

日本語としての漢字と情報

奥 雅 雄

序 章

Windows Vista によって、使える漢字が増えた。この一万字以上の漢字がどのようにして決まってきたのか。約 2 千年前に出会った日本語としての漢字の歴史を概観する。明治以降の漢字の改革と情報化による技術の歴史によって、何が可能になり、今後はいったい何が重要になっていくのか。「情報とは何か」ということも検討する。

第 1 章 漢字の受入

(1) 漢字の歴史の始まり

① 漢字への最初の接触

日本での漢字の始まりは、弥生時代の一世紀頃であることが、考古学等から明らかになってきた。文章ではなく、シンボルに近い形で文字の神秘性に触れた時期である。情報伝達という実用性はほとんどない。

文字のある出土物（1 世紀頃）

「山」の刻印ペンダント	日本最古の漢字。鹿児島県種子島広田遺跡出土。
金印（漢委奴国王）	1784 年福岡県志賀島出土。
「貨泉」の刻印中国の銅銭	長崎県シゲノダン遺跡出土。

② 刀剣の銘

4～5 世紀の歴史は空白期としてわからないことが多い。鉄剣から統合の歴史が伺える。

名 称	場 所	文字数	発見年
稻荷山古墳出土鉄剣	埼玉県行田市	115 字	1975 年
江田船山古墳出土鉄刀	熊本県玉名郡和水町	75 字	1873 年

(2) 外交の始まり

大陸への派遣が 600 年に始まる。遣隋使は三回以上派遣されたが、618 年に唐が建国して名称が遣唐使となった。

隋、唐から都市計画、建築、政治、宗教など、多くのことを学んだ。10 年以上に

および専門家としての知識・技能を習得して帰る場合もあった。

外交文書として漢字が必要であったので、事務官としての官僚が育ち、文字を理解する人々が登場してきた。続く奈良時代も遣唐使が見てきた長安などに倣って都市が整備された。大プロジェクトでは計画や統制のための実行命令や記録が重要である。

(3) 日本の成立と記録

「日本」という国名が使われたのは天武天皇の時代といわれている。日本書紀の編纂が行われ、古事記とともに大量の情報がいまに伝えられている。

この時代からの文字情報としては次のものが伝わる。

資料名	年代	その他
古事記	712年	日本最初の歴史書
日本書紀	720年	日本最初の正史、以後、公的な正史が継続する。
風土記	713年に作成の命がでる。	出雲風土記が完本で、播磨、肥前、常陸、豊後が現存する。
万葉集	8世紀後半	20巻、4,516首を伝える。
懷風藻	751年	日本漢詩集、最古のもの、116首

このほか、数多くの木簡が出土し、新たな歴史資料となっている。

第2章 国語と明治以後の歴史

(1) 国語国字問題

明治の近代化とともに、文字表記の簡素化と国民国家としての基盤が検討された。1902年に国語調査委員会が設立され、漢字の非近代性、民主化、経済効率の点から漢字廃止論へ展開していく。ローマ字表記やかな表記など文筆家を巻き込んで大論争になった。

委員会・審議会の変遷

1900年	国語調査委員の委嘱
1902年	文部省国語調査委員会発足（1913年廃止）
1921年	臨時国語調査会発足（1934年廃止、国語審議会へ）
1934年	国語審議会設立
1949年	文部省設置法制定、この法律に基づいて国語審議会が開かれる
2001年	中央省庁の再編に伴って国語審議会廃止 文部科学省文化庁に文化審議会を設置 国語分科会に引き継ぐ

(2) 国語審議会

戦後、当用漢字をはじめ、国語改革の中心的な役割を果たしてきた。

戦後直後の成果物

1946年	9月「現代かなづかい」答申
	11月「当用漢字表」答申
1947年	「当用漢字別表」(教育漢字)「当用漢字音訓表」
1949年	「当用漢字字体表」

(3) 漢字廃止論と歴史・文化尊重論

明治期には漢字が時代遅れの認識から廃止論が起きる。ローマ字とする説やかな書きなどの改革派がいる一方で、漢字を尊重する漢字擁護論者もいて、明治期は大いに揺れていた。1931年満州事変が起き、事実上漢字の制限が困難になっていった。

