

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 15 日現在

機関番号：57701

研究種目：基盤研究(C) (一般)

研究期間：2014～2016

課題番号：26370402

研究課題名(和文) ドイツ詩テキストデータベースを用いた比喻の枠組語としての都市インフラ関連語の研究

研究課題名(英文) A study on urban infrastructure related words as a framework word of metaphor using German poetry text database

研究代表者

保坂 直之 (Hosaka, Naoyuki)

鹿児島工業高等専門学校・一般教育科・教授

研究者番号：80280501

交付決定額(研究期間全体)：(直接経費) 1,500,000円

研究成果の概要(和文)：本研究は、テキスト形式で集約したドイツ語の詩作品データを使いながら、「20世紀初頭のベルリンの都市インフラ関連語彙がポエジーの詩語たりえているか」という小さな切り口を利用して、産業化の時代の大都市イメージがどのようなものかを探った。編集方針の異なる表現主義時代の抒情詩アンソロジーのデータも電子化して、その特徴を特徴語検索でつかむことを目指した。また、テキスト検索の手法を文学研究に生かす方法の確立も目指しており、表現主義の詩人トラークルの連作編集意図を調べる作業でテキストデータベースが機能するかを確認した。

研究成果の概要(英文)：In this research, applying the small research question, "can Berlin's urban infrastructure related vocabulary in the early 20th century be poetic words?", I explored the image of the metropolis in the era of industrialization. For this purpose, I used data from German poems, that I summarized in text format. I also digitized the data of the lyric poetry anthologies of the different editorial policies and aimed to grasp their features through specific keywords search. In addition, I aim to establish a search method using text files in the literature field, and confirm how the text database work out, when I check on intension to edit poetry cycles of Georg Trakl.

研究分野：ドイツ文学

キーワード：詩 ドイツ テキスト データベース 表現主義 比喻 インフラ ベルリン

1. 研究開始当初の背景

平成 22~24 年度に実施した科研費(挑戦的萌芽研究)による「比喩の枠組モデルを用いたドイツ詩のテキストデータベース研究」では、ブラックの「枠組み(frame)」と「焦点(focus)」からなる比喩モデルに着想を得て、詩的表現の地の色(コンテキスト)を色づける「枠組語」の一覧化を長期的な目的としつつ、「枠組語」と wie (～のような)などの指標のある「媒体」との相互検索により、上記問題点を解決したテキストファイルの文学研究への利用手段の確立を目指した。この研究では同時にテキスト形式で検索処理しやすいよう整形したドイツ詩のデータ集約も進めており、作品数で約 10000 のファイルが利用可能な状態となった。

本研究ではこの集約データを実際に使いながら、「20 世紀初頭のベルリンの都市インフラ関連語がポエジーの詩語たりえているか」という小さな切り口を利用して、比喩表現を「枠組語」と「焦点」の相互影響作用と仮定した比喩文彩のモデルの検討を行うことを目指した。同時にリスト化した都市インフラ関連語を「枠組語」として扱うテキストデータベースの検索方法を実践しながら確立することを目指した。

2. 研究の目的

切り口を「都市インフラ」にした理由は、1914 年 3 月後半のトラークルによるベルリン滞在を調べたことが契機である。トラークルは Wilmersdorf 地域に 2 週間ほど滞在し、流産した妹の大出血に直面してショックを受けている。この間、ヴァルデンが主宰していた Der Sturm 詩の同人であるベルリン在住の詩人・芸術家とも接している。当時のベルリンは普仏戦争戦勝後の都市インフラ整備が連綿として続いていて街の変貌は激しく、トラークルは開通直後の地下鉄 2 号線や電化されたばかりの Kurfürstendamm の街灯を見たはずだが、それらは彼の詩語にはなかった。体調不良の身重の妹を見舞うのがベルリン行きの目的であり、結局死産した子供の本当の父親はトラークルだった、という研究もある(Sauer mann)。Anhalter 駅の偉容も目に入らないのは心ここにあらずの彼の不安のためかもしれない。だが、彼をベルリンに迎えたラスカー・シューラーの作品でも都市インフラは詩語にはならない。

電子データの検索によって、都市インフラ関連語が詩語になるケース(リリエンクローン、ベン、ホディスなど)とならないケースを目で見えて確認できる。語を検索しながら読む作業では、「使用頻度の高い語彙」を拾うだけでなく、「仮説」を即座に調べられるため、「使用されない語彙」を逆算的に確定できる。通常の「読み」では不可能な読み方が、柔軟な検索処理の助けを借りることで可

能であり、この手法を実際に応用して都市インフラという人工構造物が詩的イメージとして働く状態を捉えて、「表現主義に内在するモダニズムと反モダニズムの対立」を明らかにしつつ、同時にテキストファイル検索の文学研究への利用方法自体を探る研究を目指した。

