On the Translation Quality of Google Translate:

With a Concentration on Adjectives

[1] Abouzar Oraki

ID: 1034

English Language Department

Vali-e-Asr University of Rafsanjan,

Abstract

Translation, whose first traces date back at least to 3000 BC (Newmark, 1988), has always been considered time-consuming and labor-consuming. In view of this, experts have made numerous efforts to develop some mechanical systems which can reduce part of this time and labor. The advancement of computers in the second half of the twentieth century paved the ground for the invention of machine translation (MT). One of the most commonly used MT systems is Google Translate, which currently supports 64 languages, including Persian. In consideration of the fact that Google Translate is easily accessible, it is almost always the first MT system to which Iranian users resort for meeting their translation needs, the reason which prompted the researcher to conduct a study on the output quality of the English-to-Persian translations produced by this MT system. To narrow down the study, the researcher decided to investigate the translation of adjectives of all types, i.e. simple adjectives, adjectives of similarity, comparatives and superlatives. To this end, a test suite of 140 sentences (randomly chosen among 1400 sentences) containing all types of adjectives was prepared and translated by Google Translate. The analysis of the results revealed that Google Translate translates simple, comparative and superlative adjectives fairly accurately and naturally, while it fails to distinguish adjectives of similarity and, thus, translates them quite inaccurately.

Key Terms: machine Translation, Google Translate, adjectives, accuracy

1. Introduction

The amount of translation worldwide has increased considerably over the last fifty years and in this workplace, the role of technology, in particular, the Internet and Machine Translation has been prominent. Machine Translation, also known as MT, is a sub-category of computational linguistics which analyses the use of computer software to translate written or spoken text from one language to another.To elaborate on the MT, it should be mentioned that the first real developments in MT appeared quite a while after the Second World War, a time during which the first computers had been newly designed in the UK (Hatim and Munday, 2004).

Among all MT software, Google Translate is the most commonly used worldwide. The owners of Google Translate assert that Google Translate is a free, automatic translator. They

[1] English Language Department, Vali-e-Asr University of Rafsanjan, Iran   Email: a.oraki@vru.ac.ir

also go further to say that it currently supports translation between dozens of languages. But there has always been some skepticism about the translation quality of such software. As Newmark (1988) puts it, the concept of quality in translation is somehow "relative".

Many different translation scholars have worked on the concept of translation quality. Among such scholars Reiss, 1972, Wilss, 1977, Amman, 1990, Gerzymisch-Abrogasts, 1977 can be referred to. But, according to Cary and Jumpelt (1963, cited in Khoshsima and Rostami, 2010), defining the quality of translation was first discussed in the Third Conference of the International Federation of Translators on Quality in 1959.

Needless to say, finding out the quality of a translation is no easy job. As House (1997, cited in Khoshsima and Rostami, ibid) asserts:

> "Translation quality is a problematical concept if it is taken to involve individual and externally motivated value judgment alone. Passing any "final judgment" on the quality of a translation that fulfills the demands of scientific objectivity is very difficult indeed (p. 2)."

To find out the quality of such software, the researcher carried out the present research. In order to narrow down the research, the decision was made to carry out the research on the different types of adjectives only. The present study, in brief, has tried to answer the following question and sub-questions:

To what extent is Google Translate able to translate the followings as accurately as possible?

a)   Simple adjectives
b)   Comparative adjectives
c)   Superlative adjectives
d)   Adjectives of similarity

2. Review of Literature

As pointed earlier, it is thought that in the short term, MT is likely to be of most benefit to large corporate organizations doing a lot of translation. It is of paramount importance from different aspects. As Tauschel (2008) puts it, the social or political significance of MT emanates from the socio-political significance of translation in communities where more than one language is generally spoken. The commercial importance of MT arises from the fact that translation itself is commercially important. Scientifically, MT is interesting, because it is an obvious application and testing ground for many ideas in Computer Science, Artificial Intelligence, and Linguistics, and some of the most important developments in these fields have begun in MT. Philosophically, MT is interesting, because it represents an attempt to automate an activity that can require the full range of human knowledge — that is, for any piece of human knowledge, it is possible to think of a context where the knowledge is required.