(4) 当用漢字から常用漢字へ

① 当用漢字

戦後、当用漢字が定められ、制限を強く求めるものであった。1946年の1,850字である。「法令・公用文書・新聞・雑誌および一般社会で、使用する漢字の範囲」としている。

当用漢字表(答申)は、文化庁のホームページで見ることができる。

<http://www.bunka.go.jp/kokugo/>

② 常用漢字

常用漢字1,945字は1981年に当用漢字の後継漢字として官報に公表された。「法令、公用文書、新聞、雑誌、放送など、一般の社会生活において、現代の国語を書き表す場合の漢字使用の目安」としている。

常用漢字表は戦前にも1923年の1,962字、1931年の1,858字が公表されている。

③ 人名漢字

人名についての漢字の見直しによって983文字が追加された。姓は、制限が困難であるが、名については、常用漢字と人名漢字に制限されている。

市町村で住民や戸籍管理のために使われている文字は法務省で集約され、総数は5万字を超えている。たとえば「曾」の上は「ソ」で旧字では「曾」で「ハ」の形になっている。「藤」などでも、ソ藤、ハ藤といわれる違いがある。

戸籍統一文字情報（法務省）

<http://kosekimoji.moj.go.jp/kosekimojldb/mjko/PeopleTop/EXECUTE>

人名漢字変更歴

公表日	変更内容	文字数
1951年	92文字追加, 81年当用漢字へ7字, 仙, 尚, 杉, 甚, 磨, 齊, 龍(竜)	92
1976年	28文字追加 81年当用漢字へ1字, 悠	120
1981年	54文字追加 8字削除	166
1990年	118文字追加	284
1997年	1文字(「疏」)追加	285
2004年2月	1文字(「曾」)追加	286
6月	1文字(「獅」)追加	287
7月	3文字(「毘」「瀧」「駕」)追加	290
9月	488文字追加, 許容字体からの205文字格上 計693追加 許容字体表は廃止	983
人名に使える漢字	常用漢字1,945文字を含めて	2,928

④ 表外漢字表

常用漢字表, 人名漢字表以外の漢字について, 情報機器の発達とともに, 新しい漢字が作り出される結果となり, 混乱が生じていた。漢字出現頻度数調査が, 印刷会社と新聞社を対象に2回行われ, 2000年12月, 表外漢字1,022字の標準印刷書体と簡易慣用体22字が決められた。

次の3部首は略字体を許容

飲・飴の食偏(しょくへん)の違い,

社・祇のネ(しめすへん)の違い

迷・辻の辶(しんにょう)の違い

簡易慣用対の例(鵠は当用漢字の區⇒区として作り出される。拡張新字体という)

印刷標準	鷗	醬油	祈禱	枅	麵	蠟
簡易慣用	鷗	醬油	祈祷	桱	麵	蠟

その他は第3章に印刷書体を例示する。伝統的な面を重視している。表外漢字の168文字が変更といわれる。

文化庁(<http://www.bunka.go.jp/kokugo>)の国語施策情報システムに第22期答申の「表外漢字字体表」がある。

⑤ 新常用漢字

現在、文化庁の文化審議会国語分科会漢字小委員会において、新常用漢字表（仮称）について、2010年頃の予定で検討が進められている。

府県名等の「阪鹿奈岡熊梨阜埼茨栃媛」は常用漢字に入っていないことと、読めるだけでよい「準常用漢字表」を新設するかどうかを検討している。

文化庁文化審議会国語分科会

http://www.bunka.go.jp/kokugo_nihongo/bunkasingi/index.html

第3章 産業政策と漢字

(1) JIS コードの登場

漢字の標準化は経済産業省（旧通産省）の日本工業標準調査会で審議され、JIS規格が制定される。最初の文字は1978年のJIS C 6226で決まった。

当時はコンピュータが威力を発揮し始め、欧米（英語等）では当たり前のことが、日本では文字数が多く、カタカナで表記するのがやっとであった。マイクロプロセッサがワンチップ CPU として1971年に登場している。（Intel 4004、これが現在の Dual コアへ発展する）日本で最初のワードプロセッサは、1978年に登場した。

