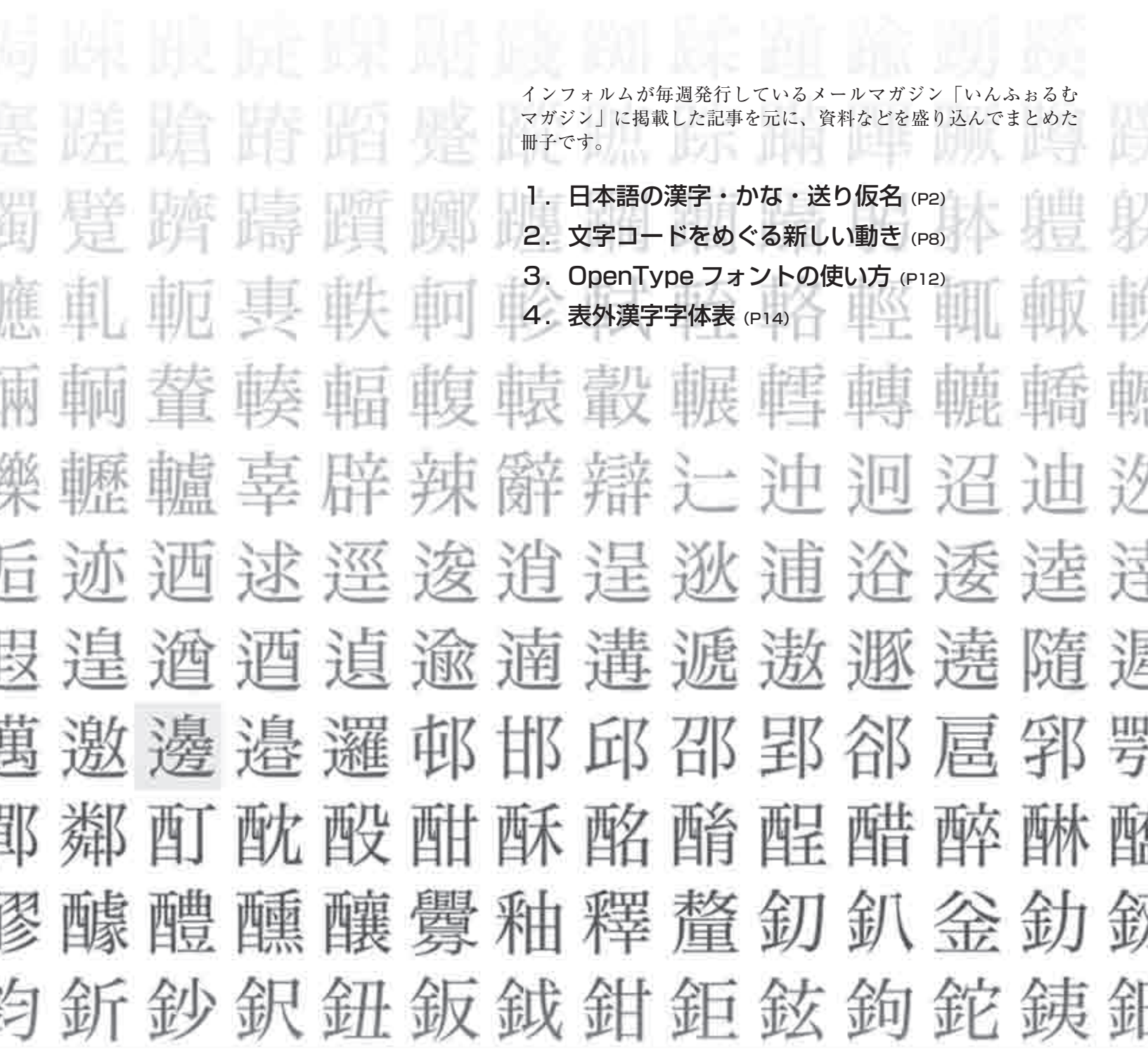


いんぷおるむマガジン 別冊 Vol.4

インフォームが毎週発行しているメールマガジン「いんぷおるむマガジン」に掲載した記事を元に、資料などを盛り込んでまとめた冊子です。

1. 日本語の漢字・かな・送り仮名 (P2)
2. 文字コードをめぐる新しい動き (P8)
3. OpenType フォントの使い方 (P12)
4. 表外漢字字体表 (P14)



表外漢字

边 辺 邊 邊

Informe

日本語の漢字・かな・送り仮名

(株) インフォーム 制作部
田村信幸

日本語の表記

書き言葉としての日本語を考える場合、漢字とひらがな、カタカナという異なる種類の文字が混在して出来上がっているという特徴を無視するわけにはいきません。たとえば欧米語を見てみると、アルファベット(アクセントなども含む)だけですべての言葉を表現しますし、アジア圏でも中国語は漢字、ヒンズー語はデーヴァナーガリー文字と、その文字数に違いはあるものの一種類の文字を使うのが一般的です。

異なる文字である漢字とかなを組み合わせさせて表記するということが、日本語表記を複雑なものにしている最大の要因と言っても過言ではないでしょう。日本語の表記を考える場合は、使われている漢字やかなそのものだけでなく、その組み合わせ方についても追求する必要があるということになります。

ここでは、現代日本語表記における漢字の制限、仮名遣いの規則、そして組み合わせ方である送り仮名のルールについて見ていきましょう。

漢字の使用制限

日本語にとって漢字はなくてはならない文字です。漢字は言うまでもなく古代に中国大陸から伝えられた文字ですが、漢文から漢字仮名交じり文まで時代によって利用形態は変わりながらも、常にわが国の文字文化の中心であり続けました。

ところが、戦後のわが国の国語政策においては、漢字の利用制限が大きな柱とされたのです。まず、昭和21年に「当用漢字表」が公布されます。これは「日常生活使用する漢字の範囲」を1,850字に絞り、公的な場面で半ば強制的に使用を制限しようとするものでした。その結果、熟語を漢字だけで表現できずに仮名と混ぜ書きしたり(「ら致」「ろ過」など)、同音の漢字で書き換える(叡智→英知など)といったことが行われ、日本語の破壊だといった批判を招くことになりました。

さらに、昭和23年内閣告示の「当用漢字音訓表」によって読み方で制限され、また、昭和24年には字体の標準を定めた当用漢字字体表が告示されるなど、一連の政策が日本語の表記に与えた影響およびトラブルは小さからぬものがあります。

そもそも、当用漢字は漢字使用廃止を視野において作られたものでした。今となっては非現実的な話ですが、わが国には幕末以来、欧米列強に伍するためには漢字ではなく、ローマ字のような表音文字を使うべきだと考える人たちがいて、

自分たちの主張を展開してきました。彼ら表音主義者は、戦前から各方面に働きかけをしていたのですが、戦後の混乱期、ついに自分たちの理想実現のための一歩を踏み出すことに成功したのです。

ただし、彼らの企てがそれ以上進むことはありませんでした。当用漢字には根強い批判が続き、最終的には国も方針を転換、昭和56年には「常用漢字表」が制定されます。文字数は1,945字と大して変わらないものの当用漢字のような強制的なものではなくあくまでも「目安」ということになりました。

