



Title	A Computational Analysis of Translator Style in Parallel Japanese Translations of Gabriele D'Annunzio's Triumph of Death
Author(s)	Camilleri, Gabriele
Citation	大阪大学, 2026, 博士論文
Version Type	
URL	https://hdl.handle.net/11094/105321
rights	
Note	やむを得ない事由があると学位審査研究科が承認したため、全文に代えてその内容の要約を公開しています。全文のご利用をご希望の場合は、大阪大学の博士論文についてをご参照ください。

The University of Osaka Institutional Knowledge Archive : OUKA

<https://ir.library.osaka-u.ac.jp/>

The University of Osaka

論 文 内 容 の 要 旨

氏 名 （ C A M I L L E R I G A B R I E L E ）	
論文題名	A Computational Analysis of Translator Style in Parallel Japanese Translations of Gabriele D'Annunzio's <i>Triumph of Death</i> (ガブリエレダンヌンツィオ著『死の勝利』の日本語並列訳における翻訳者文体の計量的分析)
論文内容の要旨	
本論文は、ガブリエーレ・ダヌンツィオの長編小説 <i>Il trionfo della morte</i> (1894) と、その日本語訳8点 (1913・1921・1927・1928・1939・1958・1961・2010年) を対象とし、一つの原作に対する多訳コーパスを構築した上で、数量的手法により訳者固有の文体と参照表現の用法を比較検討することを目的とする。対象となる作品は、生田長江による完全翻訳版 (1913年) から脇功訳 (2010年) まで約100年に及ぶ翻訳史を持つ作品であるが、特に戦前期の翻訳は、フランス語・英語・ドイツ語などの中間言語からの重訳が多く、イタリア語原典から直接訳されたものはごく少ない。また、本作品が日本近代文学にもたらした影響に関する研究は多いものの、翻訳と原作の文体的差異に着目した研究は限られ、質的な分析にとどまっている。 本論文は、こうしたギャップを踏まえ、(1) 多層注釈付き多訳コーパスの設計・構築、(2) 数量的文体分析にもとづく訳者文体の記述、(3) イタリア語―日本語間のアラインメントと参照形式選択の関係の解明、という三点を中心課題として設定する。 第1章では、本研究の対象となるダヌンツィオ『死の勝利』の日本における受容史と翻訳史を整理している。まず、1896年の『太陽』誌における紹介や上田敏の一連の評論・抄訳を通じて、ダヌンツィオが世紀末ヨーロッパ文学の重要作家として日本に導入され、とくに『死の勝利』が代表作として位置づけられた過程を跡づける。さらに、塩原事件と森田草平『煤煙』、夏目漱石『それから』との関連を通じて、同作が恋愛観・感情表現・感覚描写の面で日本近代文学に与えた影響を検討する。その後、1910年代から1920年代にかけての翻訳ブームと読者層の拡大、昭和期以降における「愛国詩人」「政治家」としてのイメージの強化を概観し、多くの日本語訳が英・仏・独語を介した重訳であったことを指摘する。とりわけ生田長江訳の初刊行と改訂 (1913・1928年)、「平明極まる大衆的」文体への自己言及、抜粋・改変を含む翻訳方針が、読者の期待とダヌンツィオ像の形成に与えた影響を検証し、平山 (2011) による比較訳読の事例を通じて、各訳者の文体的・解釈的特性が浮かび上がることを示している。最後に、こうした歴史的展開が、本論文における数量的文体分析と訳者スタイル研究の前提となることを確認して本章を結んでいる。 第2章では、本研究の理論的枠組みとして、翻訳調・訳者文体・日本語翻訳調研究に関する先行研究を整理する。まず2.1節では、Jakobson (1959)、Nida (1964)、Even-Zohar (1979) らによる翻訳理論の観点の変遷をたどりつつ、Baker (1993) やToury (1995) に代表されるコーパス翻訳研究の展開を概観し、「明示化」「単純化」「標準化」といった「翻訳の普遍性」仮説と、機械学習やPOS n-グラムを用いた「翻訳調」の検出の試みを紹介する。2.2節では、Baker (1993) の提案する「同一原作の複数訳比較」という発想を軸に、文書レベル指標や語・構文n-グラム、依存距離などを用いて訳者固有の指紋を抽出する諸研究を検討し、作者性研究との接点と限界を明らかにする。2.3節では、日本語の翻訳調 (翻訳文体) の歴史的形成を、漢文訓読から欧文直訳文体、言文一致期、戦後の流暢さ志向に至る流れの中で位置づけ、代名詞・指示詞、抽象的実体の主語、カタカナ語、女ことば、接続詞や修飾構造などの特徴を示した量的研究 (Fukuchi Meldrum 2009, 小西 2016 ほか) を整理する。その上で、本研究は日本語翻訳調一般の普遍的特徴を論じるのではなく、1913年から2010年にかけて刊行された『死の勝利』8訳という単一作品多訳コーパスを、訳者と文体を精査するための検証基盤とも位置づけることを確認する。 第3章では、本研究の方法論として、多訳コーパスの構築と前処理について述べる。まず、対象として石川 (1913) から脇 (2010) までの8訳を選定し、原作とともに『死の勝利』第1・第2部のみを分析範囲とした理由 (テキストの欠落・断簡の存在、OCR後補正や共参照アノテーションの作業量など) を明示する。その上で、国会図書館デジタルコレクション等から取得した翻訳をOCR・手作業校正し、TEI P5 (TEI, 2025) に従って段落・文・会話・内的独白・書簡、および話し手・聞き手情報をタグ付けすることで、談話の種類別サブコーパスを抽出可能な構造を整える。さらに、UniDic2UDによる形態素解析を実施し、歴史的仮名遣いに対応した「近現代口語小説」辞書と「現代書き言葉」辞書を訳の年代に応じて使い分けることで、短単位および文節レベルの分析を可能にしている。これにより、第4章以降の文体特徴抽出、文節・トークンレベルのアラインメント、共参照解析のための多層的な分析基盤が整備されたことを示	