① 日本語を扱う文字についての JIS の変遷

規格名	年 月	内 容	備 考
JIS C 6226	1978	第一水準2,965文字 第二水準3,388文字 非漢字524文字	漢字6,353文字
JIS C 6226	1983	第二水準4文字追加 字体入替	
JIS X 0208	1990	JIS C 6226を改正 人名漢字2文字追加	Unicode 原規格
JIS X 0212	1990	補助漢字文字セット 漢字5,801文字 非漢字266文字	Unicode 原規格 PC で利用できない 計6,067文字
JIS X 0213	2000 2月	新拡張 JIS コード 通称「JIS2000」「2000JIS」 第三水準1,249文字 第四水準2,436文字 非漢字659文字	JIS X 0208を含む 第三・四水準3,685文字 漢字10,040文字
JIS X 0213	2004 2月	10文字を追加 168文字のデザインの変更 通称「JIS2004」「2004JIS」 「印刷標準字体」1,022文字 「簡易慣用字体」（22文字）	漢字10,050文字 補助漢字にない文字 907文字

② JIS X 0213

正式名を「7ビット及び8ビットの2バイト情報交換用符号化拡張漢字集合」という。JIS X 0208を拡張した新しい規格で、第一水準、第二水準を含み、11,223文字で構成される。4年後に10文字増えて11,233文字となる。

③ 文字数

2004年2月改訂後の使える文字数は次の表の通り。

水 準	漢 字	非漢字	合 計	参 考 事 項
第1水準	2,965	524	6,879	常用漢字1,945字を含む
第2水準	3,390			
第3水準	1,259	659	4,354	一部は Vista で対応
第4水準	2,436			
合 計	10,050	1,183	11,233	

JIS X 0212と JIS X 0213の差異

追 加 部 分	文字数	2004年	その他	合 計
第3・4水準漢字	3,685	10		3,695
補助漢字との重複	2,741	2		2,743
第3・4水準の増加	944	8	2	954
IBM 拡張文字	47			47
追加文字数	897	8	2	907

④ 印刷標準字体

2004年には、国語審議会答申の「表外漢字字体表」をうけて JIS X 0213が改正され、文字のデザインの一部分について、168文字の変更と10文字の追加が行われた。『康熙字典』『大漢和辞典』といった既存の漢字字典のデザインを元に印刷用の字体として制定した。

経済産業省報道発表

<http://www.jisc.go.jp/newstopics/2005/040220kanjicode.pdf>

⑤ 区点コード

「区点」は94×94の表として文字の配置をしめす。1区から94区それぞれに1点から94点がある。94×94=8,836の文字の配列を表せる。区点番号を指定して文字の入力ができ、漢字辞典にも載っている。

区点を拡張して面を導入し面区点の3次元の配列とする。本のページとしてイメー

ジすると区点は1ページの表となる。「面」は各ページを示す。補助漢字は第2面に収容されている。第4水準は第2面に補助漢字と重ならないように収容された。コード間の相互の課題も配慮し、EUCなどの符号化方式も考慮して決められている。

(2) Unicode

1991年、最初のユニコードができたときには第1～2水準のJIS X 0208と補助漢字のJIS X 0212があり1990年規格が用いられて、両方ともが取り込まれた。実際にWindowsでUnicodeが使えるようになったのはWindows2000からとなった。同じ2000年に発効したJIS X 0213は、あとから文字が追加されたため使えない文字があり、Vistaまで待たなければならなかった。

① コンソーシアム

ゼロックス社が提唱して、マイクロソフト、アップル、IBM、サン・マイクロシステムズなどがコンソーシアムを作る。情報処理のシステム・機器を国際化の際の手間や効率を改善するために言語・文字を統一的に扱うことを目指す。

