

An Analysis of Translation Shifts in Subtitles:

Comparing the Act of Subtitling between English-Japanese and Japanese-English*

Miki IMAI

1. Introduction

This paper attempts to analyse translation shifts in the act of subtitling between English and Japanese from the standpoints of lexical diversity, parts of speech (PoS), and word frequency. In today's globalised society, it is impossible to avoid multicultural communications. Therefore, translation plays an important role in international communications. Due to the diversification of audiovisual media, an infinite array of multimedia content is available worldwide. This means foreign media content can shrink the distance to its viewing audiences anywhere in the world. In line with that, the demand for translation, and subtitling in particular, has been increasing. In subtitle translation, the original information is often deleted or paraphrased because of time and space constraints. Furthermore, subtitlers are required to make appropriate translations for the target culture. There are, therefore, many cases where translating original text literally would not be possible. What are the differences between the source language and the translated subtitles? It is important to understand the differences in order to understand foreign languages and intercultural communications. Thus, in this study, in order to clarify the tendency of transformation caused by the shift from original text to subtitles, an English film with Japanese subtitles and a Japanese film with English subtitles have been selected as a case example, and the transformation of lexis and parts of speech has been examined.

2. Background Research

2.1. Main Characteristics of Subtitles

Audiovisual translation—subtitling, in particular—has different features as compared to other forms of translation, such as technical translation and literal translation, in several ways. Audiovisual media contains various elements such as sound, image, and text. When subtitles are added thereto, the cognitive load of the viewer is further increased. Subtitlers, therefore,

* I am sincerely grateful to Professor Ippei Inoue of Keio University and Mr. Peter Durfee, executive editor of Nippon.com and part-time lecturer of Keio University for their insightful comments and suggestions.

strive to create readable subtitles that do not burden the viewer. The most notable characteristic of subtitling is the character limit, which depends on time and space constraints. Time constraint means that the number of characters in subtitles is limited depending on the amount of time in which the dialogue is spoken. Subtitlers must consider whether to cut any information when the speech time is too short, to translate all the dialogue information. In addition, a maximum length of two lines is recommended in subtitles; this is known as space constraint (Diaz-Cintas & Remael, 2007, pp. 82-83). For subtitles translated from English to Japanese, the current norm is the 4CPS (character per second) rule. Furthermore, up to 6 seconds per subtitle presentation time and up to 2 lines with 13 characters per line are the standard. Meanwhile, in the case of English subtitles, 12CPS and 2 lines with 40 characters per line is the standard in Japan so far, even though up to 15-17CPS is now being accepted overseas (Diaz-Cintas, 2013). Due to these limitations, translation strategies such as deletion and paraphrasing are applied in subtitling, if it is difficult to include all the original information.¹

In recent years, various studies have been conducted on subtitle translation, but many have specifically and qualitatively analysed the culture-specific items of source text (Shinohara, 2013, 2018) and aspects of politeness (Usie & Nishio, 2009). As for the quantitative approach, Hosaka, Toyota, and Shimada (2019) in their studies used text mining to examine the overall tendency of transformation between ST (source text) and TT (target text) in English subtitles of Japanese movies. Their study, however, analyses only English subtitles of Japanese films. To objectively grasp the translation shift of both directions, one must observe Japanese subtitles of English films, too, in order to get a new perspective. In this research, therefore, subtitling activity of both directions, English-Japanese (EN-JP) and Japanese-English (JP-EN), are analysed.

2.2. Translation Shift

While discussing changes in translation, Catford (1965, pp.73-82) uses the expression 'translation shift'. When translating ST to TT, various transformations are performed to make a translation appropriate for the TT. He considers two kinds of shifts: (1) shift of level and (2) shift of category. A level shift refers to what is represented by grammar in one source language and lexis in another. His analysis is mostly directed to category shifts that are

¹ However, this tendency is not applicable to fansubs, which have different characteristics (Shinohara 2018, p100).

divided into four types: (a) structural shifts (shifts in grammatical structure), (b) class shifts (shifts of parts of speech), (c) unit shifts (shifts of language units such as sentences, clauses, phrases, words and morphemes) and (d) intra-system shifts (choosing a non-corresponding term in the TT system). Although these translation shifts are limited to grammar, they have pioneered the scientific analysis of translation. This paper mainly analyses data by focusing on the category shifts.

