

## **A Corpus-Based Approach to Creativity in Non-Professional Subtitles: The Case of *Hapax Legomena***

Ebrahim Samani

Department of Language and Literature, Farhangian University, Tehran, IRAN

Fatemeh Badiozaman

Department of Foreign Languages, Higher Education Complex of Bam, IRAN

Razieh Bagheripour

Department of Foreign Languages, Higher Education Complex of Bam, IRAN

### ABSTRACT

Subtitling is a common way of breaking language barriers for audiovisual materials. A notable portion of subtitles are distributed on the internet by non-professional translators. The existing literature has demonstrated that non-professional translators try unconventional and innovative procedures in their tasks (Khoshsaligheh et al. 2019; 2020). Following this rather recent trend, this study employs the corpus-based approach (Kenny 2001; Zanettin 2014) to find out whether non-professional subtitlers come up with creative solutions for the problems of rarity and non-equivalence in the case of *hapax legomena* (i.e., words with only one occurrence through the entire corpus). The corpus in this study included English and non-professionally translated Persian subtitles of comedy movies released in the 21st century, analyzed by means of AntConc, a corpus analysis toolkit. Although it was initially assumed that amateur subtitlers and fansubbers would resort to imaginative equivalents, results revealed that these subtitlers preferred non-creative strategies to creative ones (Kußmaul 2000) due to not being trained and/or experienced in alternative translation solutions when facing troublesome items in the process of translation.

**KEYWORDS:** audiovisual translation, corpus linguistics, creativity, Kußmaul's creative strategies, non-professional subtitles

## **1. Introduction**

Nowadays, societies are replete with audiovisual material that is increasingly produced and circulated, contributing to individuals' entertainment, education, and awareness (Díaz-Cintas and Neves 2015). Among all the subcategories of audiovisual translation, Iran is mainly characterized as focusing on dubbing. Audiovisual products that do not earn the right to be dubbed<sup>1</sup> have the chance of being voiced over<sup>2</sup>; however, subtitling is not common in mainstream media (Khoshsaligheh, Ameri, Khajepoor, and Shokohmand 2019).

Furthermore, when it comes to movies and TV series, none are aired with subtitles on any TV channels, only dubbed versions are broadcast, which means that subtitled movies and TV shows are found only on the internet.

There are reasons why some viewers (or possibly the majority of the younger generation) may not choose to watch official dubbed or voiced-over products, such as the slowness of the official process to finally release programs, censoring, or the fact that not all programs are selected to be translated (Khoshsaligheh et al. 2019). For hardcore fans, these reasons will not pose any problems, since many like-minded individuals provide easy-to-use, mostly free (or low-priced) subtitles as fast as humanly possible through the internet (Gambier 2013:53). These non-official, non-professional, and mostly non-profit subtitles that are distributed via the internet are prepared by volunteers who are usually not trained in translation and/or subtitling. Amateur translators are usually not familiar with the conventions, norms, and ethics of translation, which often leads them to spontaneously performing numerous innovations in this field (Antonini, Cirillo, Rossato, and Torresi 2017). Nonetheless, all professional translators are familiar with a prominent translation universal which Toury (2012:303) called standardization—although, it has been recognized by various names: normalization, conventionalization, and conservatism; this universal refers to the fact that translations tend to use recurrent patterns of the target language rather than creative or innovative patterns. On the contrary, non-professional subtitlers do not tend towards normalization because they move away from guidelines of professional subtitling (Khoshsaligheh and Ameri 2017:18).

---

<sup>1</sup> In the dubbing process, the original dialogue track is replaced by a dialogue track in the target language that tries to adhere to the original timing, phrasing and lip synchronization (Pérez-González 2014).

<sup>2</sup> Voice-over refers to a technique in which the volume of the original soundtrack is lowered; in the meantime, one voice actor reads out the translation (Pérez-González 2019). Therefore, the original and translated track can be simultaneously heard, but the translated one is more prominent (Bogucki 2013).

It is important to point out that, according to Khoshsaligheh, Ameri, Shokoohmand and Mehdizadkhani (2020), exclusively non-professional subtitles can be found in the Iranian context. These authors classified such subtitles into three categories: quasi-professional, amateur, and fan works. In the first place, the category of quasi-professional comprises all subtitles which are produced by professional translators who cannot be classified as professional subtitlers because they are not trained specifically for this task; plus, Iran does not have its own official subtitling guidelines. Secondly, the concept of amateur subtitles refers to those which are provided by people who are probably paid a meager amount of money for subtitling pirated films, so they mostly do these works to become more acquainted with the languages in question, the act of translation, translation software, and the market of multimedia translation. Finally, the idea of fan works in this context refers to the kind of subtitles that are created by hardcore fans of a particular genre (such as anime and K-dramas), who are meant to provide other fans with an easy and fast access to their favorite entertainment. Remembering that quasi-professional subtitles are produced solely for non-fictional audiovisual material, it was important to make the above distinction, since the present article focuses on subtitles for fiction (i.e., movies and TV series). Therefore, throughout the text, the notion of “non-professional subtitles” is going to be used to refer to amateur subtitles and fansubs.

