Comparison of Features of Texts Translated by Professional and Learner Translators

Sachiko Nakamura[1]

Introduction

The history of translation studies in Japan is still very young. According to Mizuno (2006: 1), it was only in 1990s when a handful of researchers started to publish what is later regarded as translation studies. Most researchers work on studies on English to Japanese translation of literary works as an academic discipline. Extensive and full-scale studies on translation for practical purposes and those of from Japanese to English, for that matter, are yet to be seen. Despite the slow progress in development of theories on both product-oriented and process-oriented translation, demands for Japanese to English translation from an industrial community have been on the rise in such fields as manufacturing, medical science, IT and other numerous industrial and commercial sectors, thus a category of translation as commercial products has begun to burgeon.

In Japan, education and training on translation has been offered primarily by non-formal translator training schools, where course participants are often advanced adult learners of English. The foremost purpose of these schools is to foster professional translators who can cater to growing commercial needs. Those who wish to become a professional translator reasonably knock on the doors of such schools. It is only in the new millennium when under-graduate or graduate courses of Japanese universities started to offer translation programs as part of their English language curricular reform. More recently, some graduate courses have begun to organize advanced specialist programs incorporating commercial-oriented translation and interpretation education and training. Finally Japanese universities are poised to pay attention to translation for practical purposes. This paper focuses a translation program offered at a graduate course designed to foster professional translators for commercial and technical fields.

[1] Chukyo University

e-mail: sachi@mse.biglobe.ne.jp

1. Aim of the Research

The purpose of this research is to demonstrate differences of textual features, specifically of syntactic and lexical choices between professional and learner translators, thereby exploring factors affecting Japanese to English translation. Learner corpora are compiled from texts translated from Japanese into English by learner translators.   Whether problems appearing in their renditions are common or not is judged by comparing them with professional translators' corpus as well as with a large-sized corpus of the similar discourse texts.

2. Hypothesis

As a preliminary hypothesis about characteristics of learner translator corpora (hereinafter referred to as LC) vis-à-vis those of professional translator corpus (hereinafter referred to as PC), I assume learner translator corpora would present some or all of following features:
  a.   more explicit textual structures
  b.   more high frequency words than low frequency words
  c.   greater repetition
  d.   less lexical and syntactic variety
  e.   more literal translation

A concordancing software is used to analyse the followings in order to support my hypothesis:

Hypothesis 1.   Frequency of relatives: the frequency of *which* will be counted for both LC and PC.   I assume professional translators incorporate more diverse syntactic and lexical choices in their rendition.   They may use relatives to add detailed information but they will also use participle phrases and appositions instead.

Hypothesis 2.   Type-token ratio: learner translators may repeat the same words in their rendition, thus their lexical density would be lower than that of professionals. This may result in the less colourfulness, less variety of the representation.

3. Subjects

Subject learners are four advanced learners of English attending a Master's program of a university in Nagoya, Japan.   The course participants are one high-school

English teacher, taking one-year in-service recess, one housewife wishing to become a professional translator/interpreter, two immediate university graduates wishing to find a full-time position in business community. All have an experience of studying abroad or of staying some years overseas, so their English proficiency level is higher than ordinary English language learners.

4. Theoretical Framework

It has been widely believed among researchers in translation studies that whatever the language pair involved translated language tends to show greater cohesion and explicitation (Blum-Kulka, 1986; Baker, 1992, 1995; Hatim and Munday, 2004). Blum-Kulka suggests this "explicitation is a universal strategy inherent in the process of language mediation, as practiced by language learners, non-professional translators and professional translators alike" (1986:21).   Therefore, it might be all right to say that the more the number of explicitation devices (e.g. connectives, conjunctions, etc.) in some translated texts, the higher the degree of explicitation than others.

Halliday explains meaning of a non-defining nominal relative *which* as follows: "A non-defining relative clause (abridged) adds a further characterization of something that is taken to be already fully specific" (Halliday, 1994: 227).   He also explains defining *which* as an "embedded clause that functions as elaboration and expansion (op.cit.: 243)."   Although there are other cohesive devices such as appositions and participle phrases that function also to elaborate, expand and add, many learner translators tend to use a relative pronoun *which* for these purposes, sometimes excessively.   Because of the same recurring pattern, their translation products tend to be monotonous, redundant and less colourful compared with those of professionals' who use other devices as well.   Overuse of *which* functions as if it is a cohesive device, and it might result in a higher degree of explicitation.

