

国立国語研究所学術情報リポジトリ

Kanji use in seventy contemporary magazines and
Jōyō Kanji : A preliminary study for the
application of text corpora to Japanese language
policy

メタデータ	言語: jpn 出版者: 公開日: 2019-03-25 キーワード (Ja): キーワード (En): 作成者: 小椋, 秀樹, 相澤, 正夫, OGURA, Hideki, AIZAWA, Masao メールアドレス: 所属:
URL	https://doi.org/10.15084/00002186

現代雑誌70誌における漢字の使用実態と常用漢字表

—— 国語施策へのコーパス活用に向けた基礎調査 ——

小椋 秀樹

(国立国語研究所)

相澤 正夫

(国立国語研究所)

キーワード

コーパス, 現代雑誌200万字言語調査, 国語施策, 常用漢字表

要 旨

本稿の目的は、今後構築される大規模コーパスを国語施策にどのように活用していくか、そのためにはどのようなコーパスを構築する必要があるかについて、具体的な検討材料を提供することにある。そのため、現代雑誌70誌における漢字の使用実態に基づいて、常用漢字表の持つ漢字使用の目安としての機能を検証し、必要な見直しの方向を探った。

具体的には、「現代雑誌200万字言語調査」を利用して、常用漢字・表外漢字の出現状況、表外音訓・表外漢字の音訓の出現状況について報告した。その結果、上位2,000字における常用漢字の出現状況、及び表内音訓の出現状況から、調査対象となった1994年の時点では、常用漢字表が、頻度の面からも、音訓使用の面からも、漢字使用の目安として機能していたことが明らかになった。しかし、その一方で高頻度の表外漢字や表外音訓があることなどから、漢字使用の実態とは若干のずれも見られた。さらに、「現代雑誌200万字言語調査」の結果に基づいて常用漢字表を見直す場合、常用漢字に新たに追加する音訓の候補があるか、常用漢字表に新たに追加する漢字の候補があるかについて議論した。

1. はじめに

国立国語研究所は、明治時代から現代に至る日本語の全体像を把握するための大規模な言語コーパス KOTONOA の構築を、2006年度から開始した。まず2006年度から2010年度までの5か年で、現代日本語の書き言葉を対象とした「現代日本語書き言葉均衡コーパス」(以下 BCCWJ とする。)を構築する。BCCWJ は、1976年～2005年の30年間に出版された日本語の書き言葉のコーパスで、従来、国立国語研究所が語彙調査で対象としてきた新聞・雑誌のほか、一般の書籍も対象に加えている。コーパスの規模は1億語以上を目標としている¹。

BCCWJ は、国語学・日本語学といった言語研究での利用はもちろん、情報処理や国語教育、外国人のための日本語教育など、幅広い分野での活用が期待される。そのような様々な分野における活用の中でも、特に国立国語研究所が主体的に取り組むべき重要な課題として、国語施策への活用がある。国立国語研究所の主要な任務の一つに、科学的な調査研究によって現代日本語の実態を解明するとともに、その結果に基づいて、国語施策の立案に資する資料を作成し提供することが挙げられているからである。

国立国語研究所は、これまでに大規模な語彙調査・文字調査を実施し、書き言葉の実態解明を進めるとともに、その調査結果を、国語審議会に審議資料として提供してきた²。今後は、BCCWJをはじめとする大規模コーパスの活用によって、文化審議会国語分科会に対して審議資料を提供するなど、引き続き国語施策に貢献していくことを考えている。

本稿では、大規模コーパスを国語施策にどのように活用していくことが考えられるか、そのためにはどのようなコーパスを構築することが必要かについて、具体的な検討材料を提供することを目的として、国立国語研究所が実施した語彙調査・文字調査のうち、現時点における最新の調査である「現代雑誌200万字言語調査」（以下、雑誌200万字調査とする。）を取り上げ、現代における漢字使用の実態について、国語施策とのかかわりの中で見ていく。本稿で扱う雑誌200万字調査のデータはコーパスとして作成されたものではないが、比較的規模も大きく、電子化された資料という点やβ単位という言語単位で解析され³、見出し・品詞といった情報が与えられている点など、コーパスと共通する部分もある。そこで、雑誌200万字調査のデータを活用することにより、コーパスを国語施策にどのように活用していくかについての検討材料を提供することにした。

具体的には、雑誌200万字調査を基に、文化審議会国語分科会における「情報化時代に対応する漢字政策の在り方について」のための審議資料として作成した漢字頻度表（国立国語研究所 2005b）、音訓一覧表（国立国語研究所 2005c）を取り上げ、常用漢字・表外漢字の出現状況、表外音訓・表外漢字の音訓の出現状況を見ていく。また、その結果を踏まえて、現行の常用漢字表を見直すとすれば、どのような見直しが考えられるか、さらに現代における漢字使用の実態を的確に把握し、迅速に国語施策の立案に資する資料を提供していくために求められるコーパスの設計等についても言及する。

2. 現代雑誌200万字言語調査

雑誌200万字調査は、現代の書き言葉のうち、社会性と多様性とを併せ持つと考えられる媒体として雑誌を取り上げ、計量的な調査・分析によって、その実態を明らかにすることを目的として実施した調査である。この調査で、雑誌を取り上げた理由について、簡単に説明しておく⁴。

新聞は、広く社会に頒布され、ほとんどの人が目にするという意味で、社会性を持つ媒体の筆頭に挙げられるものである。しかし、言語の面から見た場合、やや多様性に欠ける面がある。例えば表記については、各社で基準を設けて統一を行っている。漢字使用に関しては、各社共通の基準として、日本新聞協会が常用漢字表を基に作成した新聞漢字表があり、使用する字種に余り差が見られない。用語の面でもいわゆる俗語を用いず、文体もほぼ「だ体」「である体」を用いるといったことが新聞という媒体に共通する傾向として指摘できる。

雑誌は、新聞と同様に社会性を持つ媒体であるが、言語の面から見た場合、出版社や雑誌・記事・執筆者ごとに独自の用字の方針を持っているとしても、新聞漢字表のような各社に共通する基準は見られない。また用語・文体についても、俗語を用いるものがあるなど、新聞よりも用語・文体に幅があることが期待される。したがって、社会性と多様性とを併せ持つと考えられる

媒体を取り上げるという調査目的からは、雑誌が適切と考えられるのである。

調査対象は、1994年刊行の月刊誌の中から、「全国の書店で販売される」「本文の内容が専門的ではなく、読者も専門的職業集団でない」等の条件を満たす雑誌（411誌）を選び、さらにそこから分野のバランスを考慮して選んだ70誌である。

雑誌200万字調査を企画した背景には、1956年刊行の雑誌90誌を対象とする調査、1966年刊行の新聞3紙を対象とする調査を実施して以降、1981年の当用漢字表の廃止・常用漢字表の実施をはじめとする国語施策の改定や情報機器の発達・普及等、書き言葉を取り巻く環境に様々な変化が起こったということがある。このような環境の変化が、書き言葉にどのような影響を与えているか、その実態を把握するために、雑誌200万字調査は企画された。

雑誌200万字調査の結果のうち、漢字調査については国立国語研究所(2002)に、語彙調査については国立国語研究所(2005a, 2006b)に、語の表記のゆれの調査については国立国語研究所(2006a)にまとめられている。

3. 漢字の頻度・音訓に関する調査資料

本稿で取り上げる国立国語研究所(2005b, 2005c)は、文化審議会国語分科会における「情報化時代に対応する漢字政策の在り方について」のための審議資料として、雑誌200万字調査の成果である国立国語研究所(2002, 2005a)を基に新たに作成したものである。