している。

第4章では、構築した多訳コーパスを用いて、8訳の文体差を数量的に比較する。地の文・主人公の会話文・ヒロインの会話文・内的独白・書簡の5サブコーパスを切り出し、文長、語彙多様度、品詞比率、文字種比率など22種の文書上の特徴と、Zeta 分析によって抽出した短単位のトリグラムを特徴量とし、主成分分析とWard 法にもとづく階層的クラスタリング、および分散分析を行った。その結果、多くのサブコーパスで、戦前訳と戦後訳のあいだに明確な分岐が見られ、とりわけ石川（1913）が高い語彙多様度・漢字率・長文志向により一貫して孤立したクラスターを形成する一方、岩崎（1958）と脇（2010）が特徴的な三連語分布を共有することが明らかになった。また、主人公の会話文と書簡では、戦後訳において一人称・二人称代名詞「ボク・キミ」の共起や「ノダ」文の使用、「テイタ」時制の頻出が確認され、ヒロインの会話文では文末表現「わ・の・のよ」による典型的な「女ことば」が戦後訳を中心に強く現れるなど、登場人物の役割語化・会話の口語化が訳者ごとに異なる形で進行していることを示した。最後に、本章の分析は訳文間比較にとどまるため、今後は同時期の日本語原文コーパスとの対照により、これらの差異が翻訳一般の傾向か否かを検証する必要性を指摘している。

第5章では、共参照分析にもとづき、原作と8訳における参照表現の分布と対応関係を検討する。まず、コーパス言語学や自然言語処理における共参照研究と、PAS（Carminati, 2002）や centering theory（Walker, Iida and Côté, 1994）によるイタリア語と日本語のゼロ照応を概観した上で、TEIを用いた登場人物の参照表現のアノテーション方針（品詞・係り受け・参照種類のタグ付け）を定める。次に、原文における主語種類の分布や談話の距離・形態素情報・遷移パターンを分析し、ダヌンツィオの用法がPASの予測とある程度整合的であることを示す。続いて、文アラインメントとアノテーションを利用したルールベースのペアリング・アルゴリズムを構築し、各訳について3,000件超の参照対を抽出する。これを用いてイタリア語の品詞別に確率を算出した結果、固有名詞は全訳にわたり高率で保持される一方、人称代名詞とゼロ照応の扱いには訳者差が大きく、石川（1913）は代名詞の省略と名詞句化が多く、三上（1921）は代名詞を多用して明示化する傾向が若干際立つことが明らかになった。さらに、10種の形態的・談話的特徴量を入力とするロジスティック回帰モデルを訳者別に学習し、省略・固有名詞・代名詞・ゼロ照応などの予測精度を比較したところ、全訳で精度約0.8前後と安定した成績を示し、参照形式選択が無作為ではなく規則的パターンに従っていることが示唆された。係数の解釈からは、省略やゼロ照応の連鎖、固有名による再アンカー、代名詞の翻訳側談話条件への依存といった一般傾向が抽出される一方、石川（1913）のゼロ重視・名詞句志向や三上（1921）の強い代名詞化など、訳者固有の傾向も浮かび上がる。