JIS X 0213の新しい文字の対応はUnicode 3.1とUnicode 3.2で行われた。

Unicodeの改訂歴

バージョン	公開日	文字数	備 考
Unicode 1.0	1991年		16ビットの文字コード
Unicode 1.1	1993年		CJK 統合漢字20,902字 ISO/IEC 10646-1と統合
Unicode 2.0	1996年	38,885	21,204文字の漢字
Unicode 3.0	2000年	49,194	6,582文字の漢字を追加
Unicode 3.1	2001年		42,711文字 JIS2004対応302文字
Unicode 3.2	2002年		JIS X 0213正式対応（1文字）
Unicode 4.0.0	2003年	96,382	71,098文字の漢字
Unicode 5.0.0	2006年		

② 国際符号化文字集合（UCS: Universal Multiple-Octet Coded Character Set）

JIS X 0221: 2001（原規格はISO/IEC 10646）で群、面、区、点で構成される文字集合の構造が定義されている。群は7ビットなので、 $128 \times 256 \times 256 \times 256$ の文字（約21億字）が定義可能になる。

正規の標準でUnicodeと協調して作業が進められている。

③ CJK 統合漢字

CJK は、中国語、日本語、朝鮮語の頭文字で、1989年頃から漢字を統合する作業が開始されていた。Unicode 1.1の CJK 統合漢字の発表から Unicode の存在を知らしめ、話題となった。1994年からベトナムが加わって、CJKV と呼ぶようになっていく。

④ サロゲートペア

第0面を BMP（基本多言語面）といい、ISO/IEC 10646の第0群第0面には256×256=65,536文字の割当が可能だが、ほはいっぱいになっている。

第2面（SIP）にある文字は UTF-16を扱う際にはサロゲート・ペアの処理が必要になる。

Unicode は当初2バイト（16ビット）で世界中の文字を表現しようとした。2バイトでは2¹⁶=65,536通りの表現ができそれぞれを文字に対応させる。登録の要求はこれをはるかに超えてしまった。この制限を外す方法がサロゲートペアで一部の文字は4バイトを使う。1,048,576文字が表記できる。多くのプログラムで16ビットの文字として処理しているので、プログラムの改造が必要になる。

⑤ 結合文字

フランス語やドイツ語のダイアクリティカルマークを使って、文字を作成する。

代表的なダイアクリティカルマーク

アクセント (アキュート)	ÁÉÍÓÚÝáéíóúý	鋭アクセント	フランス語、スペイン語、ポルトガル語
アクセント (グレイヴ)	ÀÈÌÒÙàèìòù	開口音	フランス語、イタリア語
ウムラウト	ÄÖÜäöü	母音交替	ドイツ語、スウェーデン語
オゴネク	ĄĘąę	鼻音化、長音	ポーランド語、リトアニア語
セディーユ	ÇŞşç		フランス語、トルコ語

日本語のかな「が」～「ぼ」, 「ぱ」～「ぽ」の濁音・半濁音についてもダイアクリティカルマークと同様の文字符号がある。「゛」「゜」と組み合わせる符号である。組み合わせで一体の文字として扱う。処理が煩雑になる。

⑥ 符号化方式

Unicode の符号化方式には UTF-8、UTF-16、UTF-32がある。Windows では UTF-16

③ 印刷標準字体による変更

JIS2004で変更された168文字について Vista のフォントも変更されている。
(実際のフォント MS 明朝, MS ゴシックでは122文字が変更されている)

変更された168文字 (Windows XP による互換フォント)

逢 芦 飴 溢 茨 翳 淫 迂 厥 噂 餌 襖 迦 牙 迴 恢 晦 蟹 葛 鞆 釜 翰 翫 微 祇 汲
灸 笈 卿 饗 僅 喰 櫛 屑 糸 祁 隙 倦 捲 牽 鍵 諺 巷 梗 膏 鵠 飭 叉 桼 薩 鯖 鯖
鮫 餐 杓 灼 酋 桶 薯 諸 哨 鞘 杖 蝕 訊 逗 摺 撰 煎 煽 穿 箭 註 噲 邈 掬 遜 腿
蛸 辿 樽 歎 註 瀦 抄 槌 鎚 辻 挺 鄭 擢 溺 兎 堵 屠 賭 瀨 遁 謎 灘 檣 禰 牌 這
秤 駁 箸 叛 挽 誹 樋 稗 逼 謬 豹 廟 瀕 斧 蔽 瞽 蔑 篇 婉 鞭 庖 蓬 鱒 迄 儲 餅
粉 爺 嚙 愈 猷 漣 煉 簾 榔 屢 冤 叟 咬 嘲 囁 俳 扁 棘 橙 狡 鸺 甦 疼 崇 龜 筵
箒 腓 艘 亡 虔 螢 蠅 訝 靄 靱 騙 鴉