漢字頻度表である国立国語研究所(2005b)は、国立国語研究所(2002)所収の「JISコード順漢字表」を基に作成した。この「JISコード順漢字表」は、雑誌200万字調査に出現した漢字3,584字をJISコード順に配列した上で、各漢字に対して、出現雑誌数（各漢字が出現した雑誌の数）、各雑誌における出現頻度、総合出現頻度（各雑誌における出現頻度の合計）の情報を与えた漢字表である。国立国語研究所(2005b)では、「JISコード順漢字表」で各漢字に付けられた情報のほかに、常用漢字・人名用漢字等の漢字の種別に関する情報を新たに加え、見出しの漢字を総合出現頻度順に配列し直した。

音訓一覧表である国立国語研究所(2005c)は、国立国語研究所(2005a)所収の「五十音順語彙表（自立語）」に掲げられた頻度7以上の自立語の表記に用いられた漢字2,293字を対象として、各漢字がどのような音訓で、どのような語又は語の構成要素の表記に、どのくらい用いられているのかを語例とともに示したものである。

この音訓一覧表をまとめるに当たり、雑誌200万字調査で出現した漢字3,584字すべてを対象とせず、頻度7以上の自立語の表記に用いられた漢字を対象を限定したのは、次のような理由による。漢字の音訓調査をするためには、その漢字がどのような語に使われているのかが明らかになっている必要がある。つまり、語が認定されて初めて音訓の調査が可能となる。国立国語研究所(2005c)を作成する時点で確定していた語彙調査のデータは、国立国語研究所(2005a)所収の語彙表に掲げられた頻度7以上の語のデータであった。そのため、雑誌200万字調査に出現したすべての漢字ではなく、頻度7以上の自立語の表記に使われた漢字に限定して、音訓一覧表をまとめた。そういう意味では、国立国語研究所(2005c)は、中間報告として位置付けられる。なお、こ

のような事情から、国立国語研究所(2005b)に示した漢字の総合出現頻度と国立国語研究所(2005c)に示した漢字の総合出現頻度とは、多くの場合、一致しない。国立国語研究所(2005c)を利用する際には、この点に注意する必要がある。

なお、ここで更にもう1点、国立国語研究所(2005c)を利用する際に注意すべき点について述べておきたい。

そもそも漢字表記語は、ルビが付けられている場合は別として、読みを一つに定め難い場合がある。例えば、「私の意見を述べる」という文の「私」、「重複がないようにする」という文の「重複」には、複数の読みが考えられる。つまり、「私」には「わたくし」「わたし」の2種類の読みの可能性が、「重複」にも「ちょうふく」「じゅうふく」の2種類の読みの可能性が考えられる。

語彙調査では、このような事例に対応するため、あらかじめ作業規則を立てておき、その規則に基づいて読みを定める。雑誌200万字調査では、漢字表記語「私」は、「私」が常用漢字であり、常用漢字表では訓として「わたくし」のみが認められているということから、一律に「わたくし」を採用している。「重複」については、一般に規範的とされる「ちょうふく」を採用している。

しかし、このような雑誌200万字調査の作業規則は唯一絶対のものではなく、別の作業規則を立てることも当然考えられ、別の作業規則に基づいて語彙調査を再度行った場合には、雑誌200万字調査とは異なる結果になる可能性もある。

国立国語研究所(2005c)は、雑誌200万字調査の語彙調査における作業規則によって定められた語の読みに基づいて、各漢字の音訓を集計し、一覧表にまとめたものである。したがって、別の作業規則によって語彙調査をやり直した場合、音訓一覧表に示した音訓及びその頻度も変わり得る。国立国語研究所(2005c)を利用する際には、そこに示された漢字の音訓や頻度は、一つの作業規則の下での結果であるということに注意する必要がある。

以下、本稿では、4節で国立国語研究所(2005b)を基に常用漢字・表外漢字の出現状況、5節で国立国語研究所(2005c)を基に常用漢字の音訓のうち表外音訓の出現状況、表外漢字の音訓の出現状況について見ていく。

なお、本稿では、音訓を示す際には、音は片仮名で、訓は平仮名で表記する。また、漢字とその音又は訓を併せて示す場合、例えば「国」という漢字の「コク」という音を示す場合には「コク(国)」,同じく「くに」という訓を示す場合には「くに(国)」のように示す。語の表記を示す際には「国語」のようにカギ括弧を付けて示し、語を示す際には片仮名を使って、「《コクゴ》」のように表記する。

4. 常用漢字・表外漢字の出現状況

本節では、雑誌200万字調査における常用漢字・表外漢字の出現状況について見ていく。

雑誌200万字調査で出現した漢字の異なり字数は3,584字、延べ字数は566,950字である。なお、異なり3,584字のうち、常用漢字は1,928字であった。

表1 上位2,000字における常用漢字・表外漢字の出現状況

順 位	延べ	累 積		常 用 漢 字				表 外 漢 字	
				異なり		延べ			
		頻度	割合	字数	割合	字数	割合		
1－100	203,435	203,435	35.9%	100	100.0%	203,435	100.0%		
101－200	91,181	294,616	52.0%	100	100.0%	91,181	100.0%		
201－300	59,448	354,064	62.5%	98	98.0%	58,194	97.9%	2	藤・岡
301－400	42,357	396,421	69.9%	99	99.0%	41,916	99.0%	1	阪
401－500	32,892	429,313	75.7%	100	100.0%	32,892	100.0%		
501－600	26,251	455,564	80.4%	97	97.0%	25,469	97.0%	3	伊・頃・奈
601－700	21,351	476,915	84.1%	99	99.0%	21,138	99.0%	1	誰
701－800	17,394	494,309	87.2%	98	98.0%	17,069	98.1%	2	之・呂
801－900	13,969	508,278	89.7%	97	97.0%	13,557	97.1%	3	鹿・揃・彦
901－1000	11,285	519,563	91.6%	97	97.0%	10,943	97.0%	3	弘・埼・幌
1001－1100	9,086	528,649	93.2%	94	94.0%	8,542	94.0%	6	頁・澤・須・熊・茨・智
1101－1200	7,353	536,002	94.5%	94	94.0%	6,895	93.8%	6	鍋・也・阿・狙・俺・鶴
1201－1300	5,859	541,861	95.6%	86	86.0%	5,029	85.8%	14	貼・牡・駒・那・尻・桂・宏・浩・弥・韓・袖・眉・脇・栗
1301－1400	4,783	546,644	96.4%	90	90.0%	4,313	90.2%	10	炒・嶋・柴・阜・桐・嬉・龍・菱・爽・牝
1401－1500	3,879	550,523	97.1%	82	82.0%	3,187	82.2%	18	旭・洲・枳・鎌・叩・昌・梨・淳・洒・柏・椅・溢・璧・贅・笠・篇・亀・錦
1501－1600	3,138	553,661	97.7%	77	77.0%	2,401	76.5%	23	宛・凄・衿・吾・湧・嵐・闇・葛・俱・串・謎・懂・祐・慧・忒・膳・辻・鮎・裾・沙・庄・鷹・匂
1601－1700	2,471	556,132	98.1%	72	72.0%	1,769	71.6%	28	磯・斬・蒲・馴・嘉・芦・篠・函・幡・綾・晃・亮・柿・拭・筑・餅・哉・挨・拶・菅・敦・桧・畿・鍵・蘇・綴・貌・篋
1701－1800	2,015	558,147	98.4%	61	61.0%	1,213	60.2%	39	苑・萩・綺・姑・鞍・噉・鴨・雀・諏・瘦・捉・迄・蝶・乃・剥・釜・竿・惚・諦・艶・旦・猪・叱・覗・膝・云・曾・濡・煩・梁・筐・蕉・瞳・輔・蓮・股・賭・芭・溪
1801－1900	1,648	559,795	98.7%	57	57.0%	928	56.3%	43	伎・晋・朋・呆・萩・繫・堺・蹴・噌・陀・賑・畠・茅・椎・槻・媛・靖・瞭・逢・燕・蓋・爪・吞・淀・穀・圭・辰・椿・鳩・斐・國・吊・廻・腎・舵・樋・味・枕・儲・虎・巳・稔・玲
1901－2000	1,359	561,154	99.0%	39	39.0%	523	38.5%	61	翔・嘘・歪・抒・雀・樽・馳・洛・憐・籠・唄・叶・胡・丞・富・廣・嬰・囁・腫・刺・麵・棍・稽・隙・桁・訣・訊・箸・淵・俣・溜・荏・輶・郁・樺・牙・駿・鯛・睦・蘭・李・稜・嶺・漱・潰・眩・粟・蔭・蟹・隈・緇・湘・棲・掴・這・斑・捧・妖・鷺・濁・濱
合 計	561,154	—	—	1,737	86.9%	550,594	98.1%	263	

