Journal of Applied Linguistics and Language Research
Volume 2, Issue 3, 2015, pp. 169-184
Available online at www.jallr.ir
ISSN: 2376-760X

Comparison of Google Online Translation and Human Translation with Regard to Soft vs. Hard Science Texts

Parisa Aslerasouli
Department of Foreign Language, Science and Research Branch, Islamic Azad University, Tehran, Iran

Gholam-Reza Abbasian
Assistant Professor, Imam Ali University and Islamic Azad University, South Tehran Branch, Tehran, Iran

Abstract
Machine translation (MT) is developing increasingly such that possibility of replacing Human Translation (HT) does not seem unlikely. However, it is controversial to what extent machine can replace human beings in certain areas including translation. To shed light on this issue, this study was conducted to compare the efficiency of Google Translation (GT) and HT in relation to Soft and Hard Science texts. Employing Waddington's (2001) model of Translation Quality Assessment (TQA), quality of five translation pieces of Soft and Hard Science texts translated both by GT software and HT was assessed in a bid to test three research hypotheses. The findings revealed that HT's quality was still higher than that of MT (i.e. GT); however, GT's efficiency was to some extent comparable to that of HT as far as the texts in Physics and Politics were concerned. So, it was concluded that in certain fields GT could produce roughly reliable translation if not totally as efficient as Human Translator.
Keywords: Google translate, hard science, human translation, machine translation, soft science, translation error

INTRODUCTION

Khodeir (2004, p.1) defines MT as "an automatic translation of text or speech from one language to another". "As a research and development field, MT is the oldest among the various sub-disciplines and applications of computer science to the study of natural language. MT is also a sub-discipline of computational linguistics" (Nirenburg 1939, p.3). According to Tikhmirova and Buonaiuto (2011) "the online translator was tested in March-April 2011 in English, Russian, Ukrainian, French and German. The work on the survey was continued and completed in October-December 2011, when to elaborate the results Polish was added and all language pairs were retested on the same texts" (para.2).

According to Lin, Murakami, Ishida, Murakami and Tanaka (2010), MT has been an important research topic for several decades in the area of artificial intelligence. More

Correspondence: Gholam-Reza Abbasian, Email: gabbasian@yahoo.com

and more MT services have been provided by companies with the expansion of the Internet environments, like Google1, Yahoo2, Microsoft 3 and so on. However, there is a huge gap between human and machine translators. On the one hand, translations made by machine translators always have limitations in qualities and therefore, they are not used more for translating documents with high requirement of qualities. On the other hand, finding bilingual human translators is another limitation because they are not available everywhere for any purpose at any time in the real world and translations of highly-trained bilingual individuals cost a lot in both labor and time. Moreover, MT suffers from some problem in comparison with HT in ambiguity, cultural transmit, and it needs Human intervention for betterment. Meanwhile, text type plays an important role in translations made both by MT and HT. Google Translate (GT) is one of these systems and is different from HT in terms of quality of translating different text types. It seems then inappropriate to replace HT totally by GT. Given these characteristics, the main problem relates to the nature and efficiency of MT; GT and HT as far as the quality of translation in terms of text type is concerned. By text type, it is meant Hard and Soft sciences. To address this problem, the purpose of the study is elaborated in the following section.

The current MT systems don't pay more attention to discourse and context in source language analysis, like sentence based systems. As a result, MT's quality then is seriously impaired (Zhang 2009). This study pinpoints, from quality and text type perspectives, the inadequacies of current MT systems in source language analysis compared with HT. The purpose is to analyze the features and functions of text and its influence on word and sentence meaning in text types, which is more specifically done through the analysis of the differences between MT and HT. Therefore, the present study aims at comparing MT and HT in terms of lexical, structural and contextual differences based on data gathered from translating a political text (soft science) and a physics text (hard science). Ultimately, as a Translation Quality Assessment (TQA) effort, this study focuses on comparing the quality of translations produced by human and machine as far as text type is concerned.

According to Arnold (1994), MT is one of the most interesting topics socially, politically, commercially, scientifically, intellectually, and philosophically, and one whose importance is likely to increase as the 20th Century ends and the 21st begins. As he declares in (p.4) "translation is necessary for communication, for ordinary human interaction, and for gathering the information one needs to play a full part in society". Given the evolution and development of MT and its significance in the age of technology, this study is of both theoretical and practical significance. Theoretically, sheds light on the possible role MT may play as HT does; probably replacing the latter one as the text types are concerned. Practically, the study will be effective economically if MT proves as effective as HT.