Concerning the idiosyncratic ways of using specific syntactic structures, Widdowson (1979: 44) suggests that there is a common characteristic found in the work of translators.   According to him, such tendency could represent the translator's individual behaviour.   Overuse of *which* in learner translators' texts of this study, however, may suggest that it could be a problem common to Japanese learner translators, and thus I shall regard this as a supra-individual idiolect, or a collective idiosyncrasy.

5. Data Analysis

5.1 Learner Translator Corpus (LC)

Four learners (A, B, C, D) worked on following three different categories and the corpus structure and tokens are shown in Table 1 below:

    a. Journalism texts: magazine articles on Japanese *sake* bar and *kabuki* theatre, travel guidebooks, corporate brochures

    b. Newspaper editorials

    c. Contract agreements

	Sub-corpus 1	Sub-corpus 2	Sub-corpus 3
Learner A	1,435	1,210	243
Learner B	1,438	1,222	299
Learner C	1,312	1,156	301
Learner D	1,332	1,173	341

Table 1: Structure of learner sub-corpora and word counts.

5.2 Professional Translator Corpus (PC)

Similarly, texts from the original magazines and brochures are named Professional Translators Corpus (PC).

	Sub-corpus 1	Sub-corpus 2	Sub-corpus 3
PC	1,302	1,300	307

Table 2: Structure of professional sub-corpora and word counts.

5.3 A Concordancing Software and Basic Data

The text concordancing software used for this study is called KWIC Concordance for Windows version 4.7 created by Satoru Tsukamoto. First, each student's translation texts were categorised into three genres, saved as text files, concordanced, then, word lists were produced. Each KWIC wordlist was respectively saved as KWIC files from KWIC 1 to KWIC 4, and professional translator corpus was also concordanced

and a word list was saved as KWIC 5 pro. Frequency of relatives (which, what, where, who, when, whose, etc.) was counted and the type-token ratio was also calculated as shown Table 3.

	KWIC1	KWIC2	KWIC3	KWIC4	KWIC5 pro
token	1,468	1,498	1,369	1,368	1,335
which	17	12	8	6	1
where	1	2	5	2	2
who	2	2	1	2	1
when	0	2	0	2	1
which ratio	0.0115803	0.008010	0.0058436	0.004329	0.0007490
occurrence ratio	1/105 times	1/125	1/171	1/231	1/1,335
type/token ratio	0.446866	0.442590	0.480643	0.432900	0.501124

Table 3: Frequency of relatives in corpora.

The result suggests that the use of *which* is more salient in LC compared to PC.   In fact, *which* occurs more frequently in LC while it occurs only once in PC. Furthermore, frequency of *which* ranges from once in every 105 words to 231 words in LC, whereas once in every 1,335 words in PC.   The use of other relatives does not show striking differences as *which* does.   Whether this seemingly too frequent appearance of *which* is above average, thus regarded as a deviation from generally accepted usage, can be judged by looking at the frequency of *which* appearing in large-sized corpora such as BNC or Cobuild Bank of English.   Unfortunately though, in Japan, these corpora are not readily available at reasonable costs through the Internet or some other media to be used by general practitioners such as learners and professional translators.

5.4 Target Discourse Community Corpus

In order to cope with the shortcoming of unavailability of public corpora, and to help out language users in respective target discourse community (TDC), I have compiled a TDC corpus representing language used in specific target discourse communities including journalism, politics, legalism, business, science and technology and have

named it TDC Corpus. It contains some 1.2 million words of written texts (token). For details of the corpus structures, refer to Nakamura (2002).

5.4.1 *Which* Counts in TDC Corpus

TDC corpus consists of three main categories – magazine, newspaper editorials and legalism. The frequency of *which* is counted in the magazine category and the results are shown in Table 4 below.

token	*which* count	*which* ratio	occurrence	type-token
1,170,267	2,569	0.002195225	1/455	0.036876

Table 4: Frequency of *which* in TDC corpus – magazine category.

The results show that *which* occurs 2,569 times out of total token of about 1.17 million words, which is roughly translated into a frequency of one in every 455 words.   Compared to the results in Table 3, all LC corpora have higher frequency of *which* than the TDC does.   In Chapter to follow, I shall pick up some of typical examples from LC and PC in the category of magazine articles.