上位2,000字の漢字の出現状況について、常用漢字と表外漢字とに分けて概観したのが、表1である。表1では、上位から1～100, 101～200のように100字ごとの区間に分けた上で、各区間の漢字100字の頻度の合計を「延べ」欄に示し、「累積」欄には、その区間までの累積の頻度と割合とを示した。また、各区間に含まれる常用漢字の異なり字数と延べ字数及びその割合を「常用漢字」欄の「異なり」欄・「延べ」欄に示した。各区間に含まれる表外漢字については、「表外漢字」欄に異なり字数と漢字とを示した。

表1を見ると、上位2,000字の累積頻度は561,154で、これは雑誌200万字調査で出現した漢字の延べ字数566,950の99.0%に当たる。上位の2,000字は、異なり字数では、雑誌200万字調査に出現した漢字3,584字の55.8%と半数を超える程度であるが、延べ字数では99.0%と全数近くを占めている。一方、異なり字数の約4割強を占める残りの1,584字は、延べ字数で見るとわずか1%程度を占めるに過ぎないことになる。このことから、上位2,000字が、雑誌200万字調査に出現した漢字の中でも非常によく使われる漢字の集合であると言ってよい。

4.1. 常用漢字の出現状況

上位2,000字の中に常用漢字は1,737字含まれている。この常用漢字の出現頻度の合計は550,594である。上位2,000字に占める常用漢字の割合は異なり字数では86.9%, 延べ字数では98.1%と非常に高い割合となっている。つまり、非常によく使われる漢字の集合である上位2,000字のほとんどの部分を常用漢字が占めているのである。このことから常用漢字は、よく使われる漢字の集合の中でも、特によく使われる、現代の日本語において重要度の高い漢字の集合と位置付けられる。

しかし、高頻度の常用漢字がある一方で、上位2,000字の中に含まれない、頻度の低い常用漢字もある。その数は、異なり字数で191字であり、このうち頻度10未満の常用漢字は150字ある。また、雑誌200万字調査で出現しなかった常用漢字も17字ある。その17字を、以下に示す。

謁 効 虞 儉 壘 勺 詔 鍾 斥 繕 嫡 脹
勅 朕 迭 丙 匆

雑誌200万字調査はサンプリング調査であるため、頻度の低い漢字ほど、標本に現れたか否かが偶然に左右される可能性が高くなる。それゆえ、低頻度の漢字の扱いには注意が必要である。

つまり、頻度が0であっても、今回の調査におけるサンプリングで、たまたまそのような結果になっただけであり、上記の17字についても、もう一度サンプリングをやり直せば、出現する可能性がある。

しかし、上記の17字の中には、他の漢字頻度調査でも頻度が低い漢字がある。ここで、文化庁(2000)所収の新凸版印刷調査、読売新聞調査における上記17字の出現状況を参照してみよう。文化庁(2000)は、表外漢字字体表の審議資料として作成されたものである。新凸版印刷調査は、凸版印刷株式会社が、1997年中に作成した組版データを対象として行った漢字頻度調査であり、読売新聞調査は、読売新聞社が1999年7月1日から8月31日の東京本社・中部本社管内における最終版の朝刊・夕刊紙面を対象として行った漢字頻度調査である。新凸版印刷調査は延べ

33,301,934字、読売新聞調査は延べ25,310,226字という、極めて大規模な調査である。

これら二つの調査における上記の17字の出現状況を見ると、「謁」「効」「虞」「勺」「朕」「迭」の6字が新凸版印刷調査で出現順位3,000位以下となっており、「謁」「効」「虞」「儉」「勺」「鍾」「脹」の7字が読売新聞調査で出現順位3,000位以下となっている。また、「朕」「殳」の2字は読売新聞調査では出現していない。

雑誌200万字調査と文化庁(2000)とを併せて見ると、雑誌200万字調査で出現しなかった17字の常用漢字のうち、「謁」「効」「虞」「儉」「勺」「鍾」「脹」「朕」「迭」「殳」は、再度サンプリングを行っても、出現しないか、出現したとしても極めて頻度が低いと予想される。常用漢字の中には、これら10字のように、複数の調査で低頻度のグループに属する漢字が含まれているという点には注意をしておきたい。

4.2. 表外漢字の出現状況

常用漢字にも高頻度の漢字と低頻度の漢字とがあるが、表外漢字の中にも上位2,000字の中に含まれるような、高頻度の漢字がある。

具体的に見ていくと、表外漢字の中で最も頻度が高いのは「藤」で、頻度は654である。この「藤」は現代では主に人名・地名に用いられる漢字であり、雑誌200万字調査でも654回出現したうち641回は、「安藤」「伊藤」「藤原」などの姓や、「藤ヶ丘」「藤沢」などの地名に用いられている。人名・地名以外の一般語に使われたのは、「葛藤」(3例)、植物名の「藤」(5例)など13例しかない。そのほか「岡」「阪」「伊」「奈」も、「藤」と同様に主として人名・地名に用いられる漢字である。このような主として人名・地名に使われる漢字が常用漢字表に入っていないのは、そもそも常用漢字表が固有名詞を対象としていないためである。

しかし、上位2,000字の中にある表外漢字は、人名・地名に用いられるものばかりではない。「頃」「誰」「揃」など、主として人名・地名以外のいわゆる一般語の表記に使われる漢字も含まれている。このように、一般語の表記に使われる表外漢字が高頻度の漢字として出現している点に注意したい。これらの漢字については、6節で詳しく見ていく。

4.3. 本節のまとめ

常用漢字表は、法令・公用文書や新聞・雑誌・放送など、一般の社会生活において現代の日本語を書き表す場合の漢字使用の目安であり、効率的で共通性の高い漢字を収めた漢字表である。表1を見る限り、常用漢字は効率的で共通性の高い漢字と考えられ、常用漢字表が、1981年の実施から13年経過した1994年においても漢字使用の目安としての機能を果たしていたとすることができる。しかし、その一方で、低頻度の常用漢字や高頻度の表外漢字があり、漢字の使用実態と常用漢字表との間に若干のずれが生じているということも事実である。

5. 漢字の音訓の出現状況

本節では、国立国語研究所(2005c)を基に、常用漢字の音訓のうち表外音訓の出現状況、表外

漢字の音訓の出現状況について、頻度のほか、どのような語の表記に使われているかという観点から見ていく⁵。なお、ここでは対象を、頻度7以上の自立語の中でも特に一般語に用いられた音訓に限ることとし、人名・地名に用いられた音訓は対象外とした⁶。

まず、国立国語研究所(2005c)で取り上げた音訓の異なりと延べの数を表2に示した。表2では、常用漢字の表内音訓、表外音訓、表外漢字の音訓に分けて異なり数・延べ数を示した上で、国立国語研究所(2005c)で取り上げた音訓の異なり数・延べ数を合計欄に示している。

表2 国立国語研究所(2005c)に示した漢字の音訓の数

		音の数		訓の数	
		異なり	延べ	異なり	延べ
常用漢字	表内	1,591	394,192	1,323	447,105
	表外	23	186	217	1,488
表外漢字		180	1,694	326	3,208
合 計		1,794	396,072	1,866	451,801

異なりで見ると、常用漢字の表内音訓が、音については88.7%、訓については70.9%を占めている。音では9割近くを占めるものの、訓では7割程度にとどまっている点に注意される。ただし、延べで見ると、表内音は99.5%とほぼ全数を占め、異なりで7割程度にとどまっていた表内訓も99.0%と全数近くを占めている。常用漢字表の表内音訓は、非常によく使われる音訓と位置付けることができる。

