

NDL Searchによるジャンル名の分析

浜田 秀 (天理大学文学部) †

Analysis of Genre Names Based on the National Diet Library Search

Shu Hamada (Faculty of Letters, Tenri University)

1. はじめに

ジャンル名の用法はおおむね同時代の存在として実践的に使用されるもの（実践的用法）と、歴史的な存在として客体的に使用されるもの（歴史的用法）に分けることができる。たとえば「新体詩」という用語を目次や作法書に使用するのは実践的用法、文学史の記述に使用するのは歴史的用法である。実践的用法の存在は、そのジャンル名が積極的にジャンル構築に関わっていることを示す。

NDL Search (国立国会図書館サーチ) は国立国会図書館のHPの検索サービスであるが、コーパスとして使用することで、ジャンル名の用法の変遷がたどることができる。NDL Searchの検索結果を検討すると、「口語詩」という用語が1940年代に歴史的用法へと推移したことが分かる。

2. ジャンル名の推移について

明治から大正・昭和初期にかけてはジャンルの激動期である。ジャンル名の意味するところにもまた変動があった。

現在「詩」と言えば、漢詩や短歌、俳句から区別される、口語自由詩としての詩形を通常思い浮かべる。つまり、「口語詩」「自由詩」と取り立てて言わずとも、単に「詩」と言えば通じるのである。

現在の「詩」の語義は、このプロトタイプ¹としての口語自由詩を指す用法と、漢詩や短歌、俳句を含めた詩歌の総称としての用法の二義を中心とする。

だが「詩」という用語は近世までは中国の古典詩形、すなわち漢詩を意味した。「漢詩」という語が使用されるようになったのは明治以降である。「詩」という語彙は、いつプロトタイプの用法としての「口語自由詩」という意義を獲得したのであろうか。また「口語詩」という用語は、いつその実践性を失ったのであろうか。

文学史の記述は、「口語詩」という用語は1907年(明治40年)5月に、人見東明によって使用されたのが初出であること、作品としての口語詩は実質上同年9月の川路柳虹の「塵溜」に始まることを教える。当時、五七調文語文で書かれた「新体詩」から口語自由詩を区別して表現する用語として「口語詩」「自由詩」「言文一致詩」などが乱立した。しかしながら、「口語詩」が単に「詩」と呼ばれるようになった事は、語義の変化であっても事象の変化としては捉えられないため、文学史家の注意をさほど引かないようである。

ジャンル名の変遷は、言語の変化であると同時にそれを使用する一般の人々の常識の変化を示すものでもある。本調査では近代文体成立期のめまぐるしいジャンル名の変遷に定量的にアプローチする道を探りたい。

3. ジャンル名の用法

本調査では、ジャンル名「口語詩」を (a) 歴史的用法、(b) 理論的用法、(c) 所属的

† s-hamada@tenri-u.ac.jp

¹ 認知意味論におけるプロトタイプ概念については、レイコフ(1993)、テイラー(2008)を参照。

用法の三種に分けた。

- (a) 歴史的用法とは、過去の歴史的事象としての「口語詩」に言及するものである。
例：口語詩は、1907年に川路柳虹によって始められた。
- (b) 理論的用法とは、「口語詩」の一般的性格に言及するものである。
例：口語詩は、心のままに書かれるべきである。
- (c) 所属的用法とは、当該のテキストがどのジャンルに属するのかを決定するものである。たとえば、目次や、題名の横に「口語詩」と書かれる例が相当する。
例：口語詩 哀れにもまた勇ましき古い合戦の物語 星野水裏

(b) の理論的用法は、「口語詩」という語彙が何を意味するのかについて主張するものであり、我々の語義に対する常識を構築する行為であると言える。また (c) の所属的用法は当該の作品がどのジャンルに所属するのか、その社会的コンテクストを受手に明示し、そのようなものとして受手に提示する、遂行的表現である。いずれも「口語詩」という語彙が、他のジャンルとの区別を表現するために生きて働いていることを示す。両者を「実践的用法」と一括することができる。

