ABSTRACT

The present dissertation addressed an important issue in machine translation. It aimed at analyzing and evaluating linguistic and cultural translation problems of Yemeni folklore resulted from machine translation programs (*Google Translate*, *Yandex Translate* and *Bing Translator*). It also attempted to compare the translations of the three mentioned machine translation programs to find out which program can produce a translation closer to that of the source language text in comparison with the human translator (Janet Watson). Moreover, it investigated translation strategies that are followed by the three programs in translating the Yemeni folklore into English. Besides, the study attempted to identify the reasons behind such problems. To achieve the objectives of the study in hand, the study followed a mixed qualitative and quantitative approach with descriptive, analytical and comparative methods. For qualitative analysis, a study sample of 46 examples were selected purposefully from the Yemeni work of Mutahar *Social Issues in Popular Yemeni Culture* (2002), representing the Sana'ani dialect as introduced in the famous folklore radio program *Musid wa Musida*. These excerpts were inserted into the three chosen machine translation programs. Then, the results were discussed and analyzed qualitatively to identify machine translation problems. The problems were classified into (7) linguistic problems and (4) cultural problems. For quantitative analysis, the study sample included (30) certified translators in Ibb Governorate. For data collection a five-page questionnaire was developed and distributed to the selected sample of translators. Their answers were collected, tabulated and analyzed quantitatively. The analysis showed that machine translation produced various types of problems when translating the Yemeni work of Mutahar, *Social Issues in Popular Yemeni Culture*, including linguistic and cultural problems. The findings also showed that *Bing Translator* program has less problems compared to *Google Translate* and *Yandex Translate* because it produced fewer errors (73.9%) whereas *Google's* errors rate was (78.3%) and *Yandex's* was (93.5%). The study concluded that the three mentioned programs used foreignization and literal translation strategies in the first place. However, *Bing Translator* sometimes adopted some oblique strategies such as cultural and lexical equivalents. Additionally, these programs used a ruled-based approach and statistical approach of machine translation while *Bing* program sometimes followed neural machine translation. According to the questionnaire's

results, there were many reasons that caused machine translation problems but the main ones were: cultural and structural differences between Arabic and English, deficiency of machine translation to understand culture context and lack of data resources especially in dialects. Machine translation problems were attributed to the fact that dialects are often confined to informal language. Finally, the study recommended program designers to develop machine translation and feed it with dialects resources and specialized dictionaries. Machine translation and translating dialects should be paid more attention by curriculum designers at Yemeni universities, as it should be integrated in the study plans and course specifications of translation programs.

ملخص الدراسة

تناولت هذه الدراسة موضوعًا هاماً في الترجمة الآلية هدفت من خلاله إلى تحليل وتقييم مشكلات الترجمة اللغوية والثقافية للتراث الشعبي اليمني الناتجة عن برامج الترجمة الآلية (*Google Translate*، *Yandex Translate*، *Bing Translator*). وقامت الدراسة بمقارنة ترجمات برامج الترجمة الآلية الثلاثة المذكورة آنفاً بالمترجم البشري جانيت واتسون (Janet Watson) لتحديد البرنامج الذي يستطيع إنتاج ترجمة أقرب إلى نص اللغة المصدر. كما هدفت الدراسة إلى تحديد استراتيجيات الترجمة التي تتبعها البرامج الثلاثة في ترجمة التراث الشعبي اليمني إلى اللغة الإنجليزية. بالإضافة إلى ذلك حاولت الدراسة الكشف عن الأسباب الكامنة وراء مثل هذه المشكلات. ولتحقيق هذه الأهداف، فقد أُستُخدِمَت المنهجية النوعية والكمية في البحث وأتبع الباحث طريقة الوصف والتحليل والمقارنة. وقد تم اختيار عينة قصدية مكونة من 46 مثالاً من العمل اليمني *قضايا اجتماعية في الادب الشعبي اليمني* (2002) للمؤلف اليمني عبدالرحمن مطهر. وهذه العينة تمثل اللهجة الصنعانية والتي كانت تُعرض في البرنامج الإذاعي الشعبي الشهير *مسعد و مسعدة*. وأُستخدِمَت برامج الترجمة الآلية الثلاثة المختارة في ترجمة العينة (46 مثالاً)، ومن ثم تم مناقشة النتائج وتحليلها لتحديد مشكلات الترجمة الآلية. وقد صُنِفت هذه المشكلات إلى (7) مشكلات لغوية، و(4) مشكلات ثقافية. ولغرض تحقيق أهداف الدراسة تم تصميم استبيان مكون من خمس صفحات وتم توزيعه على عينة مكونة من (30) مترجماً معتمداً في محافظة إب للإجابة عليه. وبعد ذلك تم جمع إجاباتهم وتبويبها وتحليلها كمياً. أظهرت نتائج التحليل أن الترجمة الآلية أنتجت أنواعًا مختلفة من المشكلات اللغوية والثقافية عند ترجمتها للعمل اليمني *قضايا اجتماعية في الادب الشعبي اليمني* كما أظهرت النتائج أن مشكلات برنامج *Bing Translator* كانت أقل مقارنة بالبرامج *Google Translate* و*Yandex Translate* لأن أخطائه الناتجة كانت بنسبة (73.9%) في حين كانت نسبة أخطاء *Google* (78.3%) و *Yandex* (93.5%). وخلصت هذه الدراسة إلى أن البرامج الثلاثة المذكورة استخدمت استراتيجيات التغريب والترجمة الحرفية بدرجة أساسية. ومع ذلك اعتمد مترجم *Bing* على بعض الاستراتيجيات غير المباشرة مثل المكافئات الثقافية والمعجمية و نظام الترجمة الآلية العصبية أحيانًا. كما

استخدمت هذه البرامج وخاصة (*Google* و*Yandex*) منهجًا قائمًا على القواعد النحوية و الترجمة الآلية الإحصائية. ووفقاً لنتائج الاستبيان هناك العديد من الأسباب التي أدت إلى مشكلات الترجمة الآلية من أهمها: الاختلافات الثقافية و التركيبية بين اللغتين العربية والإنجليزية، وعجز الترجمة الآلية عن فهم السياق الثقافي، ونقص مصادر البيانات خاصة في اللهجات، و تُعزى أيضاً مشكلات الترجمة الآلية إلى حقيقة أن اللهجات غالبًا ما تقتصر على اللغة غير الرسمية (غير الفصحى). وأخيراً أوصت الدراسة مصممي البرامج بتطوير الترجمة الآلية وتغذيتها بمصادر اللهجات والقواميس المتخصصة، كما يجب أن تحظى الترجمة الآلية وترجمة اللهجات بمزيد من الاهتمام من قبل مصممي المناهج في الجامعات اليمنية، حيث ينبغي دمجها في الخطط الدراسية وتوصيفات المقررات الدراسية لبرامج الترجمة.