

# 東野圭吾の小説の英訳における 文の結合と分離の分析 ——計量言語学の観点から——

大 羽 良

## 1. はじめに

翻訳者は翻訳をするとき、ソーステキストをうまく翻訳先の言語に移し替えようとするために、意識的・無意識的にいくつかの技術を忍ばせる。例えば、2010年代後半よりニューラルネットワークという枠組みを用いた機械翻訳が発達し高品質な翻訳が得られるようになったのはよく知られた事実であるが、機械翻訳された文章は人間の翻訳者が行う翻訳とは異なる点が見つかる。以下の(1)と(2)は、東野圭吾の小説を機械翻訳のDeepL<sup>1)</sup>に翻訳させた文と、実際翻訳者が翻訳した文を比較したものである。

### (1) ソーステキスト：

担当の先生は、診断書を二通書いたことをお認めになりました。ただし、二通の内容は違う。一つは本当の病名を書いたもの。もう一枚は、嘘の病名を書いたもの。 『新参者』第一章5

---

1) <https://www.deepl.com/translator>

DeepL による翻訳：

The doctor in charge admitted that he wrote two medical certificates. However, the contents of the two letters are different. One was written with the true medical name of the disease. The other one was written with a false name of the disease.

翻訳者による翻訳：

The doctor admitted to issuing two different medical certificates, one listing your mother's actual condition, the other listing a false one.

*Newcomer: 1. 5*

(2) ソーステキスト：

……久しぶりだったし、何より、あの人が今さらここへ来るとは思わなかったから、びっくりしちゃいました」 『容疑者 X の献身』 13

DeepL による翻訳：

..... I was surprised because it had been a long time and, more importantly, I didn't expect that person to come here now.'

翻訳者による翻訳：

..... It was quite a surprise, seeing him after all that time. I couldn't imagine why he would be dropping by now.'

*The Devotion of Suspect X: 13*

機械翻訳と翻訳者の翻訳結果を並べて比較すると、機械翻訳ではソースとなる日本文と翻訳された英文が一文ずつ対応するが、翻訳者の翻訳の場合、ソースとなる複数の文を一文にまとめたり、一文を複数の英文に区切って翻訳を行っている<sup>2)</sup>。

---

2) DeepL 翻訳は元の文と翻訳文が厳格に一文対一文で対応するが、Google 翻訳の場合、以下のような文章では複数文を一文にまとめる場合もあった。

ソーステキスト：

「申し訳ありませんが、本日は満室になっておりまして」ホテルマンが娘に

村上・柴田（2000）では、このような翻訳者の技術について、村上春樹が柴田元幸の翻訳ワークショップで学生の質問に答える形で説明している。

……文章を書くときにプライオリティのトップにくるものが、それぞれあるはずです。

僕の場合はそれはリズムなんです。呼吸と言い換えてもいいけど、感じとしてはもうちょっと強いもの、つまりリズムですね。だからリズムということに関しては、僕の場合によってはテキストを僕なりにわりに自由に作りかえます。どういうことかと言うと、長い文章があれば三つに区切ったり、三つに区切られている文章があったら一つにしたりとか。ここの文章とここの文章を入れ換えたりとか。

なぜそれをするかという、僕はソースのテキストにある文章の呼吸、リズムのようなものを、表層的ではなく、より深い自然なかたちで日本語に移し換えたいと思っているからです。……

村上・柴田（2000, pp.21-22）

村上の話をもとめると、ここで言及されている翻訳者の技術とは、ソーステキストが持つ作者の文体の雰囲気は保ちつつも、文をつなげたり、区切ったり、入れ替えたりして、翻訳語の自然なリズムで翻訳することである。

また Baker（1996）によると、一般的に翻訳テキストには、以下の四つの特徴が文体に現れるという。

---

いている。

『ゲームの名前は誘拐』2

Google 翻訳：

“I’m sorry, but we’re fully booked today,” the hotel manager said to his daughter.

(参考) “I sincerely apologize, but tonight we’re completely full,” the hotelkeeper was telling the girl. *The Name of the Game is a Kidnapping: 2*

明示化：ソーステキストに暗示される情報を翻訳テキストに表出すること

簡素化：ソーステキストにみられる表現を簡素化して翻訳すること

正規化：ソーステキストから翻訳される言語に合わせて変化させること<sup>3)</sup>

平準化：ソーステキストにみられる極端な文体的特徴を中庸化すること

Baker (1996) では、ソーステキストを区切って翻訳したり、ソーステキストから文の長さを大きく変化させて翻訳することは、主に簡素化と明示化に分類されるとしているが、村上の話を踏まえると、村上の目的は正規化とも考えられる。

本研究では、このような翻訳テキストに現れるソーステキストの結合・分離・並べ替え等について計量言語学的アプローチから分析・検討を行う。翻訳という行為は翻訳者の意識的行為であるが、計量言語学からのアプローチによって翻訳者の意識を超えた傾向・パターンや、複数の翻訳者に共通する傾向・パターンも捉えることが可能となり、その傾向・パターンを選択する理由を窺い知ることが可能となる。