But it should also be noted that MT output must be examined and evaluated so that corporate organizations and, on the whole, those who are intend to make use of this technology get to know the ups and downs of what they are longing for.

As Linares (2008) asserts, in the context of MT system development, evaluation methods are necessary for three main purposes:

• **Error Analysis**, i.e., to detect and analyze possible cases of error. Fine knowledge of the system capabilities is essential for improving its behavior.

• **System Comparison**, i.e., to measure the effectiveness of the suggested mechanisms. This is done by comparing different versions of the same system. It is also common to compare translations by different systems, so system developers may borrow successful mechanisms from each other. This allows the research community to advance together.

• **System Optimization**, i.e., the adjustment of internal parameters. Typically, these parameters are adjusted so as to maximize overall system quality as measured according to an evaluation method at choice.

Evaluating the MT is not an easy job. As Arnold et al (2001) put it:

> "The evaluation of MT systems is a complex task. This is not only because many different factors are involved, but because measuring translation performance is itself difficult. The first important step for a potential buyer is to determine the translational needs of her organization. (p. 157)"

Arnold (ibid) goes further to say that evaluating translation quality is not only a problem for MT. He asserts that it is a practical problem that human translators face, and one which translation theorists have puzzled over. As he asserts, for human translators, the problem is that there are typically many possible translations, some of them faithful to the original in some respects (e.g. literal meaning), while others try to preserve other properties (e.g. style, or emotional impact).

Arnold (ibid) introduces some of the methods that have been used to date. The followings are what he has mentioned:

Intelligibility:

A traditional way of assessing the quality of translation is to assign scores to output sentences. A common aspect to score for is Intelligibility, where the intelligibility of a translated sentence is affected by grammatical errors, mistranslations and untranslated words. Some studies also take style into account, even though it does not really affect the intelligibility of a sentence. Scoring scales reflect top marks for those sentences that look like perfect target language sentences and bottom marks for those that are so badly degraded as to prevent the average translator/evaluator from guessing what a reasonable sentence might be in the context.

Accuracy:

By measuring intelligibility we get only a partial view of translation quality. A highly intelligible output sentence need not be a correct translation of the source sentence. It is important to check whether the meaning of the source language sentence is preserved in the

---

[1] English Language Department, Vali-e-Asr University of Rafsanjan, Iran   Email: a.oraki@vru.ac.ir

translation. This property is called Accuracy or Fidelity. Scoring for accuracy is normally done in combination with (but after) scoring for intelligibility.

Error Analysis:

Rather than using broad indicators as guides to score assignments, you could focus on the errors the MT system makes. The technique of error analysis tries to establish how seriously errors affect the translation output.

Test Suite:

In software development, a test suite, less commonly known as a validation suite, is a collection of test cases that are intended to be used to test a software program to show that it has some specified set of behaviors. A test suite often contains detailed instructions or goals for each collection of test cases and information on the system configuration to be used during testing. To succeed in correctly translating all the sentences in a source language (SL) test suite into a target language (TL) would definitely be encouraging. However, standard test suites are rather blunt instruments for probing translation performance in the sense that they tend to ignore typical differences between the languages involved in translation.

3. **Methodology and procedures**

In order to answer the above-mentioned question and sub-questions, a test suite of 140 sentences (randomly chosen among 1400 sentences), as the raw material, containing all of the aforementioned types of adjectives was prepared and translated by Google Translate:

Table1. Sentences containing Simple adjectives

He is a **nice** man.	او مرد **خوبی** است.
They are very **tall**.	آنها بسیار **بلند** هستند.
She is very **beautiful**.	او بسیار **زیبا** است.
Ali is a **short** boy.	علی پسر **کوتاه** است.
My teacher is very **kind**.	معلم منبسیار**مهربان** است.
This room is very **dark**.	این اتاق بسیار**تاریک** است.
Jack is a **smart** student.	جک یکدانشجوی**هوشمند**است.
John is **lazy**.	جان **تنبل** است.
He is so **ugly**.	او خیلی **زشت**است.
That box is **heavy**.	که جعبه های **سنگین**است.
He has a **good** job.	او یک کار **خوب**.
The weather is **awful**.	آب وهوای**افتضاح**است.
It is not a **bad** car.	آن است که یک ماشین**بد** نیست.
That house is very **expensive**.	این خانه بسیار **گران** است.
My car is very **cheap**.	ماشین من بسیار **ارزان**است.
He has a **new** shirt.	او یک پیراهن **جدید**.
It is an **old** house.	این خانه های **قدیمی**است.
I sold my **dilapidated** car yesterday.	من ماشین **ویران** من دیروز به فروش می رسد.

