

現代日本語書き言葉均衡コーパスに基づく 外来語音の表記に関する試論

単 珊 (東京学芸大学大学院)

白勢 彩子 (東京学芸大学)

Notation of the Loanword Based on the Balanced Corpus of Contemporary Written Japanese

San Zen (Tokyo Gakugei University)

Ayako Shirose (Tokyo Gakugei University)

1. はじめに

外来語とは、「自国語に組み入れられた外国語のことで、厳密には^{レイディオウ}radio[reidiou]がラヂヲ、ラジオなどと語形をとるように、受け入れた国の音韻組織によって再構成された語形をもつ語」(杉本, 2008)を指す。現在では、外来語と言えど主として欧米系の諸言語に由来するものを指すことが多く、片仮名で書き表すのが一般的である(国語審議会, 1991)。外来語の片仮名表記については、基本的には「ア」、「カ」、「キャ」など国語の音に対応する仮名が用いられているが、これらに加え、外来語の原音に応じた、特別の仮名表記が工夫され、現在、表記のよりどころとして、平成三年(1991年)に公布された「外来語の表記」がある。この「外来語の表記」には、外来音に対応する仮名として「シエ」、「ジェ」、「トゥ」、「ドゥ」などが挙げられている。これらは、例えば「いつもそう書かなければならないことをその意味するものではない」とされ、慣用により「セ」「ゼ」などとしてよいとされている。すなわち、同一の外来音に複数の片仮名が用いられることが許容されることとなっているのが現状である。

外来語の表記には、複数の表記、いわば「ゆれ」があることになり、この問題についてはいくつかの研究がある。福盛(2010)は、「ブ」と「ヴ」など数種に対して、調査を行ない、例えば、「ブ」と「ヴ」では、比較的近年に取り入れられた外来語について「ヴ」の使用が増えていることが得られ、また、固有名詞で登録されている場合に、より多くの使用が認められることがわかった。この結果は地名・人名は原音に近い形で表記されることが多いことによると考えられた。松崎(1992)では、辞書類における外来語音の表記の「ゆれ」の実態を調査して、ゆれの類型が明らかとされ、続く松崎(1993)では、17種の辞書類の外来語表記を調査し、「ゆれ」が視覚化され、ある拍が外来語音寄りか日本語音寄りかは語により異なることが数量的に再確認されている。しかしながら、これらの研究では調査対象の範囲が限られている、辞書を調査したもので必ずしも表記の実態を捉えたものではないなどの問題がある。さらにいえば、そもそも外来語の原語の音に対し、どのような片仮名に対応するのかといった基本原則も明瞭に示されてきていないようである。

そこで、筆者らは外来音と仮名表記の対応について、基本原則を明らかとし、使用の実態を明らかにすることを目的とした研究を進めている。本稿では、原音と表記の関係についての概略を示し、これに基づき「現代日本語書き言葉均衡コーパス(BCCWJ)」を調べた結果について報告する。

なお、本稿では、「シエ」「ジェ」「トゥ」といった外来音に対応する仮名を「外来音仮名」と呼び、「ア」「カ」「キャ」など国語の音に対応する仮名を「国語音仮名」とする。

2. 原音と表記の対応関係

2.1 外来音仮名と国語音仮名の関係

現行の「外来語の表記」では、「ミルクセーキ」と「ミルクシェーキ」の二様の表記が可能となる。松崎（1992, 1993）は、外来音仮名と国語音仮名の両用がある時、二者にどのような関係があるかを調査した。これによると、1拍にあたる外来音仮名と国語音仮名は、A：1拍にあたる外来音と国語音2拍分がゆれているもの、B：母音を共通要素として、外来音と国語音がゆれているもの、C：子音を共通要素として、外来音と国語音がゆれているものの3類に関係性が整理されている。Aの例としては、「シェ」と「ジェ」、Bの例としては、「シェ」と「セ」、Cの例としては、「シェ」と「シュ」がある。このように、外来語の表記において、同一の音に対し、仮名が複数、存すると考えられるが、原語での発音に対して片仮名表記がどのようになされているかについては、明瞭でない。そこで、原音（原語の発音）と片仮名表記がどう対応しているか調べるため、以下の調査を行った。