## 2. 研究目的

本研究では、翻訳者がソーステキストから翻訳するとき、ソースとなる文を結合・分離もしくは順序の入れ替えをする行為によって、1) 翻訳テキストにどのような変化がみられるのか——全体的に文数が増減するのか、文の長さに変化がみられるのか、2) この行為はどのくらいの頻度で行われるのか、3) この行為に一般的なパターンを見出すことができるか

---

3) Baker (1996) では、実験的に書かれた文体が翻訳時には避けられやすいということは、正規化のサブカテゴリの保守化であるとしている。

という点を明らかにし、そして4) この行為はどのような目的で行われているか——簡素化や明示化のためか、それとも正規化のためかという点を検討する。

### 3. コーパスデザイン

本研究ではケーススタディとして、東野圭吾の小説を英語に翻訳した四人の翻訳者のテキストを比較する。日本のベストセラー作家である東野圭吾の小説は、複数の作品が複数の翻訳者によって翻訳されているため、東野圭吾という同じ作者によってソーステキストの文体の変動を最小限に抑えつつ、翻訳者の文体の違いを捉えることが可能となる。

比較するにあたって、ソースとなる東野圭吾の小説と対応する翻訳の本文のみをコーパス化した。使用した小説を表1に示す<sup>4)</sup>。

表1 使用した作品名

ソーステキスト		翻訳テキスト		
作品名	刊行年	作品名	刊行年	翻訳者
ゲームの名は誘拐	2002	The Name of the Game is a Kidnapping	2017	Jan Mitsuko Cash
容疑者Xの献身	2005	The Devotion of Suspect X	2011	Alexander O. Smith
新参者	2009	Newcomer	2018	Giles Murray
ナミヤ雑貨店の奇蹟	2012	The Miracles of the Namiya General Store	2019	Sam Bett

4) 実際に入力したデータは以下のとおりである。  
 『ゲームの名は誘拐』：光文社文庫 2015年5月15日35刷 / *The Name of the Game is a Kidnapping*: Vertical Inc. 2017 First edition  
 『容疑者Xの献身』：文春文庫 2008年8月15日2刷 / *The Devotion of Suspect X*: Abcus 2012 Paperback edition  
 『新参者』：講談社文庫 2013年8月9日1刷 / *Newcomer*: Abcus 2019 Paperback edition  
 『ナミヤ雑貨店の奇蹟』：角川文庫 2017年8月10日32刷 / *The Miracles of the Namiya General Store*: Yen Press June 2021 First paperback edition

ソースと翻訳それぞれのテキストを、文を単位として分割してパラレルコーパスを作成した。文を単位とするとき、日本語のソーステキストについては、句点、台詞の終わり鉤括弧『』、クエスチョンマーク・エクスクラメーションマークを文末記号であると定義した（ただし、明らかに文の中に埋め込まれたと考えられる台詞は文とはしない）。英語の翻訳テキストについては、ピリオド・クエスチョンマーク・エクスクラメーションマークを文末記号として定義した（ただし台詞で文末記号の次にクォーテーションマークがある場合は、それも文に含まれる）。

次にソーステキストの ID 番号に翻訳テキストの ID 番号を対応させた。ソーステキストが翻訳テキストに複数対応する場合、逆に翻訳テキストがソーステキストに複数対応する場合は ID を複数の ID に対応させた。対応する文がない場合は、中身が空である ID を付与した。作成したコーパスの例を表 2 に示す。

またソーステキストと翻訳テキストはそれぞれ形態素解析エンジン Juman++ (ver. 1.02)<sup>5)</sup>と品詞付与ソフトウェア Tree Tagger (ver. 3.2.5)<sup>6)</sup>を用いて単語に分離した。作成したパラレルコーパスの延べ語数の構成は表 3 の通りとなった。なお、延べ語数においては記号もすべて含めている。

日本文の場合、形態素解析エンジンによって単語の分割方法が異なり、結果としてその長さは変化する。英文に使用する品詞付与ソフトウェアも省略形や熟語等の分割方法に揺れがあるため、ソーステキストと翻訳テキストの比の値は本研究の分析内のみで扱われるもので、過去の研究結果とは比較できないことに注意する必要がある。本研究で使用したソフトウェアから得られた結果では、四作品とも翻訳テキストのほうがソーステキストよりも延べ語数が少なくなり、その割合は 83%-93% と 10%ほどの差が