- (a) 歴史的用法
 - (b) 理論的用法
 - (c) 所属的用法
- } 実践的用法

ただし (a) と (b) の区別はしばしば曖昧である。

- (1) 口語詩は俗語で書かれた詩である。その発生は1908年にさかのぼる。
- (2) 口語詩は俗語で書かれた詩である。卑俗にならぬよう注意する必要がある。

同一の文であっても、前後のコンテクストによって(1)は歴史的用法、(2)は理論的用法として区別される必要がある。

(c) の所属的用法は、いくつかの特徴を持つ。第一は、これはテキストそのものの一部ではなく、テキストに付随するという特殊な形式をもつ表現であるということである。ジュネット(2001)は、このようなタイトル・作者名・献辞・序文・挿絵・奥付など一連のテキストに伴う存在のことを「パラテキスト」と呼んだ。ジャンル名の所属的用法はパラテキストの機能の一部として考えられる。

第二は、これが文ではなく、語として提示され、あるテキストのジャンル所属を決定するという言語行為を遂行しているということである。ジュネット(2001)はタイトルにジャンル指示的機能を持つものがあることを指摘しているが、これは言語行為の一種として考えることができる。単語がそのまま言語行為を遂行するということは書記言語には豊富に見られる現象である。

第三は、これがテキストの社会的コンテクストを構成するものとして提示されるということである。テキストにアプローチする読み手は、これによってジャンルを知らされ、ジャンルの制度に従って読解を行うことになる。

第四は、これが読み手には、テキストと切り離せないものとして提示されるということである。題名とは異なり、ジャンル名は編者によって作者の意図に反したものが付けられる可能性もあるが、読者には、書籍に埋め込まれたテキストと不可分のものとして提示され、テキストをそのジャンルに所属するものとして受け取らざるを得ないということになる。

4. NDL Search とは

NDL Search (国立国会図書館サーチ) とは、7300 万件の文献情報、200 以上のデータベースと連携した統合検索サービスである。本格的な稼働は2012年1月であり、現在も発展し続けている。

NDL Search は、他の検索サービスと連携しており、ここにはNDL-OPAC、ゆにかねっと、CiNii Articles、CiNii Books、国立国会図書館デジタル化資料などが含まれる²。

NDL Search の利点の一つは、デジタル化された資料にリンクが貼られていることにより、現物を確認可能だということである。これは歴史資料を再入力したコーパスでは実現できないことである。現物を確認することで、誤記入の確認や、用法の判定が容易になる。

資料のデジタル化は急速に進んでいる。廣瀬 (2013) によれば、2012年3月末の時点で、和図書436万冊のうち90万冊、和雑誌431万冊のうち112万冊、全所蔵資料965万冊のうち225万冊のデジタル化が進んでいる。和図書については、明治期・大正期に刊行された図書のデジタル化をほぼ終え、1968年に受け入れたものまではデジタル化を完了したということである。

現在国会図書館のデジタル化資料は「国立国会図書館デジタル化資料」や「近代デジタルライブラリー」のように外部から閲覧可能なものと、館内でのみ閲覧できるものに分かっている。著作権法の改正により、2014年1月から他の図書館に送信サービスが始まる。これは国会図書館の館内閲覧資料のうち、入手困難な資料、雑誌、博士論文などについて、他の図書館から閲覧可能になるサービスである。ただし、現在も市場で流通しているものは申し出によって閲覧ができなくなる場合もある (廣瀬 2013)。

5. NDL Search の規模

NDL Search は言語分析のために設計・収集されたものではなく、狭義のコーパスには当てはまらない。これをコーパスとして使用するためには、当然検証が必要となる。ここではまずコーパスの規模について考える。なお調査結果はいずれも2013年7月のものである。

田野村(2009)によれば、「新潮文庫の100冊」が2000万字、BCCWJ2008の書籍データが6000万字である。文献情報7350万件との単純な比較は難しいが、コーパスとして使用可能な規模の言語量はあると思われる。

ところで、本調査が対象とする「口語詩」という用語についてはどうであろうか。「口語詩」で検索すると1450万字あるとされる「太陽コーパス」では0件、「少納言」でも1件しかヒットせず、これらのコーパスでは分析に耐える用例数を集めることができない。