In a bid to empirically investigate the purpose of the study and tackle the problem, three following research questions were raised followed by respective research hypotheses:

- Is there any significant difference between the quality of Human and Google online translations?
- Is there any significance relationship between text types (i.e., Soft and Hard sciences) and translation modes (i.e., Google online and Human Translations)?
- Is there any relationship among the number of translation errors, translation modes (i.e., Google on line and Human Translations) and text types (i.e., Soft and Hard sciences)?

Hypotheses:

- There is not any significant difference between the quality of Human and Google online translations.
- There is not any significance relationship between text types (i.e., Soft and Hard sciences) and translation modes (i.e., Google online and Human Translations).
- There is not any relationship among the number of translation errors, translation modes (i.e., Google online and Human Translations) and text types (i.e., Soft and Hard sciences).

REVIEW OF THE LITERATURE

As Koponen (2010) says, TQA in the context of both HT and MT is important. In the field of HT, it may be used for quality control in professional settings or in translator training, but as far as MT is concerned, developer or potential user uses it to evaluate system performance. Accordingly, "The evaluation of MT systems is a vital field of research both for determining the effectiveness of existing MT systems and for optimizing the performance of MT systems" (Dorr, Snover and Madnani 2011, p.801).

Concerning the merits and demerits of GT, Butler (2011) asserted that today GT is known as a free online application and top of third party websites offering an automated translation of the content in any of the available languages. Using GT is fast, easy and it provides adequate general content translation for over 50 languages. It gathers data and finds information on sites that were previously inaccessible due to the language barrier. However, GT can misinterpret complex structures and provides inaccurate translations while one uses it, may not be aware any errors and inadequacies. Lazzari (2006, p. 25) describes MT and HT as follow:

> In terms of quality, MT will remain inferior to HT for many years. As a consequence, the various market segments will be dominated by one of two product offerings, either HT or MT. Which product offering dominates in any particular market segment will depend on the unique characteristics and demands of that segment. HT will prevail in all areas where high quality is an absolute necessity. In contrast, MT will take over the low end of the market, and it will also dominate in new markets or market segments which emerge as a consequence of the availability of low-cost translation technology. With quality and performance improvement over time, MT will move up market.

Dubey (2011) also stated:

> As Hutchins (2003, as cited in Aiken, Park and Lindblom, 2010) declared, the use of computers for translating texts was first proposed in 1947,and the first demonstration of a translation system was in January 1954. Following this trend, "MT is used to translate large amounts of text material for rapid acquisition of content" (Bostad, 1986, p.1). MT as personal computers appeared in 1981, and *Babel Fish* as the first, free, translation service in 1997, was appeared on the World Wide Web (Yang and Lange 1998, as cited in Aiken, Park and Lindblom, 2010).

As Carbonell and Tomita (1985) said, researchers of MT for three decades have been engaged in developing highly accurate, practically useful and fully automated translation systems. Current MT systems on isolating and correcting any errors that are committed in the automated translation phase facilitate the task of a human translator by translation aides and make the best MT programs, that in this case human intervention is required after the fact. In this regard Sager (1982, p.14) says:

> Editing of MT involves a variety of different skills and functions which are not comparable to revision of human translation. It can be performed, before, during or after machine processing of texts and there is no doubt that translators have a role to play in the performance of machine-assisted translation, in the post editing of MT output from systems designed for this type of human intervention and possibly also in the selection of source texts for one or another form of machine processing.

Text types (Hard and Soft sciences)

According to Puchala (2011, p.357):

> Translation is a very broad, complex and multi-faceted phenomenon, encompassing much more factors than it seems at first glance. It is not just copying the words from the original work while changing the language, but it consists of a careful selection of appropriate phrases and expressions, combining them together in a skillful way while taking into consideration numerous aspects, one of them being the text type.

According to Trosborg (1997) text type is defined as a specific mode of discoursing or mode of presenting that fulfilling a certain rhetorical and communicative purpose, is its goal. Also, in definition of text types, Hatim and Mason (1990, 140, as cited in Puchala, 2011, p.360) therefore, defined text types as "a conceptual framework which enables us to classify texts in terms of communicative intentions serving an overall rhetorical purpose".

This study pinpointed two text types named Hard and Soft sciences. Some scholars have defined these two sciences differently. Simms (2010) defined Hard Science and Soft science as follow: The hard sciences are concerned with physical entities while the soft sciences are concerned with living entities. Another definition was suggested by Haggan (2004) when he says, Hard Sciences generally prefer a title stating the exact topic of the paper that by adding more information or context specify certain aspects. He also defined Soft Sciences as favoring a broader range of methods and attempting to engage readers by rhetorical means. Simms (2010) also recommended that typifying the extant hard (natural) sciences, such as physics and chemistry is done by the identifying and measuring their subjects and the phenomena which influence these subjects. If the subjects and the phenomena which influence these subjects do not exactly identified and measured, typify the extant soft sciences, such as life and society.