5.1. 表外音訓の出現状況

国立国語研究所(2005c)で取り上げた常用漢字の表外音訓を表3・表4として示した。表3には、国立国語研究所(2005c)の音訓一覧表に示したすべての表外音を示し、表4には頻度10以上の表外訓を示した。表3・表4の音訓の欄、語例欄において、▲印を付けた語例は常用漢字表の付表に掲げられた当て字・熟字訓などであり、▲印を付けた音訓は付表の当て字・熟字訓などで用いられた音訓である。例えば、「手」の表外音「ズ」は、常用漢字表の付表に掲げられた「上手(じょうず)」の表記に用いられた音であるので、▲印を付けた。また、その語例欄にある「上手」は常用漢字表の付表に掲げられたものであるなので、同様に▲印を付けた。

表3に示した表外音のうち頻度10以上の音を見ていくと、最も頻度の高い「ズ(手)」(頻度59)のほか、「ケ(景)」(頻度23)、「モ(木)」(頻度13)と、常用漢字表の付表に掲げられた当て字・熟字訓などに用いられた音が3種類入っている。これらは、常用漢字表の本表に示されたものではないということで表外音として扱ったが、常用漢字表に示されているものであり、表内音に準ずる位置付けのものと言うことができる。

表3 常用漢字表の本表に示されていない音の出現状況

字種	音	頻度	語 例	字種	音	頻度	語 例
手	ズ ▲	59	上手▲(59)	頭	ジュウ	9	万頭(1) 饅頭(8)
楽	ラ	26	倶楽部(26)	釈	シャ	8	釈迦(8)
景	ケ ▲	23	景色▲(23)	栄	エ	4	見栄(4)
堪	タン	17	堪能(17)	怪	ケ	3	怪我(3)
旬	シュン	17	旬(17)	個	カ	3	個所(3)
勢	セ	14	伊勢丹(14)	微	ミ	3	微塵(3)
木	モ ▲	13	木綿▲(13)	浪	ロ	3	浪漫(2) 浪漫派(1)
駄	タ	11	下駄(11)	甲	カ	2	生き甲斐(2)
参	シン	10	人参(10)	逸	イチ	1	逸早く(1)
清	シン	10	日清(10)	画	エ	1	画(1)
三	ザン	9	三昧(9)	工	グ	1	工合(1)
寿	ス	9	寿司(9)				

なお、常用漢字表の付表にある当て字・熟字訓などは、「いわゆる当て字や熟字訓のうち、慣用の広く久しいものは取り上げる。」(文化庁 1982:145)という選定基準により取り上げられたものである。したがって、付表の当て字や熟字訓などに使われた音が上位に位置するというのは、首肯できる結果である。

以上のほか、「ラ(楽)」「ス(寿)」「エ(栄)」「ケ(怪)」「ロ(浪)」「カ(甲)」「イチ(逸)」も、語例欄を見ると分かるように、いずれも当て字で用いられたものである。また、《タンノウ》の表記に用いられている「タン(堪)」も、元々は「足りぬ」が転じてできた《タンノウ》に当てられたものである。

このように、国立国語研究所(2005c)に示した表外音は、全部で23種類と少なく、頻度も低いものが多い。しかも、当て字に用いられたものが多く、語例を見てもほとんどの場合1種類にとどまっている。このことから、頻度7以上の自立語の表記に使われた常用漢字が表外音で使われることは少なく、機能という面から見ても余り重要な働きはしていないとすることができる。

次に表4から、上位10位までの表外訓の半分が常用漢字表の付表に掲げられたものであること、また、表4全体を見ても約4分の1に当たる11の訓が付表に示されているものであることが分かる。

付表に示された訓以外について見てみると、最も頻度の高いのは「ほか(他)」で、頻度は359である。そのほか、「すべて(全)」「こたえる(応)」「みる(観)」「おもう(想)」「いき(粋)」が頻度30以上で続いている。なお、▲印を付けた「たち(達)」「め(眼)」は、語例欄を見ると、付表に示された当て字・熟字訓など以外の語にも使われていることが分かる。しかも、「たち(達)」は接尾辞《タチ》での使用例が76と、付表に掲げられている《トモダチ》での使用例59よりも多く、同様に「め(眼)」も、付表に掲げられている《メガネ》での使用例22に対し、《メ》での使用例が50となっており、付表に掲げられた語例での使用例よりも多い。

表4 常用漢字表の本表に示されていない訓の出現状況（頻度10以上）

字種	訓	頻度	語 例
他	ほか	359	他(359)
達	たち ▲	135	達(76) 友達▲(59)
母	かあ ▲	74	母▲(74) [付表では「お母さん」]
眼	め ▲	73	眼(50) 眼鏡▲(22) 眼じり(1)
全	すべて	55	全て(55)
応	こたえる	54	応える(44) 手応え(8) 歯応え(2)
観	みる	39	観る(39)
父	とう ▲	36	父▲(36) [付表では「お父さん」]
想	おもう	34	想い(16) 想い出す(4) 思い出(9) 想う(5)
笑	え ▲	33	笑顔▲(33)
粋	いき	33	粋(21) 小粋(12)
奴	やつ	28	奴(28)
為	ため	24	為(24)
陽	ひ	21	陽(12) 陽差し(4) 陽射し(2) 夕陽(3)
鶏	とり	20	鶏(9) 鶏肉(11)
御	お	19	御(14) 御陰(2) 御前(3)
辛	つらい	18	辛い(18)
人	と ▲	18	素人▲(18)
創	つくる	17	創り(1) 創り上げる(1) 創りだす(1) 創り出す(2) 創る(12)
関	かかわる	16	関わり(3) 関わる(13)
浮	うわ ▲	16	浮気▲(16)
路	みち	16	小路(3) 路(6) 横路(7)
最	も ▲	15	最早(1) 最寄(2) 最寄り▲(12)
経	たつ	14	経つ(14)
旨	うまい	14	旨い(9) 旨さ(2) 旨み(3)
館	やかた	13	館(13)
腹	なか	13	お腹(13)
活	いかす	12	活かす(12)
判	わかる	12	判る(12)
描	かく	12	描き込む(1) 描く(11)
委	ゆだねる	11	委ねる(11)
獲	とる	11	獲る(10) 獲れたて(1)
兄	にい ▲	11	兄▲(11) [付表では「兄さん」]
歳	とし	11	歳(11)
姉	ねえ ▲	11	姉▲(11) [付表では「姉さん」]
生	ふ ▲	11	芝生▲(11)
素	もと	11	素(11)
放	ほうる	11	放る(11)
家	うち	10	家(10)

以上のことから、国立国語研究所(2005c)で取り上げた表外訓は、表外音に比べて種類も多く、頻度も高いということが言える。また、常用漢字表の付表に掲げられた当て字・熟字訓などに用いられたものも多いが、それ以外の一般語の表記に用いられたものもあり、それらの中には「ほか(他)」「すべて(全)」など、頻度の高いものも見られる。

5.2. 表外漢字の音訓の出現状況

表外漢字の音訓の出現状況について見ていく。ここで調査の対象とした表外漢字は、異なり字数で563、延べ字数で7,713である。

国立国語研究所(2005c)に示した表外漢字の音を表5に、訓を表6に示した。

表5 表外漢字表の音の出現状況(頻度10以上)