一方、NDLsearchによる簡易検索では322件ヒットする。yahooのサーチエンジンで引用符に入れて検索すると9740件のヒットであるから、Web全体に対しても約30分の1程度の検索結果を得られたということになる (ただし、NDL Searchによるヒットするは検索語を含む文献数であって検索語の度数ではない)。田野村(2009)によれば、Web上の日本語文書は26兆字ということであるから、単純に考えると1兆字弱のweb情報から検索するのと同様の検索結果を入手できたわけである。これはかなり効率がいいといえる。

なぜこのような現象が生じたのであろうか。それはNDL Searchが書誌情報から構成されているという特殊性に起因する。

NDL Searchで検索できるのはあくまで書誌情報であり、「文」ですらない。助詞や助動詞といった機能語の検索に向かないのは言うまでもない。しかしながら、書誌情報であるからこそ含まれるものもある。それはジャンル名、とくにその所属的用法である。

NDL Searchのデジタル情報化資料には、目次情報が含まれる。これにより、本調査で言うところの所属的用法に相当するものが検索され、ヒットするわけである。NDL Searchがジャンル名のコーパスとして使用可能な所以である。

² 詳細は国立国会図書館ホームページの「検索対象データベース一覧 <http://iss.ndl.go.jp/information/target/>」を参照されたい。

なお、国会図書館のHPによる「口語詩」の検索結果を示す。

国立国会図書館デジタル化資料 9件

NDL Search 簡易検索 322件

検索結果の内「デジタル資料」 239件

検索結果の内「国立国会図書館デジタル化資料」 237件

NDL Search 簡易検索+「すべての連携先を検索」 445件

検索結果の内「デジタル資料」 295件

検索結果の内「国立国会図書館デジタル化資料」 237件

国会図書館のHP「国立国会図書館デジタル化資料」はNDL Search内の「国立国会図書館デジタル化資料」と異なり、外部から閲覧できるもののみが検索される。ここには「近代デジタルライブラリー」が含まれる。検索結果は9件であり、送信サービス開始後はこの状況は改善されると思われるが、現時点では数が少なすぎる。

またNDL Searchで「すべての連携先」にチェックを入れると、322→445といささか件数は増える（ただし「国立国会図書館デジタル化資料」については同数である）。しかし、その内容は、インターネット資料収集保存事業（WARP）によるネット上のものや、チェックを外した検索結果と重複の多い「国文学研究資料館国文学論文目録」などであって、これをあえて含める意味はあまりない。

コーパスの内実が十分に明らかになっていない現時点では「口語詩」を調査するためには、国会図書館に赴き、デジタル化された資料を確認するのがのぞましい。総数322件に対して「デジタル資料」239件はさほど遜色のない数字である。

ところでこの「デジタル資料」と「国立国会図書館デジタル化資料」とは概念の相違がある。「デジタル資料」はデータベースの種別を意味しており、国会図書館内で全て確認できるとは限らない。たとえばCiNiiの情報も出てくるが、CiNii自体がデジタルデータと考えられているためかリンク先にpdfが存在していないものも「デジタル資料」カウントされているようである。一方「国立国会図書館デジタル化資料」とは国会図書館でデジタル化した資料である。これは全て館内で閲覧可能であるが、CiNiiでpdfのリンクを持つものも排除されてしまう。

「デジタル資料」のうち、「国立国会図書館デジタル化資料」は237件であり、それに含まれなかった2例はJAIRO、CiNii Articlesのリンクから現物の確認が可能であった。結果として239点全ての確認が可能であった。

つまり、「口語詩」という用語については、総件数に比してデジタル資料が少ない、また「デジタル資料」が常に閲覧できるとは限らないという二つの問題をそれほど考慮に入れる必要はないと思われる。

6. NDL Searchの通時的構成

歴史的検証にNDL Searchを使用するためには、年度により資料の偏りがなくどうか確認する必要がある。NDL Searchでは検索結果に対して「出版年」ごとの内訳を見ることが出来る。これを利用してその通時的構成を調べた。