---

5) <https://nlp.ist.i.kyoto-u.ac.jp/?JUMAN%2B%2B>

6) <https://cis.uni-muenchen.de/~schmid/tools/TreeTagger/>

表2 構築したパラレルコーパスの抜粋

ID	ソーステキスト	対応 ID	ID	翻訳テキスト
130010132	米沢はそこへしばしば飲みに来る客の一人だった。	230010108	230000000	φ
130010133	その店の雇われママである小代子が彼の妻だと靖子を知るのは、小代子が店を辞める直前のことだった。	230010108	230010108	Yonazawa had been a regular there and Sayoko had been the club's mama -- though Yasuko hadn't known they were married until just before Sayoko quit.
130010134	本人の口から聞かされたのだ。	230000000	230010109	'She wants to go from being the mama at a bar to the good wife at a lunch shop,' Yonazawa had told her.
130010135	「飲み屋のママから弁当屋の女房へ転身だってさ。	230010109	230010110	'Can you believe it?'
130010136	人間、わかんねえもんだなあ」	230010110	230010111	Some people never fail to surprise me.'
130010137	そんなふうには客たちは噂していた。	230010111	230010112	Rumours had begun to fly at the club, but according to Sayoko, it had been the couple's long-held dream to run a place of their own.
130010138	しかし小代子によれば、弁当屋を経営するのが夫婦の長年の夢で、それを実現するために彼女も飲み屋で働いていた、ということらしい。	230010111 230010112 230010113	230010113	She had only been working at the club to save up for that.

表3 パラレルコーパスの延べ語数の構成

翻訳者名	ソーステキスト	翻訳テキスト	翻訳 / ソース
Jan Mitsuko Cash	86768	80334	0.93
Alexander O. Smith	116567	108557	0.93
Giles Murray	111957	96026	0.86
Sam Bett	134589	111905	0.83
平均	112470.25	99205.50	
標準偏差	17079.90	12399.30	
変動係数 (CV)	0.15	0.12	

生じた。

変動係数 (Coefficient of Variation, 以下 CV) とは平均値に対するデータのばらつきの関係を相対的に評価するために用いる値であり、標準偏差を平均値で割ったものであるが、この値はソーステキスト間よりも翻訳テキスト間で小さくなっていた。これは翻訳によって各作品間の延べ語数の差が小さくなったことを示している。

## 4. 分析の結果

### 4.1. 文数

最初にテキストの文数の比較をみる。表4はソーステキストと各翻訳者が翻訳した翻訳テキストそれぞれの文数である。

表4では、ソーステキストと翻訳テキストの CV に差はみられなかった (0.11 : 0.11)。これは四作品ともソーステキストよりも翻訳テキストのほうが文数はわずかに少なくなっており、そしてそれぞれが 97% 程度であったことから窺える。

次にソーステキストと翻訳テキストを比較し、1) ソースとなる日本文と翻訳された英文が一文対一文で対応している場合 (一対一)、2) 第一節の (1) のように複数の日本文をつなげて一文に翻訳した場合 (結合)、3)

表4 テキストの文数

翻訳者名	ソーステキスト	翻訳テキスト	翻訳 / ソース
Jan Mitsuko Cash	7438	7247	0.97
Alexander O. Smith	8870	8510	0.96
Giles Murray	8173	7771	0.95
Sam Bett	10015	9850	0.98
平均	8624.00	8344.50	
標準偏差	949.34	978.18	
変動係数 (CV)	0.11	0.11	



第一節の(2)のように複数の英文で一文の日本語を翻訳した場合(分離)、4)以下の(3a)のように、複数の日本語が複数の英文で一対一ではない形で翻訳された場合(交錯)、5)以下の(3b)のようなソースの日本語と異なる順序で翻訳された場合(順序入替)、6)あらたに英文を付け加えた場合(追加)、7)ソースとなる日本語に対応する翻訳文がない場合(削除)のそれぞれの頻度を調査した。100文中に出現する頻度を翻訳者ごとに示したものが表5である<sup>7)</sup>。

### (3) a. 交錯の例

ソーステキスト：

実際、それから間もなく翔太の足が止まった。そばに一軒の家が建っている。さほど大きくもない店舗兼用の民家だった。住居部分は木造の日本建築で、間口が二間ほどの店舗はシャッターが閉じられている。

『ナミヤ雑貨店の奇蹟』第一章 1

翻訳テキスト：

It wasn't much longer before he came to a halt in front of the only house in sight – traditional Japanese home with a store on the ground floor. It was a decent size, and its living quarters were constructed out of wood. The shutter was barely ten feet wide and pulled all the way down.