THEORETICAL FRAMEWORK

"There are various rubrics in the literature of translation studies" (Khanmohammad and Osanlo, 2009, p.133). Rubrics that are more commonly used and are practical in the literature, according to Khanmohammad and Osanlo (2009) are of Farahzad (1992), Sainz (1992), Beeby (2000), and Waddington (2001). However given its comprehensive the last one was selected as the theoretical framework in order to assess the quality of both types of translations.

Among the four methods of assessment introduced by Waddington (2001) (method A, method B, method C and method D), the more well-known method is method A and it is functional in translation classes. This method is based on error analysis and possible mistakes are grouped under the following headings (Khanmohammad and Osanlo, p.136):

(i) Inappropriate renderings which affect the understanding of the source text and are divided into eight categories: counter-sense, faux sens, nonsense, addition, omission, unresolved extra-linguistic references, loss of meaning, and inappropriate linguistic variation (register, style, dialect, etc.)

(ii) Inappropriate renderings which affect expression in the target language and are divided into five categories: spelling, grammar, lexical items, text, and style.

(iii) Inadequate renderings which affect the transmission of either the main function or secondary functions of the source text.

"In each of the categories, a distinction is made between serious errors (-2 points) and minor errors (-1 point). There is a fourth category which describes the plus points to be awarded for good solutions (+1 point) or exceptionally good solutions (+2 points) to translation problems" (Khanmohammad & Osanlo, p.136).

Table 1. Serious and minor errors in Waddington's Method A, Extracted from Khanmohammad and Osanlo (2009)

Inappropriate rendering on understanding ST	Omission
	Addition

	Nonsense
	Faux sense
	Counter-Sense
Inappropriate rendering on TL	Style
	Text
	Lexicon
	Grammar
	Spelling
Inadequate rendering	Main function of ST
	Secondary function of ST
Good solution	+1 point
	+2 point

Waddington (2001, p.314, as cited in Khanmohammad and Osanloo, 2009, p.136) describes method B as “being based on error analysis and designed to take into account the negative effect of errors on the overall quality of the translations where the rater first has to determine whether each mistake is a translation mistake or just a language mistake. This is done by deciding whether or not the mistake affects the transfer of meaning from the source to the target text”. If the mistake affects the meaning, it is called translation mistake and if it doesn't, it is a language error and is penalized with -2 points. “However, in the case of translation errors, the rater has to judge the importance of the negative effect each error has on the translation, and taking in to consideration the objective and the target reader specified in the instructions to the candidates in the exam paper” (ibid). In order to judge this importance, table 2 is suggested to the rater (ibid).

Table 2. Typology of errors in Waddington’s Method B, Extracted from Khanmohammad and Osanlo (2009)

Negative effect on the words in ST	Penalty for negative effect
1-5 words	2
6-20 words	3
21-40 words	4
41-60 words	5
61-80 words	6
81-100 words	7
100 + words	8
The whole text	12

Calculating the final mark of each translation is like that of method A, in which there is a fixed number for positive points (in the case of method B this is 85). Then the total number of negative points is subtracted from 85, and finally the result is divided by 8.5. In describing Method C, Waddington (2001) believes that this method (method C) is a holistic method of assessment. “The scale is unitary and treats the translation competence as a whole, but requires the rater to consider three different aspects of the student’s performance” (Khanmohammad & Osanloo, p. 136), as shown in Table 3 below. “For each of the five levels, there are two possible marks” (ibid). “This allows the rater freedom to award the higher mark to the candidate who fully meets the requirements of a particular level and the lower mark to the candidate who falls between two levels but is closer to the upper one" (ibid).

Table 3. Scale for the holistic Method C, Extracted from Khanmohammad & Osanloo (2009)

Level	Accuracy of transfer of ST content	Quality of expression in TL	Degree of task completion	Mark
Level 5	Complete transfer of ST information; only minor revision needed to reach professional standard.	Almost all the translation reads like a piece originally written in ST. There may be minor lexical, grammatical, or spelling errors.	Successful	9,10
Level 4	Almost complete transfer; there may be one or two insignificant inaccuracies that require a certain amount of revision to reach professional standard.	Large sections read like a piece originally written in ST. There are a number of lexical, grammatical, or spelling errors.	Almost completely successful	7,8
Level 3	Transfer of the general idea (s) but with a number of lapses in accuracy; Needs considerable revision to reach professional standard.	Certain parts read like a piece originally written in ST but others read like a translation. There are Considerable number of lexical, grammatical, or spelling errors.	Adequate	5,6
Level 2	Transfer undermined by serious inaccuracies; thorough revision required to reach professional standard.	Almost the entire text reads like a translation; there are continual lexical, grammatical, or spelling errors.	Inadequate	3,4
Level 1	Totally inadequate transfer of ST content; the translation is not worth revising.	The candidate reveals a total lack of ability to express himself/herself adequately in target language.	Totally inadequate	1,2

Finally, Waddington (2001, p.315, as cited in Khanmohammad & Osanloo, 2009, p.136) defines the last method as, "A method which consists of combining error analysis Method B and holistic Method C in an appropriation of 70/30; that is to say that Method B accounts for 70% of the total result and Method C for the remaining 30%".