This study set out to explore creative strategies that non-professional subtitlers have used to cope with the problems they faced with rare or non-existent words. To this end, following Kenny’s (2001) example, the researchers found all hapax legomena— “word forms that only occur once in the corpus” (Zanettin 2014:19)—in a corpus of English-Persian subtitles, and then classified their translation strategies: these strategies could be non-creative (e.g., *deletion* and *transference*), or creative—which were classified according to Kußmaul’s (2000) creative strategies for translation. Finally, the following research questions were examined in this study: 1) What general strategies have been used in non-professional subtitles to translate hapax legomena? 2) Which of Kußmaul’s creative translating procedures have non-professional subtitlers employed for hapax legomena? And 3) When it comes to hapax legomena, do non-professional subtitlers gravitate towards creative or non-creative translation strategies?

## **2. Literature Review**

### *2.1. Theoretical Framework of the Study*

The current study was conducted under the framework of Kußmaul's (2000) creative strategies for translation. These strategies could be beneficial for translators, especially in the translation of words with no meaning, or no equivalent in the target language. The strategies and definitions are as follows:

1. Chaining of categories (Kußmaul 2000:120-121): Here, the translators do not use a completely different word (a word or an expression is called a "category" in this part), they replace the original word with a new one that is relevant, so there is a chain or link between the original category and its translation. In other words, they look at the original category from a different perspective and select a new relevant word that makes sense. Kußmaul presented an example involving proper names: in the world-renowned series of comics *Astérix le Gaulois*, by Albert Uderzo and René Goscinny, one of the characters is a druid named Panoramix, which in French evokes the sense of "visionary" and/or "wise old man", while in the English translation he is called Getafix (which refers to the distribution of magical potions); and in the German translation, he is called Miraculix (which is related to the miraculous effects of his potion). In this example, the druid's name has not been replaced by random names; the new names are just different aspects of the druid's functions and capabilities.

2. Picking out scene elements (Kußmaul 2000:121-122): "In translations we can often observe that a relatively abstract word of the source text is replaced by a more detailed and concrete expression" (Kußmaul 2000:121). For example, in *Old Possum's Book of Practical Cats*, T.S. Elliot described a cat as famous (the exact adjective famous was used); on the other hand, Carl Zuckmayer (the German translator), instead of directly using the word famous, indicated it imaginatively by writing that people could read about the cat in newspapers every day, because being featured in the newspapers is a representation (an obvious element) of fame.

3. Enlarging a scene (Kußmaul 2000:123): Translators may add some elements (i.e., words or expressions) to the scene that did not exist in the original without necessarily

changing the original meaning of the source, only expanding it. For instance, T.S. Elliot, in the same book, wrote about an old weak cat, whereas Zuckmayer translated it as the cat being at death's door. The aspect of being close to death was added to the scene to emphasize the cat's weakness.

4. Framing a scene (Kußmaul 2000:123-125): In this method, translators replace the original theme with one that is broader. As another instance, Elliot wrote this sentence: "He can pick any card from a pack" (Elliot 1961, as cited in Kußmaul 2000:123), that Zuckmayer translated to: "His card tricks are very popular" (Kußmaul 2000:124). Card tricks consist of multiple concepts, of which picking cards is only one; therefore, the translated concept is broader than the source.

5. Thinking up a new frame (Kußmaul 2000:125): The source has obviously already introduced a frame. "This is the frame people use for the kinds of mental pictures suggested by the source text. A new frame, however, would be a tailor-made frame, as it were; in lexical terms, a neologism" (Kußmaul 2000:125). For example, in an assignment, Kußmaul asked students to translate "sausages composed of mainly bread in British motorway restaurants" to German. Since the ingredients of sausages differ in the two cultures, to reach adequacy in translation, one of the students suggested a German equivalent meaning "bread sausages", which is not only new and creative, but also frames and comprises the elements mentioned in the source word (i.e., cheap, made of mainly bread).

## *2.2. Non-professional Translations*

Translation, as the practice of mediating culturally and linguistically, is sometimes performed, as mentioned above, with the help of unpaid volunteers who have not received the required training in the field. This action, which is usually mediated through technology, is referred to as "non-professional translation" (Antonini and Bucaria 2015:7). Frequently, these so-called non-professional translators are internet users who usually "operate in collaboratively structured environments" (Orrego-Carmona 2014:130).

As one aspect of these non-professional translations, subtitling has become a familiar practice in the translation of audio-visual products in contexts in which audiences are not satisfied

either with the professional and commercial translation of media or due to the lack of any translation for these materials (O'Hagan 2009; Massidda 2015). Non-professional subtitling is usually divided into “professional-amateur” or “(pro-am) subtitling” and “innovative subtitling” (Orrego-Carmona 2014). The former tries to follow the professional norms of subtitling and produces subtitles with a near-professional quality; the latter is characterized by creativity and the practice of norm-breaking regarding, for example, font and color (Orrego-Carmona 2014).