詳細検索画面で分類記号9（文学）を検索したところ、1672930件がヒットした。分類記号9の検索結果でもっとも古い「出版年」を示すものは0002（紀元2年）の仙石廬元坊『花供養』と邵寶『刻杜少陵先生詩分類集註』であるが、「詳細情報」の「出版年月日等」欄を見るとそれぞれ「寛保2（1742）序」「明暦2年（1656）」であり、出版年の「0002」は明らかに誤記入である。検索結果についてはある程度批判的に取り扱う必要があるようだ。

図1に明治初年（1868年）以来の9門全体と、そのうち「デジタル資料」の該当数の年

度ごとのヒット件数を示す(2013年は年度途中のために省略した)。なお検索時に「連携先」のチェックは外しているが、このチェックの有無で検索総数に変化はなかった。経年で見ると、以下のことがわかる。

- ・総件数は漸増の傾向にある。特に戦後は激増している。
- ・デジタル資料も増加の傾向にあるが、総件数ほどではない。1968年以降はデジタル化が進んでいないこともあり、件数が激減する。
- ・総件数・デジタル資料共に1944年、1945年は激減する。

分類記号9の検索結果は、1868～1912年で1397002件、うち「デジタル資料」152765件でおよそ10.9%となる。これは1968年以降も総件数が増加するのに比して、デジタル資料がほとんどヒットしないせいである。1868～1968で見ると、総数489124に対してデジタル資料140545件は28.7%とおおよそ3割に近くなる。1968年までの資料に関しては、戦時の2年を除き、「デジタル資料」の検索でもそれなりの検索結果が期待できると思われる。

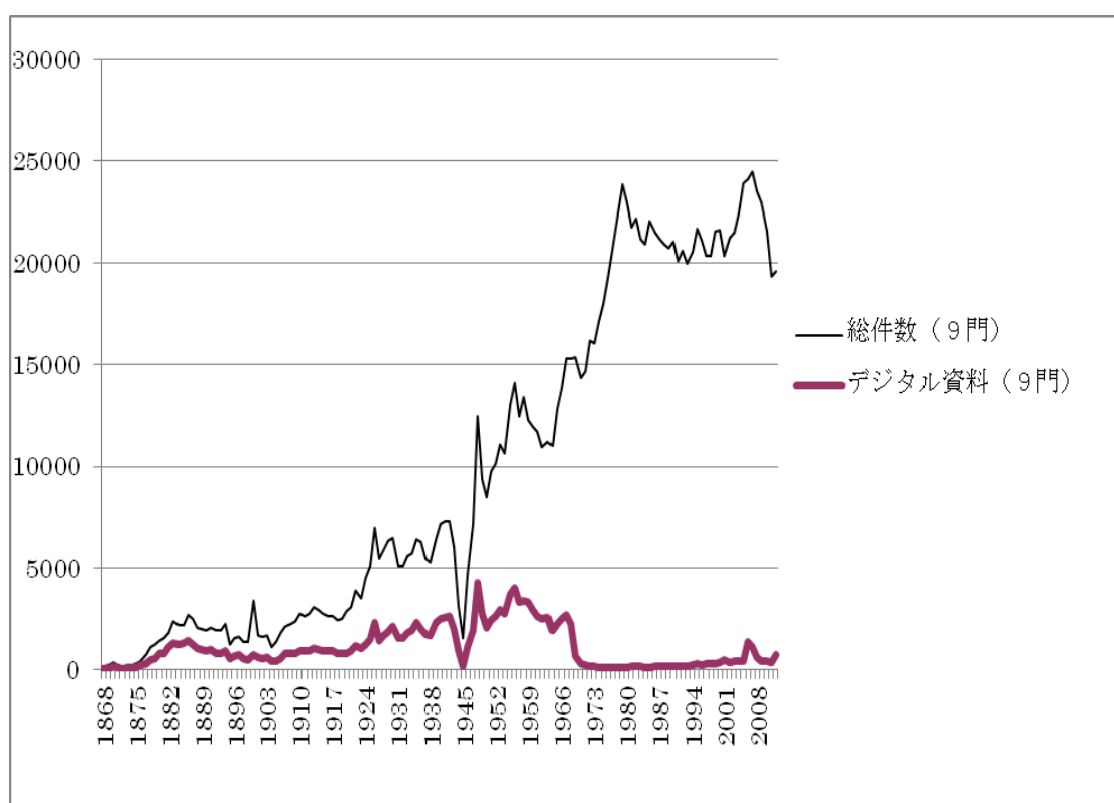


図1 NDL Searchの通時的構成(9門 1868～2012)

