

**T.C.
İSTANBUL 29 MAYIS ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
MÜTERCİM TERCÜMANLIK ANABİLİM DALI**

**MACHINE TRANSLATION POST-EDITING IN VIDEO
GAME LOCALIZATION**

**BİLGİSAYAR OYUNLARININ
YERELLEŞTİRİLMESİNDE MAKİNE ÇEVİRİSİ
SONRASI DÜZELTME**

YÜKSEK LİSANS TEZİ

Hulusi Alp SİRKINTI

**Danışman:
Doç. Dr. Senem ÖNER BULUT**

**İSTANBUL
2022**

**T.C.
İSTANBUL 29 MAYIS ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ
MÜTERCİM TERCÜMANLIK ANABİLİM DALI**

**MACHINE TRANSLATION POST-EDITING IN
VIDEO GAME LOCALIZATION**

**BİLGİSAYAR OYUNLARININ
YERELLEŞTİRİLMESİNDE MAKİNE ÇEVİRİSİ
SONRASI DÜZELTME**

YÜKSEK LİSANS TEZİ

Hulusi Alp SİRKINTI

**Danışman:
Doç. Dr. Senem ÖNER BULUT**

**İSTANBUL
2022**

T. C.
İSTANBUL 29 MAYIS ÜNİVERSİTESİ
SOSYAL BİLİMLER ENSTİTÜSÜ MÜDÜRLÜĞÜNE

Mütercim Tercümanlık Anabilim Dalı, Çeviribilim Bilim Dalı'nda 200520YL06 numaralı Hulusi Alp SİRKINTI'nın hazırladığı "Machine Translation Post-Editing in Video Game Localization" konulu yüksek lisans tezi ile ilgili tez savunma sınavı, 21/06/2022 günü (12:⁰⁰-13:⁰⁰) saatleri arasında yapılmış, sorulan sorulara alınan cevaplar sonunda adayın tezinin başarılı olduğuna oy birliği ile karar verilmiştir.

Doç. Dr. Senem ÖNER BULUT
Yıldız Teknik Üniversitesi
(Tez Danışmanı ve Sınav Komisyonu Başkanı)

Dr. Öğr. Üyesi Nilüfer ALİMEN
İstanbul 29 Mayıs Üniversitesi

Dr. Öğr. Üyesi Sinan Okan ÇAVUŞ
İstanbul 29 Mayıs Üniversitesi

BEYAN

Bu tezin yazılmasında bilimsel ahlak kurallarına uyulduğunu, başkalarının eserlerinden yararlanılması durumunda bilimsel normlara uygun olarak atıfta bulunulduğunu, kullanılan verilerde herhangi bir tahrifat yapılmadığını, tezin herhangi bir kısmının bu üniversite veya başka bir üniversitedeki başka bir tez çalışması olarak sunulmadığını beyan ederim.

Hulusi Alp SİRKINTI

21.06.2022

ÖZ

Bu çalışmanın amacı, bilgisayar oyunlarının yerelleştirilmesinde kitle kaynaklı seçenekleri tercih etmek yerine, ücretsiz makine çeviri sistemlerinin kullanılıp kullanılamayacağını sorgulamaktır. Bahsedilen amaç doğrultusunda *Kenshi* adlı bilgisayar oyununun dil dosyası içerisinde oyun içi ırklar ve bunların alt ırklarıyla bağlantılı olan 30 satır seçilmiş, bu satırlar ücretsiz olarak kullanılabilen ve nöral makine çevirisi sistemi olan Google Translate ve Amazon Translate kullanılarak çevrilmiş, elde edilen ham makine çıktısı ise TAUS “adequacy” ve “fluency” hata türleri kriterleri doğrultusunda üç gönüllü tarafından değerlendirilmiş ve sonrasında da TAUS “Human Translation Quality” kriterlerine uygun olarak ham makine çevirisi çıktısı üzerinde makine çevirisi sonrası düzeltme işlemi uygulanmıştır. Elde edilen veriler ışığında, bahsi geçen makine çevirisi sistemlerinin ham çıktılarının düzeltme olmadan kullanılamayacağı tespit edilmiştir. Google Translate ham çıktıların ancak %32’sinin, Amazon Translate ham çıktıların ise %23’ünün düzeltme gerektirmeden kullanılabilir olduğu belirlenmiştir. Buna rağmen makine çevirisi sonrası düzeltme yoluyla, çevirmenlerin saat başı kelime performanslarının en az iki kat arttığı gözlemlenmiştir. Nitekim, makine çevirisi sistemlerinin çeviri yükünü büyük ölçüde azalttığı sonucuna varılmıştır. Ayrıca, bilgisayar oyunu yerelleştirme yeterliliklerinin, makine çevirisi sonrası düzeltme işlemini gerçekleştiren kişilerin düzeltme eforları (edit effort) üzerinde belirleyici bir etkisi olduğu ve özellikle “yaratıcılık” (creativity) ve “bilgisayar oyunları ile aşinalık” (familiarity with games culture) yeterliliklerinin, düzeltme eforlarında büyük farklılıklara sebep olduğu saptanmıştır.

Anahtar Kelimeler: Kitle kaynaklı yerelleştirme, bilgisayar oyunu yerelleştirmesi, makine çevirisi, TAUS, makine çevirisi sonrası düzeltme, hata analizi, nöral makine çevirisi

ABSTRACT

The aim of this study is to explore whether free to use machine translation systems can be used to localize video games, instead of opting for crowdsourcing options. In line with this aim, 30 lines related to in-game races and their sub-races were selected from the language file of the video game called *Kenshi*. These lines were translated using Google Translate and Amazon Translate, which are free-to-use neural machine translation systems, and the raw machine translation outputs obtained were evaluated by three volunteers in accordance with TAUS "adequacy" and "fluency" error annotation criteria. Then, in accordance with TAUS "Human Translation Quality" guidelines, the volunteers post-edited the raw machine translation outputs. In the light of the obtained data, it has been determined that the raw outputs of the aforementioned machine translation systems cannot be used without post-editing. It has been determined that only 32% of Google Translate's raw machine translation output and 23% of Amazon Translate's raw machine translation output can be used without post-editing. However, by making use of machine translation post-editing, it was observed that the translators' word per hour performance doubled at minimum. Thus, it has been concluded that machine translation systems substantially decrease the translation workload. In addition, it has been determined that video game localization competences have a decisive effect on the editing efforts of the people who conduct machine translation post-editing, and especially "creativity" and "familiarity with games culture" competences cause great differences in edit efforts.

Keywords: Crowdsourced localization, video games localization, machine translation, TAUS, machine translation post-editing, error annotation, neural machine translation

ACKNOWLEDGEMENTS

I would like to first express my gratitude to my advisor Assoc. Prof. Dr. Senem ÖNER BULUT, whom I can not thank enough for her guidance, invaluable support and endless patience. Writing this thesis under her mentorship was a great opportunity for me.

I also owe many thanks to Prof. Dr. Işın ÖNER for all the knowledge she has passed onto us over the years and for her feedback and suggestions for the duration of this study and to Dr. Nilüfer ALIMEN for her generous help.

Last but not least, I would like to thank my dear wife Halise GÜMÜŞ SİRKINTI for her never-ending support.

To my daughter...

TABLE OF CONTENTS

TEZ ONAY SAYFASI.....	ii
BEYAN.....	iii
ÖZ	iv
ABSTRACT.....	v
ACKNOWLEDGEMENTS.....	vi
LIST OF TABLES	ix
LIST OF FIGURES	x
I. INTRODUCTION.....	1
CHAPTER ONE	
THEORETICAL BACKGROUND	7
2.1. Video Games Localization.....	7
2.1.1. History of Video Games Localization.....	8
2.1.1.1. Early Phase: Before the Mid-1980s	8
2.1.1.2. Growth Phase: The Mid-1980s to Mid-1990s.....	8
2.1.1.3. Development Phase: The Mid to Late 1990s	9
2.1.1.4. Maturing Phase: Early 2000 to 2005.....	10
2.1.1.5. Advancing Phase: 2005 to the Present	10
2.1.2. Competences for Video Games Localizers	11
2.2. Crowdsourcing	13
2.2.1. Crowdsourcing in the Field of Translation.....	15
2.2.2. Using Crowdsourcing in the Field of Translation: Benefits and Challenges	15
2.3. Machine Translation.....	18
2.3.1. History of Machine Translation	18
2.3.1.1. Precursors and Pioneers (1933-1956)	18
2.3.1.2. The Decade of High Expectations and Disillusions (1956-1966)	19
2.3.1.3. The ALPAC Report and its Consequences	19
2.3.1.4. The Quiet Decade (1966-1976).....	19
2.3.1.5. Operational and Commercial Systems in the 1980s.....	20
2.3.1.6. The Early 1990s	20

2.3.1.7. The Late 1990s.....	20
2.3.1.8. Since 2000.....	20
2.3.2. Machine Translation Post-Editing.....	21
2.3.2.1. “Good Enough” Quality.....	22
2.3.2.2. “Human Translation Quality”	22
2.3.3. “TAUS DQF-MQM” Machine Translation Error Annotation	24
2.3.4. Quality Evaluation Using Adequacy and Fluency Approaches	26
CHAPTER TWO	
METHODOLOGY	28
IV. RESULTS.....	30
4.1. Adequacy Report for Google Translate	30
4.2. Fluency report for Google Translate.....	32
4.3. Productivity report for Google Translate	34
4.4. Adequacy Report for Amazon Translate.....	39
4.5. Frequency report for Amazon Translate.....	41
4.6. Productivity report for Amazon Translate.....	44
4.7. “Edit Effort” Differences.....	48
4.7.1. Segment 2 Differences	48
4.7.2. Segment 3 differences:	49
4.7.3. Segment 6 differences:	50
4.7.4. Segment 15 differences:	51
V. CONCLUSION	53
APPENDICES	59
CURRICULUM VITAE.....	99

LIST OF TABLES

Table 1: Table showing motivation tiers for crowdsourcing projects	17
Table 2: Table showing ratings for adequacy evaluation for machine translation output	27
Table 3: Table showing ratings or fluency evaluation for machine translation output.	27
Table 4: Table showing “Edit Efforts” per editor for Google Translate.....	38
Table 5: Table showing “Edit Efforts” per editor for Google Translate.....	47
Table 6: Segment 2 Post-Editing Differences And Edit Efforts For Each Editor	49
Table 7: Segment 3 Post-Editing Differences And Edit Efforts For Each Editor	50
Table 8: Segment 6 Post-Editing Differences And Edit Efforts For Each Editor	50
Table 9: Segment 15 Post-Editing Differences And Edit Efforts For Each Editor	51

LIST OF FIGURES

Figure 1: Screenshot Showing Interaction Menu	2
Figure 2: Steam Community Page for <i>Kenshi</i>	4
Figure 3: Turkish Localization Project Discusssion	5
Figure 4: Early Cinematic Examples	9
Figure 5: Average Adequacy Result of Google Translate	30
Figure 6: Distribution of Adequacy Scores, Per Language for Google Translate	31
Figure 7: Distribution of Adequacy Scores, Per Evaluator for Google Translate	32
Figure 8: Average Fluency Result of Google Translate	32
Figure 9: Distribution of Fluency Scores, per Language for Google Translate	33
Figure 10: Distribution of Fluency Scores, per Evaluator for Google Translate	34
Figure 11: Average Productivity Report for Google Translate	35
Figure 12: Productivity per Editor for Google Translate	36
Figure 13: Edit Distance Graph for Google Translate	37
Figure 14: Average Adequacy Result for Amazon Translate	39
Figure 15: Distribution of Adequacy Scores, per Language for Amazon Translate	40
Figure 16: Distribution of Adequacy Scores, per Evaluator for Amazon Translate	41
Figure 17: Average Fluency Result of Amazon Translate	41
Figure 18: Distribution of Fluency Scores, per Language for Amazon Translate	42
Figure 19: Distribution of Fluency Scores, per Evaluator for Amazon Translate	43
Figure 20: Average Productivity Report for Amazon Translate	44
Figure 21: Productivity per Editor for Amazon Translate	45
Figure 22: Edit Distance Graph for Amazon Translate	46

I. INTRODUCTION

Even though the video games industry is relatively new when compared to other fields in the entertainment industry (such as film, television, radio, print, music and etc.), it has far surpassed the film and music industries in terms of global revenue generated in 2019. While the video games industry generated \$145.7b in global revenues, the film industry generated about \$42.5b and the music industry generated \$20.2b (Richter 2020). According to estimations, the video games industry will grow even more and reach a total of \$218.7b by the end of 2024 (Wijman 2020). Unsurprisingly, an industry of this size has an enormous need of localization services in order that the developed games can be sold in different regions of the world.

In order to cope with this amount of localization need, the developers have come up with different localization strategies. Video game developers either choose the outsourcing model, in which the localization process is commissioned to a specialized vendor and is the more preferred model, or they choose the in-house model, in which the developer itself is responsible for localizing the games into different languages (O'Hagan and Mangiron 2013, 118-121). Apart from these two strategies, the developers also choose “different levels of localization” for different languages, mostly according to the size of the video games market in a certain region, and these localization levels are no localization, box and docs localization (the localization of only the packaging and manual of the game), partial localization (in-game text is localized but the voice files are not) and finally, full localization (translating all assets of a game). For example, the developer of the video game *Kenshi* followed the outsourcing model for the languages the game would sell the most, and as the game doesn't include any voiceover files, it can be said that they went for partial localization. However, for the languages in which the game wouldn't sell enough to cover for the localization cost, they chose not to localize their game and to opt for a crowdsourcing model, making the language files of the game easy to access and edit so that their player base can easily

localize the game in their own languages. Yet, it can be said that choosing a crowdsourcing model didn't work out well for the non-localized languages as the crowd hasn't localized the game into any languages on their own.

Thus, the present thesis focuses on the question whether machine translation post-editing can be used as an alternative for crowdsourced video game localization projects. In order to examine this topic, the crowdsourcing case of the video game *Kenshi* is analyzed and whether or not machine translation and machine translation post-editing could be used in the localization of this video game is explored.

The developer of the video game, Lo-Fi Games, was a small studio comprising of only a few people when the game was first being developed. They had a low budget, thus, they made all communication in the game text-based. This means that, when a player is interacting with in-game non-player characters (NPCs), there are no speech audios and the player can see the interaction in some form of speech bubbles and interactive user interfaces (see Figure 1).

Figure 1: Screenshot Showing Interaction Menu



Figure 1 A screenshot from the game showing how the player interacts with NPCs

Even though there aren't any audios to be translated and dubbed in the game and that the game's video file comprises of only text, its total word count is close to 300,000, meaning that the localization workload is unignorable. The game is a role-playing game (RPG), a genre of games in which the players control a character or multiple characters through a story. In RPG games, there generally is an extensive

background lore and almost every in-game item and in-game character is interactable with, explaining why the language file of the game is this long.

Because of a lack of resources (Lo-Fi Games Forum 2020), when the game was first released, the studio had the game localized in six languages, which were Japanese, Spanish, French, German, Russian and Brazilian Portuguese. The reason why the developers chose these languages is that they thought the game would be selling the most in countries where these languages are spoken, and about 52% of the game's sales happened in countries where these languages are spoken (Mikkelson 2018). After about a year, the developers also decided to have the game localized in Korean and Chinese as the player base in these countries grew. Yet, for other languages, the developers decided to follow some type of crowdsourcing.

Even before the developers decided to opt for crowdsourcing, the player base of the game had started to ask for easy to localize or easy to mod systems as early as 2011 so that they could localize the game themselves in the form of fan/volunteer translations. In 2016, the developers started the crowdsourcing project by making the language files of the game easily accessible and translatable. They even started a forum page dedicated to this and extensively explained how the players could start localizing the game by themselves. Together with this development, the player base started translating the files themselves by coming together in forums (see Figure 2).

Figure 2: Steam Community Page for *Kenshi*

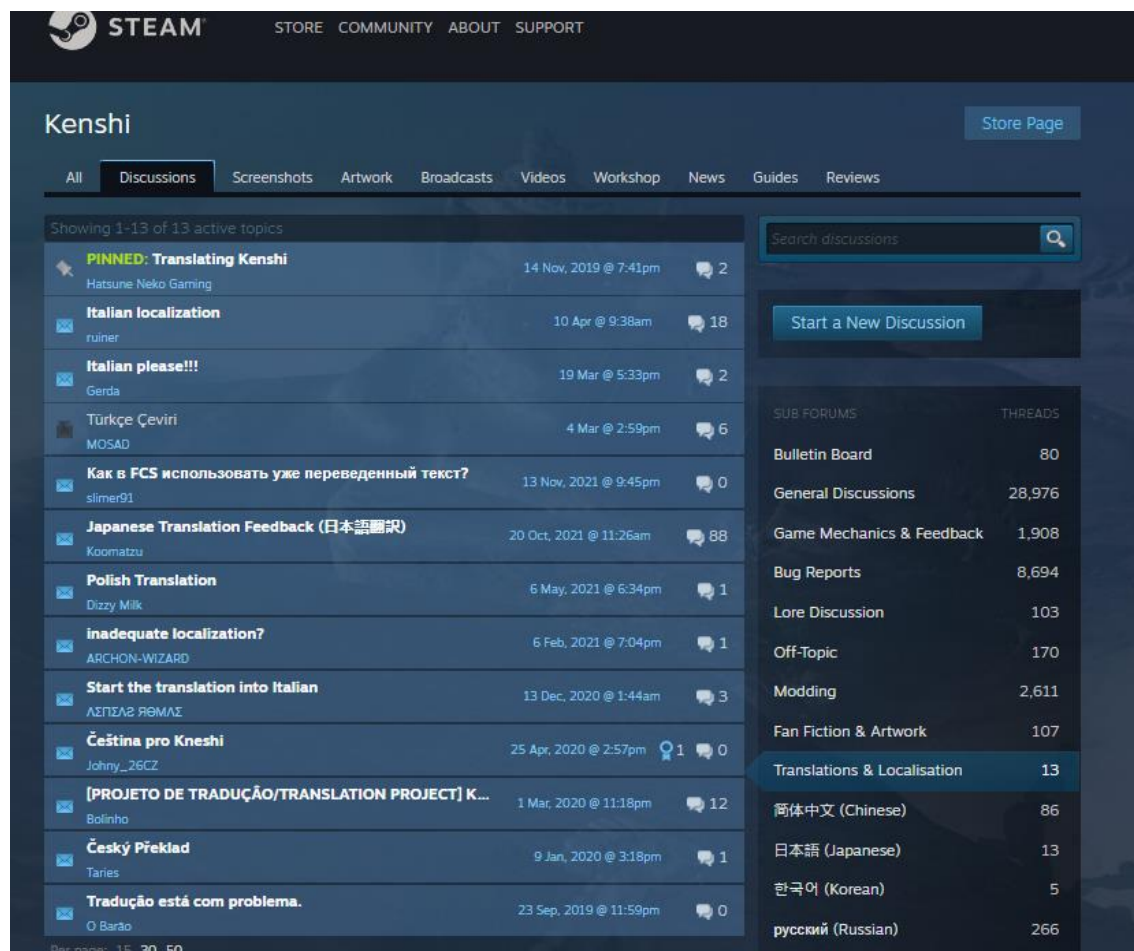


Figure 2 The player base of the video game *Kenshi* coming together in the game's Steam community page to localize the game's language files.

Source: <https://steamcommunity.com/app/233860/discussions/10/>

The player base was at first willing to localize the game on their own with no rewards in return and they started to form groups for the localization in different languages. But, none of the projects that can be seen in Fig. 2 have been completed thus far. The crowd which started working on the Turkish crowdsourcing project altogether left the project because not many people volunteered for it and the length of the game's language files makes it very difficult for only 2 people to translate the whole text (see Figure 3).

Figure 3: Turkish Localization Project Discussion

The screenshot shows a Steam Community discussion thread titled "Türkçe Çeviri" (Turkish Translation) initiated by user MOSAD on November 7, 2019. The thread has 6 posts. The main post by MOSAD explains the project's goal: to create a 100% accurate Turkish translation of the game, starting with the dialogue. It mentions that the team is currently in the process of translating the dialogue and that they are looking for help from the community. The thread includes several replies from other users, such as ManEagle, Caliburn, Prometheus™, deęiřlikerif, and ronin.exe, who are asking for updates and expressing interest in helping. A sidebar on the right lists other discussions related to localization, including "Italian localization", "Italian please!!!", "Как в FCS использовать уже пер...", and "Japanese Translation Feedback (...)".

Türkçe Çeviri

Merhaba arkadaşlar. 2 kişi oyunu çevirmekteyiz. %100 eksiksiz bir çeviri hazırlamaktayız ve çok titiz davrandığımızdan yavaş yavaş işi ilerletiyoruz ve testlerimizden sonra oyun yapımcıları ile iletişime geçip Türkçe dil çevirimizi eklemelerini isteyeceğiz. Büyük oranda ilerlemiş durumdayız ve diyalogları çevirmekteyiz son olarak. Çeviri işinde iyiyim diyen ve yazım kurallarına olabildiğince dikkat eden, çeviri ekibimize katılmak isteyen arkadaşlar bana ulaşınlar.

Showing 1-6 of 6 comments

ManEagle • 3 Dec. 2019 @ 11:17pm
İnşallah en kısa zamanda gelir MOSAD heyecanla bekliyorum :) Keşke yardımcı olabilseydim :(#1

Caliburn • 20 Jan. 2020 @ 5:22pm
Selam MOSAD, cevabını bekliyorum. 🐼
Last edited by Caliburn: 20 Jan. 2020 @ 5:22pm #2

Prometheus™ • 12 Feb. 2020 @ 11:11pm
Selamlar! Kolay gelsin. Herhangi gelişme var mı ?
Last edited by Prometheus™: 12 Feb. 2020 @ 11:11pm #3

deęiřlikerif • 22 Apr. 2020 @ 10:54pm
bir gelişme var mı? #4

ronin.exe • 3 Mar. @ 8:03pm
abi ingilizcem var ama nasıl çevireceğimi bilmiyorum nasıl yardımcı olabilirim #5

MOSAD • 4 Mar. @ 2:59pm
Originally posted by **ronin.exe**:
abi ingilizcem var ama nasıl çevireceğimi bilmiyorum nasıl yardımcı olabilirim
Kimse yardım etmediğinden ilerlemedi çeviri ve üstünden çok geçti. #6

Showing 1-6 of 6 comments

More discussions

- Italian localization 18
- Italian please!!! 2
- Как в FCS использовать уже пер... 0
- Japanese Translation Feedback (...) 88

Figure 3 The forum discussion for the Turkish localization project by the crowd. Source:

<https://steamcommunity.com/app/233860/discussions/10/1637549649096860100/>

Drawing on the above-sketched situation, the present thesis aims to test the following hypothesis: Could machine translation post-editing be used in the localization of the video game *Kenshi*, as an alternative to the crowdsourcing model the developers preferred.

In order to test the above-mentioned hypothesis, the thesis tries to answer a group of questions concerning the concepts of crowdsourcing, machine translation, machine translation post editing and video games localization: Can raw machine translation output be used to localize video games or is post-editing required? As free and neural machine translation systems, how successfully can Google Translate and Amazon Translate translate a video game's language file from English into Turkish? In accordance with TAUS's adequacy and fluency evaluation criteria, how does Google Translate and Amazon Translate fare in the translation of the language files of the video game *Kenshi*? Can machine translation post editing be used as an alternative to crowdsourcing? Can machine translation post-editing ease the workload of localizers in a video game localization project? Are video game localization competences important in post-editing the machine translation output of a video game's language file?

CHAPTER ONE

THEORETICAL BACKGROUND

2.1. Video Games Localization

As a concept, localization itself has been increasingly used in the translation industry since the 1980s as there is an ever-growing need for localization. The main reason for this is given as the incorporation of the software localization sector in the composite industry termed GILT (an abbreviation of Globalization, Internationalization, Localization and Translation), in which translation is posited in the context of globalization (Mangiron and O'Hagan 2006, 10). Within the context of GILT, there is a need for localization as target market versions of computer software are required. During the localization process, strings (lines of texts) extracted from the software are subjected to linguistic transfer and they are then recompiled in the given software environment (Mangiron and O'Hagan 2006, 11).

