

REVIEW ARTICLE

Available Online at www.jgrcs.info

GOOGLE TRANSLATION SERVICE ISSUES: RELIGIOUS TEXT PERSPECTIVE

Tariq Rahim Soomro¹, Gul Ahmad² and Muhammad Usman²

¹College of Engineering & Information Technology,
Al Ain University of Science & Technology, Al Ain, UAE
tariq.soomro@aau.ac.ae

²Department of Information Technology, SZABIST, Dubai, UAE
gulahamdms@yahoo.com, usmanoct@hotmail.com

Abstract: Google Translate service is among the widely used services of its kind and supporting translation of 64 International languages. This paper presents and explores the process of machine translation used by Google Translation Service (GTS) which translates a given text or Website to desire language automatically. The paper will explore the translation process furnished by Google Translate Service with focus on the issues of the method used by Google along with the possibilities to improve the system and better utilization of it from the user point of view. The paper will also focus on translation of religious text from English to Urdu and Arabic to Urdu issues along with issues related Quranic and hadith translations. Finally this paper will propose the solution to avoid such translations ambiguities.

Keywords: Machine Translation, Google Translation Service, Religious Translation.

INTRODUCTION

It is obvious that translation plays an important role in human communication. From ancient time until now, there have been debates about translation methods. Some scholars preferred word- for- word, whereas others prefer sentence-for- sentence. The translators have very important task of conveying the sense of the text from one language to another language. To facilitate the Internet users there are recently several Websites offer an automatic / machine translate services, which translate not only the individual words, sentences or even whole documents, but complete Websites and users of the Internet are having the wide range of choices among several machine translation services [1][2]. Google Translation Service (GST) offers, state-of-the-art free translation service and works automatically without the intervention of human translators. Currently the GTS supports translation between 64 languages.

The GTS is widely used service available to Internet users currently and allows users to translate text, documents or complete Websites. For several languages, one may see a speaker button near the translated text and by clicking this icon one can hear a machine-generated spoken version of translation. The Google translator allows translating whole documents, for example, in the form of PDF, TXT, DOC, PPT, XLS or RTF, by just clicking the “translate a document” link and submitting a file without the need for copying and pasting large blocks of text [3]. The Google has just improved their services by offering a translate gadget for Webmasters, which can integrate with the code of the Website, by inserting “Translator plug-in” into their Website to offer prospect visitors the opportunity to automatically translate their Website text into a desire language [1][3].

A machine translation (MT) system actually based on descriptions of both the source language and target language and keeping at all levels algorithm, formal grammars and vocabularies to perform translations. This machine translation process is based on following steps [1]:

- Analyze – source language text based on vocabulary, morphological and syntactical analysis
- Conversion (translation of source text to target text) and
- Synthesis – creation of text for target language based on syntactical and morphological appearance of text

All these steps in machine translation system may be interrelated closely and/or may be absent. The machine translation (MT) algorithm is usually performed by computer using application software. The text translated, by the machine translation may be then edited by users, to avoid ambiguities and mistakes [4]. This paper is organized as follow, next section will explore machine translation (MT) issues in general; section 3 will explore the problems of GTS in accordance with English to Urdu and Arabic to Urdu perspective; section 4 will also explore examples of religious translation issues; and finally concluding and suggesting few recommendations.

ISSUES IN MACHINE TRANSLATION (MT)?

According to Systran – one of the oldest machine translation company – machine translation (MT) is faster than human translation as Systran’s MT software can translate 3700 words a minute, while human can translate 2000 to 3000 words a day. The cost of MT is very lower than human translation and memory of MT software is no doubt better. The accuracy-wise human translation is very far from MT; as MT is making it sure that punctuation and spelling are correct as accurate [5]. Despite above benefits there are few issues in MT, which was conducted in study by Chang-Meadows Chinese-to-English translation in 2008. Chang-Meadows found in the Chinese particle “de” (的) an errors, resulting in confusion regarding who is doing what or to whom or who reports to whom etc., for example see following example: [6][7]



Figure: 1

In today’s networked environment more utilization of MT is desirable. According to the researchers English language will be no longer the native language of the Internet in near future as more non-English language users will use Internet [1][8][9]. Machine translation program may be divided into following categories [1]:

- a. Fully automatic
- b. 2. Machine-aided
- c. 3. Human-aided

The Google Translation Service (GTS) is no doubt widely used computer-aided translation service. While using GTS following question may arise:

- a. How efficient / deficient target language is?
- b. What are the common errors and disadvantages characterize by GTS?
- c. How GTS application functionally works?