METHOD

Participants

In order to do this study, the researcher sought the participation of thirty BA students of English Translation both male and female to translate the selected texts.

Instrumentations

The instrument needed for the research was:

Translation assessment rubric. For the purpose of this research, Waddington's rubric was used. Waddington (2001) introduces four methods of assessment (Method A, B, C and D) that in this study the second method (B) was applied. This rubric is based on error analysis and takes into account the negative effect of errors on the quality of translation. First of all, the rater should determine that each error is a translation error or language error. As mentioned in Khanmohammad and Osanloo (2009), if an error affects the transfer of meaning from the source to target text, it is called translation error, otherwise, it is a language error and penalized with -2 points. Then, it should be noted that for 1-5 negative effects on words, it penalizes 2 negative points and for 6-20 negative effects, it assigns 3 negative points and so on. In this study, the total number was 85 that according to Waddington's model the rater subtracted the negative points from 85 then divided the result by 8.5. The scores acquired from the calculations are the score assigned for each translation.

Corpus of the study

The corpus of the study was five short paragraphs of soft science (Political text) and hard science (Physics text) that are given to MT and the BA students of translation to translate them in order to determine the Quality, similarity and difference of MT and HT according to that text types (soft and hard).

Procedure

In order to do this study, sample texts, including five short paragraphs of soft science including political texts, and five short paragraphs of hard science, including texts related to physics, were given to both Google online and to the students to translate them. Then the translated texts were analyzed in the following steps:

- Quantitative evaluation based on Waddington's rubric. The researcher scored each translation both Google and human, then calculated the mean of the scores in order to answer the first research question.
- Quantitative evaluation with regards to text type. Mean of the scores of hard and soft science texts both in Google and HT were calculated in order to answer the second research question.
- Text analysis to draw frequency of errors in both Google online and HT in terms of text type in order to answer the third question.

The researcher used Waddington's second method to analyzing quantitative data that was "based on error analysis and designed to take into account the negative effect of errors on the overall quality of the translations" (Khanmohammad & Osanloo, 2009, p. 136).

Data Analysis

In order to analysis the obtained data for Research question 1, an independent samples t-test was applied to compare the quality of Human and Google online translations. A two-way ANOVA was applied to determine any significance relationship between text types and translation modes, that whether the text types affect the translation modes or not. This method was applied in answering the second research question.

In order to determine whether there is any relationship among the number of translation errors, translation modes and text types, frequency analysis through chi-square was done. Finally, inter rater reliability was estimated to probe the reliability of the collected data.

RESULTS AND DISCUSSION

Investigation of Research Question 1

The first research question was entitled "Is there any significant difference in the quality of Human and Google online translations"? An independent samples t-test was run to compare the quality of Human and Google online translations. As displayed in Table 4 the HT ($M = 9.69$, $SD = .066$) showed a slightly higher mean than the Google online translation ($M = 9.50$, $SD = .00$).

Table 4. Descriptive Statistics; Human and Google Translations

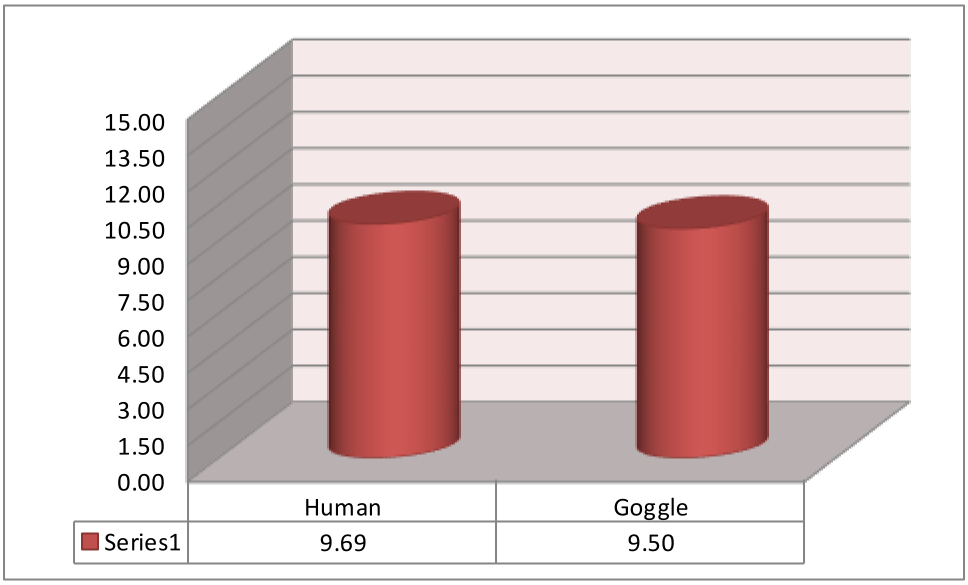
	Mode	N	Mean	Std. Deviation	Std. Error Mean
Translation	Human	60	9.69	.066	.008
	Goggle	2	9.50	.000	.000