さらに、コンピュータ社会の到来は、漢字のあり方を変えました。現在、わが国の正式な漢字表は常用漢字表ですが、JIS規格としてJIS X 0208を始めとするいくつかの規格が存在し、1万数千字もの漢字が事実上自由に使えるようになっています。さらに10万字以上の文字を収録した「今昔文字鏡」などの文字セットが登場するなど、漢字の使用制限はパソコンの普及とともに意義を失ったと言えるようです。

日本の漢字表の変遷

1946年	当用漢字表	1,850字	一般社会で使用する漢字の範囲を示したもの
1951年	人名用漢字別表	92字	常用漢字、片仮名、平仮名以外で子どもの名前に用いられる文字
1976年	人名用漢字別表	120字	
1981年	常用漢字表	1,945字	一般の社会生活における現代国語表記上の漢字使用の目安
1981年	人名用漢字別表	166字	
1990年	人名用漢字別表	284字	
1997年	人名用漢字別表	285字	
2004年	人名用漢字別表	983字	

日本語と仮名遣い

歴史的仮名遣い

一文字一文字に意味がある漢字は「表意文字」です。一方、中国から伝わった漢字を元に日本で作られたひらがなやカタカナは、文字そのものには意味がなく、音だけを表す「表音文字」です。つまり、日本語というのは、表意文字と表音文字の組み合わせで構成されている言語なのです。

表音文字である仮名を使えば、日本語の発音をそのまま表記することができます。ところが、発音というものは時代が

移ることによって変化していくものです。ある時代の発音に忠実に従った仮名の表記は、次の時代には発音と食い違いが生じてしまうのです。

常に発音にあわせて仮名を表記すると、仮名についての表記法もどんどん変えていかなければなりません。時代を超えて残る書き言葉の表記法をそれほど頻繁に変更するわけにはいきません。つまり、表音文字を使っている、仮名表記を正確に発音に合わせるのとは簡単なことではないのです。

ある特定の発音にしたがって仮名の表記法を定めることを「仮名遣い」と呼び

ます。絶えず変化を続ける発音と違い、仮名遣いは時代をまたがって使われてきました。

歴史的に使われてきた仮名遣いとしては、古事記や万葉集などの奈良時代の文献において万葉仮名で使われている「上古特殊仮名遣い」、藤原定家が古い草子本の用例を元にまとめた規則を室町時代の僧、行阿が修正したという「定家仮名遣い」、江戸時代の国学者、契沖が古い文献をあらためて研究し直して定家仮名遣いを修正した「契沖仮名遣い」を元に、本居宣長などが改訂を加えた「復古仮名遣い」などがあります。このうち、復古仮名遣いは「歴史的仮名遣い」とも呼ばれ、明治から戦前まで一般的な仮名遣いとして使われてきました。

歴史的仮名遣いは、平安時代初期の発音や表記法を元に、それ以前の発音や表記も考慮されており、現代の仮名遣いに比べると、語源や文法的に正確な仮名遣いとなっています。

時代によって変化しやすい発音でなく、変化の少ない文法や語源に基づく仮名遣いであるため、発音が変化しても仮名遣いを変更する必要がないというのが歴史的仮名遣いのメリットですが、逆に言うと、現代の発音とはかなり違って表記なので覚えるのが大変というデメリットにもなっています。

現代仮名遣い

第二次世界大戦後は、漢字の廃止を前提にした漢字制限(当用漢字表の告示)が実施されるなど、日本語のあり方について大きな変革が断行された時代であり、漢字だけでなく、漢字とともに日本語を形成する文字である仮名についても、新しい改革が行われました。

昭和21年には、「現代語音にもとづいて、現代語をかなで書き表す場合の準則」として「現代かなづかい」が内閣告示されています。なお、「現代かなづかい」は昭和61年に廃止され、代わりに強制的な色合いを薄めた「現代仮名遣い」が告示されています。この点も漢字政策

「現代仮名遣い」(昭和61年内閣告示・一部省略)

第一(原則に基づきまじり)	語を書き表すのに、次の仮名を用いる。
	直音 あいうえおかきくけこがぎぐげごさしすせそざじずぜぞたちつてとだぢづでどなにぬねのはひふへほばびぶべほびびぶべほまみむめもやゆよりるれるわを(「ぢ、づ、を」は第二に示す場合にだけ用いる)
	拗音 きゃ きゅ きょ ぎゃ ぎゅ ぎょ しゃ しゅ しょ じゃ じゅ じょ ちゃ ちゅ ちょ ぢゃ ぢゅ ぢょ ちゃにゅによひゃひゅひよびゃびゅびよびゃびゅびよみゃみゅみよりゃりゅりよ(拗音に用いる「ゃ、ゅ、よ」は、なるべく小書きにする)
	撥音 ん
	促音 っ(促音に用いる「っ」は、なるべく小書きにする)
第二(表記の慣習による特例)	長音 ア列→ア列の仮名に「あ」。イ列→イ列の仮名に「い」。ウ列→ウ列の仮名に「う」。エ列→エ列の仮名に「え」。オ列→オ列の仮名に「う」。
	助詞の「を」は、「を」と書く。
	助詞の「は」は、「は」と書く。 ※次のようなものはこの例にあたらない。いまわの際 すわ一大事 雨も降るわ風も吹くわ 来るわ来るわ きれいだわ
	助詞の「へ」は、「へ」と書く。
	動詞の「いう(言)」は、「いう」と書く。
	次のような語は、「ぢ」「づ」を用いて書く。 (1) 同音の連呼によって生じた「ぢ」「づ」 [例] ちぢみ(縮) つづみ(鼓) つづく(続) ※「いちじく」「いちじるしい」は、この例にあたらない。
	(2) 二語の連合によって生じた「ぢ」「づ」 [例] はなぢ(鼻血) そえぢ(添乳) そこぢから(底力) いれぢえ(入知恵) まぢか(間近) こぢんまり ちかぢか(近々) ちぢぢり みかづき(三日月) わしづかみ てづくり(手作) こづつみ(小包) こづつて はこづめ(箱詰) かたづく ねばりづよい つねづね(常々) ※なお、次のような語については、現代語の意識では一般に二語に分解しにくいもの等として、それぞれ「じ」「ず」を本則とし、「ぢ」「づ」を用いて書くこともできるものとする。 [例] せかいじゅう(世界中) いなづま(稲妻) かたず(固唾) きずな(絆) さかずき(杯) ときわず ほおずき みみずく うなずく おとずれる(訪) かしずく つまずく ぬかずく ひざまずく あせみずく くんずほぐれつ さしずめ ずすっぱり なかんずく うでずく ころずくめ ひとりずつ ゆうずう(融通)
	※次のような語の中の「じ」「ず」は、漢字の音読みでもともと濁っているものであって、上記(1)(2)のいずれにもあたらず、「じ」「ず」を用いて書く。 [例] じめん(地面) ぬのじ(布地) ずが(図画) りやくず(略図)
	次のような語は、オ列の仮名に「お」を添えて書く。これらは、歴史的仮名遣いでオ列の仮名に「ほ」又は「を」が続くものであって、オ列の長音として発音されるか、オ・オ、コ・オのように発音されるかにかかわらず、オ列の仮名に「お」を添えて書くものである。 [例] おおかみ おおせ(仰) おおやけ(公) こおり(氷・郡) こおろぎ ほお(頬・朴) ほおずき ほのお(炎) いきどおる(憤) おおう(覆) こおる(凍) しおおせる とおる(通) もよおす(催) いとおいしい おおい(多) おおきい(大) おおむね おおよそ
	付記 次のような語は、エ列の長音として発音されるか、エイ、ケイなどのように発音されるかにかかわらず、エ列の仮名に「い」を添えて書く [例] きれい せい(背) かせいで(稼) まねいて(招) 春めいて へい(塀) めい(銘) れい(例) えいが(映画) とけい(時計) ていねい(丁寧)