He is very **poor**.	او بسیار **ضعیف** است.
His father is **rich**.	پدر او **غني** است.
The cake tasted **sweet**.	کیک طعم **شیرین** است.
He has a **soft** bed.	او دارای یک بستر**نرم**.
Sarah is **young**.	سارا **جوان** است.
Her mother is **old**.	مادر او **قدیمي** است.
That box is **empty**.	که جعبه **خالي** است.
My box is **full**.	جعبه من**کامل** است.
The shirt has a **bright** color.	پیراهن رنگ **روشن**.
It tasted **bitter**.	طعم **تلخ** است.
I have a **gray** shirt.	من یک پیراهن**خاکستري**.
The socks are still **wet**.	جوراب هنوز**مرطوب**هستند
It is **cold** today.	امروز **سرد** است.
It was **hot** yesterday.	دیروز **گرم** بود.
The bank is **far** from here.	این بانک است**دور** ازاینجا.
The land is so **dry**.	زمین **خشک** است.
We had a **salty** soup for dinner.	ما یک سوپ **شور** برای صرف شامبود.

As it can be noticed from the above table, all of the adjectives (all of the 35 adjectives which equals with %100) have been translated almost accurately. The following chart is representative of what has happened.

Figure 1. Sentences containing simple adjectives

Table 2. Sentences containing comparative adjectives

He is **older**than me.	او از من**بزرگتر** است.
He is **younger** than me.	او **جوان تر** از من است.

[1] English Language Department, Vali-e-Asr University of Rafsanjan, Iran   Email: a.oraki@vru.ac.ir

English	Persian
He is **taller** than his brother.	او **بلندتر** از برادرش است.
This lake is **deeper** than that one.	این دریاچه **عمیق تر** از آن یکی است.
Her mother is **happier** than her father.	مادر او ، **شادتر** از پدرش است.
This road is 2 kilometers **longer** than the previous one.	این جاده ۲ کیلومتر **طولانی تر** از قبلی است.
His father is **richer** than my father.	پدر او **غنی تر** از پدر من است.
This house is **cleaner** than that one.	این خانه **پاک تر** از آن یکی است.
My blanket is **softer** than that of yours.	پتو من **نرمتر** از آن است که از مال شما است.
Ali's radio is **louder**.	رادیو علی **بلندتر** است.
This room is **darker** than that one.	این اتاق **تیره تر** از آن یکی است.
That box is **lighter** than this one.	که جعبه **سبک تر** از این یکی است.
I came **earlier** than you this morning.	من **زودتر** از شما آمد این صبح است.
He usually comes much **later** than me.	او معمولا **بعد از** من می آید.
I bought a **cheaper** book yesterday.	من خریدم یک کتاب **ارزان تر** دیروز.
I have got **more** money than my sister.	من پول **بیشتری** از خواهر من.
This apple is **bigger** than that pear.	این سیب **بزرگتر** از آن گلابی است.
Your handwriting is **worse** than your brother.	دست خط شما **بدتر** از برادر خود است.
I have found a **harder** object.	من یک جسم **سخت تر** را پیدا کرده اند.
His income is **smaller** this year.	در این سال درآمد او **کوچکتر** است.
Ahwaz is **hotter** than Chabahar.	اهواز **گرمتر** از چابهار.
Shahrekord is **colder** than Tabriz.	شهرکرد **سردتر** نسبت تبریز است.
We had a **greater** success this time.	ما این زمان را یک موفقیت **بیشتری** داشتند.
My father has rented a **faster** car.	پدر من تا به یک ماشین **سریع تر** اجاره.
I prefer **shorter** hair.	من ترجیح می دهم مو **کوتاه تر** است.
This road is **wider** than that road.	این جاده از آن جاده ها **گسترده تر** است.
Dancers are mostly **slimmer** than others.	رقصنده ها اغلب **باریک تر** از دیگران است.
Your cake is **sweeter** than your ice-cream.	کیک شما **شیرین تر** از بستنی خود را است.
I have learned a **better** method for teaching languages.	من یک روش **بهتر** برای آموزش زبان آموخته اند.
He is **quicker** than me in playing.	او در بازی **سریع تر** از من است.
Jack is **smarter** than me in the class.	جک **دقیق تر** از من است در کلاس.
English is **easier** than math.	انگلیسی **راحت تر** از ریاضی است.
He wants to buy a **slower** car.	او می خواهد برای خرید یک ماشین **کندتر** است.
They bought things at **lower** prices.	آنها همه چیز را با قیمت **پایین تر** خریداری شده است.
Their wages are **higher** this week.	حقوق و دستمزد آنها **بالاتر** این هفته است.