The results of the independent samples t-test ($t\ (59) = 22.17$, $P < .05$, $R = .95$ which represents a large sample size) as showed in table 5, indicated that there was a significant difference in the quality of Human and Google online translations. Thus the first null-hypothesis **was rejected**. The HT revealed a statistically significant higher mean than the Google translation.

Table 5. Independent t-test; Human and Google Translations

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. 2-tailed	Mean Diff.	Std. Error	95% Confidence	
								Lower	Upper
assumed	5.80	.019	4.016	60	.000	.188	.047	.094	.281
not assumed			22.17	59.00	.000	.188	.008	.171	.204

It should be noted that the assumption of homogeneity of variances was not met (Levene's $F = 5.80$, $P < .05$). That is why the second row of Table 5, i.e. "Equal variances not assumed" was reported above.



Graph 1. Human and Google Translations

Furthermore, Graph 1, in line with the respective statistics, illustrates the mean scores of both Human and Google system in the translation of the material and as discussed before HT displayed a slightly higher mean than GT.

Investigation of Research Question 2

The second research question was entitled "Is there any significance relationship between text types (i.e., Soft and Hard sciences) and translation modes (i.e., Google online and Human Translations)"? A two-way ANOVA was run to probe any significance relationship between text types (i.e., Soft and Hard sciences) and translation modes (i.e., Google online and Human Translations). Before discussing the results of the two-way ANOVA it should be noted that the assumption of homogeneity of variances was met (Levene's $F = 1.92$, $P > .05$). Thus there is no need to do any corrections on the results of the two-way ANOVA.

Table 6. Homogeneity of Variances Assumptions; Translation by Modes and Text Types

F	df1	df2	Sig.
1.921	3	58	.136

The main results (Table 7) include three F-values the first of which was discussed above. There is a significant difference between the quality of the Human and Google translations ($F(1, 58) = 15.65$, $P < .05$, Partial $\eta^2 = .21$ which represent a large effect size). These results further support the conclusions made above when discussing the first research question.

Table 7. Two-Way ANOVA; Translation by Modes and Text Types

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Mode	.068	1	.068	15.656	.000	.213
Type	.000	1	.000	.008	.930	.000
Mode * Type	.000	1	.000	.008	.930	.000
Error	.252	58	.004			
Total	5811.613	62				

The types of texts did not have any significant effect on the quality of the Human and Google translations (F (1, 58) = .008, P > .05, Partial η^2 = .00 which represents a weak effect size). As displayed in Table 8 the physics (M = 9.59) and politics (M = 9.59) showed almost the same means on the translation.

Table 8. Descriptive Statistics; Translation by Types of Texts

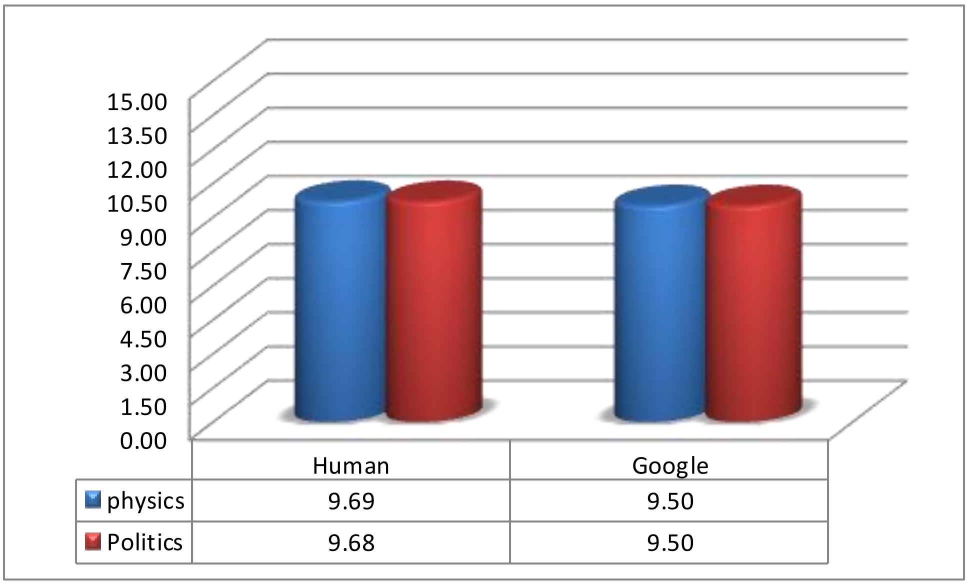
Major	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
physics	9.596	.034	9.529	9.663
Politics	9.592	.034	9.525	9.659