3. 研究の方法

1) 今研究期間ではデータベースの拡充よりも今まで整形して集約してきたデータを用いながら有効な検索手法を確立することを優先した。集めたファイルを必要に応じて変形して利用しやすくする作業、無理なくできる程度の詩データの拡充にとどめる。

2) 検索作業は、テキストエディタの Grep 検索機能に加えて、Perl を使ったフォルダをまたぐ 2 語同時検索のプログラムを併用して行う。

3) 都市インフラ関連語彙のリストについては、tr、sort などの簡単な unix のコマンドを使った単純な操作を組み合わせて作成する。この研究では手法を高度化するよりも、誰でも実行可能な単純な方法を確立して、ノウハウを文学研究者(または文学読者)に提供することを目指している。単純な方法を使うが、それによって何を発見できるかが文学研究としての価値になる。

4) テキストファイルを収めるフォルダの束ね方を変えて、作品群を一塊の語彙群にして観察する方法は有効である。例えば、トラークルの第 2 詩集『夢の中のセバステアン』はそれぞれ独自のタイトルでくくられた 4 つの連作詩(Zyklus)と 1 つの散文詩で構成されている。各 Zyklus 毎に機械的ソートを組み合わせると、第 2 連作だけ他の連作と詩語を共有していないことがわかる。同じ語が繰り返し自己模倣されるようにも見えるが、語彙の共通性がないことは機械を使わない通常の読みでは把握しにくい。本研究ではすでに集めた各時代のデータも使い、さまざまな検索範囲を仮設して語彙や有標の比喩イメージの傾きを調べる。

5) 作業の記録や検索の方法などと合わせて、整形したテキストデータ資料をインターネット上に公開する方法を探る。利用しているデータの中には商用のもの(DVD 化されたパッケージ商品など)もあって使用条件が一律ではない。著者の著作権の切れたデータとはいえ万一のことを考えて法的なトラブルを避ける措置を考えるが、この問題についてこの研究期間中に専門家の意見を聞いて対応を検討する。

6) 表現主義抒情詩のアンソロジー、大都市抒情詩のアンソロジーを使って上記3)の方法で都市インフラ関連語彙のリスト化を進める。また、コンコーダンスを使った統計計算も利用していく。

4. 研究成果

1) ドイツ詩テキストデータベース(=A)の整備についてはDVD版のデータ(Daten Deutscher Dichtung. www. Zeno. org. / Deutscher Lyrik von Luhter bis Rilke. Digital Bibliothek. / [Gutenberg.de] Edition12.)を取り込んで拡充し、また重複していたファイルの整理、文字コードのunicodeへの統一化を進めた。今研究期間中に詩のファイル総数は11,541、総語数は5,731,424となり、大きさとしては詩語に特化したコーパスとして比較作業に使えるものとなったため、各アンソロジーごとのデータなどと比較する作業(特徴語検索)が可能になった。

今研究期間以前に電子化・整理した表現主義時代の詩のアンソロジー(B = Fünfzig Gedichte des Expressionismus. Stuttgart 2002. C = Arbeitstexte für den Unterricht. Gedichte des Expressionismus. Stuttgart 1991.)の他に、20世紀初頭に出版されたアンソロジーのデータ(D: Möller, Heinz(Hg.): Großstadt Lyrik. 1903 Leipzig. / E: Pinthus, Kurt(Hg.): Menschheitsdämmerung. 1920 Hamburg. F: Hübner, Oskar/Mögelin, Johannes(Hg.): Im steinernen Meer. 1910 Berlin.)を入手し、OCR処理して電子ファイルとして整理してある。特にD、Fは自然主義vs.反自然主義モダニズムという対立する編集方針で編まれている、この編集方針には同時代のベルリンの政治動向も映し出されている。これらのデータを利用して今研究期間後の研究では20世紀初頭のアンソロジー編集方針に見られる都市インフラへの好悪の感情の調査が可能である。

2) ドイツ詩テキストデータベース(=A)を比較対象先として表現主義アンソロジーテキストデータ(=B、C)をコンコーダンス(AntConc)を使って特徴語検索し、表現主義時代の詩の特徴語としてBlut(血)、dunkel(暗い)、Meer(海)、rot(赤い)、Schatten(影)、Städte(都市)、Straße(街路)を確認した。これらの詩語を含むテキストをB、Cから拾い出し、同様にAにおけるこの詩語を含むテキストを集めて比較して、20世紀始めの都市イメージが「山々」「森」などのありふれた自然景物や神話的伝統世界と切り離された不快感につながるものであることを確認した。しかしながら表現主義時代の詩では「空」「星」などの世界を俯瞰的に捉える自然形象は多く歌われ、また「海」「河川」