As it can be seen from the above table, Google Translate has translated comparative adjectives in the same way as it did in the case of simple adjectives. For the second time, all the 35 adjectives (%100 of the adjectives) have been translated in the most accurate way for a machine. The following chart is a clear representation of the above-mentioned data.

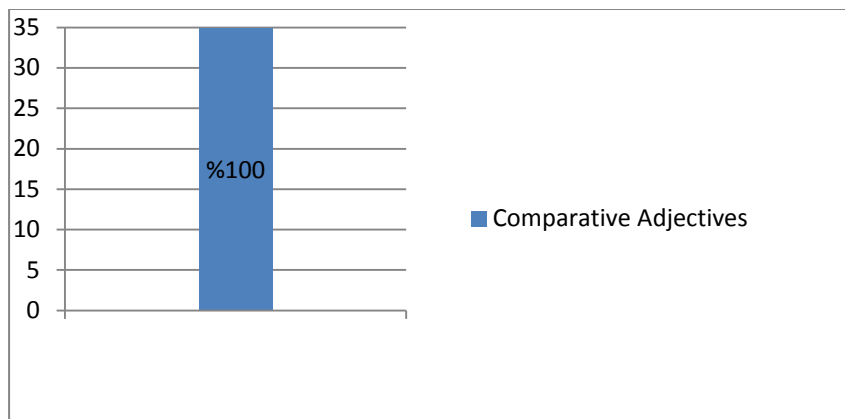Figure 2. Sentence containing comparative adjectives

Table3. Sentences containing superlative adjectives

He is **the best** student in the English class.	او **بهترین** دانش آموز در کلاس زبان انگلیسی است.
She is **the smartest** woman in the world.	او **باهوش ترین** زن در جهان است.
They are **the most expensive** cars in Europe.	آنها اتومبیل های **گران قیمت ترین** در اروپا هستند.
Ahwaz is **the warmest** city in Iran in the winter.	اهواز **گرمترین** شهر در ایران در فصل زمستان است.
Shahrekord is **the coldest** city in Iran in the winter.	شهرکرد **سردترین** شهر در ایران در فصل زمستان است.
He has **the cheapest** computer.	او دارای **ارزان ترین** کامپیوتر است.
This is **the heaviest** box in the shop.	این **سنگین ترین** جعبه در مغازه ها است.
Those are **the lightest** boxes in the shop.	کسانی که **سبک ترین** جعبه در مغازه.
They work **the least**.	آنها **حداقل** کار می کنند.
We work **the most**.	ما **بیشتر** کار می کنند.
Chabahar has **the most wonderful** beaches in Iran.	چابهار دارای سواحل **فوق العاده ترین** در ایران است.
He is **the kindest** teacher among others.	او **مهربانترین** معلم در میان دیگران است.
She is **the ugliest** of them.	او **زشت ترین** آنها است.
He is **the most handsome** teacher in the institute.	او معلم **خوش تیپ ترین** در موسسه است.
This was **the happiest** time ever.	این **شادترین** زمان تا کنون بود.
This was **the hardest** job I have ever done.	این **سخت ترین** کار من تا کنون انجام داده بود
She has **the softest** skin.	او **نرمترین** پوست.
Studying is **the easiest** way to learn.	بررسی **ساده ترین** راه برای یادگیری است.
This must be **the most tasteless** cake.	این باید **بی مزه ترین** کیک باشد.
They are aware of **the most recent** methods.	آنها از روش **اخیر** آگاه هستند.
He is **the closest** friend of mine.	او **نزدیک ترین** دوست من است.
These plants can grow in **the most frigid** environments.	این گیاهان می تواند در محیط های **بسیار سرد ترین** رشد می کنند.
This chef makes **the most delicious** food in the	این سرآشپز می سازد غذای **خوشمزه ترین** در منطقه است.