3. Method

3.1. Data

This study uses subtitle data from two films – one is a Japanese film with English subtitles titled *The Wow-Choten Hotel* (2006), and another is an English film with Japanese subtitles named *Love Actually* (2004). *The Wow-Choten Hotel* is a comedy drama in which a variety of unexpected things happen to the employees and guests of a hotel on New Years' eve. *Love Actually* is widely known as a heart-warming romantic comedy set in London just before Christmas, and features a large cast. Both films use the same method of 'ensemble casting', which comprises of multiple principal actors. Since various stories of various characters are alternatively described in both the movies, it is considered to be suitable for observing how it appears in subtitles. Both analyses used Japanese subtitles and English subtitles recorded in the DVDs.

3.2. Analytical Procedure

In this research, three points have been focused on to compare the changes between the original sentence and translated subtitles of the English and Japanese film – lexical diversity, parts of speech, and word frequency. Firstly, in order to compare lexical diversity, the token (the total number of words) and the type (the number of distinct words) of each are calculated. Next, the tendency of PoS conversion is analysed by examining the frequency of appearance in the original text and translated subtitles. The PoS picked up in this research are nouns, adjectives, adverbs, and verbs. Then, the top 10 frequent vocabularies of each PoS is taken up and compared. The analytical tool KH Coder² has been applied for the word count and morphological analysis.

² KH Coder is a free software for quantitative content analysis or text mining developed by Koichi Higuchi. It is also utilised for computational linguistics (<https://kncoder.net/>).

4. Results and Discussion

4.1. Lexical Diversity

Table 1 shows the total number of tokens and types for the original text and translated subtitles of each film. With respect to the total number of tokens (N), the translated subtitles of both films decreased as compared to the original text, with a 35% decrease in *The Wow-Choten Hotel* (English subtitles) and a 33% decrease in *Love Actually* (Japanese subtitles). This indicates that there is no big difference in the ratio of word reduction between EN-JP and JP-EN subtitling. On the other hand, a large difference has been found in the total number of types (V). For *Love Actually*, there is no large change in the number of types whereas *The Wow-Choten Hotel* shows a decrease in the number of words in the English subtitling as compared to the original Japanese text.

Based on these results, Table 1 shows the lexical diversity as Type-Token Ratio (TTR).³ In both films, the TTR value is higher in TT than in ST, but *Love Actually* has a higher increase in TTR as compared to *The Wow-Choten Hotel*. In sum, this figure infers that Japanese subtitles in English films become richer in vocabulary through translation as compared to English subtitles in Japanese films.

The Wow-Choten Hotel

	V	N	TTR
ST (Japanese)	2,422	16,703	0.145
TT (English)	1,691	10,761	0.157

Love Actually

	V	N	TTR
ST (English)	1,845	11,871	0.155
TT (Japanese)	1,854	7,905	0.235

(N= number of tokens, V= number of types, TTR= Type-Token Ratio)

Table 1: Tokens and Types

4.2. Parts of Speech

Table 2 indicates the number of tokens, types, and TTR for each PoS —nouns, adjectives, adverbs, and verbs.

³ TTR is a value obtained by dividing the number of types (V) by the number of tokens (N): $TTR = V/N$, and is mainly seen as an indicator of vocabulary richness (Kujirai, 2012).