There was not any significant interaction between types of texts and modes of translation (F (1, 58) = .008, P > .05, Partial η^2 = .00 which represents a weak effect size). As displayed in Table 9 the HT showed slightly higher means on both physics (M = 9.69) and politics (M = 9.68) translations than the Google whose means are 9.50 on both physics and politics translations. Thus it can be concluded that the second null-hypothesis **was not rejected**. There was not any significance relationship between text types (i.e., Soft and Hard sciences) and translation modes (i.e., Google on line and Human Translations).

Table 9. Descriptive Statistics; Interaction between Types of Texts and Modes of Translation

Mode	Major	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
Human	physics	9.692	.012	9.668	9.716
	Politics	9.683	.012	9.659	9.707
Goggle	physics	9.500	.066	9.368	9.632
	Politics	9.500	.066	9.368	9.632

In addition, Graph 2, in line with the respective statistics, illustrates the mean scores of both Human and Google systems with regards to text types (physics and Politics) that show a slightly higher means on both text types translated by Human.



Graph 2. Translation by Modes and Text Types

Investigation of Research Question 3

The third research question was entitled "Is there any relationship among the number of translation errors, translation modes (i.e., Google on line and HTs) and text types (i.e., Soft and Hard sciences)"? An analysis of chi-square was run to probe any relationship among the number of translation errors, translation modes (i.e., Google on line and HTs) and text types (i.e., Soft and Hard sciences). The results (χ^2 (1) = .10, P > .05) (Table 10) indicated that the third null-hypothesis was not rejected.

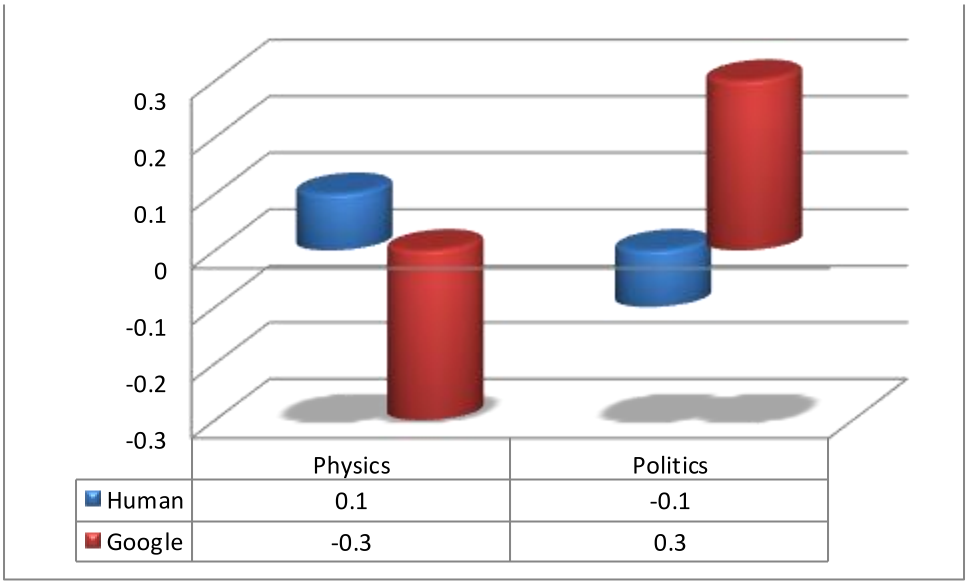
Table 10. Chi-Square; Human and Google Translation Errors (Soft and Hard Sciences)

	Value	df	Asymp. Sig. (2-sided)
Continuity Correction[b]	.108	1	.743
a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 38.16.			
b. Computed only for a 2x2 table			

Table 11 displays the frequencies, percentages and standardized residuals for the Human and Google translation errors on soft and hard sciences. The firmer two indices are descriptive ones based on which no inferences can be made; however, the latter index – Std. Residual – is an inferential statistic. Any Std. Residual beyond the ranges of +/- 1.96 indicates significant differences between the frequencies of translation errors. Since none of the Std. Residuals were beyond +/- 1.96 it can be concluded that there was not any significant relationship among the number of translation errors, translation modes (i.e., Google on line and Human Translations) and text types (i.e., Soft and Hard sciences).