*The Miracles of the Namiya General Store: 1. 1*

### b. 順序変更の例

ソーステキスト：

自分の女を職場に置いてちゃあいけません。一緒にいたいのなら結婚すればいいじゃないですか。あなたは独身なんだ。

『新参者』第七章 2

---

7) 分母となる文数は、分離と削除はソーステキストの総文数、それ以外は翻訳テキストの総文数としている。

表5 翻訳者が翻訳時に行った行為の頻度(100文中)

翻訳者名	一対一	結合	分離	交錯	順序入替	追加	削除
Jan Mitsuko Cash	94.54	3.50	1.01	0.03	0.19	0.11	0.05
Alexander O. Smith	77.90	10.24	5.76	0.15	0.68	2.90	1.96
Giles Murray	75.22	10.62	5.98	0.15	0.81	1.12	3.39
Sam Bett	76.31	8.84	8.58	0.15	0.78	1.15	1.85
平均	80.99	8.30	5.33	0.12	0.62	1.32	1.81
標準偏差	7.88	2.85	2.73	0.05	0.25	1.00	1.18
変動係数(CV)	0.10	0.34	0.51	0.45	0.40	0.76	0.65

翻訳テキスト：

I don't mean to criticise, but you should know better than to bring your girlfriends into the workplace. You're a single man now, for goodness' sake. If you like her, go ahead and marry her. *Newcomer: 7. 2*

一対一をみると、Jan Mitsuko Cashの翻訳テキストの約95%を占め、彼女の翻訳は対訳調であることがわかる。他の三者の翻訳は傾向が類似しており、一対一が75%程度、結合は10%程度、分離は約6%、交錯はほとんどなく、順序入替は100文中で一度行う程度である。よって以下では主として結合・分離を扱うこととする。

追加と削除については、それぞれ明示化と簡素化を目的として行われると考えられるため、本研究では取り扱わず他の項目とのCVの値の比較のみ用いる。この二つのCVは0.76と0.65であり、他の項目の値より高かった(0.10-0.50)。これについては、追加ではAlexander O. Smithの値(2.90)が高く、削除ではGiles Murrayの値(3.39)が高い一方、Jan Mitsuko Cashは二つの値(0.11, 0.05)はどちらも低く、各翻訳者に個性がみられたためである。

## 4.2. 文の長さ

次にソーステキストと翻訳テキストの文の長さをみる。表6は四作品の全文の平均長を示したものである。文の長さの単位は単語とし、どちらのテキストでも記号はカウントしていない。

第三節で述べたように、使用するソフトウェアによって得られる語数は異なるため、ソーステキストと翻訳テキストの文の長さの値は本研究のみで用いられるべきものである。今回の分析においては、四作品とも翻訳テキストの文の長さの比はソーステキストとほぼ同じであった(0.98-1.03)。

ソーステキストと翻訳テキストそれぞれの文長を四作品間で比較するために分散分析を行った。結果、ソーステキストでは条件間に有意な差が得られた( $F(3, 34490) = 10.577, p < 0.001, \eta^2 = 1.00$ )。分散分析によって有意な差があることがわかったため、それぞれの作品についてボンフェローニ補正を用いて二群間で $t$ 検定を行った。Cohen(1988)で提唱された効果量 $d$ とともにその結果を表7に示す。『ゲームの名は誘拐』と『容疑者Xの献身』および『新参者』の間に有意な差が得られた( $p < 0.0001$ )。しかしサンプルサイズに影響されないCohenの $d$ の値はどれも小さく、作

表6 文の長さの平均

翻訳者名		ソーステキスト	翻訳テキスト	翻訳 / ソース
Jan Mitsuko Cash	平均	9.69	9.46	0.98
	標準偏差	5.93	5.65	
Alexander O. Smith	平均	10.14	10.48	1.03
	標準偏差	6.23	6.83	
Giles Murray	平均	10.17	10.17	1.00
	標準偏差	6.26	6.39	
Sam Bett	平均	9.93	9.74	0.98
	標準偏差	6.14	6.26	
合計	平均	9.98	9.96	
	標準偏差	6.15	6.33	

表7 二ソーステキスト間の文長の比較

(\* :5% 有意  $p < 0.0083$ , \*\* :1% 有意  $p < 0.0017$ )

二 群	$p$ 値	Cohen's $d$
ゲームの名は誘拐・容疑者 X の献身	** < 0.0001	0.08
ゲームの名は誘拐・新参者	** < 0.0001	0.08
ゲームの名は誘拐・ナミヤ雑貨店の奇蹟	0.0093	0.03
容疑者 X の献身・新参者	0.7742	0.004
容疑者 X の献身・ナミヤ雑貨店の奇蹟	0.0166	0.03
新参者・ナミヤ雑貨店の奇蹟	0.0084	0.04

表8 二翻訳テキスト間の文長の比較

(\* :5% 水準  $p < 0.0083$ , \*\* :1% 水準  $p < 0.0017$ )

二 群	$p$ 値	Cohen's $d$
Jan Mitsuko Cash · Alexander O. Smith	** < 0.0001	0.16
Jan Mitsuko Cash · Giles Murray	** < 0.0001	0.12
Jan Mitsuko Cash · Sam Bett	* 0.0025	0.05
Alexander O. Smith · Giles Murray	* 0.0034	0.05
Alexander O. Smith · Sam Bett	** < 0.0001	0.11
Giles Murray · Sam Bett	** < 0.0001	0.07

