

---

 書 評
 

---

## 最近出版されたロシア語文献のコンコーダンスについて

岡 本 崇 男

## 1. 五つのコンコーダンス

『ドストエフスキー「罪と罰」コンコーダンス (文脈つき用語索引)』(安藤厚, 浦井康男, 望月哲男編—以下 C0 と略す) が北海道大学スラブ研究センターから発行されたのは1994年のことであった。この労作にかんしては既に本誌第27号(1995)の書評で大きな評価が与えられただけでなく,<sup>1</sup> ロシア本国を含む海外の学術雑誌でも取り上げられており世界的に大きな反響を呼んだ。

このコンコーダンスが評価された理由の一つは、韻文ではなく散文作品を対象にしたことであろう。かつて僅かながら存在したロシア文学作品のコンコーダンスはレールモントフやマンデルシュタムの韻文が対象になっており、散文作品を扱ったものはなかった。しかし、韻文は言語活動における特殊な領域であるため、韻文作品のコンコーダンスもまた汎用性のある言語資料とは言い難い。もっとも、『プーシキン辞典』や『シェフチェンコのロシア語作品辞典』のように特定の作家の多様なジャンルの作品群に現れる語彙についてそれらの使用例を挙げる「作家辞典」は、語彙の意味を知る以外に、文中における語彙の使用例を見ることができるので、ある種コンコーダンス的な利用が可能である。しかし、一般にコンコーダンスでは一定の原則にもとづいて選別された語彙(または語形)を辞書順に配置して、対象となるテキストに出現するこれらの語形を含む断片が網羅的に提示される。この意味で作家辞典はコンコーダンスとしては不完全である。したがって、『罪と罰』に続く散文作品のコンコーダンスの登場が待たれていたわけであるが、このたび①島田陽・浦井康男編『「チェヴェンゴール」コンコーダンス』(東京国際大, 1996年—C1), ②浦井康男編『「大尉の娘」コンコーダンス』(福井大, 1997年—C2), ③浦井康男編『「ペテルブルグからモスクワへの旅」コンコーダンス』(北大, 1998年—C3) が相次いで発表された。

選定されたテキストは、それぞれ18世紀後半(『ペテルブルグからモスクワへの旅』), 19世紀前半(『大尉の娘』), 20世紀前半(『チェヴェンゴール』)の

作品である。これらの作品が素材として選ばれたのは、すべての業績の編者として名前を挙げられている浦井氏が C2, C3 の序文で述べているように、コンコードダンス作成の背景に、近代ロシア文章語の成立と変化を観察するという目標があることが理由となっている。

また、今年になって④中條直樹編「『ノヴゴロド第一年代記』(シノド本)コンコードダンス」(名古屋大, 1998年—C4)と⑤中條直樹編「『原初年代記』(ラヴレンチー年代記)コンコードダンス」(名古屋大, 1998年—C5)も発表された。中世ロシア語テキストを対象とした同種の業績はこれらが最初ではないかもしれないが、<sup>2</sup> 中世ロシア年代記のコンコードダンスが公表された例は極めて稀ではないかと思われる。

本稿では1996年から今年にかけて相次いで発表されたこれら五つのコンコードダンスを話題の中心に、コンコードダンスの意義と問題点について考えてみたい。

## 2. コンコードダンスの認知度

「コンコードダンス」は少なくともある程度の語彙数を持つ英和・独和・仏和辞典では「用語索引」あるいは「語句索引」という解釈が与えられている(英・仏語 concordance, 独語 Konkordanz)。また「(聖書などの)用語索引」と説明されることも多い。実際、インターネットを通じて利用できるアメリカのニューヨーク公共図書館 The New York Public Library のカタログ<sup>3</sup>で concordance および konkordanz をキーワードとして検索するとあわせて約900のタイトルが列挙されるが、これらのうち半数以上が新約・旧約聖書を対象にしており、文学作品では圧倒的に英語の韻文が多い。テキストの使用言語は様々であるが、英語が最も多くこれにフランス語、ドイツ語そしてヘブライ語が続いている。スラブ語テキストにかんしては Baratynskii, Batiushkov, Tiutchev, Akhmatova, Pushkin の名が見られるが、これらロシア詩人の作品のコンコードダンスも主として英語圏で作成されている。そして、ロシア語散文テキストを対象にしたコンコードダンスとして検索リストに含まれるのは C0 のみである。また、ドイツ語圏ではカントやヴィトゲンシュタインの著作のコンコードダンスが出版されていることが注目される。いずれにせよ、英独仏語圏ではコンコードダンスが出版物の一つのジャンルとして認知されていることがわかる。