字種	音	頻度	語 例	字種	音	頻度	語 例
呂	ロ	133	風呂(133)	輜	リョウ	14	車輛(14)
阪	ハン	39	阪急(24) 阪神(15)	牡	ボ	13	牡馬(13)
椅	イ	37	椅子(37)	蝶	チョウ	13	蝶(5) 蝶々(8)
贅	ゼイ	35	贅沢(34) 贅肉(1)	憐	レン	13	可憐(13)
璧	ヘキ	33	完璧(33)	溪	ケイ	13	溪谷(8) 溪流(5)
彗	スイ	32	彗星(32)	稽	ケイ	12	稽古(12)
洒	シャ	32	洒落(24) 洒落れる(8)	訣	ケツ	12	秘訣(12)
篇	ヘン	29	短篇(6) 長篇(1) 篇(22)	繡	シュウ	12	刺繡(12)
倶	ク	26	倶楽部(26)	輯	シュウ	12	輯(11) 特輯(1)
芯	シン	25	芯(25)	牝	ヒン	12	牝馬(12)
挨	アイ	24	挨拶(24)	駒	ク	11	産駒(11)
拶	サツ	24	挨拶(24)	須	ス	11	恵比須(2) 必須(9)
綺	キ	18	綺麗(18)	膳	ゼン	11	膳(11)
伎	キ	17	歌舞伎(17)	噌	ソ	11	味噌(8) 味噌汁(3)
斬	ザン	16	斬新(16)	瞭	リョウ	11	明瞭(11)
桂	ケイ	15	桂(15)	嬰	エイ	10	嬰(10)
昧	マイ	15	曖昧(6) 三昧(9)	嘩	カ	10	喧嘩(10)
抒	ジョ	15	抒情(15)	喧	ケン	10	喧嘩(10)
伊	イ	14	伊勢丹(14)	楚	ソ	10	清楚(10)
爽	ソウ	14	爽快(14)	貼	テン	10	貼付(10)
				勿	モチ	10	勿論(8) 勿体(2)

表5を見ると、最も頻度の高い表外漢字の音は「ロ(呂)」で、頻度133と、唯一頻度が100を超えている。以下、頻度30台の音が6種類、20台の音が5種類あるものの、頻度10台のものが29と、表に示した音の約7割を占めている。このことから、表外漢字の音は、全体として頻度は余り高くないと言うことができる。また、語例欄を見ると、最も頻度の高い「呂」でも「風呂」の

表6 表外漢字の訓の出現状況（頻度20以上）

字種	訓	頻度	語 例
頃	ころ	235	いま頃(2) 今頃(2) 頃(171) 近頃(10) 手頃(21) 日頃(21) 身頃(8)
誰	だれ	210	誰(209) 誰れ(1)
俺	おれ	74	俺(74)
狙	ねらう	74	狙い(17) 狙う(57)
揃	そろう	74	勢揃い(6) 揃い(9) 揃う(59)
揃	そろえる	57	品揃(1) 品揃え(11) 揃える(33) 取り揃える(8) 取揃える(4)
鍋	なべ	49	鍋(49)
嬉	うれしい	45	嬉しい(45)
貼	はる	41	貼(1) 貼り(4) 貼る(36)
炒	いためる	41	炒めもの(4) 炒め物(3) 炒める(34)
牡	おす	36	牡(36)
菱	ひし	36	三菱(36)
尻	しり	35	尻(28) 目尻(7)
鹿	か	33	鹿(10) 鹿毛(13) 馬鹿(9) 馬鹿さ(1)
宛	あてる	32	宛(21) 宛先(11)
溢	あふれる	32	溢れる(32)
懂	あこがれる	32	懂れ(23) 懂れる(9)
眉	まゆ	32	眉(32)
袖	そで	31	袖(15) 長袖(9) 半袖(7)
謎	なぞ	31	謎(31)
牝	めす	30	牝(30)
叩	たたく	29	叩く(29)
闇	やみ	25	無闇(1) 闇(24)
匂	におう	24	匂い(18) 匂う(6)
爽	さわやか	23	爽(1) 爽やか(22)
脇	わき	23	脇(23)
噂	うわさ	22	噂(22)
捉	とらえる	22	捉える(22)
湧	わく	22	湧く(22)
凄	すごい	21	凄(17) 凄さ(2) もの凄(1) 物凄(1)

1語のみであり、表5全体を見ても、語例欄に2種類以上の語例が挙げられているものは11種類で、表5に掲げた音の4分の1程度にとどまる。以上のことから、表外漢字の音は、全体として余り頻度が高くなく、造語力も総じて低いと言うことができる。

次に、表外漢字の訓のうち頻度20以上のものを示した表6を見ていく。表6に掲げた訓のうち最も頻度が高いのは「ころ（頃）」で、頻度235である。2位の「だれ（誰）」も210で、頻度が200を超えている。以下、「おれ（俺）」「ねらう（狙）」「そろう（揃）」が頻度74、「そろえる

(揃)」が同57,「なべ(鍋)」が同49で続いている。表6は、頻度20以上の訓に限定したものではあるが、それでも訓の種類は30ある。表5に示した表外漢字の音のうち頻度20以上のものが12種類であったのと比べると、表外漢字の訓には頻度の高いものが多いと言うことができよう。また、語例欄を見ても、複数の語例が挙がっている訓が多い。このようなことから、表外漢字は音で用いられるよりも、訓で用いられることが多い、つまり和語の表記に用いられることが多いと考えられる。

5.3. 本節のまとめ

国立国語研究所(2005c)に示した音訓の延べ数に占める表内音訓の割合から考えると、常用漢字表は、漢字の音訓の使用という面から見ても、おおむね漢字使用の目安として機能していると考えられる。しかし、その一方で、特に表外訓について頻度の高いものが見られることや、表外漢字の訓についても頻度の高いものが見られることから、現代における漢字使用の実態とは若干のずれが生じているということも指摘することができる。

6. 常用漢字表の見直しに向けて

以上、雑誌200万字調査の対象となった1994年時点では、常用漢字表が頻度の面から見ても、音訓使用の面から見ても、漢字使用の目安として機能していたということ、しかしその一方で高頻度の表外漢字があること、高頻度の表外訓があることなどから、現代における漢字使用の実態とは若干のずれが生じていることを指摘した。

文化審議会国語分科会では、2005年3月の文部科学大臣の諮問「情報化時代に対応する漢字政策の在り方について」を受け、常用漢字表の見直しに関する審議を継続している⁷。その背景には、常用漢字表が現代における漢字使用の実態と合わない面があるという共通認識がある。そして、そのような認識は、4節・5節で報告した漢字使用の実態からも裏付けられるものである。

本節では、4節・5節を踏まえて、常用漢字表を見直す場合に、常用漢字に追加する音訓の候補はあるか、あるとすればどのような音訓か、常用漢字表に追加する漢字の候補はあるか、あるとすればどのような漢字で、その漢字にはどのような音訓を示せばよいのかについて考察する。

検討を行うに当たっては、改定後の常用漢字表(以下、新常用漢字表(仮称)とする。)の性格がどのようなものになるのか、字種の選定がどういう方針の下に行われるのかについて確認しておく必要がある。そこで、文化審議会国語分科会が文化審議会総会(第42回, 2007年2月2日)に報告した「国語分科会漢字小委員会における今期の審議について」を基に、現時点における同分科会の考え方を見ておく⁸。

常用漢字表の基本的な性格にかかわることとしては、準常用漢字(仮称)の設定が挙げられている。準常用漢字(仮称)とは、「読めるだけでいい漢字」(読めて、意味の分かる漢字)のことである。ただし、準常用漢字(仮称)の設定は、「新常用漢字表の字数を検討していく過程で、その総字数との関係で、改めて考えていくべき課題」という位置付けで、「総字数がかなり多くなれば、準常用漢字の設定を検討する」とされている。つまり、現時点で、準常用漢字(仮称)

の設定が決まっているわけではなく、場合によっては検討課題となり得るということである。そういう意味では、現行の常用漢字表の基本的な性格を大きく変更するような方向性は、今のところ示されていないと言える。