と歩調を合わせているようです。

このかなづかいの特徴は、現代の音韻に基づいた表記を大幅に取り入れているところでは。

発音に合わせるために、たとえば従来「ゐ」としていた表記は「い」、「ゑ」は「え」、「を」は「お」、「くわ」は「か」、「ぐわ」は「が」に、また、「ぢ、づ」→「じ、ず」、「は、ひ、ふ、へ、ほ」→「わ、い、う、え、お」といったように変更されました。

音を伸ばす場合も「いう、いふ、ゆふ」が「ゆう」、「あう、わう、あふ、ほう」が「おう」、「かう、くわう、かふ、こふ」が「こう」、「さう、さふ」が「そう」になるなどの変更が行われています。

これによって、たとえば従来は「ゐど」だった井戸が「いど」となり、「くわじ(火車)」が「かじ」、「さうち(掃除)」が「そうじ」、「えうりやう(要領)」が「ようりやう」になるなど、言葉の表記が大幅に変わりました。

例を見ると実感されると思いますが、この現代かなづかいはかなり現代の文章に浸透しており、今さら歴史的仮名遣いに戻すのは難しいでしょう。しかしながら、文学者などを中心に、文法や語源に忠実な歴史的仮名遣いに戻すべきという意見も根強く存在します。

現代仮名遣いの問題としては、「を」→「お」、「は」→「わ」、「へ」→「え」という原則があっても、助詞の「を」「は」「へ」はそのままであったり、「ぢ」→「じ」、「づ」→「ず」の規則についても、二語の連合や、同音の連呼によって生じた「ぢ」「づ」はそのまま、といった例外規則があり、完全な表音主義でも語源主義でもない中途半端さが挙げられるでしょう。

ちなみに、二語の連合で生じた「ぢ」「づ」はそのままという例外ルールでも、「せかいじゅう」などは例外のさらに例外ということになります。例外的なルールが多いため、分かりやすい仮名遣いを目指した現代仮名遣いが、結局のところ複雑にならざるを得なかった点も批判の対象にされているようです。

送り仮名

送り仮名の変遷

戦前の本を見てみると、文章の書き方が今とかなり違っていることが分かります。漢字は今の台湾などで使っている繁体字のような古い形をしていますし、仮名遣いは「てふてふ」や「おはやう」といった、いわゆる歴史的仮名遣いです。

漢字や仮名遣いほどは目立たないものの、日本語の文章で欠かせない存在として、主に訓読みの漢字の後に付く「送り仮名」があります。送り仮名によって漢字の読み方も変わってくるなど、重要度は低くないのですが、漢字や仮名遣いよりもバラツキが大きく、同じ文章中で統一されていないケースもまま見られます。

印刷物を作る場合に欠かせない表記の統一を考えると、送り仮名の問題は避けて通れません。今度はこの送り仮名に注目してみましょう。

送り仮名も戦前と戦後では変化が見られます。戦前は、送り仮名が今よりも少なめでした。ただし、送り仮名のルールが厳密に決まっていたわけではないようです。

送り仮名が少ないと、漢字の読み方が二通りあった時にどう読むのか迷うことにもなります。新しい時代に即した送り仮名のルールを定めようということで「送りがなのつけ方」が国語審議会によってルール化され、昭和34年に内閣告示されました。

しかし、この送り仮名のルールには批判が多く、昭和40年代になると再検討が行われ、48年には以前の告示を廃止して「送りがなのつけ方」が改めて告示されることになりました。これが現代の送り仮名における基本ルールです。その後、昭和56年に改訂があり、現在に及んでいます。

送り仮名の基本ルール

それでは、内閣告示の送り仮名のルールについて具体的に見ていきましょう。

といっても全ては紹介できないのでごく基本的なルールだけです。

動詞、形容詞、形容動詞といった活用がある品詞(用言)の場合、活用語尾を送り仮名にします(ただし、語幹が「し」で終わる形容詞は「し」から送る。同様に、語尾の前が「か」「らか」「やか」になる形容動詞はこれらの音から送る)。

一方、名詞には基本的に送り仮名はつけません。また、副詞、連体詞、接続詞は、最後の音節を送ります。

複数の語がつながった複合語や派生語は元の語に準拠して送ります。名詞だと、用言から名詞になったものや「さ」「み」「げ」などの接尾語が付いたものは元の語と同じように送らなければなりません。

数多い例外と許容範囲

ここまで読んできた方の中には、「おかしいぞ」と思われた人がいるかもしれません。上記のルールに当てはまらない語が少なくないのです。

実は、ここで述べたルールはあくまでも基本的なルールであり、実際には例外が数多く存在します。たとえば、「明るい」(原則だと「明い」のはず)、「同じだ」(同だ)、「幸せ」(幸)、「自ら」(自)など。要するに、一応のルール化はできるものの、それを押し通すにはあまりにも慣用と食い違いが多すぎるのです。そのため、内閣告示の「送りがなのつけ方」を見ても、基本ルールである「本則」の後にはたいてい「例外」の項があり、覚えきれないほど膨大な例が載っています。

もっとも、言語は使いながら覚えるものであり、文法などは知らなくても日常に支障はありません。送り仮名にしても例外を含めて自然と私たちの頭に入っているわけで、覚えたままを使ってもたいへい問題はなかったりするものです。

仕事で文章を扱う私たちにとって一番問題なのは、二つ以上の“正解”がある場合です。内閣告示には「本則」「例外」の他に「許容」という項目があります。これは、本則の形以外に慣用として行わ

れている形で、これを使ってもよいとされるものです。

「許容」の例として挙げられているものとしては「表す・表わす」「現れる・現われる」「行う・行なう」(いずれも前の語が本則、後が許容)など、よく使われ

る語がかなり含まれています。

要するに、どちらを使ってもよい、と国がいわばお墨付きを与えているわけです(もっとも「送り仮名のつけ方」自体、「よりどころ」であって「専門分野や個人」の表記にまで及ぼそうとするものでは

ない」とはじめに断り書きをしている)。結局のところ、送り仮名の使い方については確固としたルールは存在しないと考えたほうがいいのかも知れません。それだけに十分な注意が必要です。