Example 1 (Japanese *sake* bar)
```
LC1a
Popular chopped gristle (only 220 yen each), which was picked up on
NHK cooking show, is one for each because it is precious.
LC1b
"Nankotsu no tataki" or lightly chopped gristle costs only ¥220 per
skewer, which was introduced on NHK's cooking show, is so popular and
rare that a limit per person is one skewer.
LC1c
Famous dish 220 yen "nankotsu tataki" which is chopped gristle once
introduced NHK cooking program, because its so precious you can only
order one. (sic.)
Pro1
The popular "nankotsu no tataki" or lightly roasted chopped chicken
gristle (¥220/skewer), once introduced in NHK's cooking show, is
scarcity and limited to one skewer per person.
```

In LC, three out of four learners insert *which was* and *which is* in the middle of subsequent clause. This type of usage is also seen in conversations (Long 1961:314), suggesting it a more or less informal style and leading to possible redundancy. On the other hand, Pro1 uses *once introduced*, a participle phrase, instead of *which was*. The Pro1 is thus successful in presenting an articulated, concise journalistic style.   Following examples 2 to 7 show a repeated use of *which* in learner texts and no use in professional text.

Example 2 (Kabuki)
```
LC2
Countless stage settings were newly invented in Japan; for instance,
"Mawari Butai (Turning stage)", designed for an instant change, and
"Seri", which lifts a Kabuki actor from the bottom to the stage.  Kabuki,
which was in the pursuit of the entertainment for the ordinary people,
had created the world of spectacles long before the Hollywood did.
Pro2
The revolving stage, a gimmick that changes scenes instantaneously,
or seri, a lift that raises actors, making them appear from the
understage, and other numerous innovations were the spectacular world
first brought about via Kabuki.  In effect, Kabuki was a pioneer in
entertaining the public, a long time before Hollywood.
```

This learner LC2 uses *which* twice in the text while Pro1 does not use it at all. Instead, Pro1 uses an apposition and present participle constructions.   Repetitive use of non-defining relative pronoun gives a somewhat colloquial and redundant impression to readers.   The Pro2 text develops from specific to general, a textual structure quite different from LC.   Furthermore, Pro2 text presents an articulated prosodic rhythm.

Example 3 (Travel guidebook)
```
LC3
Ueno Castle is known as a castle which Toudou Takatora, a great castle
constructor, expanded.  The highest stone wall in Japan which is about
30 meters, double walled moat, and excessively robust castle tell you
how hard times the Sengoku era was.  You can visit nearby Ninjya
residence and Ninjya resource centre.
Pro3
```

```
Ueno Castle was expanded from its original form by master castle
builder Todo Takatora.  Surrounding the castle are 30 m high stone
walls – the tallest in Japan – and double moats.  The robust castle
walls stand as a reminder of the turbulent warring states period.  An
Iga-style Ninja House and Ninja Museum are nearby.
```

This is a typical description in a travel guidebook intended to be read by the general public. Common characteristics of these documents are of lighter mood, faster pace, and often some fun elements are embedded in the texts. Repetitive use of *which* in LC3 presents rather heavier, slower, monotonous and even serious impression. Pro3 text, on the contrary, keeps an airy mood and diversity. No complex structures being used, Pro3 is more reader-friendly.

Example 4 (Corporate brochure)
```
LC4
With our product development capability for vehicle installation which
Auto Networks Technologies have nurtured, we always are heading for
one-up society and create new possibilities to go ahead of the area
of cutting-edge technology, which is based on the fundamental research
of Sumitomo Electric Industries and the evaluation level of Research
& Evaluation Center.
Pro4
Basic research is conducted by Sumitomo Electric Industries, and
testing technologies are provided by the Research & Evaluation Center.
Supported by these basic technological research capabilities and
automotive product development capabilities nurtured in Auto Networks
Technologies, we create new potential in many areas of cutting-edge
technology.
```

This document is taken from a corporate brochure. The possible reader of documents of this kind could be visitors such as potential clients. So readers are not necessarily the general public, but more often than not, are specialists sharing the same discourse community knowledge and experience. Reader-friendliness still needs to be achieved. Reasonably some terminologies are involved in these texts, however, they are not the issue here. What draws my attention is the length of sentences in LC4. Because this learner translator adds extra information by connecting clauses with a relative pronoun *which*, the sentences result in longer and complex.