According to the literature on machine translation problems, there are some semantic level problems; also there are some types of lexical or syntactic ambiguities. It is also expected that these available programs probably would be confused by such terms particularly in Arabic text also technical terms may create problems too [1]. According to [10][11] the first level of error are classified in following categories:

- a. Missing words
- b. Word order
- c. Incorrect words
- d. Unknown words and
- e. Punctuation errors

ISSUES RELATED TO GTS

Authors found and recognize some problems, while translating from English into Urdu as well from Arabic into Urdu. Let’s first have a look at the case of English into Urdu translation. Consider the following example, both normal text and religious text to Urdu from English and Arabic.



Figure: 2



Figure: 3



Figure: 4



Figure: 5

Translate from Arabic into Urdu “إن شاء الله“:



Figure: 6

Translate from Arabic into Urdu “ما شاء الله“

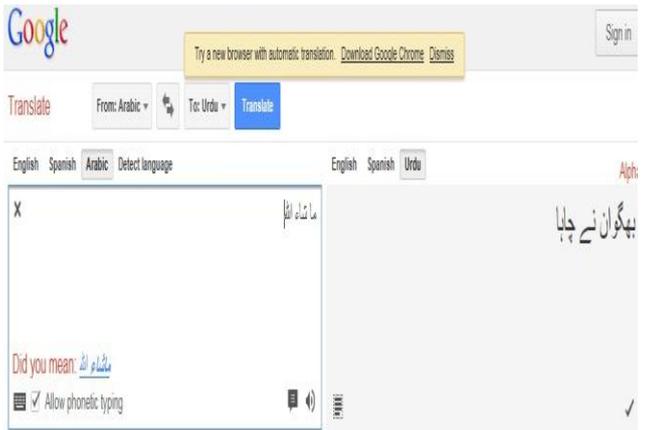


Figure: 7

Translate from Arabic into Urdu “جزاك الله خير“:



Figure: 8

Authors found that GTS is also facing similar issues, for example, “Missing words”, “Word order”, “Incorrect words”, “Unknown words” and “Punctuation errors” issues.

EXAMPLES OF QURANIC / HADITH TEXT

Quranic text and Hadith text is most important for Muslims all over the world. Here some examples of Quranic and Hadith will be presented. Both Quranic and Hadith need the best and most experienced translators and any such problem raised above may causes extremely serious problems [1][8].

Examples from Holy Quran:

The holy Quran or (sometimes spelled as Koran) is a book from Allah to His believers known as Muslims. This book was revealed on Prophet Muhammad (PBUH) in stages over 23 years. The text in Quran is regarded as words of Allah by Muslims. Therefore, it is expected by all Muslims that translation of Quranic text should be flawless [1][8]. Here are few examples:

“إِنَّا أَعْطَيْنَاكَ الْكُوثُرَ“:



Figure: 9

“فَصَلِّ لِرَبِّكَ وَأَنْتَ حَرٌّ“:

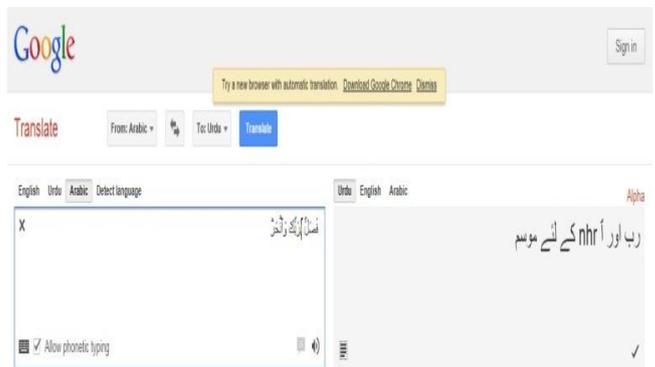


Figure: 10

Both examples from Quran are not properly generating correct translation.

Example form hadith (saying of the Prophet (PBUH)):

The Hadith is second important text for Muslims after Quranic text. Hadith is normally known as saying of Prophet Muhammad (PBUH) or recoding of His acts and traditions. Therefore, it is expected by all Muslims that translation of Hadith text should be flawless [1][8]. Here are few examples:

“إن الله رفيق يحب الرفق في الأمر كله“:



Figure: 11

“إن الدنيا سجن المؤمن وجنة الكافر”:

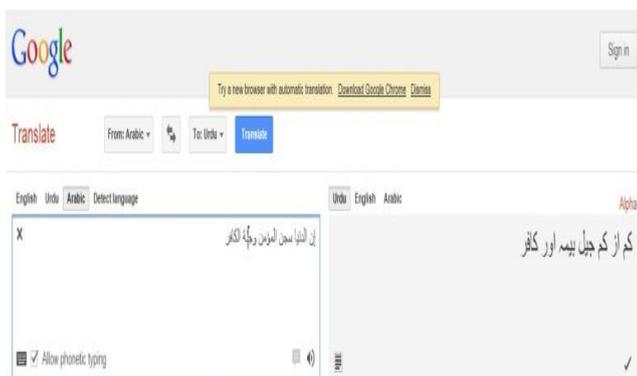


Figure: 12

Both examples from Hadith are producing unacceptable translation.