Another broader categorization is “guerrilla translation” which is offered by Dwyer (2017) and includes both fan translation (fansubbing) and non-fan subtitling. The non-fan translation is often found in the realm of pirated products and is replete with mistranslations and typos, which leads to low-quality subtitling. This kind of subtitling is profit-driven as it usually aims at providing audiovisual products which evade “media regulation and censorship” (123). Meanwhile, fan translation, which is not profit-oriented, is usually provided through the internet by volunteers who may not have sufficient training in translation and/or subtitling and intend to share the content with other interested viewers. They are usually unaware of norms and ethics in translation and may indulge in numerous innovations and creativity (Antonini, Cirillo, Rossato, and Torresi 2017:7). In other words, creative subtitling occurs when subtitlers reject the notion that certain guidelines are at work everywhere, and instead initiate both linguistically and stylistically creative processes (McClarty 2013:593).

Until the 1990s, the concept of creativity in translation was only limited to “rather marginal comments” (Wilss 1988:110). According to Wilss, the reason was that the concept was far from being conceptualized, measured, described, or weighted. It has been discussed that there is a connection between creativity and expertise. For example, according to Thomä (2003), professionals, as rated by three experts, demonstrated higher performance than students in translation, especially in translation into their mother tongue, compared to translation into a foreign language. Riccardi (1998:175) states that, as one of the differences, experts possess a broader repertoire of strategies than novices and can combine these strategies as flexibly as possible.

### *2.3. Corpus Studies*

Corpus linguistics was primarily introduced into the field of linguistics by structuralists; thus, corpus linguistics can be described as the study of language based on examples of real-life language use (McEnery, Xiao, and Tono 2006). A corpus comprises of a collection of texts often covering thousands of words that are stored in an electronic database (Baker 2006:50). Kacetyl and Frydrychova-Klimova (2015:365) then defined corpus as an electronic database of authentic materials such as transcripts of spoken language. Furthermore, Flowerdew (2011) mentioned some criteria for defining a corpus, including that a corpus must contain naturally occurring data, must represent a particular language or genre, and must be designed with a specific linguistic purpose in mind.

Corpora have been used for a plethora of academic purposes, including the production of language teaching materials, suitable tools for language learning, translation studies, and dictionaries (Baker 2006; Reppen 2010; Flowerdew 2011). Furthermore, there are different types of corpora, and the type of each corpus depends on the purpose for which it has been created. On the one hand, corpora can either be general (created from a variety of texts) or specialized (designed for a specific purpose rooted in special texts) (Hunston 2022:14); moreover, corpora can be monolingual or multilingual, the latter being particularly useful in the field of Translation Studies.

The most significant study looking at lexical creativity in translation through the lens of the corpus approach is Kenny (2001). She examined hapax legomena (words with a frequency of one) and low-frequency collocations in a parallel corpus of contemporary German-language fiction alongside their English translations. She stated that creativity was usually manifested on the levels of orthography, derivational morphology, unexpected meaning associations, and similar levels. She identified creative lexical items in original texts and verified the accuracy of these identifications by consulting native speakers, lexicographical sources, and a general reference corpus of German. Then, she examined the corresponding English translations of these selected items to investigate/observe/analyze whether translators had utilized normalization. According to her findings, a large number of items that may seem to have been creatively translated were in fact caused by lack of attention and care during the process of translation.

However, in some cases, compensation strategies have been used to moderate the non-equivalence problem. Khoshsaligheh, Ameri, Shokoozmand, and Mehdizadkhani (2018) aimed at studying translation strategies that non-professional translators employed to subtitle taboo items into Persian. They analyzed a corpus of internet-mediated fansubbing and concluded that fansubbers tended to maintain taboo features of the original language in their Persian translation. They argued that the origin of the fansubbing phenomenon in Iran is a reaction to the dominant doctrine that promotes modesty and chastity (meaning the doctrine brought forward by the principles of the Islamic-based revolution).

### **3. Methodology**

#### *3.1. Corpus of the Study*

Before the process of creating the corpus could be started, the scope of the research had to be delimited. The researchers decided that after this delimitation, they would draw a sample from all the American comedy movies released in the twenty-first century. Considering this research first began in 2019, the sample was drawn from movies released between the years 2000 and 2019. Wikipedia offered a list of 875 American comedy movies that were released in these 19 years. This list may not be comprehensive, but it could at least be used as a starting ground for sampling. Nevertheless, the researchers aimed to include movies from approximately two decades, hoping that they could achieve the highest possible level of diversity. The rationale behind the selection was that the comedic nature of these texts would make it more likely for innovative language to appear, and therefore the chances of encountering cases of complex hapax legomena would be higher. According to Krejcie and Morgan's (1970, as cited in Chuan and Penyelidikan 2006:2; Rajendran and Shah 2020:5) chart, if the population<sup>3</sup> size is more than 850 and less than 900, the sample size should comprise of 269 cases. Thus, the sample to make up this corpus consisted of English and Persian subtitles for 269 movies. Ultimately, the Persian subtitles for some of these movies were not available, so their names were removed from the list.

---

<sup>3</sup> Here, the term population refers to the 875 comedy movies from which the sample was drawn.



