり役に立つと感じた ②少しは役に立つと感じた ③どちらともいえない ④あまり役に立たないと感じた ⑤まったく役に立たないと感じた
英語以外の 外国語→ 日本語	()語	①かなり役に立つと感じた ②少しは役に立つと感じた ③どちらともいえない ④あまり役に立たないと感じた ⑤まったく役に立たないと感じた
	()語	①かなり役に立つと感じた ②少しは役に立つと感じた ③どちらともいえない ④あまり役に立たないと感じた ⑤まったく役に立たないと感じた

◆機械翻訳利用に関するその他の感想や、全体を通して何かコメントなどございましたらご記入ください。