[1] English Language Department, Vali-e-Asr University of Rafsanjan, Iran   Email: a.oraki@vru.ac.ir

region.	
I saw **the most disgusting** thing in that restaurant last night.	من **نفرت انگیز ترین** چیزی که در این رستوران در شب گذشته دیدم.
He took **the most tortuous** route through back streets.	او **پر پیچ و خم ترین** مسیر را از طریق خیابان های پشت گرفت.
This is **the most sinuous** path.	این **گمراه کننده ترین** راه است.
He has **the most money** among us.	او **بیشتر** پول در میان ما.
He is **the most splendid** painter.	او **پر زرق و برق ترین** نقاش است.
This is **the most important** matter now.	این در حال حاضر **مهم ترین** است.
Jack is **the most talkative** person here.	جک **پر حرف ترین** فرد در اینجا است.
These are **the most aromatic** herbs in the region.	این گیاهان **معطر** در منطقه هستند.
Izeh is **the most mountainous** region in Khouzestan.	ایذه **کوهستانی ترین** منطقه در خوزستان است.
I have taken up **the most difficult** task.	من گرفته تا **سخت ترین** کار است.
He was **the most vociferous** opponent of Conservatism.	او حریف **پر سر و صدا ترین** محافظه کاری بود.
This is **the funniest** joke I have ever heard.	این **جالب ترین** شوخی که تا کنون شنیده است.

For the third time, a great number of the adjectives have been translated accurately according to the above table. 33 out of 35 adjectives (%94) have been translated accurately. The following Figure represents the above-mentioned data.

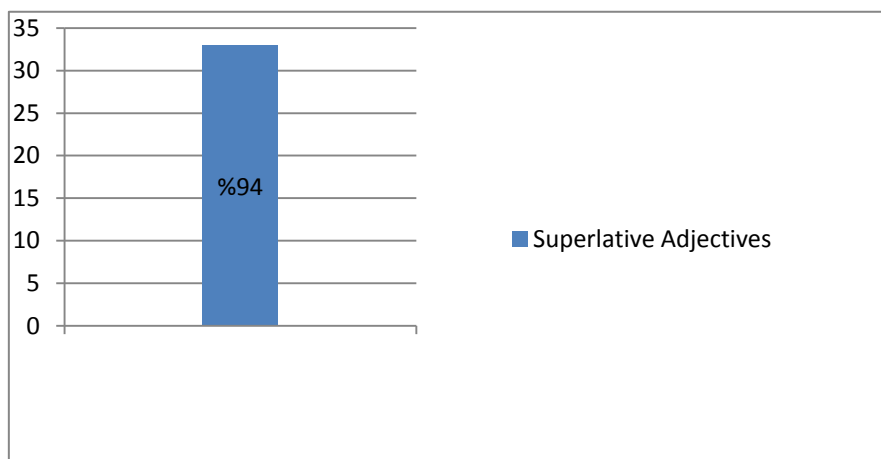Figure 3. Sentences containing superlative adjectives

Table 4. Sentences containing adjectives of similarity

|	
He is not **as smart as** his brother.	او **به عنوان هوشمند به عنوان** برادر خود نیست.
She is not **as talkative as** she used to be.	او **به عنوان پر حرف** او **به عنوان** استفاده می شود نیست.
Nick is **as brave as** his father.	نیک **وشجاع به عنوان** پدرخوداست.
Jack is not **as stupid as** he looks	جک **به عنوان احمق** که او به نظر می رسد نیست
This matter is **as important as** that one.	این **مهم است که یکی از مهم** است..