Furthermore, for the sample to be representative of the whole population, a random probability sampling method had to be used. The software Random Number Generator version 1.4 was employed:

1. It covers an option for *intervals*, so for this purpose, the logged numbers were the interval from 1 to 875.
2. Only whole numbers were employed in this study; therefore, *digits after decimal points* were set to 0.
3. The number of digits to generate was set to 269 because that is the number of movies the sample contained, so the researchers could know which movies the sample would cover.

In summary, this corpus consisted of subtitles for 269 movies, that is, English subtitles for 269 movies and their corresponding 269 Persian subtitles. The whole corpus comprised of 6,012,457 words.

### *3.2. Data Collection and Analysis*

The collection and analysis of these data was carried out in four phases. To allow for easy access and possible future replicability, every action taken by the researchers is here explained in detail. The first phase entailed the construction of the corpus. To do so, all the required subtitles were downloaded from websites<sup>4</sup> such as zirnevisha9.com, subdl.com, and subscenes.ir. The next step was extracting plain texts from subtitles which required cleaning and eliminating symbols and time codes. The deletion of these tokens individually would be a too highly labor-intensive task for it to be performed reliably; fortunately, the *Find and Replace* feature in Microsoft Word came to the researchers' aid. Each subtitle file contained three factors standing in the way of becoming plain text; these factors (all of which will be deleted) are displayed in bold in the following example from the movie *21 & over*, released in 2013:

**5**  
**00:00:51,185 --> 00:00:54,586**  
<i>Just ask any of the old-time pit bosses,  
they'll know.</i>

---

<sup>4</sup> These three websites are majorly used to download subtitles for numerous movies and TV series in many languages.

1. `<i>` at the beginning of each subtitle block and `</i>` at the end: These were very simple to remove; all that was necessary was to enter each symbol individually in the *Find what* box and to enter nothing in the *Replace what* box.
2. The number on top of each block of subtitle text: At this point, the researchers turned to the *Special* drop-down menu in *Find and Replace*. To remove these numbers, the option *any digit* was entered in *Find what* box; however, this would delete all the numbers present in the file, even those that are meant to be retained. To solve this problem, after *any digit*, a *paragraph mark* should be added. Lastly, to prevent the time codes from colliding with the line of the subtitle texts, a paragraph mark should be entered in the *Replace with* box. Here is what the ultimate code looked like:

Find what: `^#^p`

Replace with: `^p`

3. Time codes: The same *Special* drop-down menu was used for this as well. To put it simply, each number in the time code was substituted by *any digit*, every other symbol or punctuation remained intact. This code was to be replaced by nothing, so nothing was entered in the *Replace with* box. The following was what the code looked like:

Find what: `^#^#:^#^#:^#^#,^#^#^# --> ^#^#:^#^#:^#^#,^#^#^#`

Replace with: (leave empty)

It is interesting to note that the second and third step can be accomplished at the same time, but the authors decided to separate them for the sake of clarity. Nonetheless, the code below is the result of this combination:

Find what: `^#^p^#^#:^#^#:^#^#,^#^#^# --> ^#^#:^#^#:^#^#,^#^#^#^p`

Replace with: (leave empty)

The second phase was to recognize hapax legomena by means of AntConc version 3.5.8 (Anthony 2019) for Windows as a corpus analysis toolkit for the English plain texts. The researchers needed to extract plain texts from English subtitles because AntConc cannot

process subtitle files, and because the extra information present in subtitles would hinder the analysis. This software has a function tab called *Word*, which makes a list of all the words in the corpus and reports the frequency of their appearance; accordingly, when one moves to the end of the list, one can find words with only a single frequency (i.e., the hapax legomena). AntConc yielded 13,654 words of this kind.

The third phase consisted in collecting a sample from this list. Again, according to Krejcie and Morgan's chart, for populations<sup>5</sup> of 15,000 or less, a sample of 375 would be appropriate, which was achieved through the same random number generator software to pick 375 random words by taking the same steps as before.

The fourth and final phase required locating the equivalents to the sampled English hapax legomena in Persian subtitles, and identifying what sort of strategies were employed to translate them. When one clicks on the intended word in AntConc, it is possible to know which file (i.e., which movie) it came from. This way, the researchers could consult the English subtitle file to which every word belonged, search for the word, and find the time code associated with it because a time code determines at exactly what time in the movie the subtitle appears. At this moment that the time code was known, the researchers opened the corresponding Persian subtitle and searched for the time code. The Persian equivalent for the desired hapax legomena was found in the text associated with this time code. However, sometimes the translator may have rearranged word order or progression of information, so the meaning of the intended word could be found in a different time code, perhaps a few seconds before or after. This is how all the Persian equivalents were located; thus, the researchers could effortlessly move on to recognizing translation strategies.

#### **4. Results and Discussions**

In the beginning, when all the hapax legomena were analyzed, it turned out some of the items on the list were proper nouns whose method of rendition did not matter because proper nouns are not normally translated, they are just transliterated as they are. Table 1 displays the number of proper nouns as opposed to other words.