In an analysis of nouns and verbs, there is not much difference in the nouns of the Japanese original text and English subtitles, while there is a difference in the ratio in the case of English original text and Japanese subtitles. There is a 10.2% increase in nouns when the English original text is converted to Japanese subtitles. On the other hand, the rate of verb change shows the same tendency in both EN-JP and JP-EN, wherein English has about 10% more verbs than Japanese. In summary, these suggest the following – the reason why the number of nouns do not change in JP-EN is because what is expressed in Japanese by ‘nouns + particles’ is replaced by ‘nouns + verbs’ in English. Furthermore, in JP-EN translation, it can also be said that the information of nouns is not largely deleted through translation. In contrast, the reason for the increase in the number of nouns and the decrease in the number of verbs in EN-JP is presumed to be because what is expressed in English as verbs is expressed in Japanese as nouns.

With regard to adjectives and adverbs, it is believed that there is a correspondence between the change rate in EN-JP and JP-EN even though the number of words in the case of Japanese subtitles is significantly reduced. Furthermore, differences in lexical diversity in adjectives and adverbs are observed. While the adjective TTR goes up in JP-EN, no major change is seen in EN-JP. As for adverbs, TTR values are lower in JP-EN, while they are significantly higher in EN-JP. It can be inferred that in the case of subtitling into English, the vocabulary of adjectives becomes rich, and in the case of subtitling into Japanese, the vocabulary of adverbs becomes rich.

The Wow-Choten Hotel

ST (Japanese)					TT (English)				
		N	V	TTR		N	V	TTR	
Noun	17.2%	2,866	1,057	0.37	Noun	18.2%	1,955	801	0.41
Adj/Adj V	3.6%	594	168	0.28	Adj	5.6%	604	254	0.42
Adv	4.9%	812	210	0.26	Adv	8.3%	898	123	0.14
Verb	13.3%	2,223	487	0.22	Verb	23.5%	2,529	337	0.13
Total		16,703	2,422		Total		10,761	1,691	

Love Actually

ST (English)					TT (Japanese)				
		N	V	TTR		N	V	TTR	
Noun	18.3%	2,197	884	0.40	Noun	25.3%	1,997	1,039	0.52
Adj	6.5%	767	321	0.42	Adj/ AdjV	4.6%	367	142	0.39
Adv	9.3%	1,107	136	0.12	Adv	4.4%	346	134	0.39
Verb	20.9%	2,483	301	0.12	Verb	11.7%	926	321	0.35
Total		11,871	1,845		Total		7,905	1,854	

(N = number of tokens, V = number of types, TTR = Type-Token Ratio)

Table 2: PoS Frequency

4.3. Word Frequency

In order to analyse the complete picture regarding the transformation of the appearance of vocabulary, the word frequency of ST and TT of each film was extracted according to the PoS. Table 3 shows the results of extracting the top ten frequent words.

In *The Wow-Choten Hotel* (JP-EN), the word common to the top ten nouns in ST and TT is ‘hotel’. Although this is a reasonable result considering the film is set in a hotel, the other words are completely different. What can be inferred from this is that those who watch subtitles in English may have different ways of capturing the nuances and atmosphere of the film. For example, happenings related to deer and ducks are drawn with comicality in this story, and the interesting thing is that while the ST has many ‘deer’ (鹿), the subtitles have many ‘ducks’. This suggests that the way of communicating humour may also be different between JP and EN. Moreover, this film is set at the year’s end, and that has been emphasized in English subtitles, since the word ‘year’ is the most frequent noun used. With regard to *Love Actually* (EN-JP), the words common to the top two nouns in ST and TT are the same – ‘Christmas’ and ‘love’. The film is often mentioned as one of the best films for Christmas, with Christmas and love as its main topics. This is reflected in the Japanese subtitles. Some commonalities in adjectives are seen in ST and TT, which implies that Japanese subtitles appropriately express the essence of information in ST despite the character limit.