字種の選定に関しては、「基本的に一般社会においてよく使われている漢字（出現頻度数の高い漢字）を選定していく。この場合、最初に3,000字～3,500字程度の漢字集合を特定し、そこから絞り込むという作業過程を考えていくこと」とし、この過程では「教育等の様々な要素はいったん外して、とにかく日常生活でよく使われている漢字を漢字出現頻度数調査によって機械的に選ぶ」のを基本とすることが述べられている。もちろん「単に個別漢字の頻度分布だけでなく、様々な要素を総合的に勘案」していくとも述べているが、まずは実態調査から得られる使用頻度のデータを字種選定の基本にしていこうという方針がうかがわれる。

以上の新常用漢字表（仮称）の基本的性格、字種選定の方針を踏まえ、本節では、次のような立場から、常用漢字表の見直しについて検討を行う。

- ① 新常用漢字表（仮称）でも、現行の常用漢字表の基本的な性格が大きく変更されないという前提で検討を行う。
- ② 主として頻度の面から検討し、必要に応じて漢字の機能等の面からも検討を加える。

なお、本節で述べる常用漢字表の見直しに関する考え方は、あくまで雑誌200万字調査の結果、つまり、1994年に刊行された月刊誌70誌を対象とした調査の結果に基づくものである。そのため、別の調査結果を基に検討した場合、また別の調査結果と雑誌200万字調査の結果とを併せて検討した場合には、本節で述べる見直しに関する考え方とは異なる考え方が導き出されることも十分にあり得る。

6.1. 常用漢字への音訓の追加

現行の常用漢字に新たに音訓を加えるとすれば、どのような音訓が候補になり得るか検討する。

まず、音の追加候補について検討する。ここで第一に注意しなければならないのは、国立国語研究所(2005c)に示した漢字の音の延べ数に占める常用漢字の表内音の割合が99.5%と、極めて高いことである。また、表外音の使用実態に目を向けた場合、表3から、種類が少なく、大部分が当て字などに使われたものであること、また頻度もそれほど高くなく、更に造語力という点でも大体1種類の語の構成要素にしかっていないことが分かる。

このような調査結果から、現行の常用漢字への追加候補になる音は、基本的にはないと考えられる。強いて挙げるとすれば、「タン（堪）」「シュン（旬）」の二つであろう。

一方、訓について見てみると、国立国語研究所(2005c)に示した漢字の訓の延べ数に占める常用漢字の表内訓の割合は99.0%で極めて高いが、それでも表外訓の種類は多く、中には高頻度のものも見られる。

頻度から見れば、「ほか（他）」「みる（観）」「おもう（想）」「つくる（創）」が追加候補となろう。しかし、ここで注意しなければならないのは、「みる（観）」の同訓異字「見」、「おもう

「想）」の同訓異字「思」,「つくる（創）」の同訓異字「作」「造」が既に常用漢字表に入っていることである。仮に「観」に「みる」という訓を追加したとすると、既に常用漢字表に入っている「見」との間で、書き分けをする必要が生じる。しかし、これらの漢字について、だれもが理解でき、それに基づいて漢字の書き分けができるような簡明な基準を作ることは、必ずしも容易なことではない。

また、「ほか（他）」については、この訓を新たに追加した場合、既に常用漢字表に入っている音「タ（他）」との読み分けの問題が生じる。例えば、「その他」や「他の問題」といった場合の「他」を「ほか」と読むか「タ」と読むかということが、確定しにくくなるのである。

以上のことから、「ほか（他）」「みる（観）」「おもう（想）」「つくる（創）」といった訓を追加候補とすることについては慎重な配慮が必要である。

書き分けの問題が余り生じないと考えられる訓としては、「たち（達）」「すべて（全）」「こたえる（応）」「かかわる（関）」がある。ただし、これらの訓が追加候補となるか否かを検討する際には、《タチ》《スベテ》《コタエル》《カカワル》の表記の実態について確認しておく必要がある。具体的には、例えば接尾辞《タチ》の表記について、「達」という漢字で書かれる場合と、平仮名で書かれる場合のどちらが多いかということである。国立国語研究所(2006b)を基に、《タチ》《スベテ》《カカワル・カカワリ》《コタエル・コタエ》について、漢字表記と仮名表記との用例数を調べた結果は、次のとおりである。

《タチ》	達(76)	たち(806)
《スベテ》	全(55)	すべて(306)
《カカワル・カカワリ》	関(16)	かかわる・かかわり(62)
《コタエル・コタエ》	応(54)	こたえる・こたえ(16)

この結果を見ると、《コタエル》を除き、仮名表記の方が漢字表記よりも頻度が高い。したがって、雑誌200万字調査の結果からは、「たち（達）」「すべて（全）」「かかわる（関）」が追加候補になるとは言い難い。ただし、雑誌200万字調査は、1994年の月刊誌を対象とした調査であるため、既に13年が経過した現在では、漢字表記の割合が高くなっていることも予想される。「たち（達）」「すべて（全）」「かかわる（関）」を追加候補とすべきか否かについては、改めて現在の漢字の使用状況を調査した上で判断する必要がある。

なお、「応（こたえる）」については、1994年時点で漢字表記が仮名表記の約3倍となっており、当時既に漢字表記が定着しつつあったと思われる。また、新聞では2001年から「期待にこたえる」という場合の《コタエル》の表記に、振り仮名なしで「応」を使うようになっており、現在では更に「応」を使った表記が定着している可能性がある。このような状況に加え、同訓異字の「答」との書き分けもさほど困難とは考えられないことから、「こたえる（応）」は追加候補になり得ると考えられる。

6.2. 表外漢字の追加

次に、常用漢字表に新たに漢字を追加するとすれば、どのような漢字が候補となり得るかにつ

いて検討していく。

単純に頻度だけで考えるとすれば、表1に挙げた上位2,000字の中に現れた表外漢字は追加候補となるだろう。しかし、追加に当たっては、固有名詞にしか使われないのか、一般の語の表記にも使われるのか、幅広い分野で使われるのかといったことも問題となる。また、音訓の観点から言えば、音についてはどのくらいの漢語の表記に使われるのか（別の言い方をすれば、どのくらいの漢語を作り出す力（造語力）があるのか）ということが問題となるし、訓については既存の常用漢字との書き分けの可否ということも問題となる。

使用の範囲という点から言うと、表1の漢字の多くは、人名・地名でよく使われることから、高い頻度になっているとすることができる。例えば、先にも述べたが、「藤」の654例のうち、98.0%に当たる641例は、人名・地名に使われたものである。頻度が高いという点で言えば、確かに「常用」であり、しかも多くの人が読める漢字でもあると思われるが、こういう人名・地名に主として使われる漢字を常用漢字表に入れる必要があるかどうかは検討する必要がある⁹。

次に、音訓の使用実態から、表外漢字の追加を考えてみたい。

まず音の面、つまり漢語を書き表すという面から追加候補となり得る表外漢字があるか考えてみる。表5を見ると、表外漢字の音は、全体として余り頻度が高くなく、また語例もほとんどの音が1種類であり、表外漢字の造語力は余り高くないと考えられる。このことから、漢語とのかわりで追加候補となる表外漢字はないとすることができる。

ただ、強いて挙げるとすれば、次の7字について追加候補とすることも考えられよう。

呂（ロ）	椅（イ）	贅（ゼイ）	璧（ヘキ）	挨（アイ）
拶（サツ）	須（ス）			

「呂」は《フロ》,「椅」は《イス》,「贅」は《ゼイタク》,「璧」は《カンベキ》,「挨」「拶」は《アイサツ》,「須」は《ヒッス》の表記に用いられるが、それぞれの語について、国立国語研究所(2006b)を基に漢字表記と仮名表記・交ぜ書きとの頻度を比較してみると、次に示すように、漢字表記の方が仮名表記・交ぜ書きを上回っている。

《フロ》	風呂(133)	ふろ(8)
《イス》	椅子(37)	いす(2) イス(2)
《ゼイタク》	贅沢(34)	ぜいたく(6)
《カンベキ》	完璧(33)	完ぺき(3)
《アイサツ》	挨拶(24)	あいさつ(6)
《ヒッス》	必須(9)	仮名書き・交ぜ書きの例は無い。