Table 1. Frequencies of Proper Nouns and Normal Words

---

<sup>5</sup> At this point, the expression population represents all the 13,654 hapax legomena that were found.

Type	Frequency	Relative frequency
Proper nouns	47	12.5%
Words	328	87.5%
Total	375	100%

Considering the opening objectives (i.e., insufficiency or abundance of creativity in translating hapax legomena), it was time to see whether there was a gravitation towards innovative or uninventive procedures.

Table 2. Frequencies of Creative and Non-Creative Strategies

Strategy	Frequency	Relative frequency
Creative	102	31%
Non-creative	226	69%
Total	328	100%

Table 2 shows how non-professional subtitlers often resorted to non-creative solutions. Initially, this may sound unusual because non-professional subtitlers—more specifically fansubs—have been known to make innovations (Antonini, Cirillo, Rossato, and Torresi 2017). However, as it was pointed out (Riccardi 1998; Thomä 2003), it can be interpreted that only professional translators, due to their training and experience, are familiar with alternative and somewhat creative techniques; techniques that non-professional translators are not familiar with, which is why they rely on easy, more predictable methods.

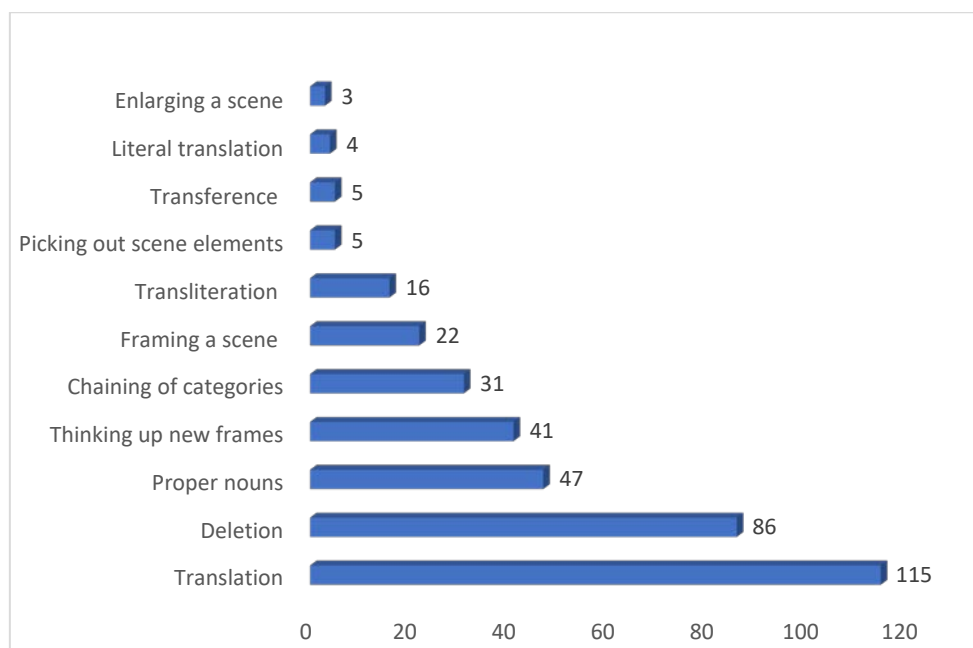
When it came to rendering words other than proper nouns, the analysis indicated several more strategies besides Kußmaul’s categorization, including: a) deletion, b) transference or borrowing (either with the English alphabet, or Persian), and c) literal translation—or calque. Table 3 and Figure 1 both indicate the frequencies of each strategy used in the sample.

Table 3. Frequencies of employed strategies in non-professional subtitles regarding the hapax legomena in English

Strategy type	Strategy	Frequency	Relative frequency
Non-creative	Translation	115	35%
	Deletion	86	26%
	Transliteration	16	4.75%
	Transference	5	1.5%
	Literal translation	4	1%
Kußmaul’s creative strategies	Thinking up new frames	41	12.5%
	Chaining of categories	31	9.25%

	Framing a scene	22	6.5%
	Picking out scene elements	5	1.5%
	Enlarging a scene	3	0.75%
Total		328	100%

Figure 1. Frequencies of the strategies used by non-professional translators



According to Table 3 and Figure 1, *Translation* was the most popular strategy. This result was expected; when translators—regardless of their professional or non-professional status—encounter an unfamiliar word, their first step is to consult a dictionary, and if the equivalent offered by the dictionary seems appropriate, translators are usually inclined to use it.

Next, *deletion* was quite high in frequency as well, which was anticipated. Non-professional subtitlers, amateur subtitlers, and especially fansubbers, are not experienced translators. Therefore, when they came across troublesome words, they preferred the easier option and thus deleted the terms that were leading to difficulties.

Among Kußmaul’s strategies, *thinking up a new frame* was the most frequent. The logic behind this could be that finding a novel and fresh construct would be easier than finding a similar notion, because there are no rules, no limitations, when picking a new construct.

Ultimately, the other strategies had varying degrees of frequency. Table 4 represents strategies analyzed in this study along with examples from the context of the study.