DISCUSSION & FUTURE WORK

Though a small samples of translations, provided in above sections, gave very clear idea of seriousness of the problem related to GTS. The examples shows normal words from English to Urdu generate incorrect translations; and religious text translation from Arabic to Urdu also generates unacceptable translation. Also the examples form Quranic and Hadith text clearly shows that translation generates by GTS is not satisfactory form not only religious point of view but from linguistic point of view. These automatic translation services are in their initial stages and providing not adequate translations. These incorrect generated results play a negative role towards these facilities. No doubt GTS online translators can give some idea about the meaning of the words and sentences, but their various flaws mentioned above prove that they are far from replacing professional translators. Authors are suggesting few recommendations for users and for these service providers to improve these translation issues or to avoid any confusion if any occur.

- User should not blindly trust these translations and should use their common sense, which may resolve translation ambiguities sometimes.
- For religious translation, use authentic human-translated Website, rather than automated Websites.
- These translation services may also refer authentic translations of religious text and improve their database to adopt those words, which are more appropriate and acceptable in religious translations
- These translation services may adopt different languages vowels to avoid any cultural ambiguities
- These translation services should focus on grammatical structure of several languages, otherwise they will be treated like dictionaries
- These translation services also update their database for Urdu language, as some of the translation are from Hindi language not Urdu and note that both Urdu and Hindi are separate languages
- These translation services also need to cooperate with linguistic professional of all languages to cope

with these issue, some of which are easily improvable by small programming changes and/or by improving and/or updating each language databases

REFERENCES

- Ahmed Abdel Azim ElShiekh, 2012, Google Translate Service: Transfer of Meaning, Distortion or Simply a New Creation? An Investigation into the Translation Process & Problems at Google, English Language and Literature Studies, Vol. 2, No. 1, March 2012
- Bell, Roger, 1992, Translation and Translating. Longman Group Ltd, London, UK
- Google Translate, <http://support.google.com/translate/?hl=en>, Retrieved January 2013
- Machine Translation, <http://encyclopedia2.thefreedictionary.com/Machine+Translation>, Retrieved January 2013
- Hannah Slavik, 2002, Translation and Interpretation, <http://www.diplomacy.edu/language/Translation/machine.htm>
- Jennifer DeCamp, What is Missing in User-Centric MT? The MITRE Corporation, http://www.mitre.org/work/tech_papers/tech_papers_09/09_2809/09_2809.pdf, Retrieved January 2013
- Shin Chang-Meadows, 2008, MT Errors in CH-to-EN MT Systems: User Feedback, Proceedings from the Association for Machine Translation in the Americas (AMTA 2008), Cambridge, MA
- Hatim, B., & Mason, J, 2003, 4th edition, Discourse and the Translator, Longman, Singapore, "Hello, World" Wired, http://www.bbc.co.uk/religion/religions/islam/texts/quran_1.shtml, Retrieved December 2012,
- Tariq Rahim Soomro, Ghassan Al-Qaimari, Internationalized Domain Name System (IDNS): Future of the Internet, 2007, Asian Journal of Information Technology 6 (10): 998-1002, 2007, ISSN: 1682-3915, DOI: [ajit.2007.998.1002](http://www.medwelljournals.com/fulltext/ajit/2007/998-1002.pdf), www.medwelljournals.com/fulltext/ajit/2007/998-1002.pdf
- David Vilar, Jia Xu_, Luis Fernando D'Haro and Hermann Ney, 2006, Error analysis of statistical machine translation output, In Proc. of the Fifth Int. Conf. on Language Resources and Evaluation (LREC), http://hmk.ffzg.hr/bibl/lrec2006/pdf/413_pdf.pdf
- Ariadna Font Llitjós, Jaime G. Carbonell, and Alon Lavie, 2005, A framework for interactive and automatic refinement of transfer-based machine translation, In Proc. Of the 10th Annual Conf. of the European Association for Machine Translation (EAMT), Budapest, Hungary, May 2005, <http://www.cs.cmu.edu/~aria/Papers/EAMT-2005-Font-Llitjos.pdf>