2.2 原音に基づく表記調査

調査は、『コンサイスカタカナ語辞書 第4版』を資料とし、前述の松崎（1992）で指摘された外来音仮名とそれを国語音で表記する場合の仮名のいずれかが含まれ、「ア」「イ」で始まる見出し語を調査対象とした。これらの語について、語の原綴りと原発音記号を調べた。調査の対象となった語は合計2,036語であった。うち、国語音仮名でのみ表記される発音があり、これらは分析対象外とした。データの結果を表1、表2に示す。表1、2の「分類」にある「c」、「v」は、それぞれ、子音、母音を指す。「cv」は、子音+母音の音節、「c_{vv}」は子音+二重母音の音節、表2の「c.v/c」は、子音音節の後に母音、あるいは子音が続く音節、「cc.v/c」は二重子音の後に母音、あるいは子音が続く音節、「c#」とは、子音で終わる場合を示している。[]内に原音の音声記号を示し、ハイフンで対応する仮名を示した。

表1に示すのは、原音に対し仮名表記が1種のみ、単独に対応する関係である。「cv」では[da]に対する「ディ」、[dø]に対する「ドゥ」、「c_{vv}」では[fəu]に対する「フォ」などがある。仮名を軸にして捉えると「ドゥ」は[dø]、[du]の2種の原音に対し、用いられていることがわかる。

表2に示すのは原音に対し仮名表記が2種以上、複数に対応する関係である。最大で7種の仮名で表記される音があることが示されている（[ta]に対する「チ」、「ツ」、「テ」、「ト」、「チュ」、「ティ」、「トゥ」）。対応する仮名の数は2~7種の範囲にあるものの、2種の仮名が用いられる音が最も多いことがわかった。2種の仮名が用いられるのは例えば[tu]に対する「ツ」、「トゥ」などである。

表1と表2を比較すると総数に大きな差はない（「出現数」の総数：表1が22、表2が19）。このことは必ずしも外来音仮名で表記される音が複数の表記を持つわけではないことを意味する。表1のように音と仮名が一对一に対応するものは原音からみて「ゆれ」がなく、一方、表2のように複数あるものは原音からみて「ゆれ」があるといえる。

表 1 単独に対応する仮名

分類	原音と仮名の対応	出現数
cv	[da]—ダイ [dø]—ドウ [du]—ドゥ [fa]—ファ [fæ]—ファ [fe]—フェ [fɛ]—フェ [fə]—フィ [sə]—シュ [ʃu]—シュ [tʃe]—チェ [tʃu]—チュ [dʒe]—ジェ [dʒu]—ジュ [dj]—ディ [dju]—デュ [wɛ]—ウエ	17
cvv	[fəu]—フォ [tʃuə]—チュ	2
c.v/c	[dju]—デュ	1
cc.v/c	[fy]—フュ	1
c#	[ʒ]—ジュ	1

表 2 複数に対応する仮名

対応する 仮名の数	原音と仮名の対応	出現数	
2	cv	[tu]—ツ・トゥ [fə]—ファ・フォ [fə]—ファ・フォ [va]—バ・ヴァ [ve]—ベ・ヴェ [tʃə]—チュ・フュ [dʒi]—ジ・ジェ [tju]—チュ・トゥ [we]—ウエ・クエ	11
	cvv	[ʃei]—シエ・シュ	
	c.v/c		
	cc.v/c		
	c#	[ʃ]—シエ・シュ	
3	cv	[fi]—ヒ・フィ・フェ [se]—セ・ゼ・チェ [dʒə]—ジ・ゼ・ジェ	5
	cvv		
	c.v/c	[d]—ジ・ディ・ド [s]—ズ・セ・シュ	
	cc.v/c		
	c#		
4以上	cv	[di]—ジ・デ・ティ・ディ [ti]—ズ・チ・テ・ト・ド・ティ [tə]—チ・ツ・テ・ト・チュ・ティ・トゥ	3
	cvv		
	c.v/c		
	cc.v/c		
	c#		