第6章では、文・トークン・文節という三層のアラインメント分析を通じて、各訳の構造と語彙選択を検討する。まず複数の文アラインメントのアルゴリズムを比較し、VecAlign（Thompson & Koehn, 2020）が唯一、分析に適している精度を示すことを確認した上で、分割・結合種類の分布と長さ比から、石川（1913）による極端な分割傾向や三上（1921）の結合志向といった訳者差を示しつつ、多くの場合は情報量を保った再分節であることを示す。また未訳文の分布を調べ、矢口（1927）が書簡や挿話をまとまりごとと削除する「抄訳」的性質をもち、石川（1913）・三上（1921）・矢口（1927）・生田（1928）の省略箇所がHarding（1898）やHornblow（1896）の英訳と高い一致を示す点から、中継翻訳の影響と検閲上の配慮を数量的・質的に裏付ける。次に、伊日間の単語（トークン）レベルの対応について検証するために、データセットを作成し、MGIZA++（Brown et al., 1993; Och & Ney, 2003）、fast_align（Dyer, Chahuneau and Smith, 2013）、SimAlign（Jalili Sabet et al., 2020）やAWESOME-align（Dou and Neubig, 2021）を比較することで、統計的モデルの限界と、多言語の大規模コーパスと一部のデータで継続学習することによりF1 $\approx$.65まで性能向上が可能である一方、句読点や助詞列・複合述語周辺に体系的誤りが残り、文体分析用の単位としては語より文節が適切であることを指摘する。最後に、大規模言語モデルにより8訳間の文節アラインメントを構築し、生田（1928）訳を中心とする整列語彙表からエントロピー距離とネットワークを算出することで、石川（1913）の孤立性、戦前訳・戦後訳双方への架橋点としての生田（1928）の中心性、および「最終」、「欲求」などの語の使用といった特徴を抽出する。これにより、生田が自ら標榜した「平明極まる大衆の文体」の下でも、ダヌンツィオ的な重々しさと終末感を保つ訳者文体が、遠読の手法からも実証的に把握しうることを示している。

最後の第7章では、本研究の射程と限界を明示し、その先の課題と展望を簡潔に示している。本論文は、『死の勝利』8訳からなる多訳コーパスを構築し、文体・アラインメント・共参照という三層の分析を通じて、訳者固有の文体と参照戦略の差異を実証的に描き出した点に特色がある。単一作品を精緻に追跡することで、年代差だけでは説明しきれない訳者行動の分岐と収斂を明らかにし、日本語翻訳文体研究および翻訳史研究に対して、新たな方法論的枠組みと再検討のための基盤資料を提示した。

論文審査の結果の要旨及び担当者

氏 名 (Camilleri Gabriele)			
	(職)	氏 名	
論文審査担当者	主 査	准教授	ホドシチェク ボル
	副 査	教授	田畑 智司
	副 査	教授	三宅 真紀

論文審査の結果の要旨

本論文は、ガブリエーレ・ダヌンツィオの長編小説『死の勝利』（1894）と、その日本語訳8点（1913・1921・1927・1928・1939・1958・1961・2010年）を対象とし、一つの原作に対する多訳コーパスを構築した上で、数量的手法により訳者固有の文体と参照表現の用法を比較することを目的とした研究である。本論文は以下の構成となっている。

第1章では、ダヌンツィオ著『死の勝利』の日本における受容史と翻訳史を整理している。まず、1896年の『太陽』誌における紹介や上田敏の一連の評論・抄訳を通じて、ダヌンツィオが世紀末ヨーロッパ文学の重要作家として日本に導入され、とくに『死の勝利』が代表作として位置づけられた過程を跡づけている。さらに、塩原事件と森田草平『煤煙』、夏目漱石『それから』との関連を通じて、1910–20年代の翻訳ブームと読者層の拡大、多くの日本語訳が英・仏・独語を介した重訳であったことを指摘している。とりわけ生田長江訳の初刊行と改訂（1913・1928年）、「平明極まる大衆的」文体への自己言及、抜粋・改変を含む翻訳方針が、読者の期待とダヌンツィオ像の形成に与えた影響を検証している。最後に、こうした歴史的展開が、本論文における数量的文体分析と訳者スタイル研究の前提となることを確認して本章を結んでいる。