Table 11. Frequencies, Percentages and Std. Residuals; Human and Google Translation Errors (Soft and Hard Sciences)

			Major		Total
			Physics	Politics	
Mode	Human	Count	221	198	419
		% within Book	52.7%	47.3%	100.0%
		Std. Residual	.1	-.1	
	Google	Count	40	40	80
		% within Book	50.0%	50.0%	100.0%
		Std. Residual	-.3	.3	
Total		Count	261	238	499
		% within Book	52.3%	47.7%	100.0%



Graph 3. Std. Residuals; Human and Google Translation Errors (Soft and Hard Sciences)

Graph 3, shows the standardized residuals for Human and GT errors on Soft and Hard sciences. It showed 0.1 for HT in Physics but -0.3 for GT, and in Politics it was estimated -0.1 for HT and 0.3 for GT.

Inter-Rater Reliability

To make sure of the reliability of the data collected through the instruments, inter rater reliability was estimated. For this purpose, Pearson correlations was run to probe the inter-rater reliability between the two raters who rated the political, physics, translation made by human and Google as follows:

A: There is a significant inter-rater reliability ($R(30) = .83$, $P < .05$) between the two raters who rated the human and Google translations of political texts.

Table 12. Inter-Rater Reliability; Political Texts

		Google
Human	Pearson	.835**
	Sig. (2-tailed)	.000
	N	30
**Correlation is significant at the .01		

B: There is a significant inter-rater reliability (R (30) = .89, P < .05) between the two raters who rated the human and Google translations of physics texts.

Table 13. Inter-Rater Reliability; Physics Texts

		Google
Human	Pearson Correlation	.890**
	Sig. (2-tailed)	.000
	N	30
**Correlation is significant at the .01 level (2-tailed)		

Summary of findings

This study was carried out on both translations (human and machine) to investigate their qualities. The sample consisting of 30 undergraduate translation students both male and female was selected randomly. Then two text types one from Hard (Physics) and the other one from Soft (Politics) were given to the participants. The data gathered in this study were analyzed both qualitatively and quantitatively based on Waddington's (2001) Model B.

The findings indicated that:

- There is statistically a significant difference in the quality of HT and MT (i.e. GT) in favor of HT.
- Mode of translation affects its quality but text type does not have any significant effects on translation quality.
- No statistically significant relationship exists among translation errors and translation modes.

CONCLUSION

Observing the difference between the means in both HT (M= 69) and GT (M= 9.50) it is concluded that HT has a slightly higher mean than the GT. Then, the first hypothesis was rejected, because there was a significant difference in the quality of both translation modes in which t= 22.17, where shows a large sample size. Second hypothesis was not rejected because as estimated in chapter four, quality of HT and GT was F= 15.65, P< .05 where shows significant difference, while it was F= .008, P> .05 in relation between text types and translation modes, but that shows a weak effect size. So, we conclude that type of texts does not have any significant effect on the quality of HT and GT.

The results obtained from chi-square show that (X^2 (1) = .10, P> .05) there isn't any relationship among the number of errors, translation modes and text type. As shown in table 11, none of the Std. Residuals were beyond +/− 1.96 and it can be concluded that there isn't any significant relationship among the number of translation errors, translation modes and text types.

Automatic MT is a vast and open ended area. We assume that students' genre familiarity, content and formed schemata can affect the quality of their translation. Other studies may select samples, who, are familiar with the genre features of the text and have the necessary content and formed schemata about the text and compare the result with those who have no familiarity with the text. Or two texts one related to students' field one which is different, can be given to the same group.

REFERENCES

Ablanedo, J., Aiken, M., & Vanjani, M. (2007). Efficiency of English to Spanish Automatic Translation. *International journal of information and operations management education*, *2*(2), 194-210.

Aiken, M., & Balan, S. (2011). An Analysis of Google Translate Accuracy. *Translation journal, 16* (2), 12-34.

Aiken, M., Ghosh, K., Wee, J., & Vanjani, M. (2009). An Evaluation of the Accuracy of Online Translation Systems. *Communications of the IIMA*, *9* (4), 69-84.

Aiken, M., Park, M., & Lindblom, T. (2010). Integrating Machine Translation with Group Support Systems. *International journal of business and management*, *5* (5).

Al kabi, M. N., Hailat, T. M., Al shawakfa, E. M., & Alsmadi, I. M. (2013). Evaluating English to Arabic Machine Translation Using BLEU. *International journal of advanced computer science and application*, *4* (I), 67-71.