この結果を見ると、上記の7字は、特定の語の表記にしか使われないという問題はあるものの、常用漢字表への追加候補として検討に値すると言えよう。ただ、あくまで1994年の月刊誌70誌を対象とした調査の結果であることを考え、改めて現在の漢字使用の状況を調査した上で判断する必要があるだろう。

次に訓の使用実態から、追加候補となり得る表外漢字があるか考えてみる。これは、訓の面、つまり和語を書き表すということから追加候補となり得る表外漢字について検討するということ

である。まず、「誰」「狙」「尻」「袖」「謎」「闇」「脇」の7字は、2001年から新聞において振り仮名なしで使っていることから、追加候補となるであろう。また、これら7字以外にも「頃」「揃」「憧」は、国立国語研究所(2006b)を基に《コロ》《ソロウ・ソロエル》《アコガレル》の表記を見た場合、

《コロ》	頃(235)	ころ(179)
《ソロウ・ソロエル》	揃(131)	そろう・そろえる(58)
《アコガレル》	憧(32)	あこがれる(12)

のように、漢字表記の方が仮名表記よりも頻度が高く、漢字表記が定着していると考えられること、「頃」「揃」「憧」を追加しても書き分けの上で問題となることはないことから、追加候補になると考えられる。

6.3. 本節のまとめと補説

以上、本節では、4節・5節を踏まえ、常用漢字表をどのように見直していくかについて議論した。表外音で追加候補になり得るものはほとんどなく、また表外漢字の追加に関しても音とのかかわりで追加候補になり得るものは余りなかった。一方、表外訓については、同訓異字の書き分けや1994年の時点では仮名表記の割合の方が高いといった点で問題はあるものの、追加候補になり得る漢字があった。また、表外漢字に関しては訓の使用実態から「だれ(誰)」「ころ(頃)」「ねらう(狙)」など10字を追加候補になり得るとして指摘した。

これらの表外音訓・漢字は、文化審議会国語分科会が示している「とにかく日常生活でよく使われている漢字を漢字出現頻度数調査によって機械的に選ぶ」という字種選定の方針を踏まえ、主として頻度の面から検討したものである。しかし、現行の常用漢字表の字種選定において、漢字の頻度のほか造語力や使用分野の広さといった観点も含めて総合的に判断が行われたことから分かるように¹⁰、本節で指摘した表外音訓・漢字を新常用漢字表(仮称)に入れるべきか否かを決定するためには、頻度を中心とした検討だけではなく、造語力等も含めた総合的な観点からの検討が必要である¹¹。また、既に述べたように、雑誌200万字調査が1994年の月刊誌70誌を対象とした調査であることを考え、改めて現在の漢字使用の状況を把握する必要もある。これらについては、今後の課題としたい。

ところで、本節の議論から、現行の常用漢字表は、漢字の音の使用、主として漢語の表記という面から見ると、現代の日本語でよく用いられる漢語を表記するための音は、既にほとんど取り込んであり、漢字使用の目安としての役割を果たしていると考えられる。しかし、その反対に漢字の訓の使用、主として和語の表記という面から見た場合は、表外訓・表外漢字の訓で高頻度のものがあり、音に比べて見直しの余地があると考えられる。

この背景には、当用漢字表・常用漢字表の字種選定の基準がかかわっていると考えられる。当用漢字表では、漢字選定の目安として「訓だけのものもしくは、おもに訓だけを使うものは省く」(文部省 1953: 2)ということが挙げられている。漢字制限の立場から、和語は基本的に平仮名表記でよいという考え方があったものと思われる。その後、常用漢字表では「異字同訓はな

るべく避けるが、漢字の使い分けのできるもの及び漢字で書く習慣の強いものは取り上げる。」(文化庁 1982:145)という字種選定の方針が取られたが、それでも、現行の常用漢字表を見ると、常用漢字の38.0%に当たる739字には、音しか示されていない¹²。交ぜ書きを避けるということから、漢語の表記とのかかわりで追加される漢字は多かったのであろうが、和語の表記に関しては、漢語の表記ほど問題として強く意識されていなかった可能性がある。

しかし、雑誌200万字調査の結果を見ると、頻度の高い表外訓や表外漢字の訓が見られるなど、一部の和語において、常用漢字表の範囲に収まらない形での漢字表記が定着しつつある傾向も見取れる。したがって、現代の日本語における和語の表記の在り方、つまり、できる限り仮名で書くのか、漢字で書くことも認めるのか、という根本的な問題についても改めて検討する必要がある。

7. 国語施策へのコーパスの活用

以上、本稿では、出現頻度・音訓使用の二つの面から、現代雑誌70誌における漢字使用の実態を把握するとともに、その結果を基に常用漢字表の見直しの可能性について議論を行った。最後に、本稿のまとめとして、国語施策、とりわけ漢字政策の立案に資する資料を作成・提供していくためには、どのようなコーパスが必要かという観点から、要点を幾つか指摘しておく。

漢字の出現頻度調査は、調査対象となったテキストから漢字を取り出し、単字として扱っていく調査である。一方、音訓調査は、調査対象となったテキストの文脈情報を用いて語の認定をした上で行うものであり、漢字を語や文脈とかかわらせて扱っていく調査と言える。つまり、同じ漢字を対象とした調査でも、前者は文字という観点からの調査であり、後者は語彙という観点からの調査とすることができる。

本稿では、文字と語彙という二つの観点から、現代雑誌70誌における漢字使用の実態を明らかにしたことになるが、このような調査ができたのは、雑誌200万字調査のデータが β 単位という言葉単位で解析され、見出し・品詞といった情報が与えられていたことによる。漢字の使用実態を明らかにするためには、雑誌200万字調査のように、当初から語彙調査までを含めて企画された大規模な書き言葉の実態調査が、必要不可欠であることが改めて確認されたと見えよう。

ただ、雑誌200万字調査のデータを、漢字政策の立案に資する資料等の作成に活用することを考えた場合、幾つか問題もある。例えば、次のことが挙げられる。

(1) 調査対象が雑誌のみであること。

(2) 雑誌の中でも一般読者(専門的職業集団でない読者)を対象としたものに限定しているため、専門分野における漢字の使用実態が必ずしも十分に把握できていないこと。

常用漢字表の前書きには、その適用範囲として法令、公用文書、新聞、雑誌、放送が挙げられている。したがって、まず、雑誌以外にも対象を広げていく必要がある。前書きにも示されている新聞のほか、国民一般が目にするものとして書籍も対象に加えていく必要があろう。新聞・雑誌・書籍を中心に、なるべく広い範囲の出版物を取り上げていくことが求められるのである。

また、漢字表作成のためには、一般の社会生活でよく用いられる漢字を明らかにしていくこと

が重要であるが、専門分野にも目を向ける必要がある。例えば、常用漢字である「膳」は、ふだんの生活では余り目にすることのない漢字である。実際、文化庁の新凸版印刷調査でも「膳」の出現順位は3,095位と、常用漢字の中ではかなり低い順位となっている。しかし、このことから直ちに「膳」を重要性の低い漢字と結論付けることはできない。なぜなら、「膳」は《コセキトウホン》の表記に用いられることから、戸籍制度において重要な漢字と位置付けられる漢字であり、しかもその戸籍制度が国民の日常生活とも深いかわりを持つからである。したがって、雑誌を対象とするとしても、専門家を対象とした雑誌も加えていく必要があるし、更に言えば、法律・白書など、専門分野の語彙・漢字が用いられていると考えられる媒体にも対象範囲を広げていく必要がある。

ところで、漢字政策の中でも最も重要なことは、常用漢字表のような、一般の社会生活における漢字使用の目安としての漢字表を作成し、実施することである。常用漢字表を見れば分かる通り、漢字表では、字種を示すとともに、表に掲げた各漢字に対して、どのような音訓で用いるかも示している。したがって、漢字表を作成するためには、まず、よく使われる漢字の集合を定めるために、漢字の出現頻度調査を行うことが必要であり、さらに高頻度の漢字集合に含まれた各漢字がどのような音訓で用いられているのかを明らかにするために、漢字の音訓に関する調査が必要である。漢字政策の立案に資する資料を提供していくためには、文字と語彙という両面からの漢字調査を実施できる体制が必要だということになる。