品間に大きな差があるとはいえないだろう。

翻訳テキストでも条件間に有意な差が得られた ( $F(3, 33374) = 41.000, p < 0.001, \eta^2 = 0.00$ )。そのため、それぞれの二翻訳テキスト間でボンフェローニ補正を用いて  $t$  検定を行い、その結果を効果量  $d$  とともに表8に示す。すべての翻訳テキストの間に5%水準で有意な差が得られた ( $p < 0.0083$ )。サンプルサイズに影響されないCohenの $d$ の値は、小さい効果とみなされる値(0.2)より小さいが、ソーステキストでみられた効果量よりはわずかに大きくなっており、翻訳によって作品間に差が生じたことが窺える。

翻訳者がソーステキストの複数の日本語をつなげて一文の英文に翻訳したとき、翻訳者が選んだ日本語と翻訳した英文の長さを示したものが表9である。

表9からは四作品ともに、1) 二文の日本語をつなげて翻訳し一文の英

表9 つなげる前の日本文とつなげた後の英文の長さ

翻訳者名		日本文の長さ	英文の長さ
Jan Mitsuko Cash	平均	7.11 ( $N=526$ )	13.21 ( $N=254$ )
	標準偏差	8.61	5.70
Alexander O. Smith	平均	8.96 ( $N=1797$ )	15.78 ( $N=871$ )
	標準偏差	8.59	6.81
Giles Murray	平均	9.25 ( $N=1792$ )	15.85 ( $N=825$ )
	標準偏差	8.08	5.85
Sam Bett	平均	8.70 ( $N=1797$ )	15.41 ( $N=871$ )
	標準偏差	8.27	5.87
合計	平均	8.80 ( $N=5804$ )	15.46 ( $N=2821$ )
	標準偏差	5.31	7.51

文にした場合がほとんどであること（英文の  $N$  が日本文の  $N$  のほぼ半分である）、2）翻訳者がつなげようとした日本文は、ソーステキスト全文の平均の文長よりわずかに短い文であること（9.98 : 8.80）、3）翻訳者がつなげて翻訳した英文の長さは、期待される値（元となる日本文を2倍し、表6の翻訳/ソース（ $\approx 0.99$ ）を掛けた長さ = 17.42）よりは短くなっていること（15.46）がわかる。

次に翻訳者がつなげようとして選んだ日本文の長さ、つなげて翻訳して得られた英文の長さが翻訳者間で差があるかどうかを確かめる。各翻訳者がつなげようとした日本文の長さを比較するために分散分析を行った結果、翻訳者間に有意な差が得られた（ $F(3, 5800) = 22.923, p < 0.001, \eta^2 = 0.01$ ）。分散分析によって有意な差があることが分かったため、ボンフェローニ補正を用いて二群間で  $t$  検定を行った。効果量  $d$  とともにその結果を表10に示す。

Jan Mitsuko Cash が選んだ文は他の三者が選んだ文の長さより有意に短く、効果量  $d$  も小さい効果（0.2）から中程度の効果（0.5）の間であった。Giles Murray と Sam Bett の選んだ文の長さの間にも有意な差はみられたが、

表 10 二翻訳者間での、つなげるために選ばれた日本語の長さの比較

(\* :5% 水準  $p < 0.0083$ , \*\*:1% 水準  $p < 0.0017$ )

二 群	$p$ 値	<i>Cohen's d</i>
Jan Mitsuko Cash · Alexander O. Smith	** < 0.0001	0.37
Jan Mitsuko Cash · Giles Murray	** < 0.0001	0.44
Jan Mitsuko Cash · Sam Bett	** < 0.0001	0.34
Alexander O. Smith · Giles Murray	0.1178	0.05
Alexander O. Smith · Sam Bett	0.1390	0.05
Giles Murray · Sam Bett	* 0.0020	0.10

表 11 二翻訳者間での、つなげて翻訳された英文の長さの比較

(\* :5% 水準  $p < 0.0083$ , \*\*:1% 水準  $p < 0.0017$ )

二 群	$p$ 値	<i>Cohen's d</i>
Jan Mitsuko Cash · Alexander O. Smith	** < 0.0001	0.38
Jan Mitsuko Cash · Giles Murray	** < 0.0001	0.39
Jan Mitsuko Cash · Sam Bett	** < 0.0001	0.32
Alexander O. Smith · Giles Murray	0.8365	0.01
Alexander O. Smith · Sam Bett	0.3226	0.05
Giles Murray · Sam Bett	0.2345	0.06

効果量  $d$  は小さかった (0.10)。

翻訳者がつなげて翻訳した英文の長さが翻訳者間で差があるのかをみるために、分散分析を行った。結果、翻訳者間で有意な差が得られた ( $F(3, 2815) = 8.944, p < 0.001, \eta^2 = 0.01$ )。分散分析によって有意な差があることがわかったため、ボンフェローニ補正を用いて二群間で  $t$  検定を行った。効果量  $d$  とともにその結果を表 11 に示す。