「送り仮名の付け方」(昭和48年内閣告示、56年改正・一部) ※活用あり：動詞・形容詞・形容動詞、活用なし：名詞・副詞・連体詞・接続詞

単 独 の 語	通則1 (活用あり)	本則 活用のある語(通則2以外)は活用語尾を送る。 [例] 憤る 承る 書く 実る 催す 生きる 陥れる 考える 助ける 荒い 深い 賢い 濃い 主だ	例外 (1) 語幹が「し」で終わる形容詞は「し」から送る。[例] 著しい 惜しい 悔しい 恋しい 珍しい (2) 活用語尾の前に「か」「やか」「らか」を含む形容動詞は、その音節から送る。[例] 暖かだ 静かだ 明らかだ 滑らかだ 柔らかだ (3) 次の語は次に示すように送る。明らか 味わう 哀れむ 慈しむ 教わる 脅かす(おどかす) 脅かす(おびやかす) 食らう 異なる 逆らう 捕まる 群がる 和らぐ 揺する 明るい 危ない 危うい 大きい 少ない 小さい 冷たい 平たい 新たに 同じだ 盛んだ 平らだ 懇ろだ 惨めだ 哀れだ 幸いだ 幸せだ 巧みだ
	通則2 (活用あり)	本則 活用語尾以外の部分に他の語を含む語は、含まれている語の送り仮名の付け方によって送る(含まれている語を〔 〕の中に示す)。 [例] (1) 動詞の活用形又はそれに準ずるものを含むもの。動かす〔動く〕 浮かぶ〔浮く〕 勇ましい〔勇む〕 晴れやかだ〔晴れる〕 及ぼす〔及ぶ〕 起こる〔起きる〕 当たる〔当てる〕 恐ろしい〔恐れる〕 (2) 形容詞・形容動詞の語幹を含むもの。怪しむ〔怪しい〕 確かめる〔確かだ〕 重たい〔重い〕 (3) 名詞を含むもの。汗ばむ〔汗〕 先んずる〔先〕 春めく〔春〕 男らしい〔男〕 後ろめたい〔後ろ〕	許容 右の語は、()中に示すように活用語尾の前の音節から送ることができる。 [例] 表す(表わす) 著す(著わす) 現れる(現われる) 行う(行なう) 断る(断わる) 賜る(賜わる) (注意) 語幹と活用語尾との区別がつかない動詞は、例えば、「着る」「寝る」「来る」などのように送る。
	通則3 (活用なし)	本則 名詞(通則4を適用する語を除く)は、送り仮名を付けない。 [例] 月 鳥 花 山 男 女 彼 何	例外 (1) 右の語は最後の音節を送る。辺り 哀れ 勢い 幾ら 後ろ 傍ら 幸い 幸せ 互い 便り 半ば 情け 斜め 独り 誉れ 自ら 災い (2) 数をかぞえる「つ」を含む名詞は、その「つ」を送る。[例] 一つ 二つ 三つ 幾つ
	通則4 (活用なし)	本則 活用のある語から転じた名詞、及び活用のある語に「さ」「み」「げ」などの接尾語が付いて名詞になったものは、もとの語の送り仮名の付け方によって送る。 [例] (1) 活用のある語から転じたもの。動き 仰せ 恐れ 薫り 曇り 調べ 届け 願い 当たり 代わり 向かい 狩り 答え 問い 祭り 群れ 憩い 愁い 憂い 香り 極み 初め 近く 遠く (2) 「さ」「み」「げ」などの接尾語が付いたもの。暑さ 大きさ 正しさ 確かさ 明るみ 重み 憎しみ 惜しげ	例外 次は、送り仮名を付けない。 謡 虞 趣 水 印 頂 帯 畳 卸 煙 恋 志 次 隣 富 恥 話 光 舞 折 係 掛(かかり) 組 肥 並(なみ) 巻 割
	通則5 (活用なし)	本則 活用のある語から転じた名詞、及び活用のある語に「さ」「み」「げ」などの接尾語が付いて名詞になったものは、もとの語の送り仮名の付け方によって送る。 [例] (1) 活用のある語から転じたもの。動き 仰せ 恐れ 薫り 曇り 調べ 届け 願い 当たり 代わり 向かい 狩り 答え 問い 祭り 群れ 憩い 愁い 憂い 香り 極み 初め 近く 遠く (2) 「さ」「み」「げ」などの接尾語が付いたもの。暑さ 大きさ 正しさ 確かさ 明るみ 重み 憎しみ 惜しげ	許容 読み間違えるおそれのない場合、右の()中のように送り仮名を省ける。 [例] 曇り(曇) 届け(届) 願い(願) 晴れ(晴) 当たり(当り) 代わり(代り) 向かい(向い) 狩り(狩) 答え(答) 問い(問) 祭り(祭) 群れ(群) 憩い(憩)
複 合 の 語	通則6	本則 複合語(通則7適用語を除く)の送り仮名は、漢字それぞれの音訓を用いた単独の語の送り仮名の付け方による。 [例] (1) 活用のある語 書き抜く 流れ込む 裏切る 聞き苦しい 薄暗い 心細い 軽々しい 気軽だ (2) 活用のない語 石橋 目印 田植え 夜明かし 入り江 寒空 愚か者 行き帰り 長生き 粘り強さ 有り難み 待ち遠しさ 乳飲み子 次々 休み休み	許容 読み間違えるおそれのない場合は、右の()の中に示すように、送り仮名を省くことができる。 [例] 書き抜く(書抜く) 聞き苦しい(聞き苦しい) 田植え(田植) 夜明かし(夜明し) 入り江(入江) 売り上げ(売上げ・売上) 乗り換え(乗換え・乗換) 有り難み(有難み) 待ち遠しさ(待遠しさ) 立ち居振る舞い(立ち居振舞い・立ち居振舞) 呼び出し電話(呼出し電話・呼出電話) (注意)「こけら落とし(こけら落し)」のように、前又は後ろの部分を仮名で書く場合、他の部分は単独の語の送り仮名の付け方による。
	通則7	複合の語のうち、右のような名詞は、慣用に従って、送り仮名を付けない。 [例] (1) 特定の領域の語で、慣用が固定していると認められるもの。⑦地位・身分・役職等の名。頭取 取締役 ⑧工芸品の名に用いられた「織」「染」「塗」等。(博多)織 ⑨その他。書留 請負 売上(高) (2) 一般に、慣用が固定していると認められるもの。木立 試合 場合 番組 日付 物語 織物 (注意) (1)「(博多)織」「売上(高)」などのようにして掲げたものは、()の中を他の漢字で置き換えた場合にも、この通則を適用する。(2) 通則7を適用する語は例に挙げたものだけではない。通則7を適用するか判断し難い場合には通則6を適用する。	
	付 表 の 語	「常用漢字表」の「付表」に掲げてある語のうち、送り仮名の付け方が問題となる以下の語は以下のようにする。 ① 次の語は次に示すように送る。なお()の付いた語は()中のように送り仮名を省くことができる。浮つく お巡りさん 差し支える(差支える) 五月晴れ(五月晴) 立ち退く(立退く) 手伝う 最寄り ② 次の語は送り仮名を付けない。息吹 棧敷 時雨 築山 名残 雪崩 吹雪 迷子 行方	

文字コード問題を理解する

田村信幸

日本の文字コード

コンピュータで文字を扱う上で、もっとも重要になるのが文字コードでしょう。文字コードとは、文字にコンピュータで扱える符号（コード）を当てはめた体系です。同じ文字コードを使えば、異なる環境でも同じ文字をやり取りすることが可能です。文字コードは便利で欠かすことができないものですが、その反面、トラブルの要因になる可能性も小さくありません。