The Wow-Choten Hotel

ST (Japanese)								
	Noun	freq.	Adj	freq.	Adv	freq.	Verb	freq.
1	人	55	いい	73	もう	61	する	272
2	先生	37	ない	63	そう	53	なる	138
3	お願い	32	よい	24	ちよつと	52	言う	84
4	ホテル	31	悪い	18	どう	36	ある	71
5	仕事	30	よろしい	16	すぐ	20	いる	60
6	鹿	27	早い	13	まだ	16	思う	56
7	あと	26	痛い	6	むっちゃ	16	やる	51
8	支配人	22	忙しい	5	これから	13	行く	29
9	自分	21	寒い	4	いつも	11	分かる	27
10	話	21	小さい	4	さっき	10	待つ	24

TT (English)								
	Noun	freq.	Adj	freq.	Adv	freq.	Verb	freq.
1	year	47	sorry	22	not	257	be	645
2	way	32	new	20	here	53	do	216
3	time	23	good	14	too	42	have	120
4	room	21	big	13	just	36	get	89
5	thing	21	happy	13	so	32	go	81
6	hotel	20	okay	12	now	30	know	46
7	duck	19	cold	11	really	28	let	45
8	party	19	ready	11	back	24	want	45
9	guest	17	last	10	then	19	come	42
10	home	17	quick	10	always	17	say	41

Love Actually

ST (English)								
	Noun	freq.	Adj	freq.	Adv	freq.	Verb	freq.
1	Christmas	66	good	35	not	234	be	747
2	love	53	sorry	26	just	85	do	193
3	God	30	great	14	so	56	have	126
4	baby	28	more	13	right	50	know	82
5	girl	23	sure	13	here	47	get	65
6	sir	23	little	12	well	46	go	62
7	time	22	fine	11	very	36	want	56
8	night	21	nice	11	now	34	come	53
9	day	20	bad	10	really	30	think	52
10	something	19	big	10	never	26	make	37

TT (Japanese)								
	Noun	freq.	Adj	freq.	Adv	freq.	Verb	freq.
1	クリスマス	56	いい	38	そう	32	する	97
2	愛	19	よい	22	本当に	17	言う	32
3	首相	17	ない	19	どう	14	いる	30
4	アメリカ	14	悪い	17	すぐ	11	ある	24
5	英国	13	欲しい	15	もう	11	なる	23
6	女	13	楽しい	8	もちろん	11	分かる	21
7	電話	12	うれしい	7	なぜ	7	待つ	17
8	ハバ	10	ずばらしい	7	とても	6	来る	17
9	プレゼント	10	悲しい	6	いつも	5	見る	16
10	人	10	美しい	6	もっと	5	知る	15

Table 3: Top 10 Frequent Words

In addition, the frequency of character name appearances in each film is calculated as shown in Table 4. As mentioned in the previous section, both films use the same method of ensemble casting, which comprises of multiple principal actors. Therefore, there are many sequences in which the names of various characters are mentioned in the films. Let us see how it appears in the subtitles.

On the whole, JP-EN and EN-JP both do not mention character names as much, when converted to subtitles. The decrease rate in mentioning character names is less in EN-JP (*Love Actually*), with the rate of decrease being about 40% in JP-EN and about 26% in EN-JP.

In *The Wow-Choten Hotel*, the most frequently mentioned names are Shindo, Mutoda/Senator/Sensei, Tokugawa/Dai-Sensei, and Kenji, but the number of mentioned names decrease when converted to subtitles. This suggests that the name of the character is replaced by a personal pronoun in English subtitles. On the other hand, in the case of conversion to Japanese subtitles in *Love Actually*, there is almost no reduction in the number of mentioned names except for Billy. In the case of Japanese subtitles, since the personal pronouns are often omitted, the proportion of character names do not decrease. Therefore, it can be inferred that viewers tend to recognise character names in Japanese subtitles more than in English subtitles.