という流れ、変化、広い世界への繋がりイメージも特徴的であったことを語彙検索で確認した。

アンソロジーデータ(B~F)からBahn(鉄道・軌道)、Bahnhof(駅)、Eisen(鉄)、Berlin(ベルリン)などのキーワード検索でファイルを集約して特徴語を検索した。Städte(都市)はdunkel(暗い)、Erde(大地)、rot(赤い)、Feuer(火)と、Meer(海)はfern(遠い)と共起する傾向があり、Meerはまた大都市の建物の連なりを示す比喩イメージとして多用されている。都市インフラ関連語は黒い色、赤い炎、騒音、速度などのイメージと結びついている。これらを詩語として使う詩人は多数派ではなく、ベン、シュタードラー、ポルト、ハイム、プラス、ケルナーなどである。総じて肯定的・未来的文脈で現れないのは産業化・都市化を非人間的と見る時代の傾向を反映してのことであろう。

これらについては口頭発表・論文報告した。

3) テキストデータの検索を自由に行うためには、語彙検索でヒットしたファイルを特定のフォルダに一括でコピーするなどの処理が必要になる。数百のファイルを手作業で操作することは不可能である。正規表現を使ったファイルパスの処理、tr、sort、cpなどのunixコマンドでの処理など、本研究で必要なスキルについても口頭発表・論文での言及で報告した。

4) テキスト検索の手法を使ってトラークルの第2詩集『夢の中のセバステアン』を調べると散文詩を除く4つの連作(Zyklus)ごとの使用語彙の傾きがはっきり観察できる。色彩語ではrot(赤い)は第2連作に多用されるが、silbern(銀の)は第2連作では避けられている。blau(青い)、schwarz(黒い)はどの連作でも均等に使用されているがweiß(白い)は第3連作を特徴づける色彩語である。Legend(伝説)は第1連作でのみ多用され、Sonnenblume(ひまわり)、Schmerz(痛み)は第2連作の特徴語、第3連作の特徴語はSchwester(妹)である。第4連作では第1連作で複数回現れた「ヒヤシンス」が現れる。この傾きから詩篇の編集において第1連作を詩人自身を主人公にした追憶、第2連作はSonja、Afraという女性を中心にした詩篇、第3連作は妹のイメージを意識した連作、第4連作は詩篇を円環として閉じるまとめの連作とする構成意図があったと推察される。また、トラークルの詩作法は表現主義の影響下にあるが、都市インフラ関連語彙には一切注意が向けられていない。

これらについては作品解釈も含めて口頭発表・論文報告した。

5) 整形したデータの著作権処理については専門家に調査を依頼し、KOTONOHA(現代日本語書き言葉均衡コーパス)の事例にならって

著者の著作権がない作品データであっても著作権者（出版社）と交渉して公開の承諾を得る方針が妥当、という結論になった。それを受けて出版社と交渉して公開のためにさらにデータを整える準備を進めた。

5. 主な発表論文等

〔雑誌論文〕(計3件)

保坂 直之、ドイツ詩テキストデータベースを使った表現主義時代の詩の特徴語の検出、高専ドイツ語教育、査読有、第16号、2017年3月、pp. 3-9.

保坂 直之、1913年秋のトラークルと秋の連作詩、トラークル研究、査読有、第13号、2016年10月、pp. 1-22.

保坂 直之、連作構造から見たトラークルの『カスパー・ハウザーの歌』、トラークル研究、査読有、第11号、2014年10月、pp. 54-76.

〔学会発表〕(計4件)

保坂 直之、テキスト検索で見る『夢の中のセバスティアン』第2連作後半の構成、トラークル協会2017年春季研究発表会、2017年5月27日、永福和泉地域区民センター(東京都・杉並区)

保坂 直之、『夢の中のセバスティアン』第2連作の構成と「パンとぶどう酒」の詩、トラークル協会2016年春季研究発表会、2016年5月28日、草加市文化会館(埼玉県・草加市)

保坂 直之、『冬の夕べ』の位置で考える連作「孤独な人の秋」の構成意図、トラークル協会2015年春季研究発表会、2015年5月30日、南大塚地域文化創造館(東京都・豊島区)

保坂 直之、テキスト検索で見る『夢の中のセバスティアン』の連作形式、トラークル協会2014年秋季研究発表会、2014年10月11日、京都テルサ(京都府・京都市)

〔その他〕

(論文掲載ホームページ)

<https://g-trakl.jimdo.com/>

6. 研究組織

(1)研究代表者

保坂 直之 (HOSAKA, Naoyuki)

鹿児島工業高等専門学校・一般教育科・教授

研究者番号：80280501

(2)研究協力者

カタリーナ・ガンドル (GANDL, Katharina)