現在、日本における最新の文字コードの規格は2004年に制定された「JIS X 0213:2004」です。この稿ではこのJIS X 0213:2004規格について見ていきますが、その前に、これまで日本の文字コードがどのように作られてきたのかをざっと眺めておきましょう。

1) 初めての漢字コード

初めのころ、コンピュータでは漢字を使うことができませんでした。漢字の数はアルファベットなどに比べてはるかに多く、それを扱うだけの能力がなかったからです。

コンピュータの性能が上がるにしたがって、各メーカーはそれぞれ独自の文字コードを作り、漢字への対応を進めていきましたが、情報交換という観点で考えれば日本語共通の文字コードがぜひとも必要です。

そこで、JIS原案委員会は標準的に使われる漢字6,349字を選定、さらに漢字以外の文字453字を加えて6,802字の文

字コード「JIS C 6226」を1978年に作りました。これが後に「JIS X 0208」と呼ばれることになる、日本語におけるもっとも基本的な文字コードです。

当時はパソコンが登場したころであり、多くの文字を扱えないコンピュータも少なくありませんでした。そこでJIS C 6226 (JIS X 0208) では、6,349字の漢字を第一水準と第二水準に分け、第二水準はオプションとして扱えるようになっていました。

JIS規格は制定や改正されたときから5年が経過するまでに見直しされなければならぬと工業標準化法で定められています。文字コードの規格も、5年という期限には必ずしも収まらないながら、数年ごとに見直しが行われています。

JIS X 0208が制定後初めて改正されたのは1983年のこと。この時に、その後の日本語文字コードの運命を左右する大きな問題が起きました。JISの規格票には文字コードの例示字体が印刷されます

が、この例示字形のうち二百数十字が変更されたのです。さらに、同じ意味で形が異なる二つの文字（たとえば壺と壺、檜と桧など）が収録されていたものに関して、そのうちの22組のコード番号が入れ替えられました。

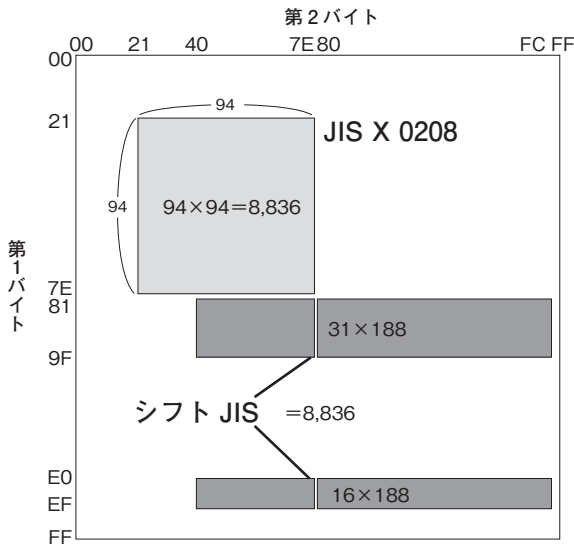
この二つの改正は、パソコンにおける日本語に大きな影響を与えるものでした。字体の変更では、それまでの正字と呼ばれる一般的な文字に替えて印刷ではまず使われなかった略字体を採用するという方針が強く打ち出されました。よく知られているのは、森鷗外の「鷗」の字（「鷗」に変更）などですが、ニーズの高い字体から低い字体に変更されたわけですから、抵抗も大きく、メーカーによっては83JISを採用せずに78JISを使い続けるという動きもありました。

2) 文字数の追求

JIS X 0208は、使用頻度の高い6,000以上の漢字を収録した本格的な日本語文

日本の文字コード（漢字コード）の歴史

制定年	規格	文字数	概要
1978年	JIS X 0208 (JIS C 6226)	6,802字	日本最初の漢字を含む文字コード。第一水準と第二水準に分かれる。
1983年	JIS X 0208	6,877字	例示字体を大幅に変更し、コードポイントの入れ替えなど、後の混乱を招いた改正。
1990年	JIS X 0208	6,879字	2字を追加。
	JIS X 0212	6,067字	俗に「補助漢字」と呼ばれる文字コード。すべてUnicodeに収録。
1997年	JIS X 0208	6,879字	収録文字に変更なし。ただし、文字の由来を調査、外字を否定するなど問題を整理した。
2000年	JIS X 0213	11,223字 (0208以外は4,344字)	97JISを踏まえ、典拠の確かな文字を収録。JIS X 0208の文字を内包する文字集合。新たな漢字は第三水準および第四水準となる。
2004年	JIS X 0213	11,233字 (0208以外は4,354字)	表外漢字字体表に従い、例示字体を変更、さらに包摂基準を変更して10字を追加。



JIS X 0208では、2つのバイトを重ねて1つの文字を指定する(例: 84×24)。第1バイトを「区」、第2バイトを「点」と言い、それぞれ1～94ある。この区点による表し方を区点コードと呼ぶ(10進法)。実際には他で使っていない番号を使うため、左図ようになる。

シフトJISは、第1バイトをASCIIなどで使われていない領域だけに限定することで、文字コードの切り替えの手間を省くという仕組み。ただし第2バイトの領域を広くしているため、環境によっては不都合が生じる。いずれも8,836字まで収録可能。なお、0000～FFFFまですべての領域を使うのがユニコード(Unicode 1.1)。

字コードであり、これだけでも日常使われる日本語の多くが表現できます。ただし、人名や地名をきちんと表すためにはまだ十分とは言えません。また、学術出版物で使われる文字などを表す場合も不足するのは否めません。

JIS X 0208が作られた当初からこういった声は存在しており、コンピュータの性能が向上してくるにつれて、より多くの文字をサポートしてほしいという要求も高まっていきました。

JIS X 0208の二度目の改定は1990年に行われましたが、その時、JIS X 0208とは別に新しい文字コードとしてJIS X 0212が作られました。俗に言う「補助漢字」です。補助漢字は、JIS X 0208に収録されていない6,000字以上の文字を収録しており、これを使えば略字体しか使えなかった「鷗」などの問題も解決します。

なぜJIS X 0208で文字を追加するのではなく新たな文字コードを作ったかという、JIS X 0208にはこれだけの文字を収録する余地がなかったからでした。JIS X 0208は、2バイトの文字コードであり、二つのバイトを掛け合わせて文字を特定します。ただし、1バイトに付き94字しか使えないという制約があり、 $94 \times 94 = 8,836$ 字までしか文字は入りません。

1983年の時点でJIS X 0208にはす

で7,000字近い文字(1990年時点で6,879字)が収録されていますが、あと2,000字近くは追加できるはずですが、しかし、実際には空いている部分にフォントメーカーなどが外字を入れていたため、それほど多くの文字は入れられません。そこで、別の文字コードとして補助漢字が作られたわけです。

しかし、日本のパソコンではシフトJISという別の文字コードが一般的に使われるようになっていました。シフトJISは、文字の配列はJIS X 0208をそのまま使うものの、 94×94 の2バイト掛け合わせではなく、 47×190 という変則的なコードです。1バイト目をASCII文