しかし、日本やロシアでは「コンコードダンス」が語学文学系の専門家のあいだでもいまだに定着した概念とは言えないようである。たとえば、広辞苑(第

四版)でも見出し語として採用されておらず、定まった訳語さえ存在しない。ロシアの場合も事情は同じである。たとえば、オックスフォードの英露辞典で concordance を引くと、「Указатель (библейских изречений и т.п.)」とある。しかし、「указатель слов, словоуказатель, лексический указатель」等の語句を含む出版物はふつう語彙索引(word index)であって「文脈つき索引」ではない。ロシアでは現在のところ英語またはフランス語の音転写形と思われる *конкорданс* が好んで使用されているが、いまだに新語辞典等に採用されるには至っていないようである。

### 3. コンコーダンスの形式

コンコーダンスは特定の語形が出現する箇所を指示するだけでなく、実際どのような環境で使われているのかを具体的に示す点で、語彙索引と大きく性格を異にしている。韻文作品のコンコーダンスで頻繁に採用されているのは、ある語形が含まれている行全体を提示する方法であり、聖書であればその語形の意味を特定するのに必要かつ最小限の文脈(ふつう単文または複雑文の構成節)が提示される。一方、本稿で取り上げた五つのコンコーダンスには KWIC (Key-Word-in-Context) 形式が採用されている。この場合、キーワードとなる語形が中央に置かれ、その前後の環境がキーワードの左右に配置される。そして、いずれの形式にも文脈を提示する上での量的な制約が課せられており、ふつう一行を越えることがない。本来「文脈」というものに量的な基準がない以上、一行に収まる「文脈」にどれほど信頼を置けるのかという疑念が生じても不思議はない。特に KWIC 形式のコンコーダンスからは、機械的に並べられたテキストの断片を見せられているような印象を受けがちである。しかし、特定の語彙または語形と他の語彙または語形との規則的な結び付きを調べるためには KWIC 形式のほうが勝っている。

キーワードの選択方法には作成者の意図が強く反映される。つまり予めキーワードを絞り込んでおくか、あるいはテキストに現れる全ての語形の使用例を提示するかは、ひとえにコンコーダンス作成の目的にかかっているのである。汎用性のある言語資料としては前者の方法が望ましく、今回取り上げた五つはすべて網羅的である。もっとも、C2, C3 は辞書の見出し語のもとに語形の使用例を配置する方法を取っている (lemmatized concordance) のにたいして、C4, C5 は語形そのものが見出し語となっている。どちらも見出し語には出現頻度が付けられているが、特定の語彙の変化形を一覧できるという点で辞書の見出

し語を採用したほうが便利かもしれない。<sup>4</sup> C1 は見出し語がないが巻末に語形の出現頻度一覧表を付けている。なお、文法情報を知る手がかりとなる逆引き索引がどのコンコーダンスにも付けられていないのは惜しまれる。

#### 4. コンコーダンス作成にまつわる問題点と課題

最近になってロシア語テキストのコンコーダンスが相次いで発表されたのは、パーソナルコンピュータの処理能力の急速な向上と文科系の研究者のあいだでもコンピュータ利用が一般化したことが関係している。しかし、同種のコンコーダンスの作成・公表を将来にわたって継続していくためには検討すべき課題がいくつか残っている。