Basically being computer software, video games are also subjected to localization in order to be sold in different markets around the world and video games localization is not a new concept in the field, but, the localization processes of video games have changed drastically since the release of the first games in the world. When some of the first video games with commercial success like Namco's *Pac-Man* (1980) or Nintendo's *Duck Hunt* (1984) and *Tetris* (1989) are analyzed, it can be clearly seen that these games didn't contain many content to be localized, except for a few words or sentences. However, together with the implementations of new technologies, the localization of modern video games is very different than the video games released in the 1960s, 70s and 80s. Apart from the graphical enhancements and the incorporation of realistic sounds, cinematics and other improvements, most video games also contain a story to be followed, which is relayed to the players through in-game texts, dialogues, above mentioned cinematics and etc. All of these elements need to be localized so that the video games can be released in different markets, meaning that the workload of the

localizers increased greatly. The developments in video games localization can be better understood when the issue is divided into different historical periods.

2.1.1. History of Video Games Localization

In 2009, Ryoichi Hasegawa, who worked as a localizer in Sega and Sony, published a report named ゲームロカライズの歴史とこれから (geemu rokaraisu no rekishi to kore kara - Game Localization's History and Future) and divided the history of video games localization into five historical periods; Early phase, Growth phase, Development phase, Maturing phase and Advancing phase (as cited in O'Hagan and Mangiron 2013, 46).

2.1.1.1. Early Phase: Before the Mid-1980s

According to Ryoichi Hasegawa (2009), the first historical period in video games localization is the "Early" phase, a period before the 1980s when not only video games localization but also video games development was at its beginning (as cited in O'Hagan and Mangiron 2013, 46). During this period, games like *Pong* (1972), *Space Invaders* (1978) and *Asteroids* (1979) were developed. In terms of localization, these games included only a few words and no dialogues, cinematics etc. The names of the games or the game texts like "HIGH SCORE" and "GAME OVER" were already used in English and thus, no localization was needed for the American market, however, even with the in-game texts being English, sometimes changes had to be made for cultural reasons, like it happened with the game *Pac-Man* (Kohler 2005, 24 and 212). The original name of the game was *Puck-Man*, but, fearing that teenagers playing the game would vandalize the game's name in America as it resembles a much used curse word, the name of the game was changed to *Pac-Man* for the American market.

2.1.1.2. Growth Phase: The Mid-1980s to Mid-1990s

The second important phase in the history of video games localization Hasegawa (2009) mentions is the period from mid-1980s to mid-1990s (as cited in O'Hagan and Mangiron 2013, 52). During the mid-1980s, cinematics and written conversations started to be included in video games (see Figure 4). This meant that starting during this period, the translatable content was subjected to localization and the contents' suitability for the overseas markets needed consideration (O'Hagan and Mangiron 2013, 53).

Figure 4: Early Cinematic Examples



Figure 4 On the left, a scene from the video game *The Legend of Zelda* (1986) showing a monologue from the game and on the right, a scene from the video game *Ultima 4* (1985) showing a cinematic at the beginning of the game

2.1.1.3. Development Phase: The Mid to Late 1990s

Being the third historical phase mentioned by Hasegawa (2009) in video games localization, the mid to late 1990s saw the first step towards modern gaming and game development (as cited in O'Hagan and Mangiron, 2013, 55). This phase is very important because it was during this period of time that new gaming systems like the Sony PlayStation or the Sega Dreamcast were first released. Together with these consoles, two important things happened:

- CDs with a capacity of 640mb, which was almost 100 times more than the maximum capacity of ROM cartridges, started to be used. As a result, the graphics, cinematics and sounds used in the games saw huge improvements and the lengths of in-game texts increased substantially.
- European markets started to be served with localized games in their own languages. It is during this time that we can really start seeing localization processes taking place in the industry. However, there are also many examples to poor localizations made during this period.

(O'Hagan and Mangiron 2013, 55).

2.1.1.4. Maturing Phase: Early 2000 to 2005

During the maturing phase as mentioned by Hasegawa (2009), there are again two important developments concerning video games localization:

- In the previous generation consoles, the texts were mostly included in pictures. In this phase however, the texts were encoded in ASCII (American Standard Code for Information Interchange) format. This means that for each character, there was a numerical representation so that computer can understand the characters in the Latin alphabet (American National Standards Institute 1975).
- The second development in this period was the implementation of human voiced dialogues, cinematics and etc. which had to be included in the localization process, bringing video games within the influence of audiovisual translation. (O'Hagan and Mangiron 2013, 58).

2.1.1.5. Advancing Phase: 2005 to the Present

In the final phase, in terms of video games localization, the first thing mentioned is the sim-ship (simultaneous shipment) model in which video games are released on the same day in all the markets (O'Hagan and Mangiron 2013, 60). This means that by the time the game's development is finished, the localizations should all be ready as well in every language the game will be published in so that the game can be released in all the markets on the same day. To ensure that, the localizers have to work on the localization while the game is still being developed. Not having access to the whole game and localizing it part by part is a problem for the localizers. Such a situation is explained by Frank Dietz (2006) as follows:

- On the translator's side this parallel development means working with a text that is, despite all assurances to the contrary, still fluid, and sometimes requires frantic re-writing and re-translating during the last few days before the game's ship date.

(Dietz 2006, 126)

2.1.2. Competences for Video Games Localizers

The competences translators must possess are a new research area in the field of translation and recent studies on the topic revealed that translators need to possess five main competences in total to be able to work efficiently and sufficiently as a translator in the industry. These competences are listed as language and culture competence, translation competence, technology competence, personal and interpersonal competence and service provision competence (European Commission 2017, 4).

Just as there are required competences in the field of translation, translators who want to specialize in the field of video games localization need to have certain competences. These competences are determined by Carmen Mangiro Hevian (2009) as follows:

- Knowledge of general software terminology and specific games platform terminology: This competence is about having general software terminology and games platform terminology knowledge (Mangiro Hevian 2009, 311). A translator who wants to specialize in the field of localization should be able to recognize terminology such as save, delete, cancel, press and other system messages, warning messages, help messages and etc.
- Familiarity with the specific features of screen translation: A translator who wants to specialize on the field of video games localization should be familiar with features of dubbing and subtitling as well (Ibid, 312).
- Mastering of natural and idiomatic language: Being able to use the correct form of language while localizing is important so that the players can enjoy the game as it is meant to be enjoyed (Ibid, 312). For example, in *Warhammer 40K Dawn of War* series, one of the races in the game, the orcs, speak an ungrammatical version of English. Some lines said by these orcs are “Dem Marine boyz iz here” (The Marine boys are here), “Skwad got dere morale back!” (Squad got their morale back!) and “Ev’ryone knowz red wunz go fasta!” (Everyone knows red ones go faster!). In this game, localizing such orc lines grammatically correctly would not be appropriate and the localizers should be aware of this.

- Creativity: Apart from knowing how to use the language, localizers also need to be creative so that the text they produce helps the players get engaged with the story. As in the *Warhammer 40K Dawn of War* series examples above, apart from being aware how to use the language, the localizers also need to be creative while localizing the content. As Carmen Mangiro Hevian (2009) points out, the localizers can completely change lines to include humor even if it that is not the case in the original version of the game or give accents to characters and etc. These are all examples of transcreation, where the translators introduce new elements that are not available in the original.
- Cultural awareness: During the localization process, the localizers are not only responsible for translating words from one language into another, they must also be able to recognize culturally sensitive issues and make the necessary changes (Ibid, 313).
- Familiarity with games culture: Even though most video games have different playstyles, stories, interfaces and etc., there are some items that are used commonly in every video game. Terms like HP, MP, Level, Attribute, Skills can be given as an example to common items in RPG games. Due to the nature of the genre, these terms are included in almost every RPG released. Thus, during the localization process, localizers must pay attention to how such terms/items are used in other games and conduct the localization process accordingly (Ibid, 315). The importance of this competence is so great that Minako O'Hagan (2009) claims:
 - ...domain-knowledge possessed by fans may in some cases compensate for the lack of formal translator training (O'Hagan 2009, 13).
- Familiarity with global pop culture: The last competence video game localizers need is mentioned as familiarity with global pop culture by Carmen Mangiro Hevian (2009). This competence is about having knowledge on global pop culture as in video games, there sometimes are intertextual elements borrowed from other games, books, comics, movies and etc. and localizers should have knowledge about these.

In the context of the present thesis, some of these competences stand out. The video game *Kenshi* is an open-world, sandbox RPG with a deep lore in the background. The game itself doesn't contain any cinematics or voiced dialogues. The only way the player interacts with non-player characters (NPC) is through dialogue boxes in which the players can read the dialogues and choose the options they want. Volunteers who want to localize the game can localize the lines containing such dialogues, the names of in-game items, their descriptions, books and journals in the game that contain information about the lore and etc. The game itself handles how these are displayed. This means that, among the above-mentioned competences, “knowledge of general software terminology and specific games platform terminology” and “familiarity with the specific features of screen translation” are not very relevant to the localization process in question. However, all the other competences are applicable for the volunteers wanting to take part in the localization process, which means that even though the volunteers participating in the crowdsourcing localization projects would highly probably have “familiarity with games culture”, “familiarity with pop culture” and “cultural awareness” as competences, which would compensate for their amateurism greatly, whether they have “mastering of natural and idiomatic language” and “creativity” as competences could be questioned. It can be speculated that as a result of the crowd lacking some competences, the quality of the localization may not be as high as if it were to be localized by translators specializing on video games localizations.

On the other hand, if machine translation post-editing was to be the preferred localization method, unless a customized translation machine is used, the machine translation output would be of a poor quality as the listed competences couldn't be applied for machine translation. The competences of the post-editors can be questioned, but, the low quality of the machine translation output would also mean that in order to achieve human translation quality, the post-editors may need to edit a great deal of the machine translation output.

2.2. Crowdsourcing

The term “crowdsourcing” was first used in an article in the *Wired Magazine* by Jeff Howe (2006). In the article, Jeff Howe mentions that in the past employers turned to China or India for cheap labor. However, with the advancements in communication and

information technologies, it doesn't matter where the laborers are. They can be one block away from you or on the other side of the world. It doesn't matter. What matters is that the workers have access to the internet. Thanks to the internet, the employers now have the chance to employ anyone, anywhere around the world and make use of their experience and information they have. What's more, these advancements also enabled hobbyists, part-timers and dabblers to have a market for their efforts. Howe (2006) gives Amazon's *Mechanical Turk* as an example to this:

Amazon Mechanical Turk is a Web-based marketplace that helps companies find people to perform tasks computers are generally lousy at—identifying items in a photograph, skimming real estate documents to find identifying information, writing short product descriptions, transcribing podcasts.

The tasks given to people on *Amazon Mechanical Turk* are very simple tasks required the use of human intelligence and the participants are paid only a few cents or a few dollars per task. At the end of Howe's article, it can be clearly seen that even back in 2006, people were already taking advantage of crowdsourcing to have tasks they normally paid for thousands of dollars completed for a few dollars only.

At a later date, in one of his blogs, Howe explained the term "crowdsourcing" as follows:

Crowdsourcing is the act of taking a job traditionally performed by a designated agent (usually an employee) and outsourcing it to an undefined, generally large group of people in the form of an open call (Howe n.d).

Apart from Howe's definition of crowdsourcing, there are many other definitions used for the term. In order to come up with a common definition to be used, Estellés and Gonzáles (2012) proposed the following definition by integrating different definitions from different researchers:

Crowdsourcing is a type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task, of variable complexity and modularity, and in which the crowd should participate bringing their work, money, knowledge and/or experience, always entails mutual benefit. The user will receive the satisfaction of a given type of need, be it economic, social recognition, self-esteem, or the development of individual skills, while the crowdsourcer will obtain and utilize to

their advantage what the user has brought to the venture, whose form will depend on the type of activity undertaken. (Estellés and González 2012, 197)

2.2.1. Crowdsourcing in the Field of Translation

Just like how there are many different definitions of what crowdsourcing is in a general sense as mentioned above, there are also many different definitions for crowdsourcing in the field of translation studies. In this thesis, Minako O'Hagan's (2011) definition of crowdsourcing is accepted as the main definition because O'Hagan focuses on fan translations and their variants to define the term. In the paper "Community Translation: Translation as a social activity and its possible consequences in the advent of Web 2.0 and beyond", O'Hagan describes crowdsourcing as:

Key common characteristics shared in the above terms highlight that it is translation performed voluntarily by Internet users and is usually produced in some form of collaboration often on specific platforms by a group of people forming an online community (O'Hagan 2011, 14).

Apart from offering a definition for crowdsourcing in the field of translation, O'Hagan also divides crowdsourced translation into two as solicited and unsolicited translation. As an example to solicited translation models, O'Hagan gives Facebook's model. Facebook's model is solicited because there was a call for translation from the company itself and that's why the crowd did the translations. Another example to solicited translation models would be Greenpeace's *Save the Arctic* campaign (Yanar Torbalı 2019). Unsolicited translation on the other hand, is when the crowd comes together without a call and start translating content. O'Hagan gives fan translations of Japanese anime as an example for this type of crowd translations (O'Hagan 2011, 13).

2.2.2. Using Crowdsourcing in the Field of Translation: Benefits and Challenges

Together with the spread of Web 2.0 applications and websites all over the internet, crowdsourcing has become popular in many different industries. One of those industries is the translation industry. Crowdsourcing translation projects have three advantages for companies (Anastasiou and Gupta 2011, 639):

- Multilingual support
- Quick solution
- Monetary benefits

A good example to these benefits can be Facebook's crowdsourcing model. By creating and patenting their own crowdsourcing translation model under the name *Community Translation on a Social Network*, Facebook was able to have its crowd, the people using the social networking website, translate the content into languages other than English in a very short time and with a high quality (O'Hagan 2011, 12). In March 2008, the whole website was translated into French in 24 hours for free, with the collaboration of 4000 French users (Anastasiou and Gupta 2011, 639). As a result, it can be said that Facebook's example covers all three advantages mentioned above.

All the above-mentioned benefits of crowdsourcing translation projects on one hand, this method also poses some challenges and isn't without limitations. Especially, three issues are posing a serious problem in crowdsourcing translation projects. These issues can be listed as:

- **Quality:** In crowdsourced translation projects, the translation quality is good because the translation is made by humans. However, the members of the community taking part in such projects are mostly people with no translation background and they are non-professionals in the field. So, the output the non-professional crowd creates shouldn't be expected to be as high quality as the output a professional translator creates.
- **Motivation:** Motivation is another issue companies should think about when deciding to opt for crowdsourcing translation projects. This issue can directly result in the delay or even non-completion of the project and without motivation, crowdsourcing wouldn't be a feasible option. It is through motivation that a dedicated crowd can be maintained as part of the project and future projects. In crowdsourcing projects, the motivation the crowd seeks is not mainly financial but psychological. The crowd is mostly looking for knowledge, recognition and reputation, charity work, emotional attachment and fun.

- **Public Management and Control:** The final issue common in crowdsourcing translation projects is public management and control. In crowdsourced translation projects, there has to be someone from within the company and also somebody from within the community to oversee the community so that collaboration among the crowd members can be ensured and that the project doesn't fall under the leadership of an individual or a group of individuals from the crowd, which would cause the rest of the crowd to lose interest. Moreover, from the companies' perspective, control can also be a problematic issue as not every company has the resources to describe tasks to the crowd, put their project online and then manage the data they get from the crowd. (Anastasiou and Gupta 2011, 7, 8)

The issue of motivation is further examined by Jiménez-Crespo (2017) and the motivations of the crowd are separated into three tiers and other subsections in accordance with the studies mentioned in the paper. The following table is formed by bringing together the mentioned information:

Table 1: Table showing motivation tiers for crowdsourcing projects

Tier 1	Making information in other languages accessible to others.
	Helping the organization with their mission, or a belief in the organization's principles.
	Intellectual development
Tier 2	The desire to practise the second language.
	Professional motivations related to the need to gain translation. experience or increase one's reputation.
Tier 3	The desire to support lesser-known languages.
	The satisfaction of completing something for the good of the community.
	For fun.
	To be part of a network.

Source: Jiménez-Crespo, "Translation Crowdsourcing: Research Trends and Perspectives" p. 203

While tier one includes the intrinsic motivations and those are the most common results in the studies mentioned in the paper, tier two includes both intrinsic and extrinsic motivations. Meanwhile, tier three motivations appeared consistently the least in the studies.

2.3. Machine Translation

2.3.1. History of Machine Translation

In our day, even though any person with access to the internet can freely use a variety of machines to translate any text they want, it took the technology almost a century of development and improvement to become what it is today. It is without a doubt that in the future, human translators will have difficulties surviving in the industry because of the ever-increasing quality of MT (Öner and Öner Bulut 2021, 103; Alimen and Öner Bulut 2020, 1049). Historically, it was even before the first computers were invented that the first ideas for machines to automatically translate texts were patented and since then, the technology keeps on improving in line with technological developments. The history of machine translation in this part of the thesis is explained in line with W.J. Hutchins' classification of historical periods in machine translation. In 1994, William John Hutchins described the whole process of development in detail and divided it into five sections according to historical context. These five sections are:

- Precursors and pioneers (1933-1956)
- The decade of high expectations and disillusion (1956-1966)
- The ALPAC report and its consequences
- The quiet decade (1966-1976)
- Operational and commercial systems in the 1980s

(Hutchins 1994)

In 2014, in another unpublished paper, W.J Hutchins summarized all the steps he described in 1994 and also added three new sections concerning newer developments in the technology. These three new steps are:

- The early 1990s
- The late 1990s
- Since 2000

(Hutchins 2014)

2.3.1.1. Precursors and Pioneers (1933-1956)

This period starts in the 1930s, a time when computers weren't invented yet. Two people from different backgrounds came up with ideas to automatically translate texts and patented their ideas. However, not long after these ideas, computers were invented

and computer-based translation machines were started to be used. Code breaking methods used during the Second World War formed the basis of these machines and later on in the 1950, many US universities started research and in 1954, as a result of collaboration between IBM and Georgetown University, the first public demonstration was given. During this demonstration, a Russian text was translated into English. Thanks to this public demonstration, funding for MT drastically increased in the US and also inspired MT projects throughout the world (Dostert 1957, 4).

2.3.1.2. The Decade of High Expectations and Disillusions (1956-1966)

During this period, most research in machine translation was made on English and Russian language pairs as a result of the rising tensions due to the Cold War. IBM Corporation installed a system for the US Air Force to produce translations and this system was used until 1970 when it was replaced by Systran. Before that, Georgetown University had installed Russian-English systems at Euratom in Italy and at the Oak Ridge National Laboratory of the US Atomic Energy Commission in 1964 (Hutchins 1994, 9). It is mentioned that even though the quality of the produced translations was low, it was satisfying for the recipients as the system was able to rapidly produce information (Hutchins 2014, 2).

2.3.1.3. The ALPAC Report and its Consequences

In 1966, the Automatic Language Processing Advisory Committee (ALPAC), which was formed by US government sponsors of machine translation systems, published a report. In this ALPAC report, it was announced that MT was slower, less accurate and twice as expensive as human translation (ALPAC 1966, 28). As a result of this report, MT research in the USA came to a halt. However, this did not mean that MT research halted all over the world. On the contrary, research on MT continued mostly in Canada, France and Germany (Hutchins 2014, 2).

2.3.1.4. The Quiet Decade (1966-1976)

Even though the ALPAC report brought MT research in the USA to a stop, that was not the case in Canada and Europe. In Europe, together with increasing cooperation between the existing members of the European Economic Community (EEC) and other countries joining the community, there was an ever-growing need for the translation of scientific, technical, administrative and legal documentation from and into all the Community languages (Hutchins 1994, 11). Thus, after the USAF started using Systran,

the Commission of the European Communities also started using it in 1976 and in Canada, a system called “Meteo” started to be used for translating weather reports due to both English and French being official languages (Hutchins 2014, 2).

2.3.1.5. Operational and Commercial Systems in the 1980s

Systran is explained to be the most successful operational system back in 1994 as it started to be used by USAF in 1970 and by the European Communities 1976 and only the system used by USAF translated more than 100,000 pages a year with an accuracy of 95% (Hutchins 1994, 13). During this period, together with the advancements in computer technology, a wide variety of new MT systems also emerged and the number of countries investing in the technology also increased.

2.3.1.6. The Early 1990s

The end of the 1980s is mentioned to be a major turning point for the technology due to four important events that happened; First of all, IBM conducted research on a system using statistical methods and published the results. Secondly, Japanese groups began using corpora-based systems. Another important event was that translation memory systems like Trados were released in the market and finally, research on speech translation was started (Hutchins 2014, 4).

2.3.1.7. The Late 1990s

The 1990s saw an increase in the usage of MT and translation aids not only by governmental bodies but also by large corporations as well. Moreover, there had also been a dramatic increase in the sales of MT software for personal use as well and MT became a mass-market product (Hutchins 2014, 4).

2.3.1.8. Since 2000

At the start of the 2000s, the focus of research was on statistical machine translation and example-based machine translation. However, in time, the focus of research became neural machine translation and the quality of machine translation output has increased greatly as a result (Turovsky 2017). Google Translate can be given as an example to this. When it was first launched in 2006, Google Translate made use of statistical machine translation but in 2016, neural machine translation started to be used. This meant that instead of translating pieces of sentences, the system translated whole sentences at a time (Turovsky 2017).

2.3.2. Machine Translation Post-Editing

As previously mentioned, the quality of machine translation tools has increased since neural machine translation became the focus of research in the field. However, although MT quality has been improved in many language pairs, it is still not possible to use the direct machine translation output, both for professional purposes and for the general internet users. As a result of that, instead of directly using machine translation output, professional translators have been using machine translation to pre-translate a text and then apply post-editing on the raw MT output so that the texts reach a certain level of quality. According to two different researches mentioned by Maarit Koponen (2016), it was shown that this method greatly increased the speed at which professional translators worked, increasing the number of words translated per day by 2,000 to 3,500 words according to one research and up to 5,000 words according to the other one but these numbers highly depended on the language pairs, translator experience and the quality of raw machine translation (Koponen 2016, 132, 134). In another study, it was found out that translation students benefited from post-editing in terms of saving processing time and reducing cognitive effort when compared with from-scratch translation (Jia, Carl and Wang 2019, 79).

Due to post-editing becoming more and more integrated in the industry with each passing day, guidelines on how to provide post-editing services were formed up. The two examples of such guidelines are TAUS's (Translation Automation User Society) guidelines on post-editing and the guidelines provided by the International Standard ISO 17100. During the study, TAUS's post-editing guidelines have been preferred as required documentation is easier to reach.

According to TAUS's post-editing guidelines, there are two types of post-editing qualities. These are "good enough" quality and "human translation quality" and guidelines for each of these is different from the other's. The post-editors aim for different levels of quality by following these quality guidelines and the "good enough" quality post-editing and "human translation quality" post-editing of the same machine translation output may be very different from each other.

2.3.2.1. “Good Enough” Quality

“Good enough” quality is the lower quality option of post-editing machine translation output. This type of post-editing is recommended for higher quality MT output because if this is applied on poor quality MT output, the resulting text may not be up to publishable quality standard. It is enough if the text is comprehensible and accurate. The style is not important, the text may sound like it was generated by a computer, the syntax might be unusual, grammar may not be perfect but it is important that the message is accurate (Massardo et al. 2016, 16, 17).

The following are the guidelines given in the book to achieve “good enough” quality:

- Aim for semantically correct translation
- Ensure that no information has been accidentally added or omitted.
- Edit any offensive, inappropriate or culturally unacceptable content.
- Use as much of the raw MT output as possible.
- Basic rules regarding spelling apply.
- No need to implement corrections that are of a stylistic nature only.
- No need to restructure sentences solely to improve the natural flow of the text.

(Ibid, 17)

2.3.2.2. “Human Translation Quality”

This quality level is defined as being comprehensible, accurate, stylistically fine even if it may not be as good as it can be achieved by a native speaker, syntax is normal, grammar and punctuation are correct. When this quality type is preferred, the post-edited text reaches a quality similar to a text translated by a human. There may be no need for this level of post-editing quality if the MT output is already of good quality (Ibid, 16, 18).