表2にあるように、原音に対し、仮名表記が複数ある場合、音と仮名がどのような関係にあるかについて、音の側面から規則性があるかどうか検討した。原語の発音を仮名の発音と比較し、母音が同じかどうか、子音が同じかどうかの観点により整理した。例えば、原語の発音記号が[va]であるものと対応する仮名表記「バ」の音声は、[ba]である。母音が同じ[a]であるが、子音はそれぞれ[v]、[b]と異なっているといえる。

[va]とバ[ba]の関係は、子音が非共通で母音が共通する「子音非共通型」に分類できる。現在、この「子音非共通型」について、音源（有声・無声）・調音点・調音様式の面から、相違点を整理し、詳細に分析している。結果としては、例えば、[dʒi]とジ[dzi]のように硬口蓋性の有無で関係づけられるものも多く見られるなど法則性があることが示唆されている。子音共通で母音が非共通であるものは「母音非共通型」に分類でき、[tu]とトゥ[tuu]、[di]とデ[de]などの例がある。これについても母音の発音の特徴から整理することを試みている。

3. コーパス調査

3.1 調査概要

上記のように、原音と片仮名表記の対応関係を把握したが、この結果に基づいて、BCCWJにおける外来音表記の用いられ方を調査し、表記と発音間のゆれの現状を把握したい。上記では、原音と片仮名表記の対応関係を「子音非共通型」、「母音非共通型」として捉えた。これらのタイプのうち、今回は「子音非共通型」の「ヴァ」、「母音非共通型」の「トゥ」を抽出し、BCCWJにて検索し、結果を整理する。前述の仮名を含む語のうち、出現語数をもっとも多かった語の上位5位から、それぞれ「子音非共通型」2件、「母音非共通型」1件を抽出した。各自の語彙素でBCCWJにより検索し、得た結果をジャンル別の出現頻度から、相違点を検討する。

3.2 調査結果と分析

外来音仮名「ヴァ」の検索結果より、語数をもっとも多かった上位五位のうち、「ヴァ/バイオリン」と「ヴァ/バイオレット」の二語を抽出し、それぞれのジャンル別の出現頻度を表3と表4にまとめた。

表3に示したように、「ヴァ/バイオリン」という語においては、「ヴァ」の表記を用いる語が415語あった。それに対して、「バ」の表記を用いる語は293語であった。「ジャンル」により各語の出現傾向を見ると、「ヴァイオリン」については、第五位までの語で出現頻度が全体の73.7%を占め、その他のジャンルは26%程度である。つまり、「ヴァイオリン」の出現頻度が、第五位までのジャンルに偏っているといえる。特に「芸術・美術」のジャンルに多く出現している(43.3%)。

一方、「バイオリン」では、ジャンル中、第五位までの語の合計が全体の52.2%を占め、その他のジャンルは47.8%となっている。「ヴァイオリン」に見られたような第五位までに集中して出現する傾向はないといえる。第五位までの各ジャンルについては、「文学」19.5%、「芸術・美術」11.6%、「社会科学」8.2%、「エンターテインメントと趣味」6.8%、「関東地方」6.1%であり、「文学」での出現頻度が高いものの、いずれかのジャンルに偏って出現する結果とはなっていない。つまり、「バイオリン」は平均的に分布しているといえる。

以上、まとめると、「ヴァイオリン」については、「芸術・美術」や「文学」という特定のジャンルに傾いて出現する傾向がみられる。これに対し、「バイオリン」は各ジャンルに平均的に分布されていることがわかった。