つなげる日本語を選ぶときと同様に、Jan Mitsuko Cash がつなげて翻訳した英文は他の三者の英文の長さより有意に短く、効果量  $d$  も 0.2 から 0.5 の間であった。Jan Mitsuko Cash と他の三人を比較すると、複数の日本語をつなげて一文の英文に翻訳するとき、彼女はより短い日本語を選び、より短い英文を書くことがわかった。

翻訳者がソーステキストの日本語の一文を区切って翻訳し、複数の英文に翻訳するときの日本語と英文の長さを示したものが表 12 である。

表 12 からは四作品ともに、1) 日本語の一文を、英文の二文に区切って翻訳する人が多いこと（日本語の  $N \times 2 \approx$  英文の  $N$ ）、2) 翻訳者が区切ろうとする日本語の文長の平均は、ソーステキスト全体の文長の平均のおおよそ 1.5 倍であること（9.98 : 15.70）、3) 翻訳者が区切って翻訳した英文の長さは、予測される値（ソーステキストの長さ  $\div 2 \times 0.99 = 7.77$ ）より長くなっていること（8.54）がわかる。

表 12 区切られる前のソーステキストの日本語と区切られて翻訳された英文の長さ

翻訳者名		日本語の長さ	英文の長さ
Jan Mitsuko Cash	平均	15.35 ( $N=74$ )	8.00 ( $N=150$ )
	標準偏差	8.61	5.70
Alexander O. Smith	平均	15.70 ( $N=471$ )	9.43 ( $N=961$ )
	標準偏差	8.59	6.81
Giles Murray	平均	15.98 ( $N=599$ )	8.99 ( $N=1251$ )
	標準偏差	8.08	5.85
Sam Bett	平均	15.51 ( $N=761$ )	8.55 ( $N=1594$ )
	標準偏差	8.27	5.87
合計	平均	15.70 ( $N=1905$ )	8.54 ( $N=3956$ )
	標準偏差	8.31	6.11

表 13 二翻訳者間での、区切って翻訳した英文の長さの比較

(\* :5% 水準  $p < 0.0083$ , \*\*:1% 水準  $p < 0.0017$ )

二 群	$p$ 値	<i>Cohen's d</i>
Jan Mitsuko Cash ・ Alexander O. Smith	0.0148	0.23
Jan Mitsuko Cash ・ Giles Murray	0.0507	0.17
Jan Mitsuko Cash ・ Sam Bett	0.2680	0.10
Alexander O. Smith ・ Giles Murray	0.0989	0.07
Alexander O. Smith ・ Sam Bett	** 0.0006	0.14
Giles Murray ・ Sam Bett	0.0513	0.07

翻訳者が区切って翻訳しようとしたソーステキストの日本語の長さ  
と、区切って翻訳した英文の長さをそれぞれ翻訳者間で比較した。まず各  
翻訳者が区切って翻訳しようとした日本語の長さを比較するために分散分  
析を行った。その結果、翻訳者間に有意な差は得られなかった ( $F(3, 1901)$   
 $= 0.408, p = 0.75, \eta^2 = 0.00$ )。もちろん第二種の過誤を犯す可能性があるの  
ではっきり述べることはできないが、データ数 ( $N = 1905$ ) の大きさを踏  
まえると、翻訳者が区切って翻訳しようとする日本語を選ぶとき、その文  
の長さは翻訳者間で差はないといえるかもしれない。

次に翻訳者が日本語の一文を区切って複数の英文に翻訳したとき、得ら  
れた英文の長さを、翻訳者間で比較するために分散分析を行った。その結  
果、翻訳者間で有意な差が得られた ( $F(3, 3952) = 5.293, p = 0.001, \eta^2 =$   
 $0.00$ )。分散分析によって有意な差があることがわかったため、ボンフェ  
ローニ補正を用いて二群間で  $t$  検定を行った。効果量  $d$  とともにその結果  
を表 13 に示す。

Jan Mitsuko Cash と Alexander O. Smith が翻訳した英文の長さについて、  
効果量  $d$  は 0.23 と小さな効果が認められたが、有意な差はなかった。一  
方 Alexander O. Smith と Sam Bett が翻訳した英文の長さに有意な差があっ  
たが、効果量  $d$  の値は低い。よって分離して翻訳した文の長さも翻訳者間  
で大きな差はみられないといえる。

#### 4.3. 結合とカンマ、分離と読点

最後に、翻訳者が翻訳するとき区切ることを決めた日本語に読点が含ま  
れている頻度と、ソーステキストの日本語をつなげて翻訳した英文にカン  
マが含まれている頻度を、全体のテキストにそれぞれ読点・カンマが含ま  
れている頻度と比較した。その結果がそれぞれ表 14 と表 15 である。