For human translation quality, the following guidelines are given:

- Aim for grammatically, syntactically and semantically correct translation.
- Ensure that key terminology is correctly translated and that untranslated terms belong to the client’s list of “Do Not Translate” terms.
- Ensure that no information has been accidentally unacceptable content.

- Edit any offensive, inappropriate or culturally unacceptable content.
- Use as much of the raw MT output as possible.
- Basic rules regarding spelling, punctuation and hyphenation apply.
- Ensure that formatting is correct.

(Ibid, 18)

As a result of the above-mentioned facts concerning the efficiency of machine translation, the necessity of applying different levels of post-editing to raw machine output and machine translation being used only with post-editing in the translation industry, it can be said that even though new technologies are being implemented in machine translation (like the transition from rule based machine translation and statistical machine translation to neural machine translation which provides a much better translation output), raw machine translation itself can't be used for texts to be published (Doherty and O'Brien 2014, 40). One of the main research questions the present study aims to answer is whether raw machine translation without the use of any post-editing can be used in the localization of video games texts. To determine whether raw machine translation can be used or not in video games localization, error annotation has been conducted on the raw machine translation output from Google Translate and Amazon Translate within the scope of the thesis. The reasons why these two machine translation tools are chosen are; they are both free to use and they both make use of neural machine translation. The error annotation of the video game texts showed how efficient which machine translation tool is in translating the texts contained in the video game *Kenshi* and whether the company could have made use of free machine translation tools instead of opting for some kind of crowdsourcing translation project from among the player base of the game in question or whether machine translation-post editing is required and whether machine translation post-editing could have been used in order to decrease localization costs.

In line with TAUS's quality evaluation guidelines, there are two options for machine translation quality evaluation both of which are human evaluation methods. The first option is to conduct human error annotation using TAUS DQF-MQM Error Typology. TAUS DQF-MQM error annotation is very detailed and the raw machine output is analyzed under 8 main categories and a total of 33 subcategories. This method

is much more detailed than the other method but it is not the preferred quality evaluation method because studies have shown that the more categories are included in human quality evaluation methods, the more discrepancies are observed among different annotators, which may alter the results of the quality evaluation as a whole (Bayerl and Paul 2011, 720).

The other method is “Quality Evaluation Using Adequacy and Fluency Approaches”. In this method, the raw machine output is evaluated only in two categories, which are adequacy and fluency. Even though TAUS DQF-MQM Error Typology is much more detailed and can provide more elements for analysis, the preferred quality evaluation methods in the present study are the adequacy and fluency approaches as it has been proved that inter-annotator agreement increases when typologies with fewer categories are used, meaning that the results obtained from the tasks the volunteers complete will be more reliable (Stymne and Ahrenberg 2012, 1789).

2.3.3. “TAUS DQF-MQM” Machine Translation Error Annotation

During the study, the quality of the machine translation output of the text has been evaluated by making use of the TAUS DQF-MQM error typology evaluation template. The TAUS DQF-MQM error typology is human annotation of machine translation errors in order to assess the quality and provide input into improvement efforts (Lommel et al. 2015, 4). DQF stands for Dynamic Quality Framework which was developed by TAUS and its members. DQF contains different tools for evaluation of translation quality, one of which is the error typology. On the other hand, MQM stands for Multidimensional Quality Metrics and it was developed as part of an EU funded project called the QTLaunchPad (TAUS 2015).

When they were first developed, the DQF and MQM error typologies were developed separately from each other. However, both of the models turned out to be very similar to each other and the developers of MQM and DQF agreed to make changes to both frameworks in order to bring them to harmony (Lommel et al. 2015, 4). The harmonization process was due to be complete in the summer of 2021 and TAUS DQF-MQM error annotation typology is a result of that harmonization. The tables containing the “Harmonized DQF-MQM Error Typology” and the “Severity Levels for Error Categories” can be found in “Appendix A” and “Appendix B”.

As it can be seen in “Appendix A”, there are a total of 8 main error types. These error types are also divided into numerous subcategories so that the error annotation can be more detailed:

- **Accuracy:** The errors in this category are mostly about the errors the machine translation tool makes during the translation process and as a result of these errors, there are discrepancies between the source text and the machine translation output. The subcategories in this category are addition, omission, mistranslation, over-translation, under-translation, untranslated segments and improper exact TM match.
- **Fluency:** This category is mostly about linguistic errors. Even though the source text is translated, these errors prevent the machine translation output from being understood. The subcategories in this category are punctuation, spelling, grammar, grammatical register, inconsistency, link/cross-reference and character encoding.
- **Terminology:** Terminology errors are the errors that come to be when a term or terms are translated differently in different segments or are translated with a term that is not expected. Thus, there are two subcategories in this category, inconsistent with termbase and inconsistent use of terminology.
- **Style:** The errors within this category all concern the stylistic problems within the machine translation output. Even though these errors may not effect the comprehensibility of the text, they may for example prevent or make it harder for the main message of the text from being conveyed. The subcategories in this category are awkward, company style, inconsistent style, third-party style and unidiomatic.
- **Design:** Design errors are errors concerning the overall design and format of the text. The machine translation tools may or may not change these in some cases. The subcategories are length, local formatting, markup, missing text, truncation/text expansion.
- **Locale Convention:** A good example to this error type would be the wrong usage of punctuation in the target language. The errors falling within this category do not adhere with locale-specific mechanical conventions. There

are a total of six subcategories in this category; address format, date format, currency format, measurement format, shortcut key and telephone format.

- Verity: If this type of errors are found in the machine translation output, it means that the output contains information not included in the source text and thus, the statements within the text are contradicting the world of the text. There is only a single subcategory here, which is culture-specific reference.
- Other: All types of errors that do not fit in the above-mentioned categories fall within this category. As a result, there are no subcategories.

Apart from the error categories, TAUS DQF-MQM error typology also categorizes the machine translation output error according to their severity levels. The severity levels indicate the severity of the error and how much it affects the overall quality of the machine translation output. There are a total of five different severity levels ranging from critical to kudos (see Appendix B).

2.3.4. Quality Evaluation Using Adequacy and Fluency Approaches

Adequacy and fluency evaluation can be used for assessing the quality of machine translation output. This method of quality evaluation can be preferred over the TAUS DQF-MQM error typology approach as the number of discrepancies between human evaluators decrease sharply when compared to evaluation methods involving numerous error categories. It is also less costly and time consuming to implement than an error typology approach (TAUS 2017, 4).

Adequacy refers to how much of the meaning from the source text is expressed in the target translation. On the other hand, fluency refers to the following qualities of the output:

- the grammatical quality,
- the correctness of spellings,
- adherence to the common use of terms, titles and names,
- acceptability and interpretability by a native speaker

(TAUS 2017, 5)

The machine translation output is rated from 1 to 4 for both qualities. The following tables show what each rating means for adequacy and fluency.

Table 2: Table showing ratings for adequacy evaluation for machine translation output

Adequacy	
1- None	3- Most
2- Little	4- Everything

Source: TAUS, Quality Evaluation Using Adequacy and Fluency Approaches, 5

Table 3: Table showing ratings or fluency evaluation for machine translation output.

Fluency	
1- Incomprehensible	3- Good
2- Disfluent	4- Flawless

Source: TAUS, Quality Evaluation Using Adequacy and Fluency Approaches, 5

The lines selected for the quality evaluation and machine translation post-editing tasks of *Kenshi*'s language files are about in-game races the players can choose to become while creating a new character. These lines were chosen so that all the lines in the tasks can be fitted in a single context, which is the in-game races context. In order to ensure consistent quality for the duration of the tasks:

- the human evaluators are native Turkish speakers,
- they are near-native English speakers,
- they either have previous experience with this quality evaluation method or,
- have received materials and guidelines with examples.

CHAPTER TWO

METHODOLOGY

In order to test the hypothesis mentioned in the introduction part, the methodology followed during the present study is as follows:

30 lines from *Kenshi*'s "gamedata.pot" file, the file containing all in-game texts, are chosen. All of these lines are within the same context; they are all shown to the player during the character creation part of the game, which is the first thing the players do when they start a new game. The chosen lines are about the races the player can choose while creating a new character. They include race names, sub-race names and race descriptions. The source text can be found in "Appendix C".

These selected lines are translated from English into Turkish using two different machine translation tools: Google Translate and Amazon Translate. Both of these machine translation tools are neural machine translation (NMT) tools and are free to use. These qualities are important for the study due to different reasons. As mentioned previously, the success rate of translation machines using neural machine translation tools are higher than those using other systems such as statistical machine translation tools, rule-based machine translation tools and etc. It is also important that these machine translation tools are free to use because the focus of the study is to determine whether machine translation and machine translation post-editing can be used as an alternative to crowdsourcing localization projects, an option companies prefer when they do not want to spend money on localization.

After the selected lines are translated using the selected machine translation tools, three volunteer translators took part in four different tasks to determine how successfully these two machine translation tools translated the source text into Turkish and whether post-editing was needed or not. Before the volunteers started doing the tasks, their consents were taken and a briefing file prepared to inform them on the tasks they were required to do and how they were required to do them was shared with them.

All three volunteers hold a bachelor's degree in translation studies, have experience working as translators and have previous experience with machine translation and machine translation post-editing. The names of the machine translation tools were not shared with the volunteers so that there would be no favoritism and the volunteers graded the raw machine translations and conducted the post-editing tasks without the knowledge of which machine translation tool they were grading and post-editing.

The volunteers error annotated the raw machine translation outputs in line with TAUS adequacy and fluency approaches, then, if post-editing was needed, they post-edited the machine translation outputs in line with the TAUS "Human-Translation Quality" guidelines. All of these tasks were assigned to the volunteers on the TAUS Dynamic Quality Framework (DQF) tool. After the tasks were assigned to the volunteers, they received e-mails containing links which directed them to the tasks on the DQF tool website.

When the tasks have been completed by the volunteers, TAUS DQF tool automatically analyzed the data and prepared reports on the results of the tasks. Adequacy and fluency task reports include the overall score given by the volunteers for the two raw machine translation outputs and also include how did each volunteer marked each line. The post-editing task, which is named productivity testing on TAUS DQF tool, provided information on the average number of words processed in a given timespan and average percentage of word changes applied by each volunteer in each sentence.

IV. RESULTS

The results taken from the TAUS DQF tool are analyzed separately for each machine translation tool and for each different task. There is a productivity report, an adequacy report and a frequency for the raw MT output from each of the machine translation tools.

4.1. Adequacy Report for Google Translate

The adequacy reports generated by TAUS DQF tool start with the “Average Adequacy” score given to the raw machine translation output. This score shows the mode average calculated on an aggregated basis (TAUS 2014). The “Number of Segments” part shows all evaluated segments multiplied by the number of evaluators. In this study, as each volunteer was responsible for 30 lines, the total number of segments is shown as 90. As previously mentioned, adequacy shows how much of the meaning expressed in the source text is also expressed in the target translation. It is graded between 1 and 4, 1 meaning that none of the meaning is expressed and 4 meaning that everything from the source text is expressed in the target text.

Figure 5: Average Adequacy Result of Google Translate

Average Adequacy (Info)					Legend 4 Everything 3 Most 2 Little 1 None
Language Pair	Average Adequacy	MT Engine	Number of Segments	Number of Words	
English (United States) > Turkish	4	Not Specified	90	448	

Figure showing “Average Adequacy” for Google Translate raw machine translation output

In terms of adequacy, as it is shown in “Figure 5”, the average was calculated as 4, meaning that on average, Google Translate was able to transfer everything. The next result shown is the “Distribution of Adequacy Score, Per Language”. This result shows adequacy levels and the number of respective segments.

Figure 6: Distribution of Adequacy Scores, Per Language for Google Translate

Distribution of Adequacy Scores, Per Language [\(Info\)](#)

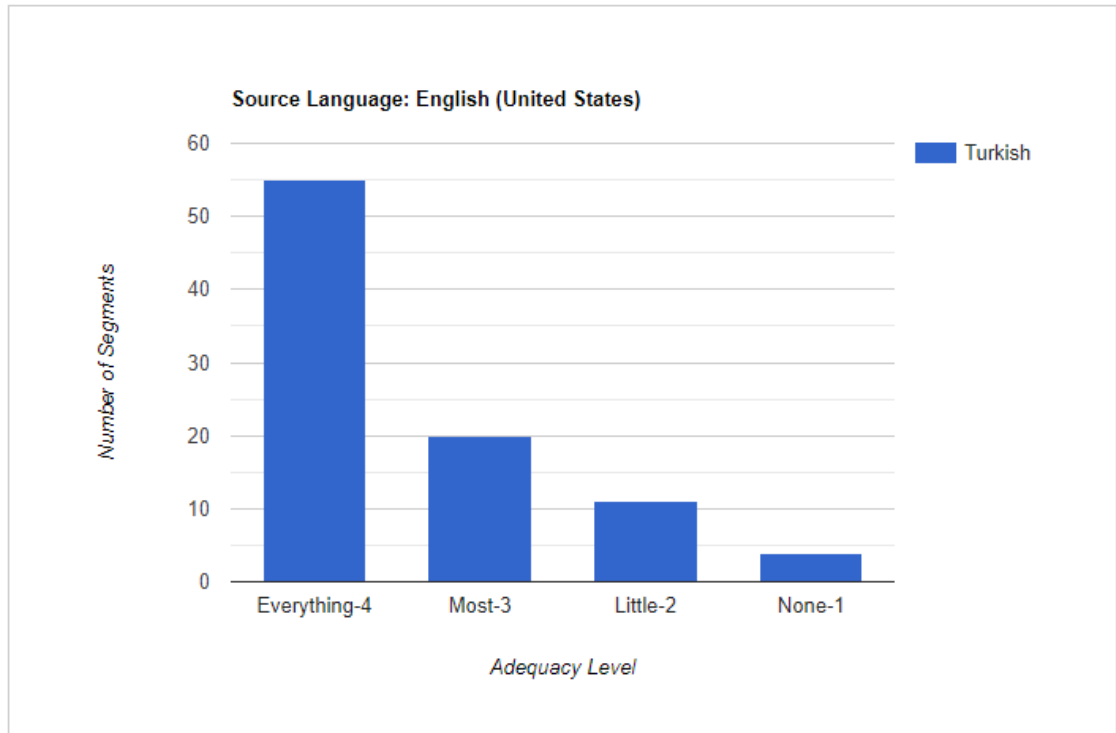


Figure showing “Distribution of Adequacy Scores, Per Language”

The results within this figure can be listed as; “Everything” 55 segments, “Most” 20 segments, “Little” 11 segments and “None” 4 segments.

The next figure shows “Distribution of Adequacy Scores, Per Evaluator”. In order to keep the volunteers’ names anonymous, their names were erased from this chart and all the other following charts. Each color indicates a different volunteer and the colors are always the same for the same volunteers in all the following figures. This chart is important to figure out whether inter-annotator agreement is high or low.

Figure 7: Distribution of Adequacy Scores, Per Evaluator for Google Translate
Distribution of Adequacy Scores, Per Evaluator

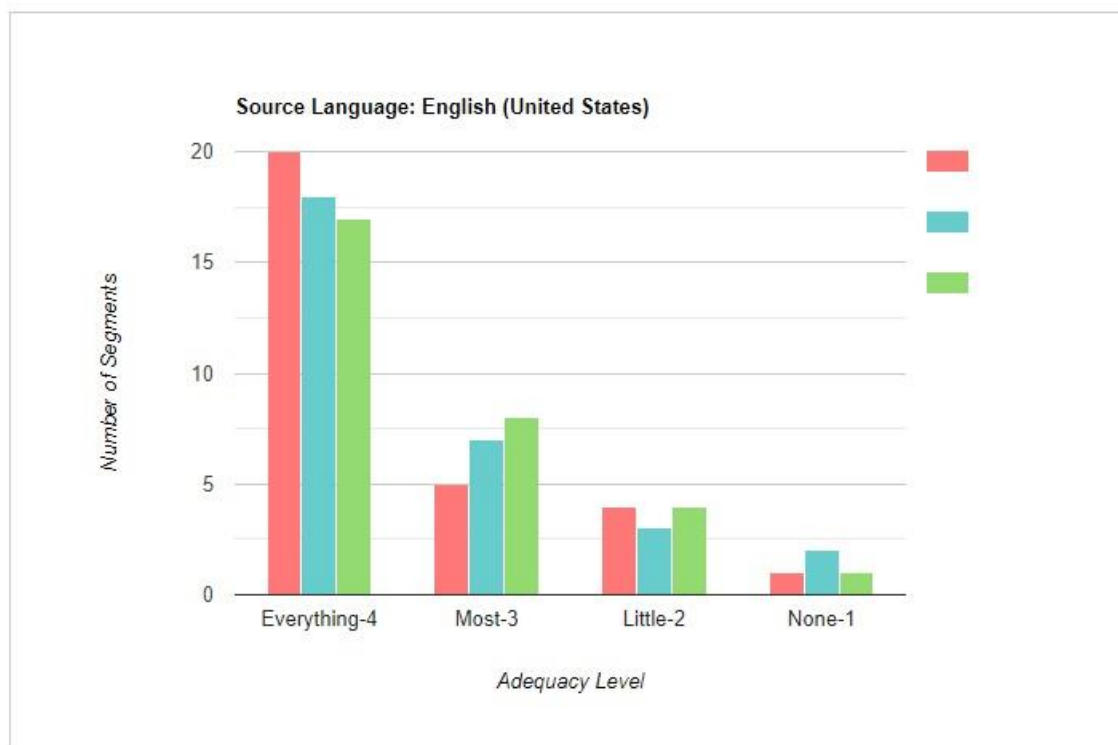


Figure showing “Distribution of Adequacy Scores, Per Evaluator”.

4.2. Fluency report for Google Translate

The fluency reports also start with an “Average Fluency” score, showing the average fluency score handed out by the three volunteers. This figure also shows the mode average calculated on an aggregated basis (TAUS 2014) and the number of segments shown are 90 as each of the three volunteers were responsible for 30 segments.

Figure 8: Average Fluency Result of Google Translate

Average Fluency [\(Info\)](#)

Language Pair	Average Fluency	MT Engine	Number of Segments	Number of Words
English (United States) > Turkish	4	Not Specified	90	448

Legend

- 4 Flawless
- 3 Good
- 2 Disfluent
- 1 Incomprehensible

Figure showing “Average Fluency” for Google Translate raw machine translation output

The fluency report for Google Translate’s raw machine translation output shows that, on average, Google Translate flawlessly translated the sentences. The next chart shows the distribution of fluency scores:

Figure 9: Distribution of Fluency Scores, per Language for Google Translate
Distribution of Fluency Scores, per Language [\(Info\)](#)

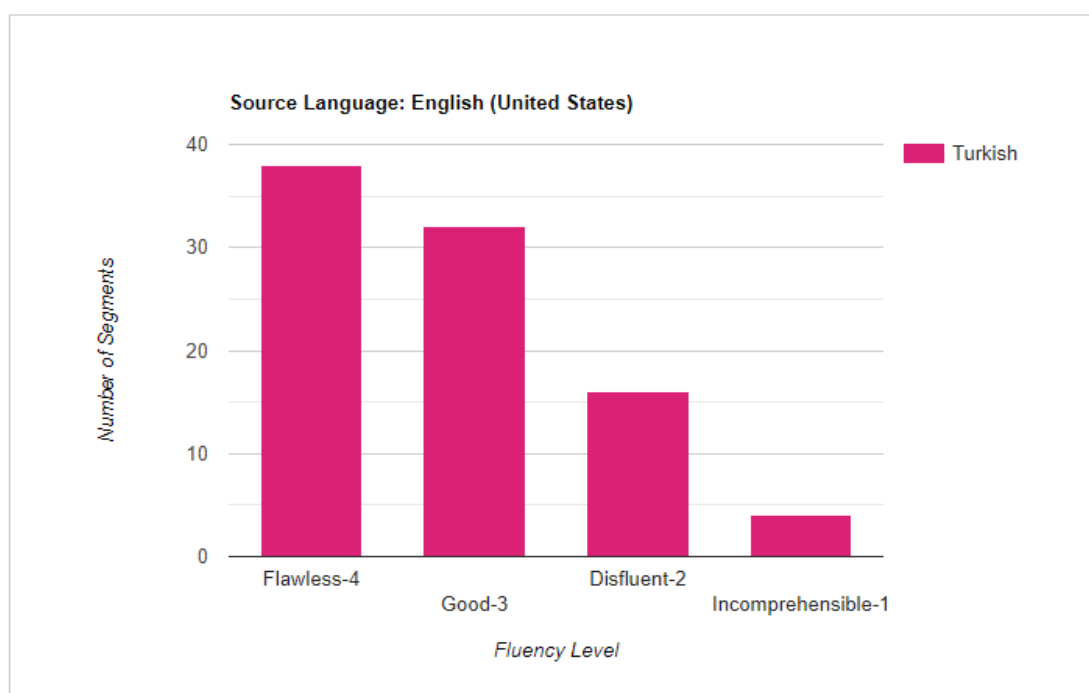


Figure showing “Distribution of Fluency Scores, per Language”.

Among 90 segments, the volunteers gave a score of 4 to 38, a score of 3 to 32, a score of 2 to 16 and a score of 1 to 4 segments. The first thing that catches attention in this chart is that the number of segments that received a score of 4 and 3 are very close to each other, this is due to low inter-annotator agreement levels, which can be seen in the next figure.

Figure 10: Distribution of Fluency Scores, per Evaluator for Google Translate

Distribution of Fluency Scores, per Evaluator

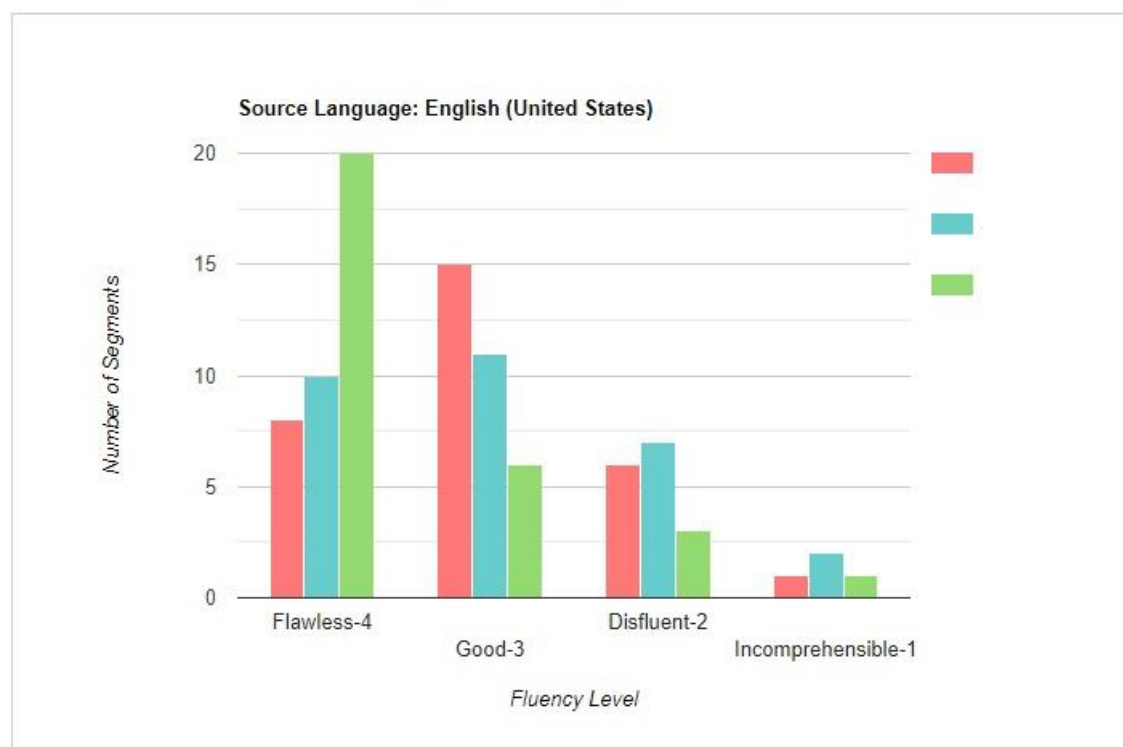


Figure showing “Distribution of Fluency Scores, per Evaluator”.