Arnold, D. (1994). Text Typology and Machine Translation: an Overview. *Language and Linguistic*, University of Essex, UK, 73-74.

Bostad, D. A. (1986). Machine Translation in the USAF. *Terminologie et traduction,* (1), 12-68.

Butler, F. (2011, February 20). Machine versus Human: Will Google Translate Replace Professional Translators? pp. 4-5.

Carbonell, J., & Tomita, M. (1985). New Approaches to Machine Translation. *Proceedings of the conference on theoretical and methodological issues in machine translation of natural languages,* New York: Colgate University Hamilton, August, pp. 60-61.

Coughlin, D. (2003). Correlating Automated and Human Assessments of Machine Translation Quality. *About MT*. Retrieved September, 20014, para.42, from http://mt-archive.info/MTS-2003-Coughlin.pdf

Dorr, B., Snover, M., & Madnani, N. (2011). Machine Translation Evaluation. *Handbook of natural language processing and machine translation*, type book chapters Publication, Springer NewYork, 801. Retrieved from http://cs.cmu.edu/GALE-book-ch5.pdf

Dubey, P. (2011). Overcoming the Digital Divide through Machine Translation. *Translation journal*, *15* (1), 1-15.

Farouk khodeir, H. (2004). *A Framework for Phrase Based SMT in the Technical Translation Domain.* Unpublished thesis, department of computer science.

Gomes, R., & Anastasiou, D. (2011). Comparison of SYSTRAN and Google Translate for English-Portuguese. *About MT,* pp. 119-135. Retrieved from http://revistes.uab.cat/Tradumatica/article/248906

Haggan, M. (2004). Research Paper Titles in Literature, Linguistic and Science: Dimensions of Attraction. *Journal of pragmatics*, *35*(2), 293-317.

Hampshire, S., & Salvia, C. (2010). Translation and the Internet: Evaluating the Quality of Free Online Machine Translators. *About quality of MT*, 206-207. Retrieved from http://raco.cat/index.php/quadernstraduccio/article/viewfile/260430

Khanmohammad, H., & Osanloo, M. (2009). Moving toward Objective Scoring: a Rubric for Translation Assessment. *JELS*, *1*, (1), fall 2009, 135-138. Retrieved from http://sid.ir/en/VEWSSID/J_pdf/1032120090108.pdf

Koponen, M. (2010). Assessing Machine Translation Quality with Error Analysis. *Electronic proceeding of the KaTu symposium on translation and interpreting studies*.

Lazzari, G. (2006). From Human Translation to MT: Dramatic Costs Reduction and Accessibility Improvement. *Human language technologies for Eroup*, April, 25.

Lin, D., Murakami, Y., Ishida, T., Murakami, Y., & Tanaka, M. (2010). Composing Human and Machine Translation Services. *Language grid for improving localization process*, 500.

Nirenburg, S. (1939). *Readings in Machine Translation: Introduction*. In Nirenburg, S. Somers, H & Wilks, Y (Eds), a Bradford Book, Cambridge, Massachusetts, London, England, the MIT Press.

Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. (2002). BLEU: a Method for Automatic Evaluation of Machine Translation. *Proceedings of the 40th annual meeting of the association for computational linguistics (ACL),* July, Philadelphia, 311-318.

Puchala, K. (2011). Text Typology and its Significance in Translation. *ZESZYT* (Vol. 69, pp. 357-360). Retrieved from http://publikationen.ub.uni-frankfurt.de/opus4/files/32477/puchala_Text_typology_2011.pdf

Sager, J. C. (1982). Types of Translation and Text Forms in the Environment of Machine Translation (MT). *ASLIB*, North-Holland publishing company, pp. 11-14

Seljan, S., Brkic, M., & Kucis, V. (2011). Evaluation of Free Online Machine Translations for Croatian-English and English-Croatian language pairs. *Information Sciences and E-society*, pp. 342-343.

Simms, J. R. (2010). Making the Soft Science hard: the Newton Model. *Systems Research and Behavioral Science,* DOI:10.1002/sres.1033

Tikhomirova, E., & Buonaiuto, N. (2011). Compatibility of Languages and Other Restrictions in the Statistical Translation by Google Translate. *About Translation*. Retrieved from http://TranslationDirectory.com

Trosborg, A. (1997). Text Typology: Register, Genre and Text Types. *Text Typology and Translation*, In Anna Trosborg (Ed), pp. 3-23.

Waddington, C. (2001). Different Methods of Evaluating Student Translations: the Question of Validity. *Meta, XLVI*, Vol. 2, pp. 311-315.

Zhang, J. (2009). Discourse, Context and Machine Translation Quality. *Shandong foreign language teaching journal*, 06, para. 1. Retrieved from http://cnki.com.cn