表 14 からは、四作品間にばらつきはあるが、全文では平均しておおよ



表 14 ソーステキストに読点が含まれる頻度

ソーステキスト	翻訳者名	全文	翻訳者が区切って 翻訳した日本語
『ゲームの名前は誘拐』	Jan Mitsuko Cash	2789 (37.1%)	63 (85.9%)
『容疑者 X の献身』	Alexander O. Smith	4488 (47.9%)	397 (86.0%)
『新参者』	Giles Murray	4655 (52.7%)	522 (84.3%)
『ナミヤ雑貨店の奇蹟』	Sam Bett	5957 (54.9%)	654 (87.1%)
	合計	17889 (48.9%)	1636 (85.9%)

注：カッコ内の数字は読点が含まれていた文の割合

表 15 翻訳テキストにカンマが含まれる頻度

翻訳者名	全文	翻訳者が日本語をつな げて翻訳した英文
Jan Mitsuko Cash	2617 (36.1%)	209 (82.2%)
Alexander O. Smith	3511 (41.3%)	689 (79.1%)
Giles Murray	2735 (35.2%)	586 (70.9%)
Sam Bett	3533 (35.9%)	672 (77.2%)
合計	12396 (37.1%)	2156 (76.4%)

注：カッコ内の数字はカンマが含まれていた文の割合

そ半分の文に読点が含まれていることがわかる。しかし翻訳者が区切って翻訳することを選んだ文には、おおよそ 86% に読点が含まれていた。

表 15 では、翻訳テキスト全体では四作品は 35% から 40% の文でカンマを含んでいるが、翻訳者が複数の日本語をつなげて翻訳した英文には平均して 76.4% にカンマが含まれていた。

## 5. 考 察

本研究で用いた東野圭吾の四小説の翻訳テキストの中では、Jan Mitsuko Cash の翻訳に一文対一文の対訳が最も多く（約 95%）、他の三者の翻訳はおおよそ 75% であった。つまりソーステキストに対して Jan Mitsuko Cash は二十文に一文程度、そして他の三者は四文に一文程度は対訳以上の翻訳技術を加えていることになる。Jan Mitsuko Cash と他の三者間のその頻度

の差は大きい。

そのような違いがあるにもかかわらず、作品間でソーステキスト対翻訳テキストの文数の比に大きな差はみられなかった。翻訳テキストの文の長さについては、翻訳者四名の間に5%水準で有意な差は得られたが、効果量  $d$  の値は翻訳者間で小さな効果であると認められる 0.2 より小さかった。意識的か無意識的かははっきりしないが、翻訳者たちは、翻訳後でもソーステキストが持つ文体の雰囲気を保とうとしているためか、文数や文の長さに変化を与えていない。よって作品間に大きな差が生じないのであろう。

ソーステキストの複数の日本文をつなげて一文の英文に翻訳する頻度は、翻訳者間で異なる (3.50–10.62 回/100 文)。そして、どの程度の長さをつなげるべき日本文として選択し、どの程度の長さの一文の英文に翻訳するかについては、Jan Mitsuko Cash が他の三者と比較して有意に短い文を選択し、有意に短い文に翻訳していた。これは Jan Mitsuko Cash の翻訳スタイルが一文対一文の対訳風であることに関係しているのではないかと思われる。彼女の文体は一文対一文の対訳の数が多いため、彼女が複数の日本文を一文に翻訳するときは、(4) で示されるような「台詞+発話者を特定する文」の文章が主となる (下線部は筆者による)。これらは今回のコーパス作成時の定義で、日本文では二文、英文では一文とカウントされることになったものである。

(4) a. ソーステキスト：

「顔相学というものがあります」おれは葛城勝俊に説明した。

『ゲームの名は誘拐』7

翻訳テキスト：

“There’s something called physiognomy,” I explained to Katsutoshi Katsuragi.

b. ソーステキスト :

「さっき高速に乗ったところ」彼女は短く答えた。

『ゲームの名は誘拐』 13

翻訳テキスト :

“Just now, he got on the expressway,” Juri replied curtly.

*The Name of the Game is a Kidnapping: 13*

表 9 に示された、翻訳者がつなげることに決めた日本文の中で、(4) にあるような台詞が含まれている割合を表 16 に示す。Jan Mitsuko Cash では他の三者と比較して台詞の割合が多い (35.4% : 15.6-26.2%) ことがわかる。また、台詞以外の文の長さの平均を表 17 に示すが、Jan Mitsuko Cash がつなげるために選んだ日本文の長さは他三者より短い (7.50 : 9.22-9.93)。これは Jan Mitsuko Cash が選んだ文には (4) の下線部で示されるような短い「発話者を特定する文」が多いためだと考えられる。