There have been big differences between the volunteers in scoring the fluency of Google Translate’s raw machine translation output. While the Green volunteer gave a “Flawless” score to 20 segments, the number of “Flawless” scores for the Red and Blue volunteers were 8 and 10 respectively. Inter-annotator agreement levels were also low for the “Good” score. According to the Red volunteer, the number of “Good” segments was 15, while this number was 11 and 6 for the Blue and Red volunteers.

4.3. Productivity report for Google Translate

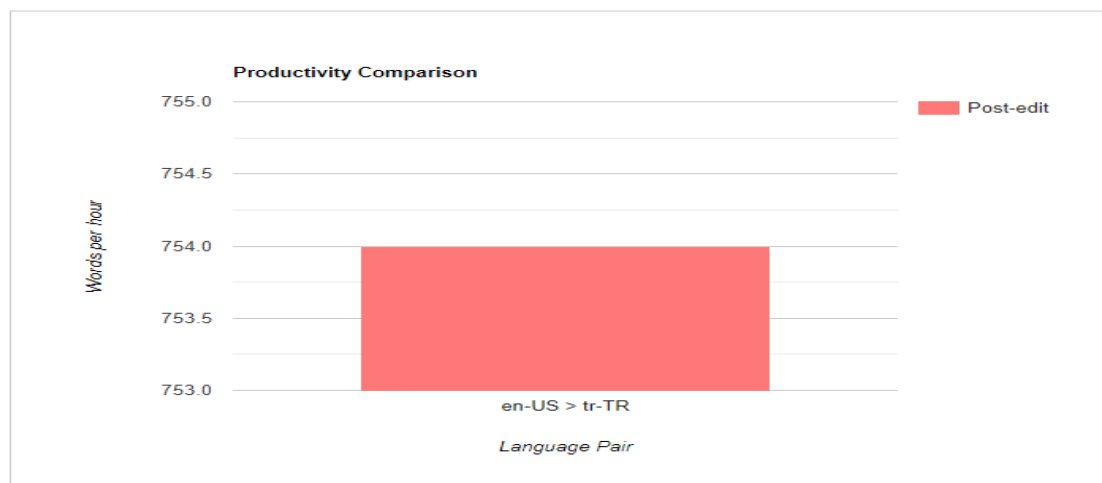
The productivity report gives information on the time to edit and post-editing effort applied. Time to edit measures the average number of words processed in a given timespan and post-editing effort measures the average percentage of word changes applied to the raw machine translation output by the volunteers (TAUS 2014).

This report starts with information on “Average Productivity”, showing how many words per hour the volunteers were able to post-edit during the task. The following figures show this information.

Figure 11: Average Productivity Report for Google Translate

Average Productivity [\(Info\)](#)

Language Pair	Number of Segments	Number of Words	Post-edit (WPH)	Time Spent for Post-edit (seconds)	MT Engine
English (United States) > Turkish	30	448	754	6,855	Not Specified



Note that charts represent an extrapolation of the results to a one hour scale

[Back to Top](#)

Figure showing “Average Productivity” of the volunteers

The “Average Productivity” result shows that, the volunteers were able to post-edit 754 word on average and the total amount of time spent for post-editing by the three volunteers is 6,855 seconds. The next figure shows “Productivity per Editor”, showing how many words per hour each volunteer was able to post-edit.

Figure 12: Productivity per Editor for Google Translate

Productivity per Editor

Evaluator Name	Post-Edit (WPH)	Time Spent for Post-edit (seconds)
	514	3134
	788	2044
	961	1677

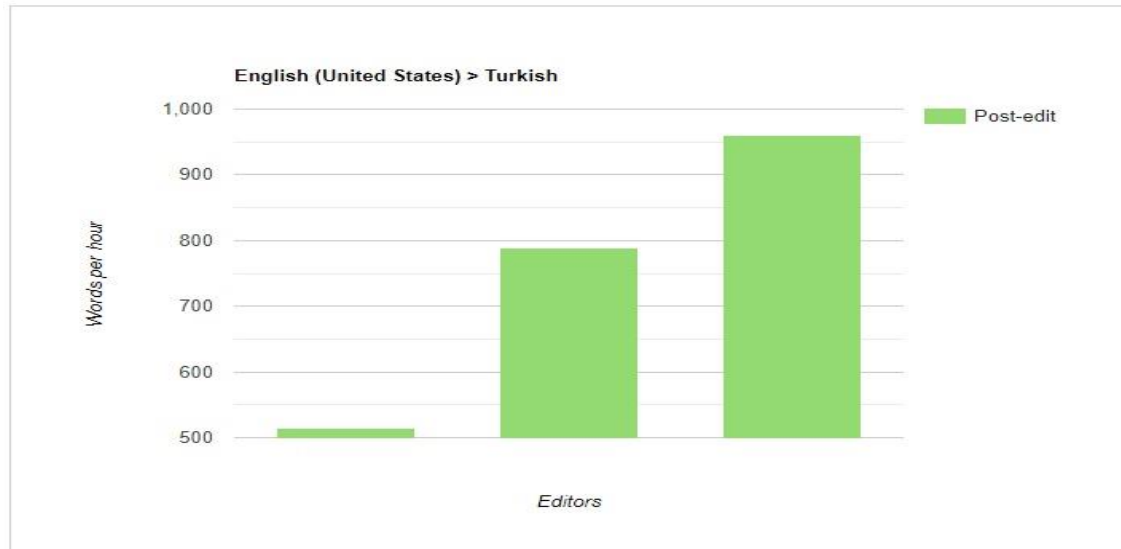


Figure showing “Productivity per Editor”

In Figure 12, it can be clearly seen that there is a large difference between the volunteers in terms of productivity. The names of the volunteers have been erased in order to protect their anonymity.

The first editor on the left, known as Red in the previous results, was able to post-edit 514 words per hour. The second editor in the middle, Blue, was able to post-edit 788 words per hour and the last editor on the right, Green, was able to post-edit 961 words per hour. As mentioned above, the average of the three editors is 754 words per hour. Even if the lowest amount of productivity is taken into account, it means that while a translator is generally expected to translate around 2000 words per day (Memsources 2015) (Robert 2013, 32), by using Google Translate and machine translation post-editing, the number of words translated per day can reach 4000 words per day.

The final result to be mentioned in terms of productivity is how many percent of total segments were edited. This result is shown via the “Edit Distance Graph”, which can be found below as Figure 13. “Edit Distance Graph” gives an overview of the post-

editing effort required to bring the raw machine translation output to the desired level of quality, which is determined as human translation quality for the duration of the study. In this graph, the percentage of total segments is displayed together with “Segment edit distance”. “Segment edit distance” shows numbers, 0 to 10, which indicate the distance of the raw machine translation output to the post-edited version of the segments. 0 means that no or almost no editing was needed while 10 means that there was no similarity between the raw machine translation output and the post-edited version of the segments (TAUS 2014).

Figure 13: Edit Distance Graph for Google Translate

Edit Distance Graph [\(Info\)](#)

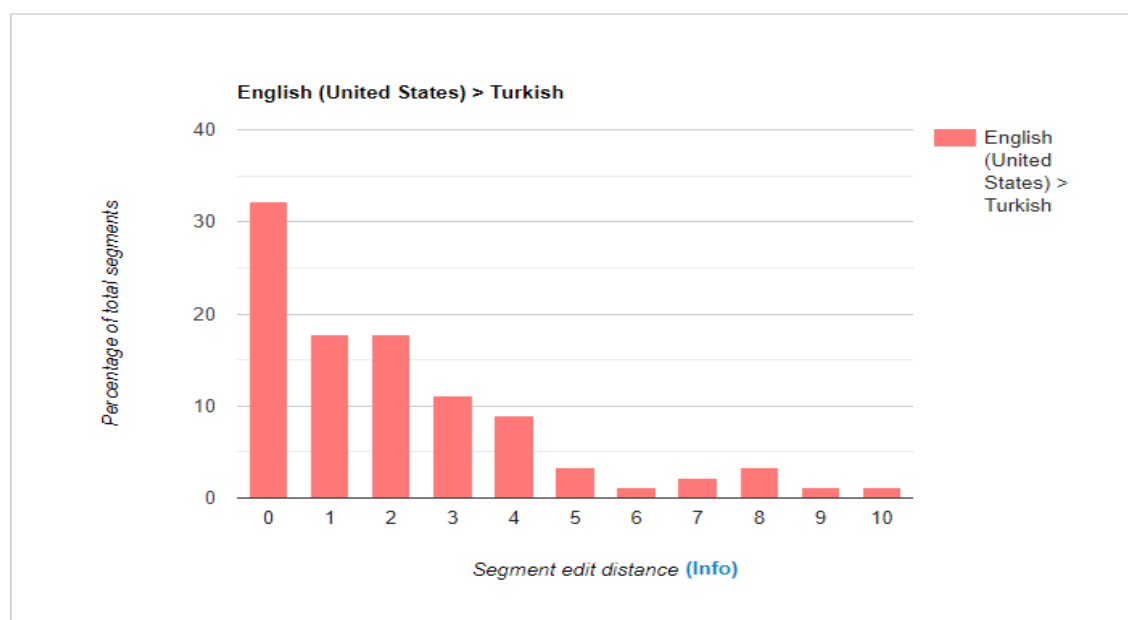


Figure showing “Edit Distance Graph”

As can be seen in Figure 13, 32 percent of total segments were subjected to no or almost no post-editing. The percentage rate of segments that were completely changed, shown as 10 in the chart, was only 1. This shows that even if post-editing was needed for the raw machine translation output of Google Translate, most of the output can also be used with no editing or by slightly editing it.

The edit efforts per editor are also given in the results and they can be found in the following table. This time, the raw machine translation output and post-editing difference is given in percentages. On TAUS DQF tool website, it is written that 100 means that the raw machine translation output and the post-editing are exactly the same

while 0 means they were completely different (TAUS 2014). However, when the results were analyzed, it was determined that the description should have been just the opposite of what is written on the website as 0 percent means that there have been no changes made to the raw machine translation output while 100 percent means that the raw machine translation output and the post-edit are completely different from each other. TAUS was contacted about this issue and confirmation has been received that there was an error on TAUS’ side. Further information on this can be found in “Appendix P”.

Table 4: “Edit Efforts” per editor for Google Translate

EDIT EFFORT	Segment ID	Red	Blue	Green
	1	0	0	0
	2	0	72	75
	3	0	58	18
	4	22	1	12
	5	100	80	91
	6	12	20	53
	7	6	38	26
	8	29	41	44
	9	0	20	20
	10	0	16	16
	11	5	3	0
	12	0	9	1
	13	16	26	46
	14	0	0	9
	15	0	75	37
	16	4	9	47
	17	11	13	0
	18	3	3	2
	19	0	70	35

Table 4: - Continue

Table showing “Edit Efforts” per editor for Google Translate

EDIT EFFORT	20	30	11	13
	21	6	31	44
	22	0	0	0
	23	0	20	32
	24	16	33	11
	25	16	29	24
	26	0	14	14
	27	2	16	0
	28	34	23	25
	29	1	20	20
	30	12	41	44

The names of the editors have been given as Red, Blue and Green in order to protect their anonymity

The “Edit Effort” differences between editors are discussed at the end of the results part of the study. The following results are about Amazon Translate’s raw machine translation output and its post-editing.

4.4. Adequacy Report for Amazon Translate

The average adequacy for Amazon Translate’s raw machine translation is also calculated as 4 (Everything), meaning that Amazon Translate was also able to translate the source text by mostly keeping the same meaning as that of the source text.

Figure 14: Average Adequacy Result for Amazon Translate

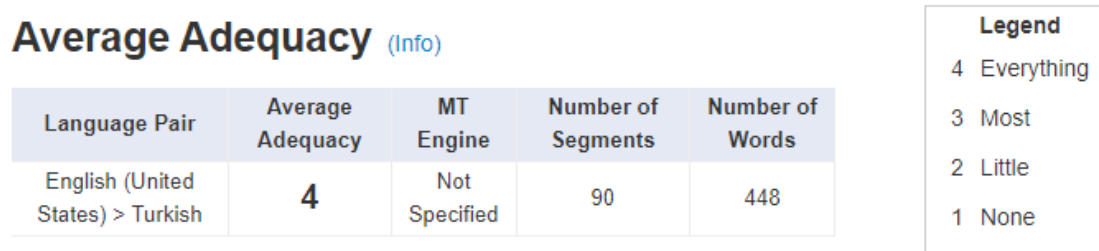


Figure showing “Average Adequacy” for Amazon Translate’s output

However, the number of segments given a score of 4, when compared to Google Translate’s results, are much lower and the number of segments that received a score of 2 are significantly higher. This situation can be clearly seen in the following figure.

Figure 15: Distribution of Adequacy Scores, per Language for Amazon Translate

Distribution of Adequacy Scores, Per Language [\(Info\)](#)

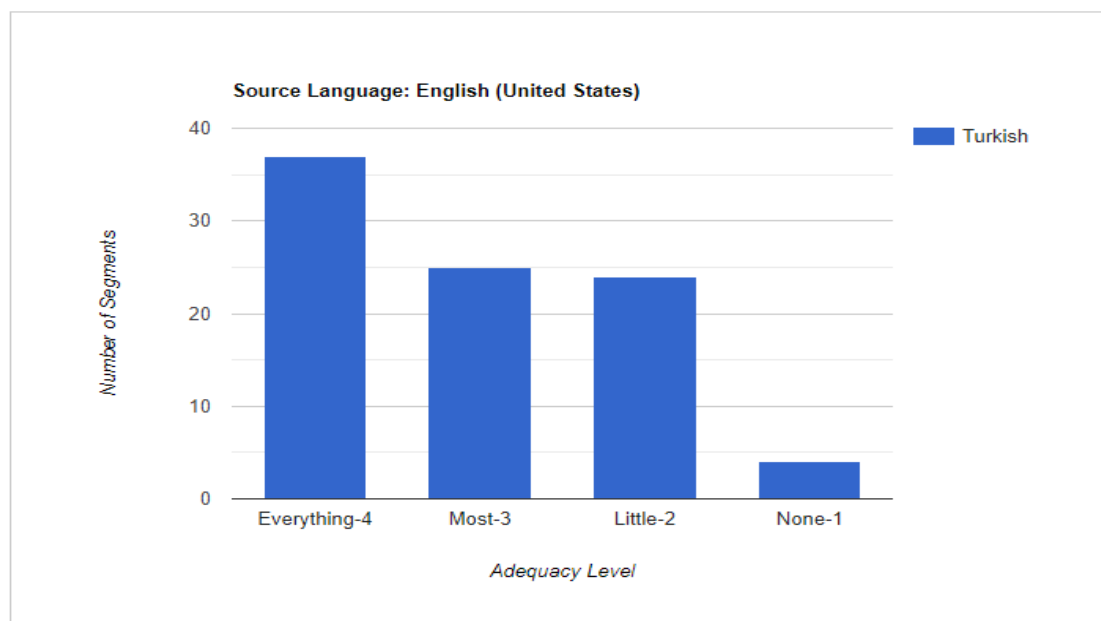


Figure showing “Distribution of Adequacy Scores, Per Language” for Amazon Translate

The volunteers gave a score of 4 to 37 segments in total. This number was 55 with Google Translate. Moreover, the number of segments that received a score of 3 and 2 are 25 and 24 respectively. The number of segments that received the same scores were 20 and 11. This means that, overall, Amazon Translate is less successful in translating the same text. On the other hand, in handing out a score of 4, inter-annotator agreement was much higher in Amazon Translate. Inter-annotator agreement mostly differed while giving scores of 3 and 2. The following figure provides evidence for this.

Figure 16: Distribution of Adequacy Scores, per Evaluator for Amazon Translate

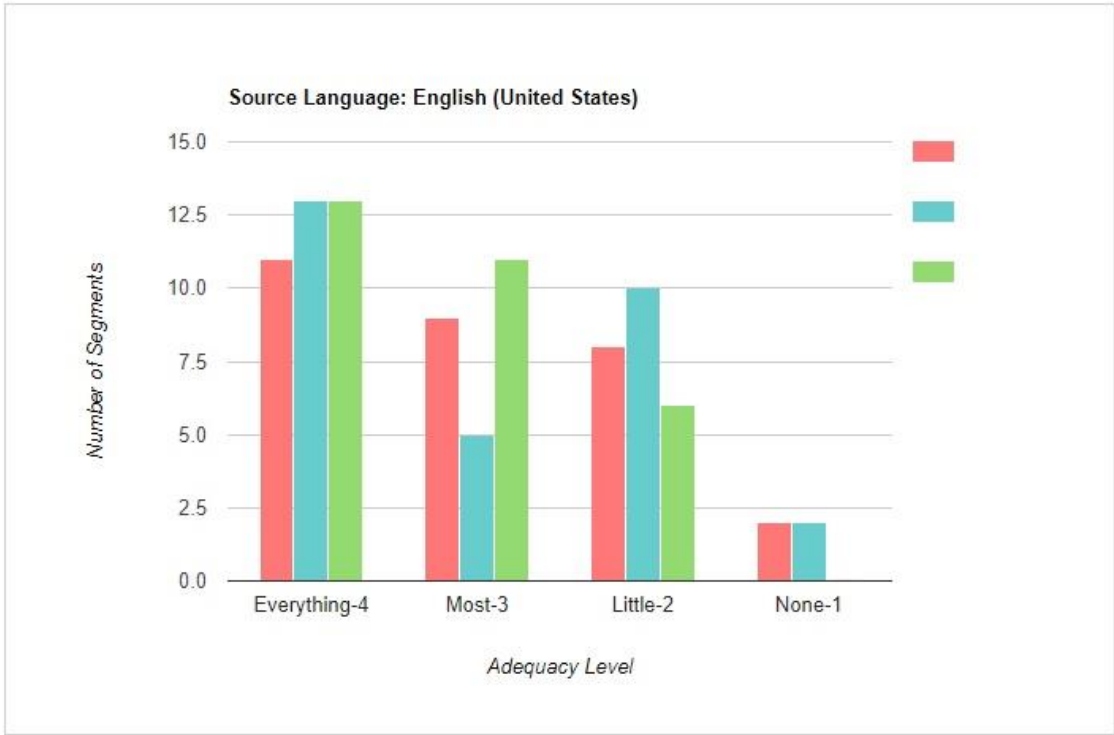


Figure showing “Distribution of Adequacy Scores, Per Evaluator”. Each volunteer is represented with a different color. The volunteer-color pairing is the same as in Google Translate results.

4.5. Frequency report for Amazon Translate

The “Average Fluency” result for Amazon Translate was calculated to be lower than Google Translate’s result. While Google Translate received an average score of 4 (Flawless), Amazon Translate’s average fluency was calculated to be 3 (Good).

Figure 17: Average Fluency Result of Amazon Translate

Average Fluency (Info)					Legend	
Language Pair	Average Fluency	MT Engine	Number of Segments	Number of Words	4	Flawless
English (United States) > Turkish	3	Not Specified	90	448	3	Good
					2	Disfluent
					1	Incomprehensible

Figure showing the “Average Fluency” score for Amazon Translate

The next figure, “Figure 18” shows “Distribution of Fluency Scores, per Language” result for Amazon Translate. The numbers of “Good-3” and “Disfluent-2” segments surpassed the number of “Flawless-4” segments in Amazon Translate’s results. Out of 90 segments, 30 segments per evaluator, a total of 22 segments received a score of “Flawless-4”, 33 segments received a score of “Good-3”, 25 segments received a score of “Disfluent-2” and 10 segments received a score of “Incomprehensible-1”. This shows that, when compared with Google Translate’s fluency results, there is a significant difference in the number of flawless and disfluent segments. The number of flawless segments is noticeably lower in Amazon Translate’s results. On the other hand, the number of disfluent segments is higher.

Figure 18: Distribution of Fluency Scores, per Language for Amazon Translate
Distribution of Fluency Scores, per Language [\(Info\)](#)

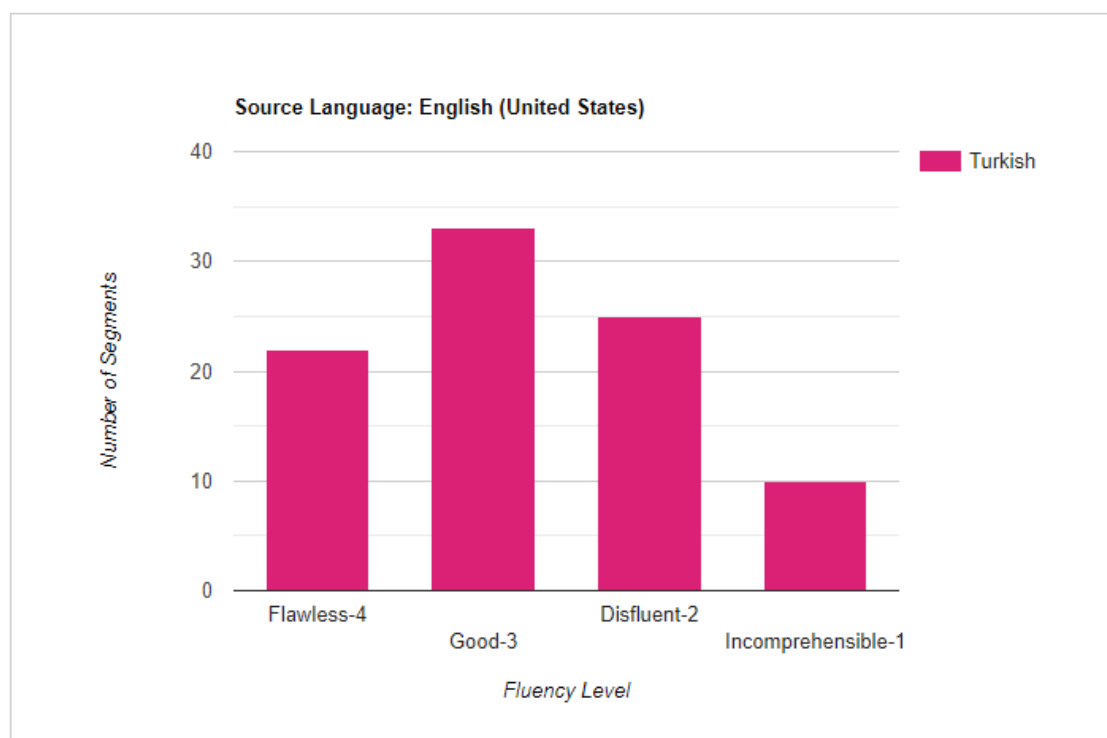


Figure showing “Distribution of Fluency Scores, per Language”

The final figure under this title shows the “Distribution of Fluency Scores, per Evaluator” (Figure 19).