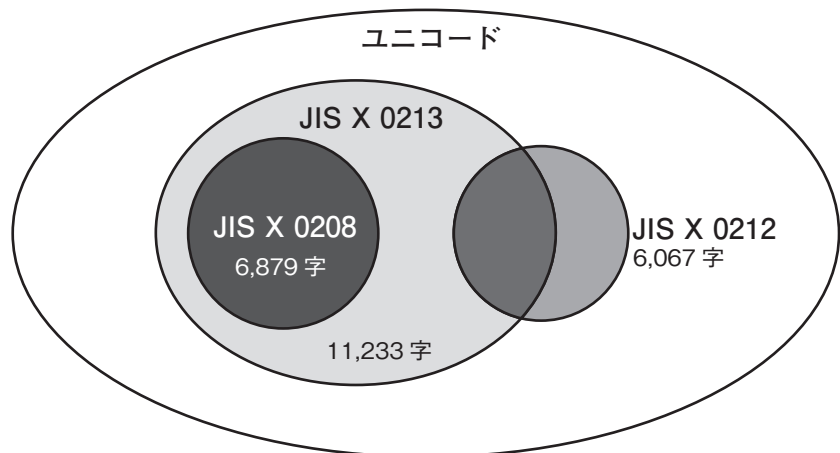
字と重ならないようにすることで、コードを切り替える必要がないというのがメリットですが、他の文字コードと共存できないという問題もあります。つまり、シフトJISを使っている限り補助漢字は使えないのです。

そのため、せっかく作られた補助漢字も、ユニコードでサポートされ、さらにWindows 98にユニコード対応のMS明朝・MSゴシックが搭載されるまで使われることはありませんでした。現在も、補助漢字を収録するフォントは少なく、活用されているとは言えません。

文字数の拡張ということでは、補助漢字のほかにもう一つの文字コードが作られています。それが2000年に作られたJIS X 0213です。JIS X 0213は、JIS X 0208の文字に加えて4,344字(後に4,354字)の文字を規定(第三水準、第四水準漢字)しています。

この規格も補助漢字と同様、ユニコードに採用されることで実際に使われるという方向をとることになりましたが、補助漢字と異なるのは収録する文字の妥当性を可能な限り検証したという点です。

1990年代以降、文字数を増やす方向で進められてきた日本の文字コード改良ですが、ここに至り文字の正当性ということが大きくクローズアップされるようになってきています。



日本の漢字に関する文字コードの関係は上図のようになる。JIS X 0208とJIS X 0212は並立する別の規格だがJIS X 0213はJIS X 0208を含んでいる。また、JIS X 0212とJIS X 0213はかなりの文字が重複している。現在、いずれもユニコードに登録されているので、JIS X 0208以外の文字を扱う場合はユニコードでの運用が前提になる。

どういった文字が必要なのか

1) 必要な文字数とは

これまで日本語の文字コードは何回も改正されてきましたが、そのたびに批判もついて回ります。そもそも文字コードについてはさまざまな意見があり、だれもが納得するものを作るのは難しいでしょう。ただ、結局のところ、問題は「文字数」および「字形」が妥当かどうかという点に集約されるようです。

日本語を記述する際にどれだけの文字が必要なのかというのは、なかなか答えが出ない問題です。JIS X 0208の第一水準と第二水準漢字だけでも日常生活で使うほとんどの漢字はまかなえますが、人名や固有名詞、あるいは古典文献などを扱うには十分ではありません。

現在、JIS X 0208以外に6,067字収録のJIS X 0212や4,354字を追加する0213があり、ユニコードを利用することで、0208とあわせてそれぞれ11,000～13,000字程度の文字を扱うことができますが、これで全ての用途がまかなえるかというところとも言い切れないのです。

たとえば異体字の問題があります。漢字には、読みも意味も同じ、しかも同じような状況で使うのに、形が微妙に異なるものが存在します。有名な例では、「高」と「はしご高」などがあります。

こういった異体字は人名や固有名詞で必要になることが多いのですが、世の中にある異体字を全て拾っていくとそれだけで膨大な数です。JIS X 0212や0213にも、0208に収録されている文字の異体字が少なからず含まれていますが、そ

れでも十分ではないのです。

ただし、情報管理という観点から考えると、使える文字が多ければよいというものでもありません。仮にこれらの文字を含む語句を検索しようとした場合、どの異体字で検索するかによってヒットしたりしなかったりということになってしまいます。「兼田」で検索したら「兼田」は引っかかりなかったというのでは、せっかくのデジタルデータのメリットが損なわれてしまうのです。

2) 包摂という考え方

微妙な違いの異体字をどんどん文字コードに取り込んでいけば、確かに情報の幅は広がりますが、それだけ情報としてのデジタルデータの質は落ちてしまいます。そこで登場するのが「包摂」という考え方です。

包摂とは、同じ意味、同じ使い方をする文字を同じものとして扱うことです。たとえば「兼」という字が「兼」を包摂しているとします。この場合、「兼」と「兼」の違いはデザイン的な差ということになり、これらの文字に相当する文字コードのコードポイントは1つだけになります。「兼」という文字を使いたければそういったデザインの書体を用意すればよいというわけです。情報管理という点で言えばこのほうが楽なのは言うまでもないでしょう。

とはいえ、現実問題として区別をつけなければならない場面は少なくありません。特に印刷物だと「包摂されているから同じ字と思ってくれ」では通用しないケースが多いのは皆さんご承知の通りで

す。

日本語文字コードを考える場合、包摂をどうするか、どこからどこまでを包摂として一まとめにするのが最大の問題と言っても過言ではありません。そのため、JIS X 0213では収録した文字の典拠や包摂基準について詳細な説明がなされています。

考えてみると、規格票の例示字体に略字体を採用して大きな問題を招いてしまった83JISにしても、例示字体は包摂されたものであるのならば、実際のフォントは旧字体でもかまわなかったわけで、文字コードだけを責めるわけにはいかないという考えも成り立ちます。

3) 文字コードから離れた異体字処理

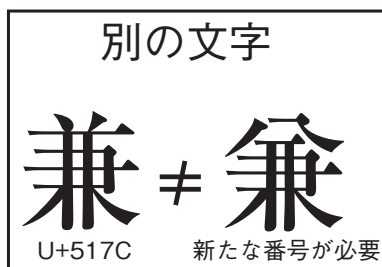
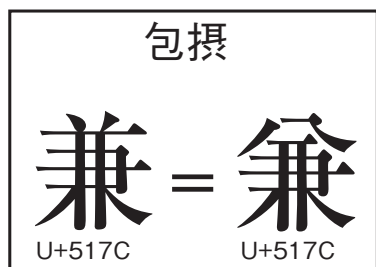
最近、この問題を別の方向から解決する技術が登場してきました。それがOpenTypeフォントです。

OpenTypeフォントは、もちろん文字コードをサポートしていますが、それとは別に異体字を認識するタグを用意しており、文字コードで扱えない異体字を扱うことができるようになっています。たとえばInDesignの字形パレットを使えば、文字コードでは包摂されている（つまり同じ字とみなされている）文字や文字コードが振られていない文字も指定することが可能です。