第2章では、本研究の理論的枠組みとして、翻訳調・訳者文体・日本語翻訳調研究に関する先行研究を整理している。翻訳理論とコーパス翻訳研究の展開を概観し、「明示化」「単純化」「標準化」といった「翻訳の普遍性」仮説と、機械学習やPOS n-グラムを用いた「翻訳調」の検出の試みを紹介している。Baker (1993) の提案する「同一原作の複数訳比較」という発想を軸に、文書レベル指標や語・構文n-グラム、依存距離などを用いて訳者固有の指紋を抽出する諸研究を検討し、作者性研究との接点と限界を明らかにしている。また、日本語の翻訳調（翻訳文体）の歴史的形成を、漢文訓読から欧文直訳文体、言文一致期、戦後の流暢さ志向に至る流れの中で位置づけ、代名詞・指示詞、抽象的実体の主語、カタカナ語、女ことば、接続詞や修飾構造などの特徴を示した量的研究を整理している。その上で、本研究は日本語翻訳調一般の普遍的特徴を論じるのではなく、『死の勝利』8訳という単一作品多訳コーパスを、訳者と文体を精査するための検証基盤と位置づけている。

第3章では、本研究の方法論として、多訳コーパスの構築と前処理について述べている。まず、対象として石川 (1913) から脇 (2010) までの8訳を選定し、一部のテキストに欠落・断簡が存在するため、第1部・第2部のみを分析対象としたことを明示している。その上で、国立国会図書館デジタルコレクションから取得した翻訳をTEIガイドラインに従って段落・文・会話・内的独白・書簡、および話し手・聞き手情報をタグ付けすることで、談話の種類別サブコーパスを抽出可能な構造を整えている。さらに、形態素解析を実施し、短単位および文節レベルの分析を可能にしている。共参照解析では係り受け解析に基づく主語省略判定を採用している。これにより、第4章以降の文体特徴抽出、文節・トークンレベルのアラインメント、共参照解析のための多層的な分析基盤が整備されている。

第4章では、構築した多訳コーパスを用いて、8訳の文体差を数量的に比較している。地の文・主人公の会話文・ヒロインの会話文・内的独白・書簡の5サブコーパスを切り出し、文長、語彙多様度、品詞比率、文字種比率など22種の文書上の特徴と、Zeta分析によって抽出した短単位のトライグラムを特徴量とし、主成分分析、階層的クラスタリング、ならびに分散分析を行っている。その結果、多くのサブコーパスで、戦前訳と戦後訳のあいだに明確な分岐が見られ、とりわけ石川 (1913) が高い語彙多様度・漢字率・長文志向により一貫して孤立したクラスターを形成する一方、岩崎 (1958) と脇 (2010) が特徴的なトライグラム分布を共有することが明らかになっている。また、戦後訳において、登場人物の役割語化や会話の口語化が訳者ごとに異なる形で進行していることを示している。最後に、本章の分析は

訳文間比較にとどまるため、今後は同時期の日本語原文コーパスとの対照により、これらの差異が翻訳一般の傾向か否かを検証する必要性を指摘している。

第5章では、共参照分析にもとづき、原作と8訳における参照表現の分布と対応関係を検討している。まず、コーパス言語学や自然言語処理における共参照研究と、イタリア語・日本語のゼロ照応に関する理論的枠組みを概観した上で、TEIを用いた登場人物の参照表現のアノテーション方針（品詞・係り受け・参照種類のタグ付け）を定めている。次に、原文における主語種類の分布や談話の距離・形態素情報・遷移パターンを分析し、ダヌンツィオの用法がゼロ照応に関する理論的予測とある程度整合的であることを示している。続いて、文アラインメントとアノテーションを利用したルールベースのペアリング・アルゴリズムを構築し、各訳について3,000件超の参照対を抽出している。これを用いてイタリア語の品詞別に確率を算出した結果、固有名詞は全訳にわたり高率で保持される一方、人称代名詞とゼロ照応の扱いには訳者差が大きく、石川（1913）は代名詞の省略と名詞句化が多く、三上（1921）は代名詞を多用して明示化する傾向が顕著であることが明らかになっている。さらに、10種の形態的・談話的特徴量を入力とするロジスティック回帰モデルを訳者別に学習し、省略・固有名詞・代名詞・ゼロ照応などの予測精度を比較した結果、全訳で精度約0.8前後と安定した精度を示し、参照形式選択が無作為ではなく規則的パターンに従っていることが示唆されている。係数の解釈からは、省略やゼロ照応の連鎖、固有名による再アンカー、代名詞の翻訳側談話条件への依存といった一般傾向が抽出される一方、石川（1913）のゼロ重視・名詞句志向や三上（1921）の強い代名詞化など、訳者固有の傾向も浮かび上がっている。