一文の日本文を分離して複数の英文に翻訳する頻度も、翻訳者間で異なるが (1.03-7.73 回 /100 文)、どの程度の長さの日本文を区切るのかにつ

表 16 表 9 の中での台詞の割合

翻訳者名	セリフ数 (N)	割合 (%)
Jan Mitsuko Cash	186	35.4
Alexander O. Smith	469	26.2
Giles Murray	406	22.6
Sam Bett	281	15.6

表 17 表 9 の中での台詞以外の文の長さの平均

翻訳者名	文数 (N)	平均	標準偏差
Jan Mitsuko Cash	334	7.50	4.42
Alexander O. Smith	1277	9.93	5.57
Giles Murray	1495	9.22	5.59
Sam Bett	1405	9.45	5.14

いては翻訳者間で有意な差を観察できなかった。また、どの程度の長さの英文に翻訳するかについても、翻訳者間で有意な差はほぼみられず、効果量  $d$  の値も小さかった。ソーステキストの一文の日本語を区切って複数の英文に翻訳するとき、文の長さについては、意識的か無意識的かはわからないが、翻訳者間で共通のパターンが存在するといえる。

また複数の日本語をつなげて英文に翻訳する場合には英文の多くにカンマが含まれ、一方、一文の日本語を区切って複数の英文に翻訳する場合にはソースとなる日本語の多くに読点が含まれることが多かった。これはカンマや読点によって分けられる句や節が翻訳するときのひとかたまりとなっていると考えられる。これは翻訳者四名に共通のパターンであったが、翻訳するにあたっては至極自然であるともいえるだろう。

最後にこれらの翻訳行為がどのような目的で行われているかを考察する。分析によって次のことが推測される。1) 文の結合・分離は、簡素化・明示化を明確に示すような文の追加・省略より頻度が高い。かつ文の追加・省略にみられた頻度には各翻訳家の個性がみられたが、文の結合・分離にはそのような傾向はみられない。これは結合・分離が簡素化・明示化とは性質が異なるためではないかと考えられる。2) ソーステキストの複数の日本語をつなげて一文の英文に翻訳したときでも、英文の長さは明示的に新たな情報が加わるほど長くはなっていないように思われる。3) 翻訳者が区切って翻訳するために選択した日本語の長さの平均は、ソーステキスト全体の日本語の長さの平均に対して 1.5 倍程度であり、文を区切って簡素化しなければならないほど長いわけではない。これらを考慮すると、ソーステキストをつなげたり、区切ったりして翻訳する技術は、村上春樹のいう通り、翻訳先での言語のリズムを自然にするための正規化が主な目的であると考えられる。もちろんこれは計量的な分析による一般化であり、個々の例については簡素化・明示化の目的で結合・分離が行われることも考え

られる。それを明らかにするために新たな分析方法を考察することについては今回の分析の範疇を超えるため、今後の課題としたい。

## 6. ま と め

Baker (2000) でも言及されているように、翻訳されたテキストには翻訳者自身の文体が潜んでいる。その中でも、ソーステキストの複数の文をつなげて一文にしたり、一文を区切って複数の文に翻訳したり、ソーステキストの文を入れ替えて翻訳する技術は、計量言語学的方法論でその特徴をとらえやすい文体的特徴であるといえる。本研究は、東野圭吾の小説を翻訳した四名の翻訳家を比較し、翻訳の際に起こる結合・分離・順序入替などの頻度とそれによって生じる結果を分析した。そして翻訳者が翻訳の際に結合・分離をする目的は Baker (1996) での正規化に分類されると考察した。

本研究で使用した四名が翻訳したテキストでは、1) 頻度としては順序入替はほとんど起こらないこと、2) ソーステキストの結合・分離は翻訳者によって頻度が異なるが、翻訳テキスト全体の文数や文の長さの平均は四名ともソーステキストに近似し、そのため翻訳者間での違いは大きくなく、各翻訳者はソーステキストの文体の雰囲気を保とうとしていると考えられること、3) 翻訳者がソーステキストとなる日本文を区切ったり、つなげたりして英文に翻訳する場合、翻訳者が選択するソースとなる日本文と翻訳した英文の長さの共通のパターンが見受けられること、そして4) 翻訳する際にはカンマ、読点を意味の区切りとして考えていることを明らかにした。これらが日本語から英語に翻訳する文芸翻訳の文体のより一般的パターンとして考えられるかどうかは今後の課題となる。

## 参考文献

- Baker, Mona (1996) Corpus-based translation studies: the challenges that lie ahead. In Somers, H. (ed.) *Terminology, LSP and Translation: Studies in Language Engineering*. In Honour of Juan C. Sager. Amsterdam: John Benjamins, 175-186.
- (2000) Towards a methodology for investigating the style of a literary translator. *Target*. 12: 2, 241-266.
- Cohen, Jacob (1988) *Statistical Power Analysis for the Behavioral Sciences*. New York: Routledge Academic.
- 村上春樹・柴田元幸 (2000) 『翻訳夜話』 文藝春秋。