7. NDL Searchによる「口語詩」調査

7.1 調査方法

NDL Searchで「口語詩」に「デジタル資料」をチェックして簡易検索し、カウントされた239文献のうち、以下のものを除いた229件を対象とした。

- ・1969年以降の8件
- ・同一年度に出版された明らかに同一の書籍2点

カウントの仕方・用法の判定は以下の基準によった。

- ・「口語詩」および「口語詩」を含む語彙（「口語詩問題」など）をすべて検索事例に含めた。「口語詩人」「口語詩集」は「口語詩+人」「口語詩+集」ではなく、「口語+詩人」「口語+詩集」であるが、内容に鑑み検索対象に含めた。
- ・書籍・雑誌の1点を1カウントとした。一つの文献内に、複数の「口語詩」を含む文献が含まれるケースも1カウントとした。複数の用法が併用している例は見られなかった。
- ・明らかに文語詩に付属しているものも件数に入れた。
- ・選集等に再録されているもの、年度を変えて再版されているものもその時点での用例として意味を持つため、件数とした。
- ・著者の死後の出版は、理論的用法のものであっても歴史的用法に入れた。
- ・ある章の題名の一部として「口語詩」が使用されている時には、著書そのものではなく、章のレベルをコンテキストとして用法を判定した。

7.2 結果

結果は表1のようになった。

実践的用法としての理論的用法と所属的用法の出現はほとんど重なっている。いずれも1939~40年を境に全く見ることができなくなる。

歴史的用法は実践的用法より10年ほど遅れて出現し、現在にいたる。現在でも使用されているという意味では「口語詩」という用語は決して古語ではないが、ジャンルを実践的に弁別・構築する機能は既に果たし終えていると言える。

歴史的用法の初出は1919年であるが、これはその年出された『抱月全集 第1巻』に「口語詩問題」（初出は1908年）というエッセイが収められたためである。死後の所収であり基準に従い「歴史的用法」に入れたが、前年の抱月の急逝を受けた出版であり、口語詩の発生からおよそ10年しか立っておらず、歴史的用法が発生したと考えるにはやや特殊な事例である。はっきりとした歴史的用法といえるのは1924年の2点以降である（白鳥省吾『現代詩の研究』「前期の口語詩論」「相馬御風の口語詩論」と福井久蔵『日本新詩史』「口語詩」）。口語詩発生から17年を経ており、それなりに現象が回顧可能となったと考えられる。

我々は現在口語自由詩を「詩」とも「口語詩」とも呼ぶことができるはずである。しかし、「口語詩」という用語は口語自由詩を一般的に論じる時には使用されず、歴史的な経緯を背景知識とした時にのみ使用されるようになってきているのである。

NDL Searchによる理論的用法の初出は1908年であり、文学史上の初出から1年しか遅れていない。これは当時口語詩をめぐる議論が激しく交わされ、文献が多く出されたためである。

驚くべきことに、用語および口語詩テキストの出現からほとんど間をおかず所属用法が出現している。NDL Searchでヒットしたのは『中学文壇』『新文林』『少女の友』といった青少年向けの雑誌である。「口語詩とは何か」をめぐる議論がされている最中に、「口語詩」という語彙はジャンル構築の力をこれらの層に対して持っていたのである。

理論的用法は1908年の3例を除き、『少女詩の作り方』「文章詩と口語詩」（水谷まさる1922）、『新しい詩とその作り方』「口語詩と文語詩との区別」（室生犀星1925）といった啓蒙的なものにみられるのが主流である。概念を巡る戦いが終わり、常識の啓発のためにこの用語が使用された。しかし、それも1940年ごろを境に全く見られなくなる。「詩」が「口語で自由に書かれたもの」という概念が常識となり、「口語」という用語で他ジャンルと弁別する必要が無くなっていったのだろう。