まず第一に、テキストをコンピュータ処理するためには、テキストが「電子化」されていなければならない。C1, C2, C3 は C0 と同じく光学的文字読み取り (OCR) によって作成された電子テキストを利用しており、C4, C5 はワードプロセッサまたはエディタに手作業で翻字入力したテキストデータを基礎に作られた。入力速度の点でも正確さの点でも OCR が勝っていることは言うまでもないが、底本が良質の活字で印刷されたものでない場合は OCR が利用できないかもしれない。電子テキストはそれ自体がテキスト・データベースとしての価値を持っているので、これを蓄積することが最も重要なことなのである。

次に検討しなければならないのはコンコーダンスを公表する形態である。一行に収まる範囲とはいえ、キーワードを文脈つきで提示すると、当然キーワードの数だけの行数が必要となる。たとえば、今回取り上げたコンコーダンスの頁数を見ると (括弧内の数字は1頁あたりの行数)、C1: 1684 (60), C2: 528 (61), C3: 668 (72), C4: 593 (83), C5: 1194 (71) となる。そして、これらのうち B5 版のものは C2 のみで他はすべて A4 版である。

このような紙に印刷するという出版形態を脱する試みとして、ロシアのペトロザヴォツク大学のザハロフ教授のもとで作成された「ドストエフスキー全作品コンコーダンス」にふれておく。これはロシア人文研究基金の助成によるプロジェクトで、インターネット<sup>5</sup> と CD-ROM という最新の媒体を利用している点が斬新である。これを利用すると、(1)使用環境を知りたい語形の先頭の文字の選択 (例えば «счастие» であれば C), (2)語形の先頭の二文字の選択 (СЧ.), (3)語形の選択 (СЧАСТИЕ) という、コンピュータソフトウェアとしては常識的な手続きを経て、その語形が出現する段落全体が表示される。また、収録されているのは完全な旧正書法テキストであり、テキストに対する厳密さ

と忠実さがコンコーダンス作成にあたっての大原則であることが伺える。段落全体が文脈として表示されることについて言えば、特定の語形を含むテキストの断片を一行に制限して提示する従来型のコンコーダンスの方が、問題の語形または語彙にまつわる表面的な使用パターンを見つけやすいかもしれない。もっとも、コンピュータソフトウェアという形態をとれば、利用者の目的に応じて文脈の長さが変化するコンコーダンスを作ることにもできるかもしれない。

(おかもと たかお・神戸市外国語大学)

## 注

- 1 井桁貞義書評, 『ロシア語ロシア文学研究』 27, 1995年, 105-108頁。
- 2 故佐々木秀夫先生が『イヴァン雷帝とアンドレイ・クループスキイ公の往復書簡』のコンコーダンスを作成されているはずであるが, 本稿執筆時にはそのコンコーダンスにかんする詳細を確認することができなかった。
- 3 インターネットで <http://www.catnyp.nypl.org/> にアクセスするとニューヨーク公共図書館の The Research Libraries On-Line Catalog を利用することができる。この図書館には稀覯本も含めて相当数のスラブおよびバルト関係の蔵書があり, 検索も高速に行える。ただし, インターネットによる検索は1972年以降に受け入れた書籍に限られる。
- 4 二つの年代記コンコーダンスが辞書的な見出しを付けなかったのは, いずれの年代記にもすでに語彙索引が公表されているからである。
- 5 [http://www.karelia.ru/~Dostoevsky/main\\_a.htm](http://www.karelia.ru/~Dostoevsky/main_a.htm)

藤沼貴著 『近代ロシア文学の原点——ニコライ・カラムジン研究』  
れんが書房新社, 1997年, 623頁

金 沢 美 知 子

二人の君主, ピョートル一世とエカテリーナ二世に代表される18世紀は, 紛れもなくロシア史における最も重要な時代のひとつであり, 歴史研究に数多くの興味深い題材を提供してきた。しかしながら文学の領域ではこれまで, ロシアの18世紀は脚光を浴びることが少なかった。むしろ18世紀が無内容な色褪せた時代だったわけではなく, その重要性は常に深く認識されていたのであり,