English	Persian
This box is not **as heavy as** that one.	این جعبه **به عنوان سنگین به عنوان** که یکی نیست.
Chabahar is not **as humid as** Bandar Abbas.	چابهار **به عنوان مرطوب** و بندر عباس نیست..
He is not **as lazy as** his partner.	او **به عنوان تنبل عنوان** شریک زندگی خود را نیست.
This part of the city is not **as salubrious as** the previous one.	این بخشی ازشهر **به عنوان گوارا به عنوان** یکی از قبلی نیست.
Ahmad is **as strong as** his older brother.	احمد است **به عنوان قوی به عنوان** برادر بزرگتر خود است.
She is **as pretty as** her sister.	او را **به عنوان زیبا به عنوان** خواهر او است.
He is not **as ugly as** you.	او **به عنوان زشت به عنوان** شما نیست.
This job is **as difficult as** that one.	این کار **به عنوان مشکل** که یکی است.
Russia is **as frigid as** Alaska.	روسیه **به عنوان بسیار سرد مانند** آلاسکا است.
This cake is **as sweet as** that one.	این کیک **شیرین است** که **به عنوان یکی.**
His car is **as fast as** that of mine.	ماشین خود را **به عنوان سریع به عنوان** من است.
He is **as slow as** he used to be.	او که **آهسته به عنوان** او استفاده می شود است.
She's twice **as old as** her sister.	او دو بار **به عنوان قدیمی به عنوان** خواهر او.
I'm almost **as good** in math **as** in science.	من تقریبا **همانطور** که در ریاضی **خوب در علم** است..
This book is not **as exciting as** the last one.	این کتاب **به عنوان هیجان انگیز به عنوان** یکی از آخرین نیست.
The cafeteria is not **as crowded as** usual.	کافه تریا **به طور معمول شلوغ** نیست.
Ramona is **as happy as** Raphael.	رامونا **به عنوان رافائل خوشحال** است.
Einstein is **as famous as** Darwin.	انیشتین **به عنوان معروف به عنوان** داروین است.
A tiger is **as dangerous as** a lion.	یک ببر **به اندازه** شیر **خطرناک** است.
Jenny's new flat isn't **as nice as** her old one.	صاف جدید جنی **به عنوان خوب به عنوان** یکی از قدیمی او نیست.
It hasn't got **as big** a garden **as** the old one.	این **به عنوانبزرگ**باغ**بهعنوان**یکی از قدیمیندارم.
It's **as good as** you can get for the price.	آن را **به عنوان خوب به عنوان** شما می توانید برای قیمت دریافت کنید
England isn't nearly **as big as** Russia.	انگلستان تقریبابه عنوانبزرگ **به عنوان**روسیهنیست.
The apple is **as light as** the orange.	سیب **به عنوان نوربهعنوان**رنگ نارنجیاست.
My son is **as fat as** my wife.	پسر مناست که **به عنوانچربیبه عنوان** همسرمناست.
The snake is **as long as** the rope.	مار **به عنوان طولانی به عنوان** طناب است.
My son is **as intelligent as** my daughter.	پسر من **به عنوان هوشمند به عنوان**دخترمناست.
The pen is **as small as** the ruler.	قلم **به عنوان کوچک به عنوان**حاکماست.
Oliver is not **as optimistic as** Peter.	الیور **به عنوان خوشبینانه به عنوان** پیترنیست.

[1] English Language Department, Vali-e-Asr University of Rafsanjan, Iran   Email: a.oraki@vru.ac.ir

The tomato soup was **as delicious as** the mushroom soup.	سوپ گوجه فرنگی **به عنوان خوشمزه عنوان**سوپقارچبود.

As the above table shows, only 2 out of 35 adjectives have been translated accurately, which makes up only % 5.7. The following figure represents what has happened in the case of adjectives of similarity.