The Wow-Choten Hotel

Name	ST (JP)	TT (EN)
Shindo	49	27
Yabe	4	4
Kenji	18	11
Seo	12	6
Hana	7	4
Tange	3	0
Mutsuko	6	0
Chairman	16	16
Mutoda/ Senator/ Sensei	34	23
Yoko	5	3
Yumi	1	1
Hotta/ Professor/ Sensei	13	7
Naomi	13	11
Tokugawa/ Dai-sensei	21	5
Cherry/ Sakura	8	8
Rub-A-Dub	14	8
	224	134

Love Actually

Name	ST (EN)	TT (JP)
David/ Prime Minister	17	21
Natalie	17	15
Jamie	8	2
Aurelia	9	8
Juliet	4	2
Peter	6	5
Mark	5	3
Sarah	5	4
Karl	6	4
Billy	21	7
Karen	4	2
Harry	2	1
Mia	3	3
Daniel	3	3
Sam	10	8
Joanna	5	5
Colin	6	6
Tony	2	2
Judy	6	3
John	5	2
	144	106

Table 4: Frequency of Character Names

5. Conclusion

A comparative examination of the act of subtitling between English-Japanese and Japanese-English demonstrates differences in (a) lexical diversity, (b) the tendency of PoS conversion, and (c) frequency of vocabulary used. The differences can be summarised as follows.

In English-Japanese subtitling, lexical diversity is retained even if the number of words in subtitles is reduced through translation. As for the parts of speech, the proportion of verbs is reduced, and the vocabulary of adverbs is diversified. Additionally, the omission of character names tends to be very small despite the character limit. On the other hand, in Japanese-English subtitling, lexical diversity is reduced as compared to Japanese subtitles. Many words are replaced by personal pronouns, the proportion of verbs increases, and the adjective lexical diversity increases. In addition, the reflection of character names in subtitles tend to be less than Japanese subtitles.

The result of this study may provide several suggestions not only with regard to the translation methods but in taking advantage of films for language education. By examining both directions of English and Japanese translations, several features of the two languages can be discovered. In the case of subtitling that translates under various constraints, the features and differences of each language would appear more clearly. However, further research is required because the amount of data collected in this study is not enough to determine the general characteristics of subtitle translation. Therefore, one would need to collect and analyse more subtitle data from various genres of films. In addition, qualitative analysis of the texts is also needed as this study only conducted a quantitative analysis. With globalisation, the demand of translation is expected to increase exponentially. Therefore, analysing various translation data from a variety of angles would be significant in the current intercultural communication society.

Bibliography

- Anzai, T. (2000). *Eigo no hassou*. [Concept of English]. Tokyo: Chikuma Shobou.
- Anzai, T., Inoue, K., & Kobayashi, A. (2005). *Honyaku wo manabu hito no tame ni* [For those who learn translation]. Kyoto: Sekai Shisou Sha.
- Catford, J.C. (1965). *A linguistic theory of translation*, London: Oxford University Press.
- Diaz-Cintas, J. and Remal, A. (2007). *Audiovisual translation: Subtitling*. New York: Routledge.
- Diaz-Cintas, J. (2013). Subtitling: Theory, practice and research. In C. Millan, & F. Bartrina (Eds.), *The Routledge hand book of translation studies* (pp. 273-287). London: Routledge.
- Hosaka, T., Toyota, T., & Shimada, M. (2019). Eizou sakuhin no nihongo wa jimaku honyaku ni yotte nani ga kawaru no ka. [What does Japanese translation of film change with subtitle translation?]. *The 43rd Meeting of the Japanese Association of Sociolinguistic Sciences (JASS) Proceedings*, 14-17.
- Kujirai, A. (2012). The quantitative condition of repeated nouns in writing: A type-token ratio analysis. *Mathematical Linguistics*, 28(6), 211-225.
- Shinohara, Y. (2013). A study of strategies for translating culture-specific items in the English subtitles in the film *Departures*. *Invitation to Translation Studies*, 9, 81-98.
- Shinohara, Y. (2018). *Eiga jimaku no honyaku gaku* [Translation studies of film subtitles]. Kyoto: Koyo Shobo.
- Ushie, Y., & Nishio, M. (2009). Politeness in English subtitles of Japanese movies. *Interpreting and Translation Studies*, 9, 253-272.