OpenTypeフォントの登場は、文字コードのあり方も変えようとしています。これまでは、包摂されたものと文字が区別できなくなり、結局使えないというのが大きな問題だったのですが、インターネットなど通常のデジタル情報レベルでは包摂された文字を使い、組版レベルで異体字を区別して使えば少なくとも印刷物での問題は解決します。

もっとも、今のところ異体字タグやInDesignの文字処理機能に問題がないわけではないので、全て解決というわけにはいきませんが、問題を解決する仕組みは一応用意されたと言ってもいいでしょう。

文字コードにおける文字数の問題につ



包摂の考え方では、字形が違って同じ字として扱うので、文字コードの番号もひとつになる。包摂されない文字となった場合は、文字コード上の番号が別に用意されることになる。

文字セット	文字数	特 徴
Adobe-Japan1-0	8,284 字	JIS X 0208 および JIS X 0201、アップル・富士通・NEC 文字セット (JIS83 ベース) のサポート
Adobe-Japan1-1	8,359 字	漢字 Talk 7.1・富士通・NEC 文字セット (JIS90 ベース) のサポート
Adobe-Japan1-2	8,720 字	Windows 3.1J 文字セットのサポート
Adobe-Japan1-3	9,354 字	半角およびプロポーショナル文字の回転字形
Adobe-Japan1-4	15,444 字	商業印刷のニーズを満たすためにアドビが選んだ文字セット。OpenType の Pro 書体
Adobe-Japan1-5	20,317 字	JIS X 0213 : 2004 をサポート
Adobe-Japan1-6	23,057 字	JIS X 0212、共同通信 U-PRESS 文字セットのサポート

いては、今のところ一段落といった感じで、これまでのような大きな追加は当分なさそうです。フォントへの収録文字(グリフ)数はまだ増え続けるかもしれませんが、OpenType フォントの仕組みを使えば必ずしも文字コードも一緒に拡張しなければならないわけではないからです。

JIS 例示字体の変更

1) 表外漢字字体表

2000年12月8日、国語審議会から「表外漢字字体表」という、一見地味な、しかし、今後のコンピュータの文字環境、そして我々 DTP・印刷業界にも大きな影響を及ぼす極めて重大な意味を持つ答申が出されました。

「表外漢字」とは聞きなれない名称ですが、これは“常用漢字表”の外の漢字という意味で、常用漢字表に収録されていない漢字の字体についてのよりどころを示すために作られたのが表外漢字字体表です。

昭和21年制定の当用漢字表、昭和56年の常用漢字表はいずれも漢字の使用範囲に一定の制約を課するという目的で作られたものです。

ところが、漢字を取り巻く状況はその当時とは一変することになりました。文字情報を伝達するツールとしてコンピュータが普及したのです。

漢字政策の見直しが進みつつあった1978年、日本で初めての漢字を含む文字コード JIS X 0208 が作られ、従来の漢字表にない文字をコンピュータを使っ

て表現することが可能になりました。

元々、漢字の習得が負担であるということから始まった漢字制限ですが、コンピュータで簡単に漢字を出せる時代になると、その意味はあいまいなものにならざるを得ません。コンピュータで利用できる文字数への要求は強く、その後、JIS X 0212 や 0213 によってさらに文字数が拡張されたのはこれまでにきてきた通りです。

一方、字体について見てみると、当用漢字表、常用漢字表にはそれぞれ字体表がありましたが、漢字表に収録されていない漢字に関しては何も定められていないという状況が長い間続いていました。

そういった状況の中、コンピュータで使う文字において実質的な字体の目安となったのが JIS 規格票の例示字体です。

特に、1983年の改正では、元々あった文字の代わりにあまり見慣れない略字体が例示字体として多く採用されたため、それを元に作ったフォントだと元の文字(字体)が使えなくなるなど、大きな混乱を巻き起こしたことで悪名高いものでした。

しかも、コンピュータの普及に伴い、JISの例示字体がコンピュータの世界にとどまらず、印刷物など社会全体の文書に影響を与える事態が起きてきました。

国語審議会が表外漢字字体表を作ったのは、このような字体に関する混乱状況を国語基本政策の観点から正し、漢字のよりどころを示そうというのが最大の目的だったと言えるでしょう。ただし、常用漢字表内の漢字については、すでに字体が定着しているという判断から見直し

は行われず、あくまでも“表外”漢字についてだけ基準を示すことになりました。

表外漢字字体表を作るにあたっては、大手印刷会社および新聞社による漢字出現頻度数調査が行われました。それによると、人名用漢字以外の常用漢字表外の文字は使われる頻度こそ少ないものの字種はかなりあり、その多くはいわゆる“康熙字典体”(中国清朝の康熙帝の命で作られた「康熙字典」は漢字辞典の模範として後世に大きな影響を与えた。康熙字典体はこの辞典で使われている字体を意味しているが、定義があいまいという批判もある)だったということです。

表外漢字字体表は、常用漢字とともに使うことが比較的多いと考えられる表外漢字1,022字を特定し、その「印刷標準字体」を示しています。簡略された文字を字体として採用した当用漢字表や常用漢字表と比べて、康熙字典体を基本にしているため、従来の印刷物で使われてきたいわゆる“正字”が中心になっています。ただし、俗字・略字といわれる文字でも、使用頻度が高く康熙字典体の正字以上に使われているもの(5倍以上の出現頻度など)については印刷標準字体に採用されました。

また、正字が収録された文字のうち、俗字・略字も一般的に使われているとみなされた22字については、略字も簡易慣用字体として収録されています。なお、この字体表に収録されなかった漢字については、漢和辞典で正字体とされてきたものを原則とするということです。

表外漢字字体表は、従来印刷物で使われてきた字体を基本に作られており、印刷業界にとっては一見受け入れられやすいもののように見えるかもしれませんが、しかし、今ある文字コードとの整合性を考えた場合、そこにはきわめて重大な問題が潜んでいました。

次に、表外漢字字体表へ対応するために改正された JIS X 0213 の問題点および今後トラブルが起きることが避けがたいと思われる新 Windows のフォント環

境について見ていきます。

2) JIS X 0213 の改正

表外漢字字体表は、当時の国語審議会による答申であり、内閣告示になった当用漢字字体表などのような強制力はありません。

ただし、表外漢字字体表をよく読むと、そこには「情報機器時に搭載される表外漢字の字体については、表外漢字字体表の趣旨が生かされること」が望ましく、文字コードの見直しがあった場合はその「趣旨が生かせる形での改訂が望まれる」と書かれています。

この表現を素直に読めば、文字コードの改訂時には表外漢字字体表に合わせて例示字体を変更するべきだということになるでしょう。実際、表外漢字字体表に対応する JIS コード改正時の通商産業省プレス発表でも「パソコンなどに搭載される字形の変更を求めるものではない」が、今回の改正によって字形が「徐々に印刷標準字形に変更されることが期待される」とあり、コンピュータの文字環境を修正しようという当局側の強い意思が感じられます。

ちょうど見直しの時期を迎えていた JIS X 0213 の改訂作業を行う JCS 委員会では、規格票の例示字体をこの表外漢字字体表に対応させることにしました。

現在、基本となっている文字コード JIS X 0208 には、表外漢字字体表の印刷標準字体と相容れない字体が例示字体として数多く採用されており、パソコンで使われるフォントもそれに準じて作られています。