Figure 4. Sentences containing adjectives of similarity
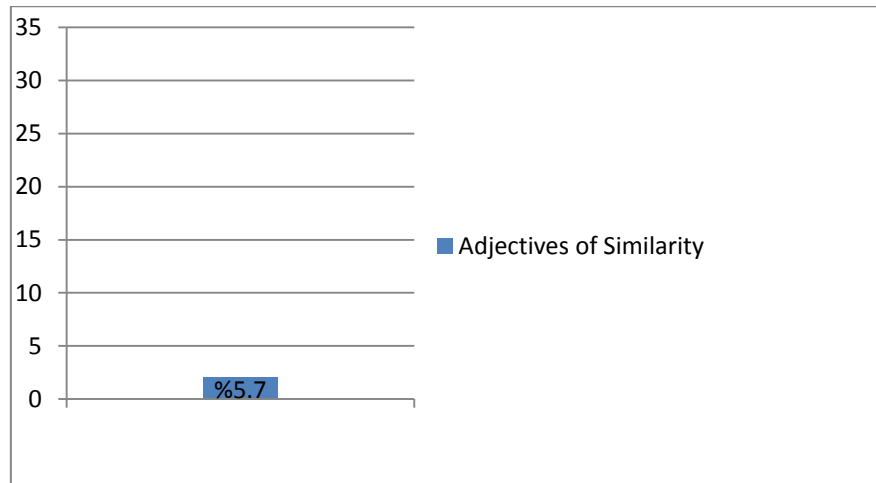
4. Discussion of the Results

The information inserted under the analysis part of each different kind of adjective, which is the representation of the data gathered in this study showed that simple adjectives were translated accurately by %100, i.e. all the 35 adjectives were translated in an accurate way. The same happened to the comparative adjectives (35 out of 35= %100). In the case of the superlative adjectives, 33 adjectives out of the 35 adjectives were translated accurately, which makes up %94. However, in the case of the adjectives of similarity, little success was achieved. In other words, out of the 35 adjectives, only 2 (%5.7) were translated almost accurately.

5. Conclusions

The present study was an attempt to find out the translation quality of Google Translate, as a widely used device, with a concentration on the different types of adjectives. In this study, all the four types of adjectives (simple adjectives, comparative adjectives, superlative adjectives, and adjectives of similarity) were translated separately by Google Translate. The analysis of the data showed that Google Translate is very successful in translating simple adjectives (%100), comparative adjectives (%100), and superlative adjectives (%94), but fails to translate adjectives of similarity as accurately as possible. Only %5.7 of the adjectives was

translated (almost) accurately. This analysis showed that the users can rely on this device for translating simple, comparative, as well as superlative adjectives, but in case of adjectives of similarity a kind of human aided machine translation (machine translation with the help of human translator) must be applied if they are to get good results.

Refrences:

Amman, M (1990) 'Anmerkungen zu einer Theorie der Übersetzungskritik und

ihrer praktischen Anwendung', TEXTconTEXT6: 55–62.

Arnold, D. and Balkan, L (2001) *Machine Translation,* USA: Blackwell Publishers.

Carry, E. and R. W. Jumpelt (1963) "Quality in Translation"*, Proceedings of the 3

Congress of the International Federation of Translators (Bad Godesberg, 1959),* New

York: Macmillan/Pergamon Press.

Gerzymisch-Abrogast, H. (1997). Wissenschaftliche Grundlagen für die Evaluierung von Übersetzungsleistungen. In E. Fleischmann (Ed.), Translationsdidaktik: Grundfragen der Übersetzungswissenschaft (pp. 573—579). Tübingen: Narr.

Hatim, B. and Munday. J (2004) *Translation*, USA: Routledge.

House, J (1997) *Translation Quality Assessment,* Germany: Gunter Narr Verlag.

Khoshsima, H. and Rostami A (2010) "ON a New Text Based Approach to Marka

Translation" *International Maritime English Conference IMEC 21*S Szczecin: Poland.

Linares, J (2008) *Empirical Machine Translation and its Evaluation,* Spain: Universitat

Polit`ecnica de Catalunya

Newmark, P (1988) *A Textbook of Translation,* New York & London: Prentice-Hall.

Reiss, K (19971/1978 translated by Errol F. Rhodes) *Translation Criticism – The

potentials and Limitations: Categories and Criteria for Translation Quality

Assessment*, 2000, Manchester: St. Jerome.

Swan, M (1995) *Practical English Usage,* England: Oxford University Press.

Tauschel, A (2008) *Linguistic Aspects in Machine Translation,* Germany: University of

Frankfurt.

[1] English Language Department, Vali-e-Asr University of Rafsanjan, Iran   Email: a.oraki@vru.ac.ir

Wilss, W (1997) *The science of Translation,* Germany: Gunter Narr Verlag.