表1 「口語詩」用法の通時的変化

年代		実践的用法		歴史的用法	年代		実践的用法		歴史的用法
		理論	所属				理論	所属	
1968	昭和43			2	1937	昭和12	1		
1967	昭和42				1936	昭和11	1		
1966	昭和41			2	1935	昭和10			
1965	昭和40			1	1934	昭和9			1
1964	昭和39			1	1933	昭和8		1	2
1963	昭和38			5	1932	昭和7			
1962	昭和37			5	1931	昭和6			1
1961	昭和36			11	1930	昭和5	1		
1960	昭和35			2	1929	昭和4	1		
1959	昭和34			1	1928	昭和3		1	2
1958	昭和33			5	1927	昭和2			1
1957	昭和32			1	1926	昭和1			1
1956	昭和31				1925	大正14	1	3	
1955	昭和30			4	1924	大正13		9	2
1954	昭和29			6	1923	大正12		4	
1953	昭和28			6	1922	大正11	1	6	
1952	昭和27				1921	大正10		7	
1951	昭和26			1	1920	大正9		7	
1950	昭和25				1919	大正8		15	1
1949	昭和24			1	1918	大正7	1	16	
1948	昭和23				1917	大正6		16	
1947	昭和22				1916	大正5		7	
1946	昭和21				1915	大正4		11	
1945	昭和20				1914	大正3		11	
1944	昭和19			1	1913	大正2		4	
1943	昭和18				1912	大正1			
1942	昭和17				1911	明治44		5	
1941	昭和16			1	1910	明治43			
1940	昭和15				1909	明治42		20	
1939	昭和14		2		1908	明治41	3	6	
1938	昭和13	1			合計		11	151	67

8. まとめ

NDL Search を使用し、1908 年から 1968 年にわたる「口語詩」の用法を 2 類 3 種に分けた上で、その推移を分析した。

理論的用法・所屬的用法といった実践的用法は、そのジャンル名が生きて働いており、他の諸ジャンルと区別するために積極的な力を持っていることを示す。

「口語詩」の使用においては、実践的用法が先行し、実践的用法・歴史的用法の共存の時期を経た上で、歴史的用法のみの時期が続く。現在の我々にとって、口語詩という用語は歴史的な文脈を前提としたものへと変化してしまっている。

ジャンル名の用法は、このようなジャンル構造の変遷を示すものである。特にジャンル名用法の推移を調べることは、ジャンルに所屬するテキストの属性の変化のみならず、ジャンルをカテゴライズする我々の知識の変化をも調べることになる。我々がいかにジャンル名を使いこなし、ジャンルを認知・構築・管理しているのか、ジャンル名の調査が教え

てくれるところは大きいと言える。

謝 辞

本調査は、平成 25 年度天理大学 学術・研究・教育活動助成費による研究結果の一部である。

文 献

国立国会図書館サーチパンフレット (日本語)

(http://iss.ndl.go.jp/information/wp-content/uploads/2013/05/pamphlet_japanese_201303.pdf)

よりダウンロード可能)

国立国語研究所編(2005)『太陽コーパス—雑誌『太陽』日本語データベース』博文館新社

ジュネット(2001)『スイユ テキストから書物へ』水声社.

田野村忠温(2009)「コーパスからのコロケーション情報抽出—分析手法の検討とコロケーション時点項目の試作—」『阪大日本語研究』21、pp.21-41

テイラー(2008)『認知言語学のための 14 章 第三版』紀伊國屋書店.

服部嘉香(1963)『口語詩小史—日本自由詩前史—』昭森社.

浜田秀(2010)「ジャンル」『認知物語論キーワード』、pp.87-94、和泉書院.

人見円吉(1975)『口語詩の史的研究』桜楓社

廣瀬信己(2013)「国立国会図書館のデジタル化資料の図書館等への送信に向けて」『図書館雑誌』107(2)、pp.86-88

レイコフ(1993)『認知意味論：言語から見た人間の心』紀伊國屋書店.

関連 URL

NDL Search 簡易検索 <http://dl.ndl.go.jp/>

国立国会図書館デジタル化資料 <http://dl.ndl.go.jp/>