Figure 19: Distribution of Fluency Scores, per Evaluator for Amazon Translate
Distribution of Fluency Scores, per Evaluator

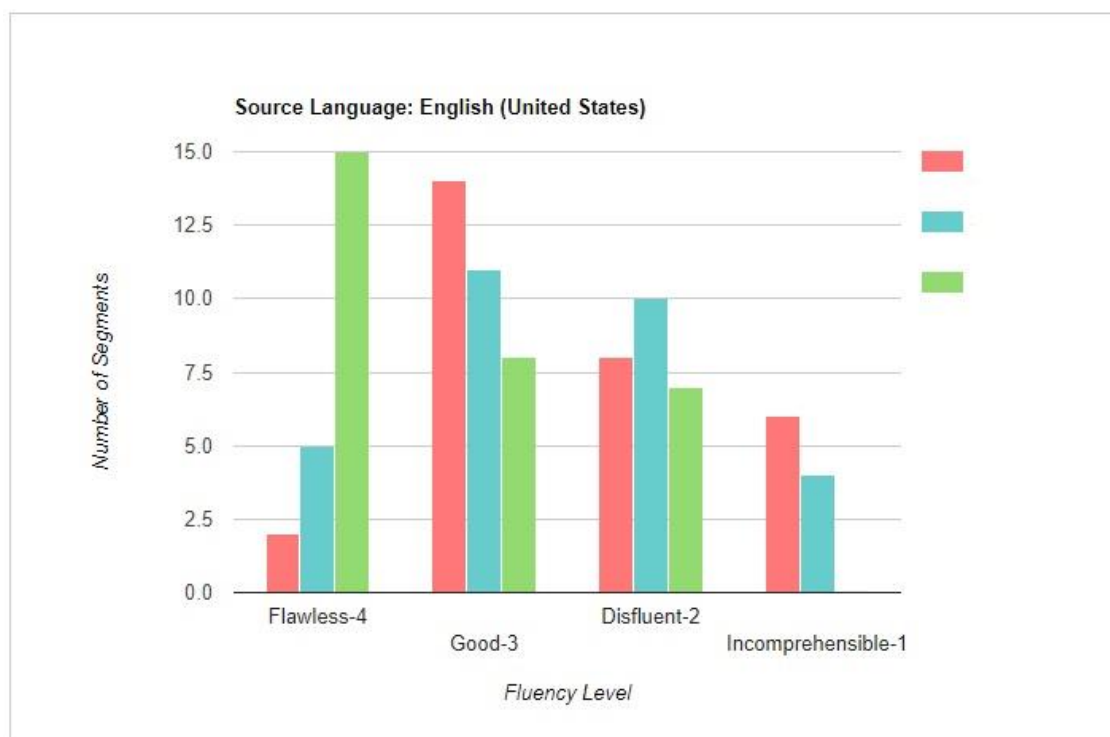


Figure showing “Distribution of Fluency Scores, per Evaluator”

When the “Distribution of Fluency Score, per Evaluator” result is examined, it can be seen that inter-annotator disagreement is much higher when compared with Google Translate. The biggest inter-annotator disagreements are for “Flawless-4” and “Good-3” scores. The evaluator represented in Green gave a score of 4 to 15 segments while the numbers of segments for the same score are 5 and 2 for the Blue and Red evaluators. There are also some significant differences between evaluators for “Good-3” scores. Red evaluator gave a score of 3 to 14 segments. This number was 11 for Blue and 8 for Green, almost half of Red. Another curious result in this distribution is that while Red and Blue evaluators gave a score of “Incomprehensible-1” to a similar number of segments, 6 and 4 respectively, Green evaluator found no segments to be incomprehensible.

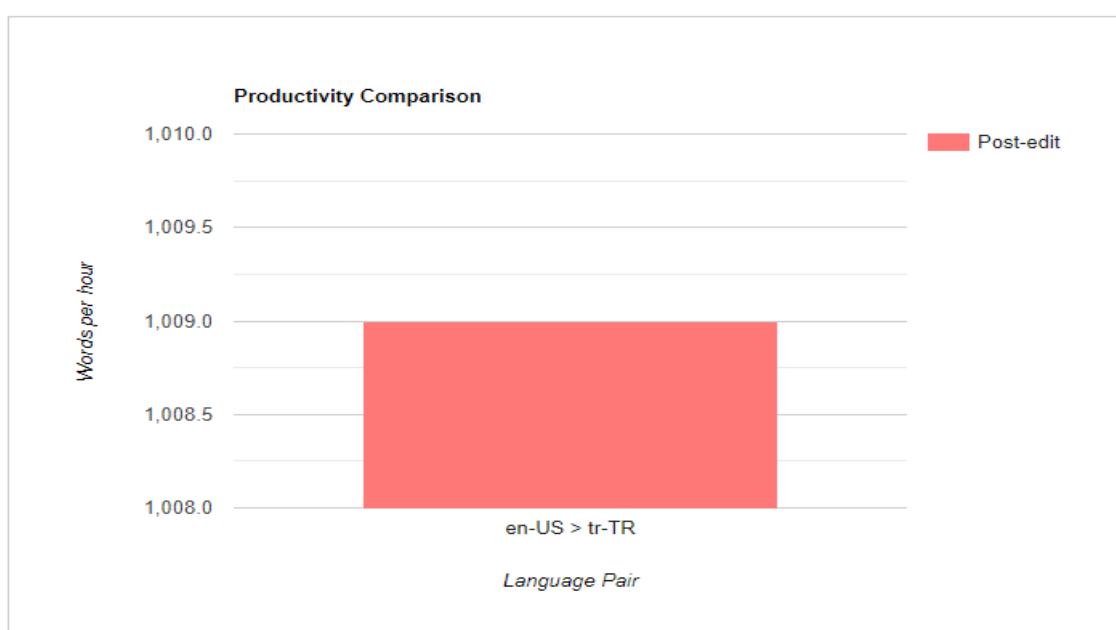
4.6. Productivity report for Amazon Translate

The productivity report for Amazon Translate also starts with information on “Average Productivity” and it provides a very curious result.

Figure 20: Average Productivity Report for Amazon Translate

Average Productivity [\(Info\)](#)

Language Pair	Number of Segments	Number of Words	Post-edit (WPH)	Time Spent for Post-edit (seconds)	MT Engine
English (United States) > Turkish	30	448	1,009	4,932	Not Specified



Note that charts represent an extrapolation of the results to a one hour scale

[Back to Top](#)

Figure showing “Average Productivity” results for Amazon Translate

As it can be seen in the figure above, even though in terms of adequacy and fluency the overall quality of Amazon Translate’s raw machine translation output was lower than Google Translate’s, the words per hour the volunteers post-edited was higher. Meaning that more segments from Google Translate’s raw machine translation output can be used without the need of post-editing, yet, if the text is to be post-edited, post-editors may benefit from Amazon Translate more as they had a higher words per hour score. The next figure is about “Productivity per Editor” and this situation becomes easier to notice.

Regardless of the productivity differences between editors, there is a visible difference between the post-editing speeds of Google Translate’s and Amazon Translate’s raw machine translation outputs. Apart from the average number of words per hour, all three post-editors have significantly increased words per hour scores. The slowest post-editor was able to post-edit 514 words per hour in Google Translate’s result but this number has risen to 808 for Amazon Translate.

However, it should also be kept in mind that, the volunteers first completed the post-editing task for Google Translate’s raw machine translation output and then completed Amazon Translate’s; ergo, the post-editors had already solved the translation problems they were going to encounter when they started post-editing Amazon Translate’s raw machine translation output and remembering how they post-edited the problematic parts, they have not spent as much time as they previously did.

Figure 21: Productivity per Editor for Amazon Translate

Productivity per Editor

Evaluator Name	Post-Edit (WPH)	Time Spent for Post-edit (seconds)
	808	1993
	988	1631
	1,232	1308

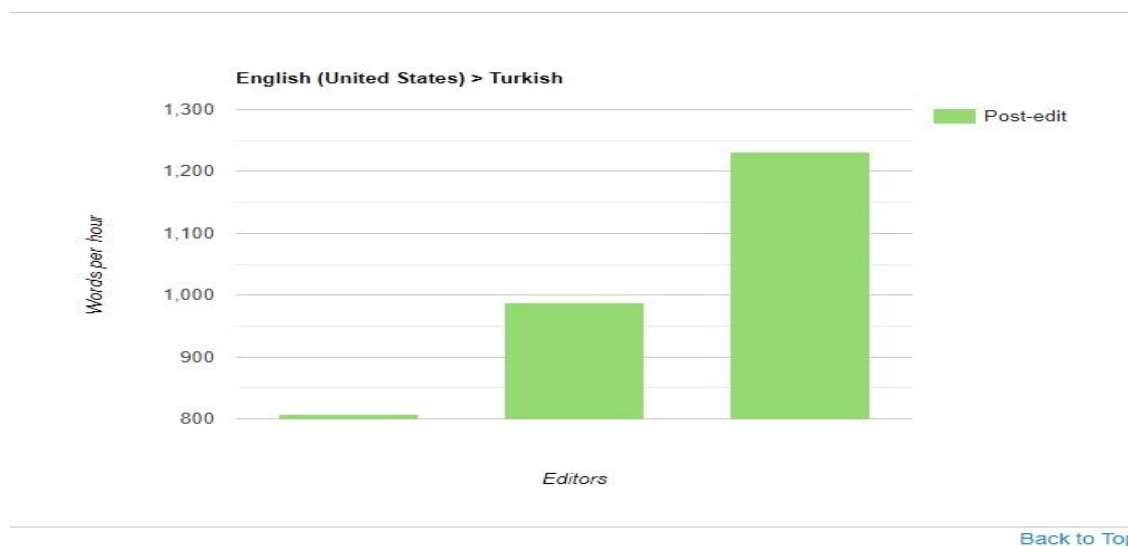


Figure showing “Productivity per Editor”

The next figure is the “Edit Distance Graph”, showing the percentage of total segments and the segment edit distances. The numbers 0 to 10 showed the relative distance of the raw machine translation output to the final post-edited form of the

segments. 0 meant that no or hardly any editing was required and 10 meant that the segments had to be translated from scratch.

Figure 22: Edit Distance Graph for Amazon Translate

Edit Distance Graph [\(Info\)](#)

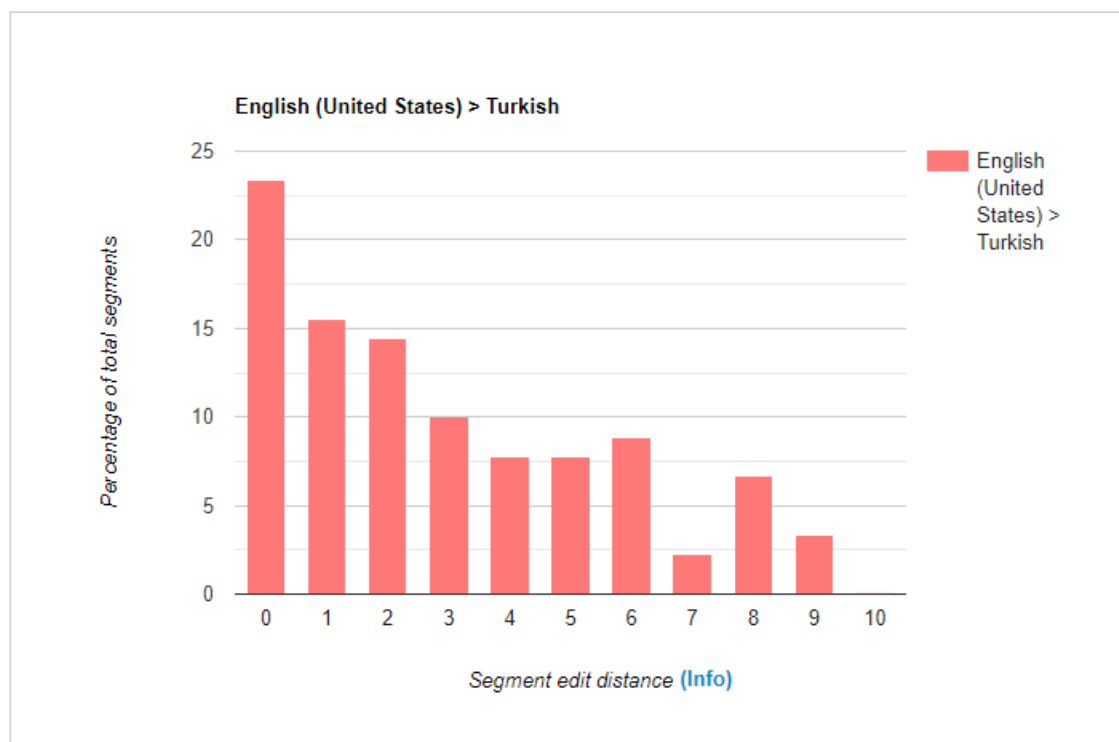


Figure showing “Edit Distance Graph”.

The graph above shows different levels of edit distance and the percentage of segments falling within each edit distance level. According to the Dynamic Quality Framework tool, 23% of segments from Amazon Translate’s raw machine translation output needed no or hardly any editing. For Google Translate, this percentage was 32%. Thus, the percentage rates of all the other edit distance levels are higher in Amazon Translate’s results. The only edit distance result Amazon Translate was better at is 10, meaning that none of the output from Amazon Translate had to be translated from scratch. The percentage for this level of edit distance level was 1% for Google Translate.

Table 5: Table showing “Edit Efforts” per editor for Google Translate

EDIT EFFORT	Segment ID	Red	Blue	Green
	1	0	0	0
	2	20	90	83
	3	6	33	17
	4	58	45	10
	5	83	91	86
	6	38	20	53
	7	47	40	21
	8	48	55	40
	9	0	0	0
	10	0	0	0
	11	25	14	15
	12	0	9	0
	13	38	34	31
	14	9	9	18
	15	6	75	37
	16	59	57	62
	17	18	13	2
	18	33	33	57
	19	5	0	35
	20	41	14	13
	21	26	22	18
	22	50	0	0
	23	64	5	65
	24	20	24	2
	25	28	23	13

Table 5: - Continue

Table showing “Edit Efforts” per editor for Google Translate

EDIT EFFORT	26	0	0	0
	27	0	11	0
	28	77	78	78
	29	57	53	65
	30	34	46	24

Table 5 showing “Edit Efforts” per editor for Amazon Translate. The names of the editors are again given as Red, Blue and Green to protect their anonymity

As it can be seen in “Table 5”, for Amazon Translate, the edit efforts of the post-editors are higher on average when compared with Google Translate. Yet, for some segments, there are still big differences between annotators. The reasons for the edit effort differences are discussed in the next part of the results.

4.7. “Edit Effort” Differences

When the productivity reports for both raw machine translation outputs are analyzed, it was determined that even though there were differences in edit distances and words per hour numbers between the two translations of the source text, it was also determined that there were some big differences between different post-editors for both results. When the edit efforts of different post-editors are analyzed, it can be seen that the main reason for the differences in edit effort are related to the previously mentioned video games localization competences, especially the “creativity” competence and “familiarity with games culture” competences. The reasons are explained in details below.

Looking at the “Edit Effort” table for Google Translate (Table 6), there are eleven segments which Red has not post-edited nothing or hardly anything at all while Blue and/or Green made significant changes. These numbers of these segments are 2, 3, 7, 9, 10, 15, 19, 21, 23, 29 and 30. Four of these segments are analyzed separately and the reason why such differences might have happened are discussed.

4.7.1. Segment 2 Differences

In segment 2, the edit efforts of Blue and Green are very similar. Blue’s edit effort was calculated to be 72 and Green’s was 75. Red’s edit effort was 0. To better see and understand these differences, after the source segments and machine translation

segments are shared, tables showing edit efforts and post-edited target texts per editor are shared for each segment.

Segment 2 source segment: “Greenlander”

Segment 2 machine translation segment: “Grönlandlı”

Table 6: Segment 2 Post-Editing Differences And Edit Efforts For Each Editor

	Post-edit	Edit Effort
Red	Grönlandlı	0
Blue	Yeşilvadili	72
Green	Yeşildiylarlı	75

Table showing segment 2 post-editing differences and edit efforts for each editor.

For segment 2, Red editor did not edit anything, thinking that the machine translation is adequate and that nothing needs to be changed. But Blue and Green were more creative and changed the machine translation to better suit a fantastic fiction context. The “Greenlander” human sub-race has better farming, cooking and laboring skills, which is shared with the players on the character creation page and the next segment is about that. It can be deducted that Blue and Green annotators tried to reflect this by using their “creativity” and “familiarity with games culture” competences. As a result, the difference in edit effort happened.

4.7.2. Segment 3 differences:

In segment 3, Blue editor has the highest edit effort, with a score of 58. Green’s edit effort is a lot lower than Blue’s, 18. Red editor did not edit anything in this segment too.

Segment 3 source segment: “Primarily from farmer cultures, they still tend to be an aggressive race and are easily led astray.”

Segment 3 machine translation segment: “Öncelikle çiftçi kültürlerinden, hala saldırgan bir ırk olma eğilimindedirler ve kolayca yoldan çıkarlar.”

Table 7: Segment 3 Post-Editing Differences And Edit Efforts For Each Editor

	Post-Edit	Edit Effort
Red	Öncelikle çiftçi kültürlerinden, hala saldırgan bir ırk olma eğilimindedirler ve kolayca yoldan çıkarlar.	0
Blue	Çoğunluğu çiftçilik kültüründen gelen bu ırk, saldırgan ve kolayca yoldan çıkarılan bir ırktır.	58
Green	Çoğunlukla çiftçi kültürlerinden gelmekte, hala saldırgan bir ırk olma eğilimindedirler ve kolayca yoldan çıkarılırlar.	18

Table showing segment 3 post-editing differences and edit efforts for each editor

Unlike the previous segment, most of the difference in this segment is caused by fluency changes to the sentence. Even though the raw machine translation output can be understood easily, fluency changes made by Green and especially Blue make it easier to read and understand.

4.7.3. Segment 6 differences:

Segment 6 contains information about one of the human sub-races, the “Scorchlanders”. Unlike the “Greenlanders” in the previous segments, “Scorchlanders” have decreased farming, cooking and laboring skills but they are better at smithing, athletics, dexterity and etc. For this sub-race, all three of the editors made some curious changes.

Segment 6 source segment: “Scorchlanders tend to value personal freedoms above all else.”

Segment 6 machine translation segment: “Scorchlanders, kişisel özgürlüklere her şeyden önce değer verme eğilimindedir.”

Table 8: Segment 6 Post-Editing Differences And Edit Efforts For Each Editor

	Post-Edit	Edit Effort
Red	İsotyalılar, kişisel özgürlüklere her şeyden önce değer verme eğilimindedir.	12
Blue	Alazvadililer, kişisel özgürlüklere her şeyden çok değer verme eğilimindedir.	20
Green	Yanıkdiyarlılar, her şeyden çok kişisel özgürlüklere değer verme eğilimindedir.	53

Table showing segment 6 post-editing differences and edit efforts for each editor

Apart from fluency changes, all three editors preferred a different translation for the sub-race. Red editor claimed that the “Scorchlander” culture was similar to the Scottish culture. In order to make the sub-race sound like “İskoçyalılar” (the Turkish word for the Scottish nation) and to also keep the meaning of “scorch”, the editor decided to post-edit the word as “İsotyalılar”, “isot” being a type of hot spice commonly used in Turkish cuisine, so, Red editor based his decisions on the “creativity” competence. The two other editors preferred editing the word to reflect the in-game attributes of the “Scorchlanders”, relying more on their “familiarity with games culture” competence.

4.7.4. Segment 15 differences:

Like the second segment, segment 15 is also the name of a sub-race within the game and has one of the biggest edit effort differences.

Segment 15 source segment: “Hive Worker Drone”

Segment 15 machine translation segment: “Kovan İşçi Drone”

Table 9: Segment 15 Post-Editing Differences And Edit Efforts For Each Editor

	Post-Edit	Edit Effort
Red	Kovan İşçi Drone	0
Blue	İşçi	75
Green	Kovan İşçi	37

Table showing segment 15 post-editing differences and edit efforts for each editor.

This segment is about a humanoid in-game bug race. These intelligent humanoid bugs are living in hives under the control of queens and each is born with a certain role to fulfill, like soldiers and workers. It can be said that they have a similar system to ants and bees. “Hive Worker Drone” is the worker class of these humanoid bugs and it is one of the sub-races of the Hive race. The editing differences in this segment were about fluency and “familiarity with games culture” competence. Google Translate kept the “Drone” word in the translation, which does not have a meaning in Turkish so, it causes a fluency problem in the sentence. Thinking that the word drone could have a distinctive meaning in the game, Red decided to keep it and did not post-edit anything. Green decided to erase the word drone and post-edit the segment as “Kovan İşçi” (Hive

Worker), keeping the word “Kovan” so as to make sure that the players can understand that this is the sub-race of the Hive race and not just any workers that may appear in-game, post-editing this segment in accordance with their “familiarity with games culture”. Blue only kept the word “İşçi” (Worker), thinking that it would be enough to explain the role of the sub-race.

As a result of these analyses, it can be said that the differences in “Edit Efforts” are mostly due to fluency concerns and video game localization competences, especially the “creativity” competence and “familiarity with games culture” competence.



V. CONCLUSION

The aim of this study was to test the following hypothesis: Could machine translation post-editing be used in the localization of the video game *Kenshi*, as an alternative to the crowdsourcing model the developers preferred. For this purpose, 30 segments from the video game's language files were selected. All of the chosen segments are about the races, sub-races and their definitions in the character creation page. The reason for that was to make sure that all the segments were within the same context. The segments were translated using Google Translate and Amazon Translate. In the next step, three volunteers were assigned six different tasks related to the raw machine translation outputs, on TAUS Dynamic Quality Framework tool. The volunteers first error annotated the raw machine translation outputs in terms of adequacy and fluency. After that, they post-edited the raw machine translation outputs in accordance with TAUS "Human Translation Quality" post-editing guidelines, so that their "words per hour" averages and "edit efforts" could be calculated.

According to the results calculated by the TAUS Dynamic Quality Framework tool, both of the raw machine outputs were unfit to be used without any post-editing. Only 32% of Google Translate's raw output can be used without any post-editing. This percentage turned out to be even lower for Amazon Translate, 23%. Despite these results, it was also proved that translators can post-edit at least 514 words per hour at the very least, meaning that using post-editing, the number of words translated by a professional translator per day can be more than doubled, significantly increasing the number of words per day they translate which would also have an effect on lowering translation and localization costs.

Another point of consideration is the importance of competences for video game localization. Post-editors with different levels of competences, especially “Creativity” and “Familiarity with Video Games” competences can visibly affect the post-edit processes, which may result in different levels of post-editing efforts and post-edited sentences very different from each other.

To conclude, this study has shown that even though raw machine translation output cannot be used as an alternative to crowdsourcing, machine translation post-editing can be an alternative. Regardless of the fact that this study is limited to only a certain part of a single video game’s language file, the findings at the end of the study have showed that even non-customized, free to use neural machine translation systems, especially Google Translate, can be very beneficial for video game localization projects in terms of saving time and money in big projects like this one. Thinking about how machine translation systems are being improved with each passing day, it is without a doubt that raw machine translation outputs will improve as well. In this regard, there is need for further research on whether machine translation and machine translation post-editing can be used in video game localization projects.

REFERENCES

- Alimen, Nilüfer, and Senem Öner Bulut. 2020. "Çevirinin Teknolojikleşmesi Bağlamında İnsan Çevirmenin Rollerini Yeniden Düşünmek: Çevirmen Eğitiminde Teknik Metin Yazarlığı ." *Turkish Studies - Language and Literature* 1047-1062.
- ALPAC. 1966. *Language and Machines: Computer in Translation and Linguistics*. Washington: National Academy of Sciences Printing and Publishing Office.
- American National Standards Institute. 1975. *ASCII Graphic Character Set*. ASCII Standards Registration, New York: American National Standards Institute.
- Anastasiou, Dimitra, and Rajat Gupta. 2011. "Comparison of crowdsourcing translation with Machine Translation." *Journal of Information Science* 37 637-659.
- Anastasiou, Dimitra, and Rajat Gupta. 2011. "Crowdsourcing as Human-Machine Translation (HMT)." *Journal of Information Science* XX 1-25.
- Bayerl, Petra Saskia, and Karsten Ingmar Paul. 2011. "What Determines Inter-Coder Agreement in Manual Annotations? A Meta-Analytic Investigation." *Computational Linguistics* 699-725.
- Dietz, Frank. 2006. "Issues in localizing computer games." In *Perspectives on Localization*, by Keiran J. Dunne, 121-134. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Doherty, Stephen, and Sharon O'Brien. 2014. "Assessing the Usability of Raw Machine Translated Output: A User-Centered Study Using Eye Tracking." *International Journal of Human-Computer Interaction* 30 40-51.
- Dostert, Leon. 1957. *Brief History of Machine Translation Research*. Eighth Annual Round Table Meeting on Linguistics and Language Studies, Washington: Georgetown University.
- Estellés, Enrique, and Fernando González. 2012. "Towards an integrated crowdsourcing definition." *Journal of Information Science* 38 189-200.
- European Comission. 2017. *European Master's in Translation Competence Framework 2017*. European Comission.