Table. 4. Examples of each strategy (note: the bold words in subtitles are hapax legomena and their Persian equivalents)

Movie title	English subtitle	Persian subtitle	Back translation	Strategy
Bandits (2001)	Gosh, no, I didn't <b>Harv</b> .	اوه , نه , نمي دونم , <b>هارو</b> .	Oh, no, I don't know, <b>Harv</b> .	Proper noun
Spy (2015)	This anti- <b>funga</b> l spray	اين اسپري ضد قارچ هم ميتونه	This anti- <b>fungus</b> spray	Translation
Ghosts of girlfriends past (2009)	Such a shvantza			Deletion
Big fat liar (2002)	Adiós, <b>sugarpuff</b>	آديوس، برف شادي	Adios, <b>snow spray</b>	Thinking up a new frame
The emoji movie (2017)	Get that <b>bozo</b> out of there!	اون دلقک رو بندازيدش بيرون!	Throw this <b>clown</b> out.	Chaining of categories
Barbershop (2002)	To me it sounds like you got a little <b>haterism</b> in your game.	به نظر من تو خيلي حق به جانب نظر دادی	In my opinion, you commented very <b>self-righteously</b> .	Framing a scene
Barbershop (2002)	Sit up there, just watch too much TV and listen to that <b>jigga</b> ray...	بشين اينجا فقط تلويزيون ببين و به جيگاری گوش کن	Sit here, just watch TV and listen to <b>jigga</b> ray	Transliteration
Knocked up (2007)	It's like <b>saran</b> wrap!	مثل پلاستيکي که رو مواد غذايي مي پيچن!	It's like that <b>plastic</b> that is wrapped around food.	Picking out scene elements
My big fat Greek wedding (2002)	You know, the word "baptism" comes from the Greek word " <b>vaticia</b> "	کلمه baptism از کلمه يونانی <b>vaticia</b> مياد	The word baptism comes from the Greek <b>vaticia</b>	Transference
The good dinosaur (2015)	This is <b>Dream-Crusher</b> .	اين روياشکن هـ	This is <b>Dream-Crusher</b> .	Literal translation
Nutty professor 2 (2000)	<b>Klumpville</b>	محله خانواده کلامپ	The <b>neighborhood of the Klump</b> family	Enlarging a scene

## **5. Conclusions**

### *5.1. Answers to Research Questions*

The present study aimed at studying strategies that non-professional subtitlers have used to cope with the problems they faced in translating hapax legomena, and whether they would use creative solutions. In resolving the first two research questions, the results revealed that non-professional subtitlers leveraged both creative and non-creative strategies. Among non-creative procedures (covering *translation, deletion, transliteration, transference, literal translation*), *translation* and *deletion* were favored the most. Additionally, although all Kußmaul's creative strategies were used to various frequencies, *thinking up new frame* and *chaining of categories* were the most recurrent. More importantly, non-creative strategies were used more often.

In keeping with all the pieces of previous research about fansubs and non-professional subtitles, in general, indeed, previous research contains many elements of creativity in the ways that subtitlers break norms by, for example, adding surtitles<sup>6</sup> to explain certain terms, adding their own words to show their feelings and reactions (e.g., Look how handsome this actor is! He shouldn't have said that, oh my God, I can't believe that happened!) or by adding explanations in parentheses. They take this sort of actions because they have not been trained in subtitling conventions, and since they are not held back by the same restrictions that professional subtitlers must adhere to, non-professional subtitlers produce creative elements. Moreover, the assumption before starting this article was that non-professional subtitlers might have a creative approach towards subtitling, they would translate hapax legomena creatively as well, but the opposite proved true.

Thus, in answer to the third research question about preference for the strategy types, non-professional translators repeatedly resorted to non-creative strategies (such as *deletion*, or *transliteration*) to translate these rarer words. This could be explained by the idea that this lack of creativity might be related to the fact that non-professional subtitlers are not actually translators. They are not trained or experienced enough to cope with the hardships of translation, and instead decide to avoid problems such as the translation of hapax legomena.

---

<sup>6</sup> Surtitles usually refers to texts shown on screens located on top of the stage in Opera performances (Secară 2018:130); however, for audiovisual files, it can also mean the text that is displayed at the top of the screen as opposed to subtitles which are displayed at the bottom of the screen (Künzli & Ehrensberger-Dow 2011:4).

Professional translators, on the other hand, know alternative methods to deal with not-straight-forward elements, such as problems that are caused by linguistic or cultural differences, due to their education or experience.

### *5.2. The Results of this Study in Comparison to other Studies*

The finding that non-professionals do not turn to creative solutions corroborates Thomä (2003) who found that there was a connection between expertise and creativity, meaning that professional translators tend to be more creative. Riccardi (1998) states that, as one of the differences, experts not only demonstrate a more comprehensive list of strategies than novices, but also have the ability to utilize these strategies in combination in a flexible way.