表4に示した「ヴァ/バイオレット」という語について見ると、「ヴァ」の表記を用いる語が115語、「バ」の表記を用いる語が38語あった。「ヴァイオレット」においては、ジャンルの第五位までで出現頻度が97.5%となり、その他のジャンルは2.5%に過ぎない。中でも、「文学」のジャンルにおける出現率をもっとも多く、91.3%に達している。つまり、「ヴァイオレット」は非常に偏って特定のジャンルに出現しているといえる。一方、「バイオレット」では、第五位までの出現頻度が76.3%で、その他のジャンルが23.7%である。上位第五位までのジャンルの中で、それぞれ、「技術・工学」26.3%、「総合」23.7%、「文学」10.5%、「分類なし」7.9%、「科学」7.9%となっている。「バイオレット」については、特定のジャンルにやや偏っているものの、「ヴァイオレット」に比しては平均的に分布しているといえる。

表3「ヴァ/バイオリン」の検索結果

語例	総語数	ジャンル	出現頻度
ヴァイオリン	415	芸術・美術	43.3(%) 180
		文学	14.5 60
		エンターテインメント	8 33
		芸術	4.8 20
		Yahoo!サービス	3.1 13
		その他	26.3 109
		バイオリン	293
芸術・美術	11.6 34		
社会科学	8.2 24		
エンターテインメントと趣味	6.8 20		
関東地方	6.1 18		
その他	47.8 140		

表4「ヴァ/バイオレット」の検索結果

語例	総語数	ジャンル	出現頻度
ヴァイオレット	115	文学	91.3(%) 105
		分類なし	3.5 4
		哲学	0.9 1
		社会科学	0.9 1
		自然科学	0.9 1
		その他	2.5 3
		バイオレット	38
総合	23.7 9		
文学	10.5 4		
分類なし	7.9 3		
科学	7.9 3		
その他	23.7 9		

1「総語数」はそれぞれ、語彙素である「バイオリン」と「バイオレット」で検索した語数のことである。

2「ジャンル」の欄では、検索結果の上位五位までを載せる。五位以下は「その他」に含めた。

3「出現頻度」の上段には、総語数に対する相対頻度をパーセンテージで示し、下段には実数を示した。

以上、まとめると、「ヴァイオレット」については、特定のジャンルに非常に偏向しており、これに対して、「バイオレット」ではいくつかのジャンルに偏るものの、「ヴァイオレット」に見られるような特定のジャンルに集中して分布する様相はないことがわかった。

続いて、外来音仮名「トゥ」の検索結果の中、語数をもっとも多い上位五位のうち、「トゥール」を抽出し、調査結果を表5にまとめた。「トゥール」に対応する「ツール」についても同様に表5に示す。

表5 「トゥ/ツール」の検索結果

語例	総語数	ジャンル	出現頻度
トゥール	107	文学	63.6(%) 68
		歴史	12.1 13
		芸術・美術	6.5 7
		総合	6.5 7
		社会科学	4.7 5
		その他	6.6 7
ツール	2972	インターネット、PC と家電	20.4 607
		総記	16.2 482
		技術・工学	12.6 374
		社会科学	5.6 166
		自然科学	4.0 117
		その他	41.2 1226

- 1 「総語数」はそれぞれ、語彙素である「トゥール」と「ツール」で検索した語数のことである。
- 2 「ジャンル」の欄では、検索結果の上位五位までを載せる。五位以下は「その他」に含めた。
- 3 「出現頻度」の上段に、総語数に対する相対頻度をパーセンテージで示し、下段には、実数を示した。

表 5 に示したように、「トゥ/ツール」という語では、「トゥ」の表記を用いる語が 107 語あり、「ツ」の表記を用いる語が 2972 語あった。「ヴァ/バイオリン」、「ヴァ/バイオレット」では外来音表記の方が多く出現する傾向が見られており、「トゥ/ツール」には異なる特徴が見受けられる。これまでと同様、「ジャンル」の出現傾向を見ると、「トゥール」では、第五位までの語の出現頻度が全体の 93.4% を占め、その他のジャンルの出現頻度は 6.6% である。加えて、上位五位のジャンル中、「文学」のジャンルに集中して出現しており、2 位の「歴史」と合わせると 75% を超える。このことから、「トゥール」もまた、特定のジャンルに偏って出現しているといえる。