④ フォント

Vista には, MS 書体 Version 5.0 (MS ゴシック, MS P ゴシック, MS UI Gothic, MS 明朝, MS P 明朝) 5 書体と, Clear Type 対応日本語フォント「メイリオ」1 書体, 計 6 書体が付属している。「メイリオ」は「MS ゴシック」に代わるサンセリフ系ゴシック体文字として JIS 2004に対応している。

⑤ Windows の旧バージョンでの対応

Windows XP 用としてマイクロソフトの Web ページ「Windows Vista」の「JIS2004 とメイリオ」に「Windows XP および Windows Server 2003向け JIS 2004 対応 MS ゴシック & MS 明朝フォントパッケージについて」でサポートを行っている。

Office ソフトでは JIS 2004文字セットのすべての漢字を扱うためには Office 2003 以降のバージョンが必要である。

⑥ Vista および Office 2007での注意事項

Vista で扱えるようになった文字は,「環境依存文字」として変換時に表示される。これをそのまま Vista 以外の PC へ送ると正常に表示されない。かな書きなど, 表現を変更することで対応する。また Office 以外でも同様だが, 以前のバージョンへファイルを引き渡すときは, 相手バージョンを意識した書き出しが必要である。

Visita には MS Internet Explorer 7 が付属して, JIS 2004に対応している。

⑦ 入力方法

マイクロソフトは MS-IME 2007を充実させて変換の効率を上げている。Justsys-

tem は ATOK 2008 をリリースして、多くの改良とともに JIS 2004 への対応も行っている。通常の変換の他に漢字の選択の容易性が問われ、利用頻度の低い文字ほど手間がかかるので辞書機能が強力な変換ソフトが求められる。文字が増加するに従って探す効率は落ちていく。

(4) その他、漢字の世界

「超漢字」は国産 OS トロン上で開発され、漢字の制約をなくした仕組みである。18万文字の文字をもつ。最近は Windows 上で使えるようになっている。

「今昔文字鏡」は漢字15万字を含む17万字あまりの文字を扱える。表示や印字はフォントを切り替える簡易的な方式ではあるが入力時の辞書などが充実している。

マック（マッキントッシュ）は、OS X の10.5（Leopard）から JIS 2004 に対応した。もともと出版や DTP に強いマックなので印刷関係は魅力的なものを多くもっている。

EUC-JIS-2004 は、UNIX での日本語を扱う文字コードで、JIS X 0213 の符号化方式のひとつとなっている。UNIX 系の OS はサーバーで利用されることが多い。

機種間の文字の違いもあるが、OS やアプリケーションのバージョンの違いが大きく、文字の制約になるので最新のものを使う必要がある。

青空文庫は著作権の切れた作品をデジタル化してウェブ上に公開している。入力をボランティア活動で行い、約7,000作品を保有している。入力時に会おう見慣れない文字を集約して、「外字注記辞書」を作成している。JIS X 0213 策定作業中には意見書と資料を提出している。

第4章 漢字のこれから

漢字の複雑さは、不利と思われた。そのために明治期の漢字廃止論まで引き起こす。経済発展の足かせを憂慮した結果だった。今、この時点で漢字を扱う上で不利なことはほとんどない。

歴史や古典などを扱う場合は、別のシステムを検討する必要もある。とくに国産のシステムは文字の追加が馴染みやすい。いずれは Unicode にも追加されることもあるが、概して敷居が高い。JIS 2004 により、ここ10年来の矛盾が解決され、日常生活においては、ほぼ望ましいシステムが実現しているのでこれらをベースに上に乗るサービスなどが実現できる。