According to the findings, translation was the strategy with the highest frequency in the study. Indeed, novice translators can be expected to have a harder time understanding the source text, clarifying vocabulary issues, devising search strategies, and solving mother tongue related issues such as spelling. Meanwhile, they are less concerned with consequences and practice dangerous creativity as a result. In other words, they are less likely to produce creative translations that deviate from literal translation and differ from the intended source message. As one of the translation universals, normalization implies (diminishing) creativity, and it seems a usual practice across many social and cultural contexts. According to Kenny (2001), many seemingly creative items cannot be categorized as truly creative translations. The findings of this study also align with Vintar (2016), who reported that there was no difference in translations and their parallel source in terms of lexical variety. In other words, the percentage of innovative and creative vocabulary is not significantly less than in original texts.

### *5.3. Implications of this Study*

As replete as the literature is with studies focusing on creative subtitles (McClarty 2013; Kapsaskis 2018; Pérez-González 2007), this study can prove significant because to the best of the researchers' knowledge, there have been little to no corpus-based works in terms of creativity in non-professional subtitles. This study is also significant because it can provide a list of strategies for professional and non-professional subtitlers alike to benefit from adding to the creativity of their final products. Iran is rather new to subtitling in a sense that it does not have its own subtitling conventions. Consequently, a solid tradition of subtitling still



needs to be established as well as attempts should be directed toward introducing Iran/Persian-specific guidelines for subtitling translation.

This study can be beneficial to translator training as well. As this exploration revealed, unexperienced beginner translators are not accustomed to creative solutions to the various difficulties they may face (the researchers have also come across this reality in their own translation training). Accordingly, translator trainers need to convey to their students that languages have numerous differences, so turning to literal verbatim approaches is often a poor choice leading to inadequate quality. Furthermore, many fan and amateur subtitlers are translation trainees; hence, it is important to teach them subtitling conventions in audiovisual translation courses.

Finally, this line of investigation into the differences between professional and non-professional translation can be further developed. If future researchers wish to take the direction of this article and expand it, they may consider analyzing other movie genres, such as TV series, K-dramas, anime, comic books, manga, webtoons, novels, and video games (because all these entertainment materials are known to be translated by non-professional translators). Future research could also adopt and analyze other theoretical frameworks about creative translation strategies, as well as benefit from alternative sampling procedures to ensure a representative sample.

## **Contact Details**

1. Ebrahim Samani

Assistant Professor, Department of Language and Literature, Farhangian University,  
Tehran, IRAN

Email: [e.samani@cfu.ac.ir](mailto:e.samani@cfu.ac.ir)

2. Fatemeh Badiozaman

Visiting Lecturer, Department of Foreign Languages, Higher Education Complex of  
Bam, IRAN

Email: [badiozaman.fatemeh@gmail.com](mailto:badiozaman.fatemeh@gmail.com)

3. Razieh Bagheripour

Visiting Lecturer, Department of Foreign Languages, Higher Education Complex of Bam, IRAN

Email: [badiozaman.fatemeh@gmail.com](mailto:badiozaman.fatemeh@gmail.com)

## References

- Ameri, Saeed and Massod KhoshsAligheh (2021) 'An Experiment on Amateur and Professional Subtitling Reception in Iran', *SKASE Journal of Translation and Interpretation*, 14(2): 2-21.
- Anthony, Laurence (2019) *AntConc* (Version 3.5.8) [Computer Software]. Tokyo, Japan: Waseda University. Available online at [<https://www.laurenceanthony.net/software>] (accessed 03 November 2022).
- Antonini, Rachele and Chiara Bucaria (2015) *Non-Professional Interpreting and Translation in the Mediam*, Frankfurt: Peter Lang.
- Antonini, Rachele, Cirillo Letizia, Linda Rossato and Ira Torresi (eds) (2017) *Non-Professional Interpreting and Translation: State of the Art and Future of an Emerging Field of Research*, Amsterdam: John Benjamins
- Baker, Paul (2006) *Glossary of Corpus Linguistics*, Edinburgh: Edinburgh University Press.
- Bogucki, Łukasz (2013) *Areas and Methods of Audiovisual Translation Research*, Frankfurt: Peter Lang.
- Chuan, Chua Lee and Jabatan Penyelidikan (2006) 'Sample Size Estimation Using Krejcie and Morgan and Cohen Statistical Power Analysis: A comparison', *Jurnal Penyelidikan IPBL*, 7(1): 78-86.
- Díaz-Cintas, Jorge and Josélia Neves (eds) (2015) *Audiovisual Translation: Taking Stock*. Newcastle, UK: Cambridge Scholars Publishing.
- Dwyer, Tessa (2017) *Speaking in Subtitles: Revaluating Screen Translation*, Edinburgh: Edinburgh University Press.
- Flowerdew, Lynne (2011) *Corpora and Language Education*, Basingstoke: Palgrave Macmillan.
- Gambier, Yves (2013) 'The Position of audiovisual translation studies', in Millán Carmen and Francesca Bartrina (eds) *The Routledge Handbook of Translation Studies*, Oxfordshire: Routledge, 45-59.