- Howe, Jeff. n.d. *Crowdsourcing*. <https://crowdsourcing.typepad.com>.
- . 2006. *The Rise of Crowdsourcing*. June 1. <https://www.wired.com/2006/06/crowds/>.
- Hutchins, William James. 2014. "ACL Anthology." *ACL Anthology Web site*. Accessed 03 11, 2022. <https://aclanthology.org/www.mt-archive.info/10/Hutchins-2014.pdf>.
- Hutchins, William John. 1994. "Machine Translation: History and General Principles." In *The Encyclopedia of Languages and Linguistics*, by Ronal E. Asher, 2322-2332. Oxford: Pergamon Press.
- Jia, Yanfang, Michael Carl, and Xiangling Wang. 2019. "How does the post-editing of neural machine translation compare with from-scratch translation? A product and process study." *The Journal of Specialised Translation* 60-86.
- Jiménez-Crespo, Miguel A. 2017. "Translation Crowdsourcing: Research Trends and Perspectives." In *Collaborative Translation From the Renaissance to the Digital Age*, by Anthony Cordingley and Céline Frigau Manning, 192-211. London: Bloomsbury Publishing.
- Kohler, Chris. 2005. *Power-up: How Japanese Video Games Gave the World an Extra Life*. Indianapolis: BradyGAMES Publishing.
- Koponen, Maarit. 2016. "Is machine translation post-editing worth the effort? A survey of research into post-editing and effort." *The Journal of Specialised Translation* 131-148.
2020. *Lo-Fi Games Forum*. June 03. Accessed July 21, 2022. <https://www.lofigames.com/phpBB3/viewtopic.php?f=17&t=13606>.
- Lommel, Arle, Attila Görög, Alan Melby, Hans Uszkoreit, and Maja Popovic. 2015. *QT21 Deliverable 3.1 Harmonised Metric*. Project Deliverable, Quality Translation 21 (QT21).
- Mangiro Hevian, Carmen. 2009. "Video Games Localisation: Posing New Challenges to the Translator." *Perspectives: Studies in Translatology* 306-323.
- Mangiron, Carmen, and Minako O'Hagan. 2006. "Game Localisation: Unleashing Imagination with "Restricted" Translation." *The Journal of Specialised Translation* 10-21.
- Massardo, Isabella, Jaap van der Meer, Sharon O'Brien, Hollowood Fred, Nora Aranberri, and Katrin Drescher. 2016. *MT Post-Editing Guidelines*. Amsterdam: TAUS Signature Editions.

- Memsource. 2015. *Memsource Blog*. August 15. Accessed June 06, 2022. <https://www.memsource.com/blog/post-editing-in-memsource/>.
- Mikkelsen, Natalie. 2018. "LOCALISATION: DID IT PAY?" *Indie Game Writing*. December 28. Accessed 04 24, 2022. <https://indiegamewriting.com/videogame-localisation-post-mortem/>.
- O'Hagan, Minako. 2011. "Community Translation: Translation as a social activity and its possible consequences in the advent of Web 2.0 and beyond." *Linguistica Antverpiensia* 11-23.
- O'Hagan, Minako. 2009. "Evolution of User-generated Translation: Fansubs, Translation Hacking and Crowdsourcing." *The Journal of Internationalization and Localization* 94-121.
- O'Hagan, Minako, and Carmen Mangiron. 2013. *Game Localization*. Amsterdam: John Benjamins Publishing.
- Öner, Işın, and Senem Öner Bulut. 2021. "Post-Editing Oriented Human Quality Evaluation of Neural Machine Translation in Translator Training: A Study on Perceived Difficulties and Benefits." *Translogos* 100-124.
- Richter, Felix. 2020. *Statista*. September 22. Accessed June 04, 2022. <https://www.statista.com/chart/22392/global-revenue-of-selected-entertainment-industry-sectors/>.
- Robert, Anne-Marie. 2013. "Vous avez dit post-éditrice? Quelques éléments d'un parcours personnel." *The Journal of Specialized Translation* 29-40.
- Stymne, Sara, and Lars Ahrenberg. 2012. "On the practice of error analysis for machine translation evaluation." *Proceedings of the Eighth International Conference on Language Resources and Evaluation*. Istanbul: European Language Resources Association. 1785-1790.
- TAUS. 2017. "Quality Evaluation using Adequacy and Fluency Approaches." *TAUS the language data network*. Accessed April 27, 2022. <https://info.taus.net/quality-evaluation-using-adequacy-and-fluency-approaches>.
- . 2014. *TAUS Dynamic Quality Framework Tool*. Accessed June 06, 2022. <https://dqf.taus.net>.
- . 2015. *TAUS the language data network*. July 31. Accessed April 17, 2022. <https://www.taus.net/qt21-project#dqf-qt21>.
- Turovsky, Barak. 2017. *Google The Keyword*. March 6. Accessed 03 27, 2022. <https://www.blog.google/products/translate/higher-quality-neural-translations-bunch-more-languages/>.

Wijman, Tom. 2020. *Newzoo*. May 8. Accessed June 04, 2022.
<https://newzoo.com/insights/articles/newzoo-games-market-numbers-revenues-and-audience-2020-2023/>.

Yanar Torbalı, Öznur. 2019. "“Kitle kaynaklı çeviri” [“crowdsourced translation”] ile çeviriye ilişkin." *RumeliDE Dil ve Edebiyat Araştırmaları Dergisi* 413-424.



APPENDICES

Appendix A

Harmonized DQF-MQM Error Typology

ID	High-level error type	Granular error type	Definition	Example
1	Accuracy		The target text does not accurately reflect the source text, allowing for any differences authorized by specifications.	Translating the Italian word 'canali' into English as 'canals' instead of 'channels'.
11		Addition	The target text includes text not present in the source.	A translation includes portions of another translation that were inadvertently pasted into the document.
12		Omission	Content is missing from the translation that is present in the source.	A paragraph present in the source is missing in the translation.
13		Mistranslation	The target content does not accurately represent the source content.	A source text states that a medicine should not be administered in doses greater than 200 mg, but the translation states that it should be administered in doses greater than 200 mg (i.e., negation has been omitted).
14		Over-Translation	The target text is more specific than the source text.	The source text refers to a boy but is translated with a word that applies only to young boys rather than the more general term.

15		Under-Translation	The target text is less specific than the source text.	The source text uses words that refer to a specific type of military officer but the target text refers to military officers in general.
16		Untranslated	Content that should have been translated has been left untranslated.	A sentence in a Japanese document translated into English is left in Japanese.
17		Improper exact TM match	An translation is provided as an exact match from a translation memory (TM) system but is actually incorrect.	A TM system returns Press the Start button as an exact (100%) match when the proper translation should be Press the Begin button.
2	Fluency		Issues related to the form or content of a text, irrespective as to whether it is a translation or not.	A text has errors in it that prevent it from being understood.
21		Punctuation	is used incorrectly (for the locale or style).	An English text uses a semicolon where a comma should be used.
22		Spelling	Issues related to spelling of words.	The German word Zustellung is spelled Zustetlugn.
23		Grammar	Issues related to the grammar or syntax of the text, other than spelling and orthography.	An English text reads The man was seeing the his wife.
24		Grammatical Register	The content uses the wrong grammatical register, such as using informal pronouns or verb forms when their formal counterparts are required.	A text used for a highly formal announcement uses the Norwegian du form instead of the expected De.
25		Inconsistency	The text shows internal inconsistency.	A text uses both app. and approx. for approximately.

26		Link/cross-reference	Links are inconsistent in the text.	An HTML file contains numerous links to other HTML files; some have been updated to reflect the appropriate language version while some point to the source language version.
27		Character encoding	Characters are garbled due to incorrect application of an encoding.	A text document in UTF-8 encoding is opened as ISO Latin-1, resulting in all upper ASCII characters being garbled.
3	Terminology		A term (domain-specific word) is translated with a term other than the one expected for the domain or otherwise specified.	A French text translates English e-mail as e-mail but terminology guidelines mandated that courriel be used. The English musicological term dog is translated (literally) into German as Hund instead of as Schnarre, as specified in a terminology database.
31		Inconsistent with termbase	A term is used inconsistently with a specified termbase.	A termbase specifies that the term USB memory stick should be used, but the text uses USB flash drive.
32		Inconsistent use of terminology	Terminology is used in an inconsistent manner within the text.	The text refers to a component as the brake release lever, brake disengagement lever, manual brake release, and manual disengagement release.
4	Style		The text has stylistic problems.	The translation of a light-hearted and humorous advertising campaign is in a serious and “heavy” style even though specifications said it should match the style of the source text.

41		Awkward	A text is written with an awkward style.	A text is written with many embedded clauses and an excessively wordy style. While the meaning can be understood, the text is very awkward and difficult to follow.
42		Company style	The text violates company/organization-specific style guidelines.	Company style states that passive sentences may not be used but the text uses passive sentences.
43		Inconsistent style	Style is inconsistent within a text.	One part of a text is written in a light and terse style while other sections are written in a more wordy style.
44		Third-party style	The text violates a third-party style guide.	Specifications stated that English text was to be formatted according to the Chicago Manual of Style, but the text delivered followed the American Psychological Association style guide.
45		Unidiomatic	The content is grammatical, but not idiomatic.	The following text appears in an English translation of a German letter: “We thanked him with heart” where “with heart” is an understandable, but non-idiomatic rendering, better stated as “heartily”.
5	Design		There is a problem relating to design aspects (vs. linguistic aspects) of the content.	A document is formatted incorrectly.
51		Length	There is a significant discrepancy between the source and the target text lengths.	An English sentence is 253 characters long but its German translation is 51 characters long.

52		Local formatting	Issues related to local formatting (rather than to overall layout concerns).	A portion of the text displays a (non-systematic) formatting problem (e.g., a single heading is formatted incorrectly, even though other headings appear properly).
53		Markup	Issues related to <i>markup</i> (codes used to represent structure or formatting of text, also known as <i>tags</i>).	Markup is used incorrectly, resulting in incorrect formatting.
54		Missing text	Existing text is missing in the final laid-out version.	A translation is complete, but during DTP a text box was inadvertently moved off the page and so the translated text does not appear in a rendered PDF version.
55		Truncation/text expansion	Truncation-text-expansion.	The German translation of an English string in a user interface runs off the edge of a dialogue box and cannot be read.
6	Locale convention		The text does not adhere to locale-specific mechanical conventions and violates requirements for the presentation of content in the target locale.	An incorrect format for currency is used for a German text, with a period (.) instead of a comma (,) as a thousands separator. A text translated into Japanese uses Western quote marks to indicate titles rather than the appropriate Japanese quote marks (「 and 」).
61		Address format	Content uses the wrong format for addresses.	An online form translated from English to Hindi requires a house number even though many addresses in India do not include a house number.
62		Date format	A text uses a date format inappropriate for its locale.	An English text has 2012-06-07 instead of the expected 06/07/2012.

63		Currency format	Content uses the wrong format for currency.	A text dealing with business transactions from English into Hindi assumes that all currencies will be expressed in simple units, while the convention in India is to give such prices in lakh rupees (100,000 rupees).
64		Measurement format	A text uses a measurement format inappropriate for its locale.	A text in France uses feet and inches and Fahrenheit temperatures.
65		Shortcut key	A translated software product uses shortcuts that do not conform to locale expectations or that make no sense for the locale.	A software product uses CTRL-S to save a file in Hungarian, rather than the appropriate CTRL-M (for <i>menteni</i>).
66		Telephone format	Content uses the wrong form for telephone numbers.	A German text presents a telephone number in the format (xxx) xxx - xxxx instead of the expected 0xx followed by a group of digits separated into groups by spaces.
7	Verity		The text makes statements that contradict the world of the text.	The text states that a feature is present on a certain model of automobile when in fact it is not available.
71		Culture-specific reference	Content inappropriately uses a culture-specific reference that will not be understandable to the intended audience.	An English text refers to steps in a process as <i>First base</i> , <i>Second base</i> , and <i>Third base</i> , and to successful completion as a <i>Home run</i> and uses other metaphors from baseball. These prove difficult to translate and confuse the target audience in Germany.
8	Other		Any other issues.	

Source: <https://www.taus.net/qt21-project#harmonized-error-typology>

Appendix B

Severity Levels for Error Categories

Definition	Description
Critical	Errors that may carry health, safety, legal or financial implications, violate geopolitical usage guidelines, damage the organization's reputation, cause the application to crash or negatively modify/misrepresent the functionality of a product or service, or which could be seen as offensive.
Major	Errors that may confuse or mislead the user or hinder proper use of the product/service due to significant change in meaning or because errors appear in a visible or important part of the content.
Minor	Errors that don't lead to loss of meaning and wouldn't confuse or mislead the user but would be noticed, would decrease stylistic quality, fluency or clarity, or would make the content less appealing.
Neutral	Used to log additional information, problems or changes to be made that don't count as errors, e.g. they reflect a reviewer's choice or preferred style, they are repeated errors or instruction/glossary changes not yet implemented, a change to be made that the translator is not aware of.
Kudos	Used to praise for exceptional achievement.

Source: <https://www.taus.net/qt21-project#harmonized-error-typology>

Appendix C

Source Text

Segment ID	Source Segment
1	Human
2	Greenlander
3	Primarily from farmer cultures, they still tend to be an aggressive race and are easily led astray.
4	The worst of them live in ignorance, knowing only greed and fear, reacting to anything different with hatred and violence, but the best of them are some of the greatest civilised scientists, engineers and warriors in the land.
5	Scorchlander
6	Scorchlanders tend to value personal freedoms above all else.
7	They don't tend to get along well with rules, regulations and religions, and as a result have a reputation as social misfits and are often found in more adventurous professions. Despite this however, they are highly creative making them natural born traders and skilled weapon smiths.
8	Honest to a fault, laughing in the face of manners, moderation and anything sensible, they make loyal friends, passionate enemies, and are great fun to drink with.
9	Hive
10	Hive Prince
11	The highest ranking in a Hive and right-hands to the Queen.
12	Nobody has ever seen an actual Queen, it is the Princes that run everything.
13	Free Princes are usually the result of losing command to a rival Prince, or exiles sent away as punishment for some kind of failure.
14	They are much more intelligent and free-willed than normal Hive types.
15	Hive Worker Drone
16	Expendible workers born to labour for the good of their Hive, they know only obedience and value service to the Hive above any personal benefits, safety or comfort.
17	If separated from the controlling influence of the Hive they usually die, but in rare cases become free individuals.
18	Bewildered by their newfound freedom and loss of a predetermined purpose they wander lost and seek meaning, but usually just find death or exploitation.

19	Hive Soldier Drone
20	Solder drones created to protect their Hive and enforce control over the workers.
21	Stronger and less intelligent than a worker, freedom-seekers are rarer and usually end up finding mercenary work.
22	Shek
23	A warrior society that values strength and hardiness above more intellectual pursuits.
24	They are sometimes a little humourless and have a reputation as dumb brutes, but this is not always the case.
25	Their external skeletal plates provide a natural armour but hinders more dexterous pursuits like sneaking around.
26	Skeleton
27	The Skeletons are a complete mystery, nobody knows where they came from and how they were made but they are suspected to be thousands of years old.
28	Fully sentient and capable of feeling sadness, anger, excitement, compassion, thrill and enjoyment, although they don't possess the ability to express anything visually.
29	This combined with their dry sense of humour often confuses the other races, whom the Skeletons see as somewhat volatile and unpredictable.
30	They have no prejudice or fear of death, and make for brutally effective warriors.

Appendix D

Google Translate Raw Machine Translation Output

Segment ID	Google Translate
1	İnsan
2	Grönlandlı
3	Öncelikle çiftçi kültürlerinden, hala saldırgan bir ırk olma eğilimindedirler ve kolayca yoldan çıkarlar.
4	En kötüsü cehalet içinde yaşar, sadece hırs ve korkuyu bilir, farklı olan her şeye nefret ve şiddetle tepki verir, ama en iyileri ülkenin en büyük medeni bilim adamlarından, mühendislerinden ve savaşçılarından bazılarıdır.
5	Kavurucu
6	Scorchlanders, kişisel özgürlüklere her şeyden önce değer verme eğilimindedir.
7	Kurallar, düzenlemeler ve dinlerle iyi geçinme eğilimi göstermezler ve sonuç olarak sosyal uyumsuzlar olarak ün kazanırlar ve genellikle daha maceracı mesleklerde bulunurlar. Ancak buna rağmen, onları doğuştan tüccarlar ve yetenekli silah ustaları yapan son derece yaratıcıdır.
8	Bir hataya kadar dürüst, görgü, ılımlılık ve mantıklı olan her şeye gülerek, sadık arkadaşlar, tutkulu düşmanlar edinirler ve birlikte içmek çok eğlencelidir.
9	Kovan
10	Kovan Prensi
11	Bir Kovadaki en yüksek rütbe ve Kraliçe'nin sağ eli.
12	Hiç kimse gerçek bir Kraliçe görmedi, her şeyi yöneten Prenslerdir.
13	Özgür Prenslar genellikle komutasını rakip bir Prens'e kaptırmanın veya bir tür başarısızlık için ceza olarak gönderilen sürgünlerin sonucudur.
14	Normal Kovan türlerinden çok daha zeki ve özgür iradelidirler.
15	Kovan İşçi Drone
16	Kovanlarının iyiliği için çalışmak için doğmuş harcanabilir işçiler, yalnızca itaati bilirler ve Kovana hizmet etmeye herhangi bir kişisel çıkar, güvenlik veya rahatlıktan daha fazla değer verirler.
17	Kovanın kontrol edici etkisinden ayrılırlarsa genellikle ölürlər, ancak nadir durumlarda özgür bireyler haline gelirler.

18	Yeni keşfettikleri özgürlükleri ve önceden belirlenmiş bir amacı yitirmeleri karşısında şaşkına dönmüş olarak, kayıp bir şekilde dolaşıp anlam ararlar, ancak genellikle sadece ölüm ya da sömürü bulurlar.
19	Kovan Asker Drone
20	Kovanlarını korumak ve işçiler üzerinde kontrol sağlamak için yaratılan lehim uçağı.
21	Bir işçiden daha güçlü ve daha az zeki olan özgürlük arayanlar daha nadirdir ve genellikle sonunda paralı bir iş bulurlar.
22	Şek
23	Daha fazla entelektüel uğraşın üzerinde güce ve dayanıklılığa değer veren savaşçı bir toplum.
24	Bazen biraz mizahsızdırlar ve aptal vahşiler olarak ün yaparlar, ancak bu her zaman böyle değildir.
25	Dış iskelet plakaları doğal bir zırh sağlar ancak gizlice dolaşmak gibi daha hünarlı arayışları engeller.
26	iskelet
27	İskeletler tam bir muamma, kimse nereden geldiklerini ve nasıl yapıldıklarını bilmiyor ama binlerce yıllık olduklarından şüpheleniliyor.
28	Görsel olarak herhangi bir şey ifade etme yeteneğine sahip olmasalar da, tamamen duyarlı ve üzüntü, öfke, heyecan, şefkat, heyecan ve zevk hissetme yeteneğine sahiptirler.
29	Bu, onların kuru mizah anlayışıyla birleştiğinde, genellikle İskeletlerin biraz değişken ve öngörülemez olarak gördüğü diğer ırkların kafasını karıştırır.
30	Önyargıları veya ölüm korkuları yoktur ve vahşice etkili savaşçılar olurlar.

Appendix E

Amazon Translate Raw Machine Translation Output

Segment ID	Amazon Translate
1	İnsan
2	Grönlander
3	Öncelikle çiftçi kültürlerinden, hala agresif bir ırk olma eğilimindedirler ve kolayca yoldan çıkarılırlar.
4	En kötüsü cehalet içinde yaşıyor, sadece açgözlülük ve korkuyu bilmek, nefret ve şiddetle farklı bir şeye tepki vermek, ama en iyisi en büyük medeni bilim adamlarından bazıları, ülkedeki mühendisler ve savaşçılar.
5	Scorchlander
6	Scorchlanders kişisel özgürlüklere her şeyden önce değer verme eğilimindedir.
7	Kurallarla iyi geçinme eğiliminde değiller, düzenlemeler ve dinler, ve sonuç olarak sosyal uyumsuzlar olarak bir üne sahiptir ve genellikle daha maceracı mesleklerde bulunurlar. Buna rağmen, son derece yaratıcıdır ve onları doğal doğuştan tüccarlar ve yetenekli silah demircileri yaparlar.
8	Bir hataya karşı dürüst, görgü kuralları karşısında gülmek, ılımlılık ve mantıklı her şey, sadık arkadaşlar edinirler, tutkulu düşmanlar, ve birlikte içmek çok eğlencelidir.
9	Kovan
10	Kovan Prensi
11	Bir Kovan"da en yüksek rütbeli ve Kraliçe"ye sağ eller.
12	Hiç kimse gerçek bir Kraliçe görmedi, her şeyi yöneten Prenslerdir.
13	Özgür Prenslar genellikle rakip bir Prens"e emir kaybetmenin sonucudur., veya bir tür başarısızlık için ceza olarak gönderilen sürgünlerin.
14	Normal Kovan türlerinden çok daha zeki ve özgür iradeli.
15	Kovan İşçi Drone
16	Kovanlarının iyiliği için emek için doğan harcanan işçiler, Kovan"a kişisel menfaatlerin üzerinde yalnızca itaat ve değer hizmetini bilirler, güvenlik veya konfor.

17	Kovanın kontrol edici etkisinden ayrılırsa genellikle ölürlür, ancak nadir durumlarda özgür bireyler olurlar.
18	Yeni keşfettikleri özgürlükler ve önceden belirlenmiş bir amacı kaybettikleri için şaşkına, kayboldular ve anlam arıyorlar, ama genellikle sadece ölüm veya sömürü bulur.
19	Kovan Asker Drone
20	Kovanlarını korumak ve işçiler üzerinde kontrolü sağlamak için oluşturulan lehim dronları.
21	Bir işçiden daha güçlü ve daha az zeki, özgürlük arayanlar daha nadirdir ve genellikle paralı asker işi bulurlar.
22	Shek
23	Güce ve dayanıklılığa daha entelektüel arayışların üzerinde değer veren bir savaşçı toplum.
24	Bazen biraz mizahsızdırlar ve aptalca vahşiler olarak ün yaparlar., ama bu her zaman böyle değildir.
25	Dış iskelet plakaları doğal bir zırh sağlar, ancak etrafta gizlice dolaşmak gibi daha hünerli arayışları engeller.
26	İskelet
27	İskeletler tam bir gizem, kimse nereden geldiklerini ve nasıl yapıldıklarını bilmiyor ama binlerce yaşında olduklarından şüpheleniliyor.
28	Tamamen duyarlı ve üzüntü hissedebilir, öfke, heyecan, merhamet, heyecan, heyecan ve keyif, görsel olarak hiçbir şeyi ifade etme yeteneğine sahip olmasalar da.
29	Bu, kuru mizah anlayışlarıyla birleştiğinde genellikle diğer ırkların kafasını karıştırır., İskeletlerin biraz değişken ve öngörülemez gördüğü.
30	Önyargıları veya ölüm korkusu yok, ve acımasızca etkili savaşçılar yapın.