このように考えた場合、2006年度から国立国語研究所で構築を開始した大規模コーパス KOTONOA の計画の重要性が改めて確認される。KOTONOA の計画の中で、まず構築される BCCWJ には、言語研究に必要な付加情報として形態論情報も付与されることが決まっている。また、BCCWJ では、雑誌以外にも新聞・書籍を対象としていること、法律・白書・専門家を対象とした雑誌等についても数百万語規模のデータを作成することから、先に挙げた雑誌200万字調査の問題点も解消される。したがって、BCCWJ は、現代日本語における漢字の使用実態を、文字と語彙との両面から明らかにしていくことのできる、重要なコーパスとなる。

BCCWJ を活用して、より大きな規模で、本稿と同様の頻度調査・音訓調査を進めていけば、より正確に漢字の使用実態を把握でき、質・量ともに十分に審議に耐え得る資料を、文化審議会国語分科会に対して提供していくことができる。

また、BCCWJ は、一般への公開が予定されている。このことは、結果として、漢字政策立案の基礎となったデータが、広く国民一般に共有されることを保証する。これにより、新たに作られた漢字表の妥当性を検証したり、またより良い漢字表を具体的に提案したりすることも可能になるのである。KOTONOA の整備計画は、科学的な調査に基づく客観的なデータを使って、国民が漢字政策、更には国語施策全般について考えるための、重要な基盤を提供する事業としても位置付けられるのである。

注

- 1 KOTONOA 計画の概要及びBCCWJの設計については、前川喜久雄(2006)、山崎誠(2007)を

- 参照。なお、BCCWJ は、Balanced Corpus of Contemporary Written Japanese の略称である。
- 2 例えば、常用漢字表作成の際には、1956年刊行の雑誌90誌を対象とする調査、1966年刊行の新聞3紙を対象とする調査に基づくデータを審議資料として国語審議会に提供している。
 - 3 β 単位については、国立国語研究所（1962：6-14）を参照。
 - 4 雑誌200万字調査の対象・方法等の詳細については、国立国語研究所（2002：1-42）を参照。
 - 5 5節以下で取り上げる漢字の音訓及びその頻度が、どういう性格のものかについては、3節を参照。
 - 6 国立国語研究所(2005c)では、人名・地名に使われた音訓は固有名詞に使われた音訓として別扱いにしているが、組織名・商品名等に使われた音訓は一般語に使われた音訓として扱っている。本稿でもこの方針を踏襲したため、表3の語例には「伊勢丹」「日清」、表5の語例には「伊勢丹」「阪急」「阪神」が挙げられている。また、人名・地名以外の固有名詞の構成要素になっている一般語に使われた音訓も固有名詞に使われたものとして扱っていない。例えば、表3に表外音「ラ（楽）」とその語例「倶楽部」が示されているのは、このことによる。
 - 7 文部科学大臣の諮問、文化審議会国語分科会・同漢字小委員会における審議の状況については、文化庁ホームページ（http://www.bunka.go.jp/kokugo_nihongo/bunkasingi/index.html）に掲載されている議事要旨等を参照。
 - 8 「国語分科会漢字小委員会における今期の審議について」は、文化庁ホームページ（http://www.bunka.go.jp/laramasi/42_soukai.html）を参照。
 - 9 「国語分科会漢字小委員会における今期の審議について」には、「固有名詞専用字ということで、これまで外されてきた「阪」や「岡」についても、出現頻度数が高ければ、今回は最初から排除はしない」とあり、新常用漢字の候補として検討対象に含めるという方針が示されている。
 - 10 常用漢字表の字種選定の基準については、文化庁(1982：145)を参照。
 - 11 「国語分科会漢字小委員会における今期の審議について」でも、字種の選定について、常用漢字表の字種選定基準を参考にすることが述べられている。
 - 12 甲斐睦朗・篠崎佳子(2006：249)を参照。

参考文献

- 甲斐睦朗・篠崎佳子(2006)「資料「常用漢字表の音訓」について」『日本語学』25-11、246-282
- 国立国語研究所(1962)国立国語研究所報告21『現代雑誌九十種の用語用字(1)』秀英出版
- 国立国語研究所(2002)国立国語研究所報告119『現代雑誌の漢字調査』
- 国立国語研究所(2005a)国立国語研究所報告121『現代雑誌の語彙調査 — 1994年発行70誌 —』
- 国立国語研究所(2005b)『現代雑誌の漢字調査（頻度表）』
- 国立国語研究所(2005c)『『現代雑誌の語彙調査』に基づく漢字音訓一覧表』
- 国立国語研究所(2006a)国立国語研究所報告125『現代雑誌の表記 — 1994年発行70誌 —』
- 国立国語研究所(2006b)「現代雑誌200万字言語調査語彙表 公開版 (ver.1.0)」(<http://www.kokken.go.jp/>で公開。)
- 文化庁(1982)『国語審議会報告書 14』ぎょうせい
- 文化庁(2000)漢字字体関係参考資料集『漢字出現頻度数調査(2)』
- 前川喜久雄(2006)「特定領域研究「日本語コーパス」— 目標、進捗状況、そして夢 —」『特定領域「日本語コーパス」平成18年度公開ワークショップ（研究成果報告会）予稿集』1-12

文部省(1953)『国語問題問答』光風出版

山崎誠(2007)「現代日本語書き言葉均衡コーパス」の基本設計について」『特定領域「日本語コーパス」平成18年度公開ワークショップ（研究成果報告会）予稿集』127-136

付 記

本稿は、国立国語研究所の平成17年度研究課題「日本語の現在：国語力の調査」、平成18年度研究課題「公共的コミュニケーションの円滑化のための調査研究と基礎資料の提供」の「文化審議会の資料作成」の成果である。

(投稿受理日：2007年1月31日)

(最終原稿受理日：2007年7月19日)

小椋 秀樹（おぐら ひでき）

国立国語研究所研究開発部門

190-8561 東京都立川市緑町10-2

ogura@kokken.go.jp

相澤 正夫（あいざわ まさお）

国立国語研究所研究開発部門

190-8561 東京都立川市緑町10-2

aiz@kokken.go.jp

Kanji use in seventy contemporary magazines and Jōyō Kanji:

A preliminary study for the application of text corpora to Japanese language policy

OGURA Hideki

The National Institute for Japanese Language

AIZAWA Masao

The National Institute for Japanese Language

Keywords

text corpus, Linguistic Survey of Contemporary Magazines, Japanese language policy, Jōyō Kanji

Abstract

The purpose of this paper is to report on the results of a preliminary study that will serve as a basis for the application of text corpora to Japanese language policy, with the particular goal of revising Jōyō Kanji, the list of kanji recommended for use in contemporary written Japanese together with their standard *on* (Sino-Japanese) and *kun* (native Japanese) readings. Based on current patterns of kanji use in seventy contemporary magazines, we assess the suitability of Jōyō Kanji as the standard of kanji use in Japanese society, and discuss necessary revisions. The details of the study are as follows:

(1) Using data from the Linguistic Survey of Contemporary Magazines conducted in 1994, we analyzed the usage frequency of Jōyō and non-Jōyō kanji, both as total frequencies and as partial frequencies categorized according to their standard and non-standard readings. Judging from the frequencies of Jōyō and non-Jōyō kanji in the list of the two thousand most-frequent kanji, Jōyō Kanji can be said to have fulfilled their function as the standard of kanji use in 1994, although some deviations such as high frequency non-Jōyō kanji and non-standard *kun* readings of Jōyō kanji were also observed.

(2) On the basis of the analysis mentioned above, we discuss how revision of Jōyō Kanji might be carried out: what kanji and kanji readings should be considered for addition to, or deletion from, the list.