- Hunston, Susan (2022) *Corpora in Applied Linguistics*, Cambridge: Cambridge University Press.
- Kacetl, Jaroslav and Blanka Frydrychova-Klimova (2015) 'English Vocabulary in Video Clips on Travel and Tourism', *Procedia-Social and Behavioral Sciences*, 182: 364-368.
- Kapsaskis, Dionysios (2018) 'Translation in the Creative Industries: An Introduction', *The Journal of Specialised Translation*, 29: 2-11.
- Kenny, Dorothy (2001) *Lexis and Creativity in Translation: A Corpus-based Study*, London: Routledge.
- Khoshsaligheh, Masood and Saeed Ameri (2017) 'Translator's Agency and Features of Non-Professional Translation of Video Games (A Case Study of Uncharted 4: A Thief's End)', *Language Related Research*, 8(5): 181-204.
- Khoshsaligheh, Masood, Saeed Ameri, Farzaneh Shokoohmand and Mehdi Mehdizadkhani (2018) 'A Socio-Cultural Study of Taboo Rendition in Persian Fansubbing: An Issue of Resistance', *Language and Intercultural Communication*, 18(6): 663-680.
- Khoshsaligheh, Masood, Saeed Ameri, Binazir Khajepoor and Farzaneh Shokoohmand (2019) 'Amateur Subtitling in a Dubbing Country: The Reception of Iranian Audience', *Observatorio (OBS)*, 13(3): 71-94.
- Khoshsaligheh, Masood, Saeed Ameri, Farzaneh Shokoohmand and Mehdi Mehdizadkhani (2020) 'Subtitling in the Iranian Mediascape: Towards a Culture-specific Typology', *International Journal of Society, Culture & Language*, 8(2): 55-74.
- Künzli, Alexander and Maureen Ehrensberger-Dow (2011) 'Innovative subtitling: A reception study', in Cecilia Alvstad, Adelina Hild and Elisabet Tiselius (eds) *Methods and Strategies of Process Research: Integrative Approaches in Translation Studies*, Amsterdam: John Benjamins, 187-200.
- Kußmaul, Paul (2000) 'Types of creative translating', in Andrew Chesterman, Natividad Gallardo San Salvador and Yves Gambier (eds) *Translation in Context: Selected Papers from the EST Congress, Granada, 1998*, Amsterdam: John Benjamins, 117-126.
- Massidda, Serenella (2015) *Audiovisual Translation in the Digital Age: The Italian Fansubbing Phenomenon*, Basingstoke: Palgrave Macmillan.
- McClarty, Rebecca (2013) 'In Support of Creative Subtitling: Contemporary Context and Theoretical Framework', *Perspectives*, 22(4): 592-606.
- McEney, Tony, Richard Xiao and Yukio Tono (2006) *Corpus-Based Language Studies: An Advanced Resource Book*, London: Routledge.

- O'hagan, Minako (2009) 'Evolution of user-generated translation: Fansubs, translation hacking and crowdsourcing', *The Journal of Internationalization and Localization*, 1(1): 94-121.
- Orrego-Carmona, David (2014) 'Using Non-Professional Subtitling Platforms for Translator Training', *Rivista internazionale di tecnica della traduzione*, 15: 129–144.
- Pérez-González, Luis (2007) 'Intervention in New Amateur Subtitling Cultures: A Multimodal Account', *Linguistica Antverpiensia, New Series–Themes in Translation Studies*, 6: 67-80.
- Pérez-González, Luis (2014) *Audiovisual Translation: Theories, Methods and Issues*, London: Routledge.
- Pérez-González, Luis (ed.) (2019) *The Routledge Handbook of Audiovisual Translation*, London: Routledge.
- Rajendran, Tamilarasu and Mohd Shah Parilah (2020) 'Students Perception on Gamification: The Use of Kahoot', *International Journal of Scientific and Research Publications (IJSRP)*, 10(05): 773-783.
- Reppen, Randi (2010) *Using Corpora in the Language Classroom*, Cambridge: Cambridge University Press.
- Riccardi, Alessandra (1998) 'Interpreting strategies and creativity', in Ann Beylard-Ozeroff, Jana Králová and Barbara Moser-Mercer (eds) *Translators' Strategies and Creativity*, Amsterdam: John Benjamins, 171-179.
- Secară, Alina (2018) 'Surtitling and captioning for theatre and opera', in Luis Pérez-González (ed.) *The Routledge Handbook of Audiovisual Translation*, London: Routledge, 130-144.
- Thomä, Sandra (2003) '*Creativity in Translation. An Interdisciplinary Approach*', PhD thesis, University of Salzburg, Salzburg.
- Toury, Gideon (2012) *Descriptive Translation Studies: And Beyond*, Amsterdam: John Benjamins.
- Vintar, Špela (2016) 'A Bird's Eye View of Lexical Creativity in Original vs. Translated Slovene fiction', *Perspectives*, 24(4): 591-605.
- Wilss, Wolfram (1988) *Kognition und Übersetzen: Zu Theorie und Praxis der menschlichen und der maschinellen [Cognition and translation: On the theory and practice of human and machine]* (Vol. 41), Berlin: Walter de Gruyter GmbH & Co KG.

Zanettin, Federico (2014) *Translation-Driven Corpora: Corpus Resources for Descriptive and Applied Translation Studies*, London: Routledge.