Example 5 (Travel guidebook)

```
LC5
A prestigious shrine where Amatsu Hikonenomikoto, a son of Amaterasu
Omikami, the ancestral goddess of the Imperial Family is worshipped.
This shrine is also well-known for the Horseback Climbing Ceremony
in which horses are ridden swiftly and attempt to climb a steep hill,
forecasting abundance or scarcity of agricultural harvests.
Pro5
This shrine worships Amatsu-Hikone-no-Mikoto, the Child of the Shinto
sun goddess Amaterasu Omikami. It is famous for a ceremony called Age
Uma Shinji in which the fortunes of the year's crops – abundant or
poor – can be discovered depending on whether a horseback rider is
able to climb a cliff.
```

Although *in which* is used at the same position in LC5 and Pro5, the subject in subsequent clauses differs. In LC5 subject is *horses* while in Pro5, it is *fortunes of this year's crops.* Of course it is not the horses that perform forecasting, so LC5 translator chose a wrong subject. By the way, *in which* in Pro5 is the only one *which* occurred in the whole PC.

6. Discussion

The purpose of this text genre – bar and restaurant information, cultural information, travel guidebook and corporate brochure – is an advertisement to the outside world. For those documents, reader-friendly textual style is preferred. Articulation, conciseness, prosodic rhythm and other journalistic textual elements are required. Too many use of *which* may run counter to the needs of such documents, because it may generate lengthy, redundant impression. It appears that *which* in learner texts, in most cases, is non-definitive usage, thus it serves as a conjunction, e.g. *and*, therefore, adding extra information to the antecedent is the main function of such usage. In translation examples above, learner translators almost always choose *which* for addition, elaboration and expansion. Because of this overuse of *which,* learner translators renditions tend to present a greater explicitation.

I suspect this tendency has something to do with English education in their secondary level. In the secondary education (junior-high and senior-high schools), much longer time is spent on relative clause learning in a grammar class. For instance, much time is spent on such exercise as rewriting from a. to b.

a. A girl is standing by the window.　She is Jane.
b. A girl who is standing by the window is Jane.

Connecting two sentences requires a higher level of syntactic knowledge and technique, so learners reasonable tend to consider b. as more advanced usage.　Also the relationship with the antecedent can be more explicitly demonstrated.　Even though where a relative clause is unnecessary, learners tend to leave it to maintain the explicitness, and perhaps there is a psychological factor urging them to demonstrate their syntactic knowledge as they were often so trained in their previous language learning experience.

7. Conclusion

The purpose of this study was to demonstrate differences of textual features between professional and learner translators.　It was hypothesised that learner translator corpora would present greater repetition which might result in more explicitness in their rendition, while professional translators would incorporate more diverse syntactic and lexical choices.　KWIC for Windows concordancing software was used to obtain statistical data: type-token ratio and frequency count of *which.*　The type-token ratios of learner corpus were lower than that of professional corpus, suggesting that the syntactic and lexical features of learner translators' renditions have less colourfulness and variety.　The computational analytic results showed that learner texts involved higher frequency of *which*, ranging from 6 times to 17 times higher than that of professional texts.　A tendency of overusing *which* was commonly found in the learner translators of this study.　Many instances showed that they did not opt for other cohesive devices but almost always chose *which* for addition, elaboration and expansion.　It is, thus, safe to say that the results obtained from computational data as well as from the inductive data both support the initial hypothesis.

However, in order to investigate whether the tendencies found among the learner translators under this study are problems common to other learner translators, I should look into a learner translator corpus of a larger size.　At the same time, in order to look into the appropriateness of lexical choices by learner translators, I should investigate lexical combinations and collocations.　I should like to deal with them in my next research.

References

Baker, M. (1992) In Other Words: A Coursebook on Translation. Oxon: Routledge.

Baker, M. (1995) 'Corpora in Translation Studies: An Overview and Suggestions for Future Research'. *Target 7.2*: pp. 223–43.

Blum-Kulka, S. (1986) Shifts of cohesion and coherence in translation in J. House and S. Blum-Kulka (eds.) Interlingual and Intercultural Communication: Discourse and Cognition in translation and Second Language Acquisition Studies. Tubingen: Gunter Narr.

Halliday, M.A.K. (1994) An Introduction to Functional Grammar Second Edition. New York: Arnold.

Hatim, B and I. Mason (1990) Discourse and the Translator. London: Longman.

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

Long, R.B. (1961) The Sentence and Its Parts. Chicago: University of Chicago Press.

Mizuno, A. (2006) Maegaki (Foreword) in Translation Studies Special Interest Group, JAIS (ed.) Honyaku kenkyu he no shotai (An Invitation to the Translation Studies in Japan), pp.1–2. Tokyo: Japan Association of Interpreting Studies

Nakamura, S. (2002) A Study of Restrictive Collocational Relationships of Synonyms: Synonym Grid as a Differentiation Tool. Master's dissertation presented to Aston University.

Widdowson, H.G. (1979) Exploration in Applied Linguistics. Oxford: Oxford University Press.