Appendix F

“Red” Editor Google Translate Productivity Report

Segment ID	Post-edited Target	Edit Effort
1	İnsan	0
2	Grönlandlı	0
3	Öncelikle çiftçi kültürlerinden, hala saldırgan bir ırk olma eğilimindedirler ve kolayca yoldan çıkarlar.	0
4	En kötülerini sadece hırs ve korkuyu bilerek ve farklı olan her şeye nefret ve şiddetle tepki vererek cehalet içinde yaşarlar, ama en iyileri ülkenin en büyük medeni bilim adamlarından, mühendislerinden ve savaşçılarından bazılarıdır.	22
5	İsotyalı	100
6	İsotyalılar, kişisel özgürlüklere her şeyden önce değer verme eğilimindedir.	12
7	Kurallar, düzenlemeler ve dinlerle iyi geçinme eğilimi göstermezler ve sonuç olarak sosyal uyumsuzlar olarak ün kazanırlar ve genellikle daha maceracı mesleklerde bulunurlar. Ancak buna 72ontro, doğuştan tüccar ve yetenekli silah ustaları olmalarıyla son derece yaratıcıdır.	6
8	Bir hataya kadar dürüst olmaları; görgü, ılımlılık ve mantıklı olan her şeye gülmeleri ile sadık bir arkadaş ve tutkulu düşmanlar olmalarının yanında birlikte demlenmek için çok eğlencelidirler.	29
9	kovan	0
10	kovan prensi	0
11	Bir Kovadaki en yüksek rütbe ve Kraliçe'nin sağ kolu.	5
12	Hiç kimse gerçek bir Kraliçe görmedi, her şeyi yöneten Prenslerdir.	0
13	Özgür Prenslar genellikle komutasını rakip bir Prens'e kaptırmanın veya bir tür başarısızlık yüzünden ceza olarak gönderilen sürgünlerin sonucunda ortaya çıkmaktadır.	16
14	Normal Kovan türlerinden çok daha zeki ve özgür iradelidirler.	0
15	Kovan İşçi Drone	0
16	Kovanlarının iyiliği için çalışmak için doğmuş fedakar işçiler, yalnızca itaati bilirler ve Kovana hizmet etmeye herhangi bir kişisel çıkar, güvenlik veya rahatlıktan daha fazla değer verirler.	4

17	Kovanın 73ontrol edici etkisinden ayrılırlarsa genellikle ölümler ancak bazen de özgür bireyler haline gelirler.	11
18	Yeni keşfettikleri özgürlükleri ve önceden belirlenmiş bir amacı yitirmeleri karşısında şaşkına dönmüş olarak yitik bir şekilde dolaşıp anlam ararlar ancak genellikle sadece ölüm ya da sömürü bulurlar.	3
19	Kovan Asker Drone	0
20	Asker dronlar Kovanlarını korumak ve işçiler üzerinde 73ontrol sağlamak için yaratılmıştır.	30
21	Bir işçiden daha güçlü ancak daha az zeki olan özgürlük arayanlar daha enderdir ve genellikle sonunda paralı bir iş bulurlar.	6
22	Şek	0
23	Daha fazla entelektüel uğraşın üzerinde güce ve dayanıklılığa değer veren savaşçı bir toplum.	0
24	Bazen biraz mizahsızdırlar ve aptal vahşiler olarak nam salmışlardır ancak bu her zaman geçerli değildir.	16
25	Dış iskelet plakaları doğal bir zırh işlevi görür ancak gizlice dolaşmak gibi daha hünerli işleri engeller.	16
26	iskelet	0
27	İskeletler tam bir muamma; kimse nereden geldiklerini ve nasıl yapıldıklarını bilmiyor ancak binlerce yıllık olduklarından şüpheleniliyor.	2
28	Görsel olarak herhangi bir şey ifade etme yeteneğine sahip olmasalar da üzüntü, öfke, heyecan, şefkat, heyecan ve zevki hissedebilirken bunlara karşı tamamen duyarlıdırlar.	34
29	Bu onların kuru mizah anlayışıyla birleştiğinde genellikle İskeletlerin biraz değişken ve öngörülemez olarak gördüğü diğer ırkların kafasını karıştırır.	1
30	Önyargıları veya ölüm korkuları yoktur ve vahşice etkili savaşçı olma yolundadırlar.	12

Appendix G

“Blue” Editor Google Translate Productivity Report

Segment ID	Post-edited Target	Edit Effort
1	İnsan	0
2	Yeşilvadili	72
3	Çoğunluğu çiftçilik kültüründen gelen bu ırk, saldırgan ve kolayca yoldan çıkarılan bir ırktır.	58
4	En kötülerini cehalet içinde yaşar, sadece hırs ve korkuyu bilir, farklı olan her şeye nefret ve şiddetle tepki verir, ama en iyileri ülkenin en büyük medeni bilim adamlarından, mühendislerinden ve savaşçılarından bazılarıdır.	1
5	Alazvadili	80
6	Alazvadililer, kişisel özgürlüklere her şeyden çok değer verme eğilimindedir.	20
7	Kurallar, düzenlemeler ve dinlerle araları iyi değildir ve sonuç olarak uyumsuzlar olarak ün kazanmışlardır ve genellikle daha maceracı mesleklere sahiptirler. Ancak buna 74ontro son derecede yaratıcıdır ve tüccarlığa, silah ustalığına doğuştan yatkınlıkları vardır.	38
8	Fazlasıyla dürüst olan; görgü, ılımlılık ve ölçülülüğe uzak olan bu ırk, sadık arkadaşlar, tehlikeli düşmanlarla çevrilidir ve onlarla içmek çok eğlencelidir.	41
9	Kovan	20
10	Kovan Prensi	16
11	Bir Kovadaki en yüksek rütbeli ve Kraliçe'nin sağ eli.	3
12	Hiç kimse gerçek bir Kraliçe görmemiştir. Her şeyi yöneten Prenslerdir.	9
13	Özgür Prensler, genellikle komutasını rakip bir Prens'e kaptırmış veya bir tür başarısızlık nedeniyle ceza olarak sürgüne gönderilmiş Prenslerdir.	26
14	Normal Kovan türlerinden çok daha zeki ve özgür iradelidirler.	0
15	İşçi	75

16	Kovanlarının iyiliği için çalışmak için doğmuş gözden çıkarılabilir işçiler, yalnızca itaat etmeyi bilirler ve Kovan'a hizmet etmeye herhangi bir kişisel çıkar, güvenlik veya rahatlıktan daha fazla değer verirler.	9
17	Kovanın kontrolcü ortamından ayrılırlarsa genellikle ölürlər, ancak nadir durumlarda özgür bireyler haline gelebilirler.	13
18	Yeni keşfettikleri özgürlükleri ve önceden belirlenmiş amaçlarını yitirmeleri karşısında şaşkına dönmüş olarak, kayıp bir şekilde dolaşıp anlam ararlar, ancak genellikle sadece ölüm ya da sömürü bulurlar.	3
19	Asker	70
20	Kovanlarını korumak ve işçiler üzerinde kontrol sağlamak için yaratılan birlikler.	11
21	Bir işçiden daha güçlü ve daha akılsız oldukları için özgürlük arayışına çıkan sayısı daha azdır ve genellikle sonunda paralı asker olurlar.	31
22	Şek	0
23	Entelektüel uğraştan ziyade güce ve dayanıklılığa değer veren savaşçı bir toplum.	20
24	Çok şaka kaldırmayan Şekler, aptal vahşiler olarak ün salmışlardır, ancak bu her zaman böyle değildir.	33
25	Dış iskelet plakaları doğal bir zırh görevi görür ancak gizlice dolaşmak gibi daha çok beceri isteyen eylemlere engel olur.	29
26	İskelet	14
27	İskeletler tam bir muammadır. Kimse nereden geldiklerini ve nasıl ortaya çıktıklarını bilmez ama binlerce yıllık olduklarından şüphelenilmektedir.	16
28	Görsel olarak herhangi bir şey ifade etme yeteneğine sahip olmasalar da, duyguları vardır ve üzüntü, öfke, heyecan, şefkat, korku ve zevk hissederler.	23
29	İnce mizah anlayışlarıyla da birleştğinde bu durum İskeletlerin biraz değişken ve öngörülemez olarak gördüğü diğer ırkların kafasını karıştırır.	20
30	Ölüme karşı önyargıları ve korkuları yoktur. Vahşi, etkili savaşçılardır.	41

Appendix H

“Green” Editor Google Translate Productivity Report

Segment ID	Post-edited Target	Edit Effort
1	İnsan	0
2	Yeşildiylarlı	75
3	Çoğunlukla çiftçi kültürlerinden gelmekte, hala saldırgan bir ırk olma eğilimindedirler ve kolayca yoldan çıkarılırlar.	18
4	En kötü durumdakileri cehalet içinde yaşar, sadece hırs ve korkuyu bilir, farklı olan her şeye nefret ve şiddetle tepki verirler, ama en iyileri de ülkedeki en medeni bilim adamlarından, mühendislerinden ve savaşçılarından bazılarıdır.	12
5	Yanıkdiylarlı	91
6	Yanıkdiylarlılar, her şeyden çok kişisel özgürlüklere değer verme eğilimindedir.	53
7	Kurallar, düzenlemeler ve dinlerle iyi geçinme eğilimi göstermezler ve sonuç olarak sosyal uyumsuzlar olarak ün kazanırlar ve genellikle daha maceracı meslekleri tercih ederler. Ancak buna 76ontro, son derece yaratıcılardır ki bu da onların doğuştan tüccar olmalarını ve yetenekli silah ustası olmalarını sağlar.	26
8	Dürüstlük abideleri olup, görgü kuralları, ılımlılık ve mantıklı her şeye karşı alaycı yaklaşırlar. Sadık arkadaşlıklar, hiddetli düşmanlıklar kurarlar ve onlarla birlikte içmek çok eğlencelidir.	44
9	Kovan	20
10	Kovan Prensi	16
11	Bir Kovadaki en yüksek rütbe ve Kraliçe’nin sağ eli.	0
12	Hiç kimse gerçek bir Kraliçe görmedi, her şeyi yöneten Prenslerdir.	1
13	Bir Prens, komutasını rakip bir Prens’e kaptırınca ya da bir başarısızlık sebebiyle sürgüne gönderilince Özgür Prens’e dönüşürler.	46
14	Normal Kovan ırklarından çok daha zeki ve özgür iradelidirler.	9
15	Kovan İşçi	37
16	Kovanlarının iyiliği için çalışmak amacıyla doğmuş, feda edilebilir işçiler. Yalnızca itaat etmeyi bilirler ve şahsi menfaat, güvenlik ya da konfordan daha çok Kovana hizmet etmeye değer verirler.	47

17	Kovanın 77ontrol edici etkisinden ayrılırlarsa genellikle ölürlr, ancak nadir durumlarda özgür bireyler haline gelirler.	0
18	Yeni keşfettikleri özgürlükleri ve önceden belirlenmiş bir amaçlarını yitirmeleri karşısında şaşkına dönmüş olarak, kayıp bir şekilde dolaşıp anlam ararlar, ancak genellikle sadece ölüm ya da sömürü bulurlar.	2
19	Kovan Asker	35
20	Kovanlarını korumak ve işçiler üzerinde 77ontrol sağlamak için yaratılan askerler.	13
21	İşçilerden daha güçlü ama daha az zeki olup, aralarında özgürlük arayışında olanların sayısı da daha azdır ve genellikle paralı asker olurlar.	44
22	Şek	0
23	Entellektüel uğraşlardan daha çok güce ve dayanıklılığa önem veren savaşçı bir toplum.	32
24	Bazen biraz mizahsızdırlar ve aptal vahşiler olarak ünleri yayılmıştır, ancak bu her zaman böyle değildir.	11
25	Dış iskelet plakaları doğal bir zırh sağlar ancak gizlice dolaşmak gibi maharet gerektiren işlere mani olmaktadır.	24
26	İskelet	14
27	İskeletler tam bir muamma, kimse nereden geldiklerini ve nasıl yapıldıklarını bilmiyor ama binlerce yıllık olduklarından şüpheleniliyor.	0
28	Görsel olarak herhangi bir şey ifade etme yeteneğine sahip olmasalar da, hislere sahiptirler ve üzüntü, öfke, heyecan, şefkat ve zevk hissedebilirler.	25
29	Bu durum, ince espi anlaşılarıyla birleştğinde, genellikle İskeletlerin biraz değişken ve öngörülemez olarak gördüğü diğer ırkların kafasını karıştırır.	20
30	Önyargıları yoktur ve ölmekten korkmazlar. Bu da onların vahşi ve etkili savaşçılar olmalarını sağlar.	44

Appendix I

“Red” Editor Amazon Translate Productivity Report

Segment ID	Post-edited Target	Edit Effort
1	İnsan	0
2	Grönlandlı	20
3	Öncelikle çiftçi kültürlerinden dolayı hala agresif bir ırk olma eğilimindedirler ve kolayca yoldan çıkarılırlar.	6
4	En kötüler sadece hırs ve korkuyu bilmeleri ve farklı olan her şeye nefret ve şiddetle tepki vermeleri ile cehalet içinde yaşarlar, ama en iyileri ülkenin en büyük medeni bilim insanlarından, mühendislerinden ve savaşçılarından bazılarıdır.	58
5	İsotyalı	83
6	İsotyalılar kişisel özgürlüklerini her şeyden üstün tutma eğilimindedirler.	38
7	Kurallar, düzenlemeler ve dinlerle iyi geçinme eğilimi göstermezler ve sonuç olarak sosyal uyumsuzlar olarak nam salarlar ve genellikle daha maceracı mesleklerde bulunurlar. Ancak buna rağmen doğuştan tüccar ve yetenekli silah ustaları olmalarıyla son derece yaratıcıdır.	47
8	Bir hataya karşı dürüst olmaları; görgü, ılımlılık ve mantıklı olan her şeye gülmeleri ile sadık bir arkadaş ve tutkulu düşmanlar olmalarının yanında birlikte demlenmek için çok eğlencelidirler.	48
9	Kovan	0
10	Kovan Prensi	0
11	Bir Kovan'daki en yüksek rütbe ve Kraliçe'nin sağ kolu.	25
12	Hiç kimse gerçek bir Kraliçe görmedi, her şeyi yöneten Prenslerdir.	0
13	Özgür Prensler genellikle komutasını rakip bir Prens'e kaptırmış veya bir tür başarısızlık yüzünden ceza olarak gönderilen sürgün gönderilmişlerdir.	38
14	Normal Kovan türlerinden çok daha zeki ve özgür iradelidirler.	9
15	Kovan İşçi Dron	6
16	Kovanlarının iyiliği için çalışmak üzere doğmuş fedakar işçiler, yalnızca itaati bilirler ve Kovan'a hizmet etmeye herhangi bir kişisel çıkar, güvenlik veya rahatlıktan daha fazla değer verirler.	59
17	Kovanın kontrol edici etkisinden ayrılırsa genellikle ölürlar ancak nadiren de olsa özgür bireylere dönüşebilirler.	18

18	Yeni keşfettikleri özgürlükleri ve önceden belirlenmiş bir amacı yitirmeleri karşısında şaşkına dönmüş olarak yitik bir şekilde dolaşıp anlam bulmayı umarlar ancak genellikle sadece ölüm ya da sömürü bulurlar.	33
19	Kovan Asker Dron	5
20	Asker dronlar Kovan'larını korumak ve işçiler üzerinde kontrol sağlamak için yaratılmıştır.	41
21	Bir işçiden daha güçlü ve daha az zeki özgürlük arayıcıları sayısı nispeten azdır ve sonları genellikle paralı asker işi olur.	26
22	Şek	50
23	Daha fazla entelektüel uğraşın üzerinde güce ve dayanıklılığa değer veren savaşçı bir toplum.	64
24	Bazen biraz mizahsızdırlar ve aptal vahşiler olarak nam salmışlardır ancak bu her zaman geçerli değildir.	20
25	Dış iskelet plakaları doğal bir zırh işlevi görür ancak gizlice dolaşmak gibi daha marifetli işlere engel olur.	28
26	İskelet	0
27	İskeletler tam bir gizem, kimse nereden geldiklerini ve nasıl yapıldıklarını bilmiyor ama binlerce yaşında olduklarından şüpheleniliyor.	0
28	Görsel olarak herhangi bir şey ifade etme yeteneğine sahip olmasalar da üzüntü, öfke, şefkat, heyecan ve zevki hissedebilirken bunlara karşı tamamen duyarlıdırlar.	77
29	Bu onların kuru mizah anlayışıyla birleştiğinde genellikle İskeletlerin biraz değişken ve öngörülemez olarak gördüğü diğer ırkların kafasını karıştırır.	57
30	Önyargıları veya ölüm korkuları yoktur ve vahşice etkili savaşçı olma yolundadırlar.	34

Appendix K

“Blue” Editor Amazon Translate Productivity Report

Segment ID	Post-edited Target	Edit Effort
1	İnsan	0
2	yeş	90
3	Çoğunlukla çiftçi kültürlerinden olan bu ırk, agresif eğilimdedir ve kolayca yoldan çıkarılır.	33
4	En kötülerini cehalet içinde yaşar, sadece açgözlülük ve korkuyu bilir, her şeye öfke ve şiddetle tepki verir, ama en iyileri bu topraklardaki en büyük bilim adamları, mühendisler ve savaşçılardan bazılarıdır.	45
5	Alazvadili	91
6	Alazvadililer kişisel özgürlüklere her şeyden çok değer verme eğilimindedir.	20
7	Kurallar, düzenlemeler ve dinlerle araları yoktur ve sonuç olarak uyumsuzlar olarak bilinirler. Genellikle daha heyecan verici meslekleri vardır. Buna rağmen, son derece yaratıcıdır ve bu da onları doğuştan tüccarlar ve demir ustaları yapar.	40
8	Fazlasıyla dürüst olan bu milletin, görgü kuralları, ılımlılık ve ölçülülükle işi olmaz. Arkadaşları sadık, düşmanlıkları büyüktür ve onlarla içmek çok eğlencelidir.	55
9	Kovan	0
10	Kovan Prensi	0
11	Bir Kovan'da en yüksek rütbeli ve Kraliçe'nin sağ eli.	14
12	Hiç kimse gerçek bir Kraliçe görmemiştir. Her şeyi yöneten Prenslerdir.	9
13	Özgür Prenslar genellikle rakip bir Prens'e rütbe kaybetmiş veya bir tür başarısızlık nedeniyle ceza olarak sürgüne gönderilmiş Prenslerdir.	34
14	Normal Kovan türlerinden çok daha zeki ve özgür iradelidirler.	9
15	İşçi	75
16	Kovanlarının iyiliği için emek vermek adına doğan, gözden çıkarılabilen işçiler, yalnızca Kovan'a itaat ve hizmet eder, kişisel menfaatlerini, güvenlik veya konforlarını ikinci planda tutar	57
17	Kovanın kontrolcü ortamından ayrırlırlarsa genellikle ölürler, ancak nadir durumlarda özgür bireyler olurlar.	13

18	Yeni keşfettikleri özgürlüklerin ve önceden belirlenmiş bir amacı kaybetmenin şaşkınlığıyla kendilerini kaybeder ve anlam arayışına girişirler ama bunun sonu genellikle sadece ölüm veya sömürüdür.	33
19	Kovan Asker Drone	0
20	Kovanlarını korumak ve işçiler üzerinde kontrolü sağlamak için oluşturulan birlikler	14
21	Bir işçiden daha güçlü ve daha akılsızdırlar. Dolayısıyla özgürlük arayanları daha azdır ve genellikle paralı asker işi bulurlar.	22
22	Shek	0
23	Güce ve dayanıklılığa entelektüel arayışların üzerinde değer veren bir savaşçı toplum.	5
24	Bazen çok da şaka kaldırmazlar ve aptal vahşiler olarak ün salmışlardır, ama bu her zaman böyle değildir.	24
25	Dış iskelet plakaları doğal bir zırh görevi görür, ancak etrafta gizlice dolaşmak gibi daha beceri isteyen eylemleri engeller.	23
26	İskelet	0
27	İskeletler tam bir gizem. Kimse nereden geldiklerini ve nasıl ortaya çıktıklarını bilmiyor ama binlerce yıllık olduklarından şüpheleniliyor.	11
28	Görsel olarak hiçbir şeyi ifade etme yeteneğine sahip olmasalar da, duyguları olan ve üzüntü, öfke, heyecan, merhamet, korku ve keyif yaşayan varlıklardır.	78
29	Bu, ince mizah anlayışlarıyla birleştiğinde, genellikle İskeletlerin biraz değişken ve öngörülemez gördüğü diğer ırkların kafasını karıştırır.	53
30	Ölüme karşı önyargıları ve korkuları yoktur. Acımasız, etkili savaşçılardır.	46

Appendix L

“Green” Editor Amazon Translate Productivity Report

Segment ID	Post-edited Target	Edit Effort
1	İnsan	0
2	Yeşildiyarlı	83
3	Çoğunlukla çiftçi kültürlerinden gelmektedirler, hala agresif bir ırk olma eğilimindedirler ve kolayca yoldan çıkarılırlar.	17
4	En kötüsü cehalet içinde yaşıyor, sadece açgözlülük ve korkuyu biliyor, farklı şeylere karşı nefret ve şiddetle farklı bir şeye tepki vermek, ama en iyisi en büyük medeni bilim adamlarından bazıları, ülkedeki mühendisler ve savaşçılar.	10
5	Yanıkdiyarlılar	86
6	Yanıkdiyarlılar her şeyden çok kişisel özgürlüklere değer verme eğilimindedirler.	53
7	Kurallarla ve dinlerle iyi geçinme eğiliminde değiller. Sonuç olarak sosyal uyumsuzlar olarak bir üne sahiptirler ve genellikle daha maceracı mesleklerde bulunurlar. Buna rağmen, son derece yaratıcıdır ki bu da onlar doğuştan tüccar ve yetenekli silah demircileri yapar.	21
8	Dürüstlük abideleri olup, görgü kuralları, ılımlılık ve mantıklı her şey karşısında gülerler. Sadık arkadaşlar, tutkulu düşmanlardır, ve birlikte içmek çok eğlencelidir.	40
9	Kovan	0
10	Kovan Prensi	0
11	Bir Kovan'da en yüksek rütbeli ve Kraliçe'nin sağ kolları.	15
12	Hiç kimse gerçek bir Kraliçe görmedi, her şeyi yöneten Prenslerdir.	0
13	Rakip bir Prens'e komutasını kaptırmanın sonucudu., veya bir tür başarısızlık için ceza olarak gönderilen sürgünlerin.	31
14	Normal Kovan türlerine göre çok daha zeki ve özgür iradelidirler.	18
15	Kovan İşçi	37
16	Kovanlarının iyiliği için çalışmak amacıyla doğan, gözden çıkarılabilir işçiler. Yalnızca itaat etmeyi bilirler ve şahsi çıkarlardan, güvenlikten ya da konfordan daha çok Kovanlarına hizmet etmeye önem verirler..	62
17	Kovanın kontrol edici etkisinden ayrılırlarsa genellikle ölürlar, ancak nadir durumlarda özgür bireyler olurlar.	2

18	Belirlenmiş amaçlarını kaybettikleri ve özgürlüklerini yeni keşfettikleri için şaşkına dönüp, kaybolmuş bir şekilde gezinirler ve yeni amaçlar ararlar, ancak ya ölüm ya da sömürü bulurlar.	57
19	Kovan Asker	35
20	Kovanlarını korumak ve işçiler üzerinde kontrolü sağlamak için oluşturulan askerler.	13
21	Bir işçiden daha güçlü ve daha az zekidirler, aralarında özgürlük arayanlar daha nadirdir ve genellikle paralı asker olarak iş bulurlar.	18
22	Shek	0
23	Entellektüel arayışlardan çok güce ve dayanıklılığa önem veren savaşçı bir toplum.	65
24	Bazen biraz mizahsızdırlar ve aptal vahşiler olarak ün yaparlar., ama bu her zaman böyle değildir.	2
25	Dış iskelet plakaları doğal bir zırh sağlar, ancak etrafta gizlice dolaşmak gibi daha hüner gerektiren davranışlara engel olur.	13
26	İskelet	0
27	İskeletler tam bir gizem, kimse nereden geldiklerini ve nasıl yapıldıklarını bilmiyor ama binlerce yaşında olduklarından şüpheleniliyor.	0
28	Hisleri vardır. Görsel olarak ifade etme yetenekleri olmasa da ve üzüntü, öfke, heyecan, merhamet, heyecan, keyif gibi duyguları hissedebilirler.	78
29	Bu durum, ince espiri anlayışlarıyla birleştiğinde, İskeletler tarafından dengesiz ve tahmin edilemez olarak görülen diğer ırkların genellikle kafasını karıştırır.	65
30	Önyargıları veya ölüm korkuları yoktur, acımasız ve etkili savaşçılar olurlar.	24

Appendix M

Ethics Committee Approval for the Study



İSTANBUL 29 MAYIS ÜNİVERSİTESİ
BİLİMSEL ARAŞTIRMA VE YAYIN ETİĞİ KURULU
(Karar Sırası)

Toplantı Tarihi: 30.05.2022

Toplantı Sayısı: 2022/05

KARAR NO: 2022/05-3: Sosyal Bilimler Enstitüsü Çeviribilim Bölümü Yüksek Lisans Programı öğrencisi Haluk Alp SİRİKİNTİ'nin "Crowdsourcing and Machine Translation Post Editing in Video Games Localizations: The case of "Kenshi" (Bilgisayar Oyunlarının Yerleştirilmesinde Kitle Kaynak Kullanımı ve Makine Çevirisi Sonrası Düzenleme: "Kenshi" vakası " konulu bilimsel çalışma ve araştırma yapmaya ilişkin talebini incelemiş olup, yapılan değerlendirme "'Crowdsourcing and Machine Translation Post Editing in Video Games Localizations: The case of "Kenshi" (Bilgisayar Oyunlarının Yerleştirilmesinde Kitle Kaynak Kullanımı ve Makine Çevirisi Sonrası Düzenleme: "Kenshi" vakası " konulu bilimsel çalışmada etik açıdan bir sakınca olmadığına oybirliği ile karar verilmiştir.