第6章では、文・トークン・文節という三層のアラインメント分析を通じて、各訳の構造と語彙選択を検討している。まず複数の文アラインメントのアルゴリズムを比較し、埋め込みベースの手法が最も安定した精度を示すことを確認した上で、分割・結合種類の分布と長さ比から、石川（1913）による極端な分割傾向や三上（1921）の結合志向といった訳者差を示しつつ、多くの場合は情報量を保った再分節であることを示している。また未訳文の分布を調べ、矢口（1927）が書簡や挿話をまとまりごと削除する「抄訳」的性質をもち、石川（1913）・三上（1921）・矢口（1927）・生田（1928）の省略箇所が既刊の英訳と高い一致を示す点から、中継翻訳の影響と検閲上の配慮を数量的・質的に裏付けている。次に、伊日間の単語（トークン）レベルの対応について検証するためにデータセットを作成し、統計ベースとニューラルネットワークベースモデルを比較することで、統計的モデルの限界を指摘している。多言語の大規模コーパスと一部のデータで継続学習することにより性能向上が可能である一方、句読点や助詞列・複合述語周辺に体系的誤りが残るため、文体分析用の単位としては語より文節が適切であることを示している。最後に、大規模言語モデルにより8訳間の文節アラインメントを構築し、生田（1928）訳を中心とする整列語彙表からエントロピー重み付け距離と翻訳類似性ネットワークを作成することで、石川（1913）の孤立性、戦前訳・戦後訳双方への架橋点としての生田（1928）の中心性、および「最終」、「欲求」などの語の使用といった特徴を抽出している。これにより、生田が自ら標榜した「平明極まる大衆的文体」の下でも、ダヌンツィオ的な重々しさと終末感を保つ訳者文体が、遠読的手法からも実証的に把握しうることが示されている。

第7章では、本研究の射程と限界を明示し、その先の課題と展望を簡潔に示している。本論文は、『死の勝利』8訳からなる多訳コーパスを構築し、文体・アラインメント・共参照という三層の分析を通じて、訳者固有の文体と参照戦略の差異を実証的に描き出している点に特色がある。単一作品を精緻に追跡することで、年代差だけでは説明しきれない訳者行動の分岐と収斂を明らかにし、日本語翻訳文体研究および翻訳史研究に対して、新たな方法論的枠組みと再検討のための基盤資料を提示したものである。

以上の諸点を踏まえ、本論文は、1913年から2010年にわたる約百年の翻訳史を対象に、合計8訳の第1部・第2部からなる多訳コーパスを構築し、文体・アラインメント・共参照の三層から包括的に分析した研究として高く評価できる。コーパス言語学と計量文体分析の手法を翻訳研究に精緻に適用し、質的解釈と数量的分析を相補的に組み合わせ、訳者固有の文体特徴と参照戦略を実証的に解明している。とりわけ、イタリア語・日本語というパラレルデータの少ない言語対における多層レベルのアラインメントは、コーパスベースの翻訳研究における先駆的な試みであり、三層のアプローチによる精密な分析を通じて明らかになった各訳者の語彙選択と翻訳ストラテジーの多様性について説得力のある客観データを提示し、新たな知見を示した意義深い研究として、その学術的貢献は明確である。

なお、本論文の構成について、第6章の技術的な側面を第4章の後に配置することで、可読性が向上する余地があるとの指摘もなされた。また、分析対象が単一作品に限定されており、得られた知見の一般化には他作品や同時期の日本語原文との対照、また通時的変化などの多次元側面を捉えるための分析手法の工夫が、それぞれ今後の課題として残されている。しかしながら、これらは本論文の方法論的貢献と実証的成果の価値を損なうものではない。

以上により、本論文を博士（言語文化学）の学位論文として価値のあるものと認める。