CJKV 統一漢字から、中国、日本、韓国・朝鮮、ベトナムの漢字文化圏を連想するが、文化を議論する以前の断絶がある。政治的、経済的、歴史認識的な課題が多すぎる。EU（欧州共同体）に比べて、まとまりのない状況になっている。

漢字文化圏は過去の歴史遺産でしかない。文字としての共通文化が実り多いもの

になるような関係ができると良いと思う。漢字の共通性よりも言語や思考の異質性が大きい。異質をよい方向に考えるならば、それぞれの得意な面を分担すれば、有利にすることができる。異文化交流でなくて、異文化共創であり、協力して進める協創が重視される時代なのだ。情報の共有が文化の始まりである。

メディアの多様化により、文字情報への依存割合は低下していくだろう。明治期には多くの訳語がつくられたが、近年はカタカナが多くなった。これらが欠陥というのではなくて、日本語あるいは日本人の適応力が高いと見た方がよい。もともとの大和言葉に対して、随・唐の時代の漢語や技術・文化の吸収は、明治・戦後の欧米からのものに匹敵する。日本人は遅れを悟ったときには猛烈に推進するエネルギーな行動力の持ち主なのだ。

効率のみではなく、精神性を支える部分では和語や漢字が重視される。写経や俳句なども廃れていない。心情は大和言葉が残り、漢字熟語や漢字略語はしっかりした信頼感を与えている。「とうふ」、「さけ」、「からおけ」、など、日本発の文化があるが、これからの時代は、漢字も一緒に伝えることが可能な時代になっている。例えば I like a 桜花. とか I go to 東京. など英文のなかに漢字や日本語を併用することも可能なので、文化の壁を乗り越えていける。

終 章

100年かけた国語改革もひとまず完成した。遅れて情報環境も整備され、基盤ができあがったと言える。これからも微調整が続けられていくことにはなるが、しっかりした土台ができあがった。土台の上に仕組みや制度が乗っかる。さらにその上に文化が築かれていく。

日本語が経済的な面で妨げになるといわれたこともあったが、なんら非論理的でコミュニケーションに妨げになるような要因は指摘できない。むしろ競争上、不利であったことを有利に転換するような知恵と組織力を持った民族であると考えられる。次の時代の人々が新しい基盤の上でどんなものを咲かせていくのか、大いに期待したい。国際社会の中で、役割をきちんと認識して、今後は国内に限定されず、世界を豊にしていく文化の創造に望みをかけたい。

<参考文献>

- 『漢字伝来』大島正二 岩波新書 2006. 8. 18
- 『国語審議会 迷走の60年』安田敏朗 講談社現代新書 2007. 11. 20
- 『漢字と日本語』朝倉漢字講座① 前田富祺・野村雅昭 朝倉書店 2005. 03. 25
- 『文字・表記』現代日本語講座⑥ 飛田良文・佐藤武義編 明治書院 2002. 05. 25
- 『異字体の世界 旧字・俗字・略字の漢字百科』小池和夫 河出文庫 2007. 7. 10

『漢字がつくった東アジア』 石川九楊 筑摩書房 2007. 4. 25

『遣唐使』 東野治之 岩波新書 2007. 11. 20

『漢字の文化史』 阿辻哲次 筑摩書房 2007. 6. 10

『旧字力, 旧仮名力』 青木逸平 生活人新書 2005. 6. 10

< web サイト >

国語審議会	http://www.mext.go.jp/b_menu/shingi/12/kokugo/index.htm
国文学研究資料館	http://www.nijl.ac.jp/
Unicode Home Page	http://www.unicode.org/
経済産業省	http://www.meti.go.jp/
日本規格協会	http://www.jsa.or.jp/
日本工業標準調査会	http://www.jisc.go.jp/
法務省	http://www.moj.go.jp/
情報処理推進機構	http://ossipedia.ipa.go.jp
青空文庫	http://www.aozora.gr.jp/
Vista 詳細資料	http://go.microsoft.com/fwlink/?LinkId=78800