ASLI GİBİDİR
30.05.2022

Prof. Dr. Ferihta M. EMECEN

Appendix N

Consent Form Example

BİLGİLENDİRİLMİŞ GÖNÜLLÜ ONAM FORMU

Sayın Gönüllü;

Makine çevirisi sonrası düzeltme yönteminin, kitle kaynaklı yerleştirme yöntemiyle karşılaştırılması amacıyla bir araştırma yapmaktayız. Yaptığımız bu çalışmaya gönüllü çevirmen olarak katılmanız istenmektedir.

Araştırmaya katılım gönüllülük esasına dayalıdır.

Kararınızdan önce araştırma hakkında sizi bilgilendirmek isterim. Elde edilen bilgiler, "Machine Translation Post-Editing as an Alternative to Crowdsourcing in Video Games Localizations" konulu tez çalışmasında kullanılacak, başka bir amaçla veya başka bir araştırmada kullanılmayacaktır.

Bu araştırmaya katılmanız dahilinde isminiz kullanılmayacak ve kimlik haklarınız korunacaktır. Bu araştırmaya katılmayı kabul etmeniz durumunda 4 adet görevi yerine getirmeniz sizden beklenecektir.

Araştırma sırasında sizden alınacak bilgiler gizli tutulacak ve sadece araştırma amaçlı kullanılacaktır. Araştırma sürecinde konu ile ilgili her türlü soru ve görüşleriniz için aşağıda iletişim bilgisi bulunan araştırmacıyla görüşebilirsiniz. Bu araştırmaya **katılmama** hakkınız bulunmaktadır. Aynı zamanda çalışmaya katıldıktan sonra dilerseniz çalışmadan **çıkabilirsiniz**. Bu formu onaylamanız, **araştırmaya katılım için onam verdiğiniz** anlamına gelecektir.

Katılımınız için teşekkür ediyorum.

Gönüllü (Islak imzası ile)

Adı-Soyadı: [Redacted]
İmzası: [Redacted]

Araştırmacının

Adı-Soyadı: [Redacted]
E-postası: [Redacted]
İmzası: [Redacted]

29 MAYIS ÜNİVERSİTESİ
SOSYAL BİLİMLERİ ENSTİTÜSÜ
ETİK KURULU

BİLGİLENDİRİLMİŞ ONAM FORMU

Bu formun amacı katılmanız rica edilen araştırma ile ilgili olarak sizi bilgilendirmek ve katılımınız ile ilgili izin almaktır.

Bu kapsamda “Machine Translation Post-Editing as an Alternative to Crowdsourcing in Video Games Localizations” başlıklı araştırma “Hulusi Alp SİRKINTI ve Doç. Dr. Senem ÖNER BULUT” tarafından gönüllü katılımcılarla yürütülmektedir. Araştırma sırasında sizden alınacak bilgiler gizli tutulacak ve sadece araştırma amaçlı kullanılacaktır. Araştırma sürecinde konu ile ilgili her türlü soru ve görüşleriniz için aşağıda iletişim bilgisi bulunan araştırmacıyla görüşebilirsiniz. Bu araştırmaya katılmama hakkınız bulunmaktadır. Aynı zamanda çalışmaya katıldıktan sonra dilerseniz çalışmadan çıkabilirsiniz. Bu formu onaylamanız, araştırmaya katılım için onam verdiğiniz anlamına gelecektir.

Araştırmayla İlgili Bilgiler:

Araştırmanın Amacı:

Araştırmanın amacı, bilgisayar oyunlarının yerelleştirmesinde kitle kaynaklı yerelleştirme yöntemlerinin yerine makine çevirisi sonrası düzeltme yönteminin daha çabuk sonuçlar verip, yerelleştirme kalitesinin daha yüksek olup olmayacağını sınamaktır.

Araştırmanın Nedeni:

Kitle kaynaklı çeviri ve kitle kaynaklı yerelleştirme projelerinin yetersiz kaldığı durumlarda, kitle kaynaklı çeviri ve kitle kaynaklı yerelleştirme yöntemlerine alternatif olarak makine çevirisi ve makine çevirisi sonrası düzeltme yönteminin kullanılabilirliğini sınamak.

Araştırmanın Süresi:

Gönüllüler araştırmaya yalnızca bir hafta boyunca katılacak olup, bir haftalık süre içerisinde 4 farklı görevi yerine getirmeleri istenecektir.

Araştırmanın Yürütüleceği Yer:

Araştırma online olarak yürütülecek olup, araştırma sırasında gerçekleştireceğiniz görevler için TAUS DQF aracını kullanmanız istenecektir.

Çalışmaya Katılım Onayı:

Katılmam istenen çalışmanın amacını, nedenini, katılacağım süreyi ve yeri ile ilgili bilgileri okudum ve gönüllü olarak çalışma süresince üzerime düşen sorumlulukları anladım. Çalışma ile ilgili ayrıntılı açıklamalar sözlü ve yazılı olarak araştırmacı tarafından yapıldı. Bu çalışma ile ilgili faydalar ve riskler ile ilgili bilgilendirildim.

Bu araştırmaya hiçbir baskı ve zorlama olmaksızın katılmayı kabul ediyorum.

Gönüllü (Islak imzası ile)

Adı-Soyadı:

İmzası:

Araştırmacının

Adı-Soyadı:

E-posta:

İmzası:

Appendix O

Briefing for Post-editors

Görev Bilgilendirmesi:

Yüksek lisans tez çalışmasında kullanılmak ve değerlendirilmek üzere *Kenshi* adlı bilgisayar oyununun dil dosyası Google Translate ve Amazon Translate kullanılarak çevrilecek, bu makine çevirisi çıktıları üzerinden; makine çevirisi hata analizi, makine çevirisi sonrası düzeltme ve makine çevirisi çıktılarının karşılaştırılması yapılacaktır. Yapılacak olan her bir görev için görev tanımı aşağıda yer almaktadır. Bahsedilen görevlerin üçü de TAUS Dynamic Quality Framework (DQF) araçları üzerinden yapılacak olup, görev tanımlarıyla birlikte bu araçları nasıl kullanacağınız da anlatılacaktır.

1. Makine Çevirisi Hata Analizi Görev Tanımı

İlk yapılacak görev makine çevirisi hata analizi olacaktır. Bu görevin amacı, kaynak metnin Google Translate ve Amazon Translate çıktılarının kalitelerinin değerlendirilmesidir. Bu görev boyunca makine çevirisi çıktıları, TAUS DQF'in "Adequacy" ve "Fluency" hata analiz kurallarına göre değerlendirilecektir. TAUS tarafından 2017 yılında yayınlanmış *Quality Evaluation using Adequacy and Fluency Approaches* adlı belgede yer aldığı kadarıyla "Adequacy" ve "Fluency" tanımları aşağıdaki gibidir:

- **Adequacy:**

"How much of the meaning expressed in the gold-standard translation or the source is also expressed in the target translation"

- **Fluency**

"To what extent the translation is "one that is well-formed grammatically, contains correct spellings, adheres to common use of terms, titles and names, is intuitively acceptable and can be sensibly interpreted by a native speaker"

(TAUS 2017, 5)

Yine aynı belgede "Adequacy" ve "Fluency" için değerlendirme kriterleri de verilmiştir:

- **Adequacy**

On a 4-point scale rate how much of the meaning is represented in the translation:

- Everything
- Most
- Little
- None

- **Fluency**

Rate on a 4- point scale the extent to which the translation is well-formed grammatically, contains correct spellings, adheres to common use of terms, titles and names, is intuitively acceptable and can be sensibly interpreted by a native speaker:

- Flawless
- Good
- Dis-fluent
- Incomprehensible

(TAUS 2017, 5)

2. Makine Çevirisi Hata Analizi Görevi için TAUS DQF Aracı Kullanımı:

- Göreve atandığınızda e-posta adresinize aşağıdaki gibi bir ileti gelecektir:

MT Evaluation Project at TAUS DQF Tools by Hulusi Alp SÄ±rkÄ±ntÄ± 

dqf@taus.net

to me ▾

Hello,

A Quality Evaluation task has been assigned to you for job Briefing Quality Eval. [English (United States) > Turkish] by Hulusi Alp SÄ±rkÄ±ntÄ±.

Clicking the link below will take you to the task information page.

On that page, click the Start button to get started.

Please read the translation and rate its quality by selecting one of the radio buttons in the Translation Quality section.

If an error typology task has been assigned, please also indicate types of error(s) you found in the translation (if any) by categorising and counting the errors.

Use the 'Info' links to view explanations of rating scales and/or error categories.

You can use the Comments field to provide clarification or explanation regarding the selected errors.

If need be, you can return to earlier segments by clicking the Previous button.

To interrupt your review, click Home.

Please click the link below to get started

<http://dqf.taus.net/workbench/obsvli4u58keij3qmttb8kog4umu4eqanv75va4l9r59ml7nt0/20621>

Make sure to keep this e-mail until you have completed all segments in this assignment. The link in this e-mail is the only way to get access to your task and continue with an unfinished assignment.

Please reply to this message if you experience difficulties accessing the assigned task.

Kind regards,

The DQF Tools Team

- Bu iletideki linke tıkladığınızda TAUS DQF araç sayfasına yönlendirileceksiniz.
- Yeni açılan sayfada görev detaylarını görebilirsiniz.
- Görev detaylarını inceledikten sonra hata analizi görevine başlamak için “BEGIN TASK” tuşuna basınız:

Project Details

Project Name	Briefing Quality Eval.
Company	-
Content Type	Audiovisual Content
Industry	Entertainment
Source Language	English (United States)
Target Language	Turkish
Number of Segments	30
Completed Segments	0.0

Please note that a broadband internet connection is required for smooth operation of the tool

BEGIN TASK

Instructions

Please read the translation and rate its quality by selecting one of the radio buttons in the Translation Quality section.

If an error typology task has been assigned, please also indicate types of error(s) you found in the translation (if any) by categorising and counting the errors.

Use the "Info" links to view explanations of rating scales and/or error categories.

You can use the Comments field to provide clarification or explanation regarding the selected errors.

If need be, you can return to earlier segments by clicking the Previous button.

To interrupt your review, click Home.

Options

Fluency: Yes

[\(More Info\)](#)

Adequacy: Yes

[\(More Info\)](#)

- Bir sonraki sayfada sırasıyla:
 - Değerlendirilecek kaynak metin sekmesini,
 - Değerlendirilecek bir sonraki kaynak metin sekmesini,
 - Değerlendirilecek makine çevirisi sekmesini,
 - Değerlendirilecek bir sonraki makine çevirisi sekmesini,
 - “Fluency” değerini,
 - “Adequacy” değerini ve
 - İsteğinize bağlı olarak yorum yazabileceğiniz kısmı görüyor olacaksınız.

Briefing Quality Eval.

Source (English (United States))	
Start	
Current	Human
Next	Greenlander

Target (Turkish)	
Start	
Current	İnsan
Next	Grönlandlı

Fluency:

☐ Incomprehensible ☐ Disfluent ☐ Good ☐ Flawless [\(More Info\)](#)

Adequacy:

☐ None ☐ Little ☐ Most ☐ Everything [\(More Info\)](#)

Comments

Characters left: 500

Filename: Kenshi Google Translate Character Creation Page Races Strings.xlsx
Segment: 1 of 30

[PREVIOUS](#) [NEXT](#)
Or Press Enter

- Makine çevirisi çıktısına birer “Fluency” ve “Adequacy” değeri seçtikten sonra “NEXT” tuşuna basarak bir sonraki sekmeye geçiş yapabilirsiniz. Bütün sekmeler bitene kadar bu adımı tekrarlayınız.
- Bu işlem, metnin Google Translate çıktısı ve Amazon Translate çıktısı için ayrı ayrı yapılacaktır.

3. Makine Çevirisi Sonrası Düzeltme Görev Tanımı

Makine çevirisi sonrası düzeltme görevi, TAUS tarafından 2016 yılında yayınlanmış *MT POST-EDITING GUIDELINES* belgesindeki “Human Translation Quality” kurallarına uygun yürütülecektir. Yukarıda bahsi geçen belgede “Human Translation Quality” aşağıdaki gibi tanımlanmıştır:

“This level of quality is generally defined as being comprehensible (i.e., an end user perfectly understands the content of the message), accurate (i.e., it communicates the same meaning as the source text), stylistically fine, though the style may not be as good as that achieved by a native-speaking human translator. Syntax is normal, grammar and punctuation are correct.”

(Massardo, ve diğerleri 2016, 18)

Makine çevirisi sonrası düzeltme görevi süresince “Human Translation” kalitesine erişmek için verilen kurallar ise yine aynı belgeye göre aşağıdaki gibidir:

- Aim for grammatically, syntactically and semantically correct translation.
- Ensure that key terminology is correctly translated and that untranslated terms belong to the client’s list of “Do Not Translate” terms.
- Ensure that no information has been accidentally added or omitted.
- Edit any offensive, inappropriate or culturally unacceptable content.
- Use as much of the raw MT output as possible.
- Basic rules regarding spelling, punctuation and hyphenation apply.
- Ensure that formatting is correct.

(Massardo, ve diğerleri 2016, 18)

4. Makine Çevirisi Sonrası Düzeltme Görevi için TAUS DQF Aracı

Kullanımı

- Göreve atandığınızda e-posta adresinize aşağıdaki gibi bir ileti gelecektir:

dqf@taus.net 9:43 PM (0 minutes ago) ☆ ↩ ⋮
to me ▾

Hello,

A Productivity task has been assigned to you for job Briefing Post-Editing [English (United States) > Turkish] by Hulusi Alp SÄ±rkÄ±ntÄ±.

Clicking the link below will take you to the task information page.

On that page, click the Start button to get started.

The system records the time it takes to translate from scratch versus post-edit machine translated content. The workbench provides a CAT tool agnostic environment to undertake evaluations.

You will be presented with sentences (source and sometimes target sentences) one after the other.

If the translation is correct, click Next. If the translation has mistakes, correct the mistakes and click Next. If the translation is missing, provide it yourself (translate from scratch) and click Next (you can also press the enter button to go to the next sentence).

Note that it is not possible to go back to previously edited segments. You can only move forward to the next segment.

To interrupt post-editing, click Pause. Make sure to pause before you have made any edits. Any changes made to the current segment will be lost upon pausing.

To resume if you did not close your browser session, click Continue on the task overview page. To resume if you did close your browser session, click the link in this assignment e-mail to get to the task information page and click Continue.

The task will resume at the segment where you left off.

Make sure not to interrupt your work or pause without clicking the Pause button. Post-editing times are recorded as you work and pausing without clicking the Pause button may result in productivity test results that are not accurate.

You should have a broadband internet connection for this task.

Please click the link below to get started
<http://dqf.taus.net/workbench/obsvli4u58kej3qmttb8koq4umu4eqanvl75va4i9r59ml7nt0j/20622>

- Bu iletideki linke tıkladığınızda TAUS DQF araç sayfasına yönlendirileceksiniz.
- Yeni açılan sayfada görev detaylarını görebilirsiniz.
- Görev detaylarını inceledikten sonra hata analizi görevine başlamak için “BEGIN TASK” tuşuna basınız:

Project Details

Project Name	Briefing Post-Editing
Company	-
Content Type	Audiovisual Content
Industry	Entertainment
Source Language	English (United States)
Target Language	Turkish
Number of Segments	30
Completed Segments	0.0

Please note that a broadband internet connection is required for smooth operation of the tool

BEGIN TASK

Instructions

The system records the time it takes to translate from scratch versus post-edit machine translated content. The workbench provides a CAT tool agnostic environment to undertake evaluations.

You will be presented with sentences (source and sometimes target sentences) one after the other.

If the translation is correct, click Next. If the translation has mistakes, correct the mistakes and click Next. If the translation is missing, provide it yourself (translate from scratch) and click Next (you can also press the enter button to go to the next sentence).

Note that it is not possible to go back to previously edited segments. You can only move forward to the next segment.

To interrupt post-editing, click Pause. Make sure to pause before you have made any edits. Any changes made to the current segment will be lost upon pausing.

To resume if you did not close your browser session, click Continue on the task overview page. To resume if you did close your browser session, click the link in this assignment e-mail to get to the task information page and click Continue.

The task will resume at the segment where you left off.

Make sure not to interrupt your work or pause without clicking the Pause button. Post-editing times are recorded as you work and pausing without clicking the Pause button may result in productivity test results that are not accurate.

You should have a broadband internet connection for this task.

Options

Level of Quality: Similar or equal to human translation

[\(More Info\)](#)

Process: MT+PE+Human

[\(More Info\)](#)

- Bir sonraki sayfada sırasıyla:
 - Kaynak metin sekmesini,
 - Bir sonraki kaynak metin sekmesini,
 - Makine çevirisi sonrası düzenleme sekmesini görüyor olacaksınız.

Briefing Post-Editing (Productivity)

Information
Required Level of Quality: [Similar or equal to human translation](#)
Content Type: Audiovisual Content
Filename: Kenshi Google Translate Character Creation Page Races Strings.xlsx
Segment: 1 of 30

Source: English (United States)
Start
Current Human
Next Greenlander

Target: Turkish
Start
Current

PAUSE

NEXT
Or Press Enter

- Makine çevirisi çıktısı üzerinde gerekli değişiklikleri yaptıktan sonra “NEXT” tuşuna basarak bir sonraki sekmeye geçiniz. Bütün sekmeler bitene kadar bu adımı tekrarlayınız.
- Makine çevirisi sonrası düzeltme işleminde sekme başına ne kadar süre harcadığınız hesaplanmaktadır ve bulguların değerlendirilmesi açısından bu görevi yaparken başka bir iş ile ilgilenmemeniz, gereken durumlarda “PAUSE” tuşunu kullanmanız rica olunur.
- Bu işlem, metnin Google Translate çıktısı ve Amazon Translate çıktısı için ayrı ayrı yapılacaktır.

Appendix P

6/14/22, 5:21 PM

Gmail - Can't download findings and misinformation about Productivity Task findings



Alp Sirkinti <[redacted]>

Can't download findings and misinformation about Productivity Task findings

5 messages

Alp Sirkinti <[redacted]>
To: dqf@taus.net

Mon, Jun 6, 2022 at 10:46 PM

Hi,

I am H. Alp Sirkinti,

I've been using the TAUS DQF tool for my master's degree thesis and I've assigned some tasks to three volunteers, which they completed. While checking the reports, I've come across two problems and I'd be really happy if you could clarify one of them.

1. The first problem I've encountered is; while trying to download Adequacy and Fluency reports, after clicking on "Download csv" or "Download xls", nothing downloads and I'm just directed to a webpage where nothing happens. <https://dqf.taus.net/reports/download/xlsx/7751/59/8726> for example.

2. Secondly, and this is the problem I'll be asking for clarification, I think there is a mistake in one of your descriptions. After the volunteers completed a "Productivity Task" (MT+PE), I was able to download the results, but, when I checked the description of csv/xlsx export, I came across the highlighted description of MT/PE difference %

Description of CSV/XLSX Export

Productivity Tasks

Once the findings are downloaded, the CSV/XLSX file can be used for more detailed analysis, for example, to assess whether your translators are over or under-editing segments.

For productivity tasks, the following information is exported:

- **Segment ID:** segment ID (same as in import file)
- **Source Segment:** source sentence (same as in import file)
- **MT Target Segment:** target sentence that was evaluated, typically the unedited machine translation output (same as in Target Segment column in import file)
- **Post-edited target:** modified target sentence after it has been post-edited using the target sentence of MT Target Segment as input
- **Segment Length:** the number of words in the source sentence (i.e. the sentence in column Source Segment)
- **Post-edit time(ms):** the time spent by the post-editor converting the target sentence provided as input (i.e. target sentence in column MT Target Segment) into the eventual target sentence (i.e. target sentence in column Post-edited target). This time is expressed in milliseconds.
- **Segment Origin:** the name of the file the source sentence originates from (same as in import file)
- **MTPE difference(%):** the similarity between the target sentence provided as input to the post-editor (i.e. target sentence in column MT Target Segment) and the target sentence that was eventually produced by the post-editor (i.e. target sentence in column Post-edited target). This similarity is expressed as a percentage value where 100 means the pre-edited and post-edited target sentence are exactly the same and 0 means they were completely different.

I believe that the description is wrong as it should have been "0% means pre-edited and post-edited target sentence are exactly the same and 100 means they are completely different", the direct opposite of what's written in the description because in the results that I've downloaded, I see a 0 next to unchanged segments while 100 is given for segments that have been completely changed. I'm also attaching the result files if you want to check them.

Thank you for your time,

H. Alp Sirkinti

3 attachments

Kenshi Makine Çevirisi Düzeltme 1_Report(Productivity)_English (United States)_to_Turkish.xlsx
9K

Kenshi Makine Çevirisi Düzeltme 1_Report(Productivity)_English (United States)_to_Turkish (1).xlsx
9K

<https://mail.google.com/mail/u/2/?ik=dc5f079727&view=pt&search=all&permthid=thread-a%3Ar8205362565963795192&simpl=msg-a%3Ar820701505...> 1/2

6/14/22, 5:21 PM

Gmail - Can't download findings and misinformation about Productivity Task findings



Kenshi Makine Çevirisi Düzeltme 1_Report(Productivity)_English (United States)_to_Turkish (2).xlsx

9K

To: Alp Sirkinti <[redacted]@taus.net>

Tue, Jun 14, 2022 at 1:30 PM

Dear Alp,

Apologies for the late reply.

The issue occurs because of an unsupported character (ğ) in the project name (Kenshi Makine Çevirisi Değerlendirme)
You can update the name of the project on the projects page and then the reports page will work.

Best regards,

[Quoted text hidden]

Alp Sirkinti <[redacted]@taus.net>

Tue, Jun 14, 2022 at 1:42 PM

Thank you for the clarification

But what about the second question? Is there a mistake in the "MT/PE difference(%)" part of the "Productivity Task" description?

Many thanks for taking the time to answer my questions,

Best,

Alp

[Quoted text hidden]

To: Alp Sirkinti <[redacted]@taus.net>

Tue, Jun 14, 2022 at 2:23 PM

Hi Alp,

Indeed, you are correct. There's an error in the description and your suggestion is accurate.

Best regards,

[Quoted text hidden]

Alp Sirkinti <[redacted]@taus.net>

Tue, Jun 14, 2022 at 4:43 PM

Hi

Thank you very much for quickly answering my questions and for your time,

H. Alp Sirkinti

[Quoted text hidden]

<https://mail.google.com/mail/u/2/?ik=dc5f079727&view=pt&search=all&permthid=thread-a%3Ar8205362565963795192&simpl=msg-a%3Ar820701505...> 2/2

CURRICULUM VITAE

ÖZGEÇMİŞ			
Adı, Soyadı	Hulusi Alp Sırkıntı		
Doğum Yeri ve Yılı			
Bildiği Yabancı Diller	İngilizce	Fransızca	
ve Düzeyi	C2	A2	
Eğitim Durumu	Başlama - Bitirme Yılı	Kurum Adı	
Lise	2005	2009	24 Kasım Anadolu Lisesi
Lisans	2009	2013	Marmara Üniversitesi
Yüksek Lisans	2020	2022	29 Mayıs Üniversitesi
Doktora			
Çalıştığı Kurum/lar	Başlama - Ayrılma Yılı	Çalışılan Kurumun Adı	
1.	2014	Devam etmekte	Fatih Sultan Mehmet Vakıf Üniversitesi Yabancı Diller Bölümü
2.			
3.			
Üye Olduğu Bilimsel ve Mesleki Kuruluşlar			
Katıldığı Proje ve Toplantılar			
Yayınlar:			
Diğer:			
İletişim (e-posta):			
Tarih	21.06.2022		
İmza			
Adı Soyadı	Hulusi Alp Sırkıntı		