「ツール」では、第五位までの語の出現頻度で 58.8% となり、各ジャンルの出現頻度がそれぞれ、「インターネット、PC と家電」20.4%、「総記」16.2%、「技術・工学」12.6%、「社会科学」5.6%、「自然科学」4% であり、「インターネット、PC と家電」での出現頻度が高いものの、いずれかのジャンルにのみ偏って出現する結果とはなっていない。つまり、「ツール」は比較的均等に複数のジャンルに分布しているといえる。

以上、まとめると、「トゥール」が特定のジャンルに傾く傾向があり、これに対し、「ツール」では、特定のジャンルへの偏りはなく分布しているといえる。

4. まとめ

本研究は、外来音の仮名表記について、表記の現状を把握するため、まず、原音と仮名表記の関係を示し、次いで、「シェ」、「チェ」といった外来音に対応する仮名の表記を調査の対象とし、外来語表記の用いられ方を、BCCWJ により検索し、ジャンルによる出現の様相の相違について、検討した。

検討の結果、次の点が明らかとなった。

1. 原音と仮名の関係を整理したところ、外来音仮名に対応する音が必ずしも複数の表記を持つわけではないことがわかった。
2. 「ヴァイオリン」については、「芸術・美術」という特定のジャンルに傾く傾向がみられた。これに対し、「バイオリン」は各ジャンルに平均的に分布されていることがわかった。
3. 「ヴァイオレット」については、特定のジャンルに非常に偏向して出現しており、これに対して、「バイオレット」ではいくつかのジャンルに偏るものの、「ヴァイオレット」に見られるような特定のジャンルへの偏向はないことがわかった。
4. 「トゥール」については、「文学」という特定のジャンルに傾く傾向があり、これに対し、「ツール」では特定のジャンルへの偏りはなく、比較的均等に分布していることがわかった。

以上をまとめると、「ヴァ」、「トゥ」の外来音仮名が用いられた場合では、いくつかの特定のジャンルに傾いて出現する傾向がみられるといえる。これに対し、国語音仮名が用いられた場合では、特定のジャンルへの偏りはなく、複数のジャンルに分布する様相があることが示された。今回、こうした分析を行うことにより、外来語の表記において、これまで把握されにくかった原音と仮名の関係を明瞭化する指針を得ることができ、また、大規模コーパスを用いることによって外来語表記における仮名の使用状況の実際を整理して、仮名による相違点を検討することができ、今後の研究の基礎的な論点を示すことができたと考えている。

上記のような点が明らかとなったが、以下の問題点も存在している。辞書に基づく表記

調査においては、収集したデータの結果が一部分の外来音仮名に偏る現象が見られ、十分なデータを得たとは言いがたい。また、コーパス調査においては、語彙素を用い、検索を行うことなど調査方法の妥当性について、検討の余地がないとはいえない。コーパス調査のデータに関しては、未整理のデータが含まれていることも今後の課題である。例えば、外来音「トゥ」の検索結果では、もともと数が多かった外来音の「トゥ」と対応する原語が「to」と「two」の両方が含まれており、今回は、「トゥ」についての分析を省略した。今後は、まず、十分なデータを得たうえ、コーパス調査の手法について十分検討した上でデータを議論する必要がある。さらには、「ジャンル」以外の項目についても、詳細な分析を進め、外来語を仮名表記することに関連・影響する要因を明らかにしていきたいと考えている。

文 献

- 福盛貴弘(2010)『基礎からの日本語音声学』東京堂出版.
- 国語審議会(1991)「外来語の表記」前文『公用文の書き表し方の基準(資料集)増補二版』, pp.197-201.
- 松崎寛(1992)「外来語音におけるゆれの類型:辞書類の表記を中心として」言語学論叢, 10:11, pp.43-56.
- 松崎寛(1993)「外来語音の表記のゆれに関する定量的研究」東北大学文学部日本語学科論集, 3, pp.83-94.
- 三省堂(2010)『コンサイスカタカナ語辞典 第4版』三省堂編修所.
- 杉本つとむ(2008)「外来語」『日本語学研究事典』pp.408-411, 明治書院.