すでに社会全体で広く使われている JIS X 0208 の例示字体を変更してしまうと、フォントベンダーはフォントを新たな例示字体に合わせて作るようになり、従来のフォントの字体との齟齬によって大混乱が生じることは 1983 年の改訂を思い起こすまでもなく明らかです。また、JIS X 0208 はそのまま ISO/IEC 10646 (ユニコード) にも収録されているので、0208 を変更すれば国内規格の改定だけ

ではすみません。

そこで JCS 委員会がとった手段は、JIS X 0208 は変更せず、JIS X 0213 で例示字体の変更と文字の追加を行うというものでした。

JIS X 0213 は、JIS X 0208 と独立した規格ではなく、正確には 0208 の文字をすべて含み、さらに 4344 字(後に 4354 字)の新たな文字を加えた規格なのです。つまり 0208 の文字は 0213 にも重複して収録されているわけです。にも関わらず、JCS 委員会は 0208 を変更せず、0213 の例示字体だけを変更することにしました。このままだと 0208 と 0213 が矛盾するのではないと思われるかもしれませんが、変更字形は包摂の範囲内なので問題はないというのが 0213 の立場です。

0208 はもっとも基本的な文字コードであり、従来の環境はこの文字コードを前提に作られています。一方、0213 は新しい規格であり、事実上ユニコード環境でしかサポートされていません。そこで、0213 だけ変更することで旧来の環境でのトラブルの危険を最小限に留めようとしたのです。

また、1997 年の JIS 改訂では包摂範囲を厳密に特定しましたが、その結果、83JIS で変更された字体はほとんどが 78JIS と包摂の関係にある(つまり同一視できる)としました。ただし、29 文字だけはあまりに異なる字体であり包摂できなかったのです(互換基準)。そこで JIS X 0213 : 2000 では、これら 29 字の 78JIS 字形を、新たに加わった 4344 字の中に収録しています。

仮に、元の 0208 の字形を 78JIS に変更するとすると、4344 字の中にある 29 字の文字とバッティング(重複)してしまい、今度は 4344 字の中の 29 字を 83JIS に変更しなければなりません。そうになると、字形の入れ替えという最悪の事態が起きてしまいます。

結局、JIS X 0208 はそのままにするしかなかったということでしょう。JIS X 0213 は 2004 年に改訂されました (JIS X 0213 : 2004) が、表外漢字字体表に対応

するために 168 字の例示字体の変更と、包摂基準の変更によって必要になった 10 字の追加がなされました。

文字化け必至の新 Windows

遅くとも 2007 年には間違いなく登場するであろうとされる Windows の新バージョン「Windows Vista」。その売りの一つが一新されたフォント環境です。

これまで、日本語版 Windows には MS 明朝、MS P 明朝、MS ゴシック、MS P ゴシック、MS UI ゴシックの 5 書体が標準フォントとしてバンドルされてきました。Windows Vista ではさらに「メイリオ」という書体が加わり、合計 6 書体すべてで JIS X 0213 : 2004 への対応が行われることになっています。

JIS X 0213 : 2004 への対応というのは、具体的には文字数の追加と字体の変更を意味しますが、特に重要なのが字体の変更です。新しい Windows の標準バンドルフォントでは、100 字以上の字体が JIS X 0213 : 2004 に合わせて変更されることになったのです。

では、どういった文字が変更されるのでしょうか。まだ変わる可能性はあるかもしれませんが、現時点では細かな変更点を入れて 122 字、たとえば、葛、樋、撰、榊、摺、揃、杓、騙、鯖、鱒、牙、襴、謎、逢、辻、迄、餌、鉛、餅といった文字が変更されることになります。

これらの文字が使われている従来の Windows で作ったテキスト・データを新しい Windows Vista 環境で開くと、自動的に JIS X 0213 : 2004 の字形に変わってしまうわけです。もちろん逆も同じです。

しかも、新しい MS フォントはフォント名は以前とまったく同じであり、アプリケーションレベルで区別をつけることはできません。

ただし、マイクロソフトでは Vista 用に従来字形のフォント、XP 用に新しい字形のフォントを提供するというので

JIS X 0213の字形変更箇所

JIS X 0213 面区点番号	2000 字体	2004 字体	JIS X 0213 面区点番号	2000 字体	2004 字体	JIS X 0213 面区点番号	2000 字体	2004 字体	JIS X 0213 面区点番号	2000 字体	2004 字体	JIS X 0213 面区点番号	2000 字体	2004 字体
1-16-9	逢	逢	1-23-9	糸	糸	1-32-89	煎	煎	1-39-55	牌	牌	1-46-91	煉	煉
1-16-18	芦	芦	1-23-23	邗	邗	1-32-90	煽	煽	1-39-71	這	這	1-46-92	簾	簾
1-16-27	飴	飴	1-23-68	隙	隙	1-32-92	穿	穿	1-39-73	秤	秤	1-47-17	榔	榔
1-16-78	溢	溢	1-23-81	倦	倦	1-32-93	箭	箭	1-39-93	駁	駁	1-47-64	屢	屢
1-16-81	茨	茨	1-23-94	捲	捲	1-33-7	詮	詮	1-40-4	箸	箸	1-49-45	冤	冤
1-16-83	鯛	鯛	1-24-3	牽	牽	1-33-25	噌	噌	1-40-32	叛	叛	1-50-55	叟	叟
1-16-92	淫	淫	1-24-16	鍵	鍵	1-33-44	邇	邇	1-40-52	挽	挽	1-50-91	咬	咬
1-17-10	迂	迂	1-24-33	諺	諺	1-34-23	揃	揃	1-40-80	誹	誹	1-51-62	嘲	嘲
1-17-25	厩	厩	1-25-11	巷	巷	1-34-29	遜	遜	1-40-85	樋	樋	1-51-83	囀	囀
1-17-29	噂	噂	1-25-28	梗	梗	1-34-60	腿	腿	1-41-3	稗	稗	1-55-49	徘	徘
1-17-34	餌	餌	1-25-49	膏	膏	1-34-93	蛸	蛸	1-41-15	逼	逼	1-57-8	扁	扁
1-18-8	襖	襖	1-25-84	鵠	鵠	1-35-9	辿	辿	1-41-21	謬	謬	1-59-89	棘	棘
1-18-64	迦	迦	1-25-89	甌	甌	1-35-14	樽	樽	1-41-31	豹	豹	1-60-84	橙	橙
1-18-71	牙	牙	1-26-21	叉	叉	1-35-23	歎	歎	1-41-32	廟	廟	1-64-36	狡	狡
1-18-86	廻	廻	1-26-71	榊	榊	1-35-80	註	註	1-41-46	瀕	瀕	1-65-17	甕	甕
1-18-90	恢	恢	1-27-7	薩	薩	1-35-85	瀦	瀦	1-41-64	斧	斧	1-65-20	甦	甦
1-19-2	晦	晦	1-27-10	鯖	鯖	1-36-29	抄	抄	1-42-35	蔽	蔽	1-65-54	疼	疼
1-19-10	蟹	蟹	1-27-12	鑄	鑄	1-36-40	槌	槌	1-42-45	瞥	瞥	1-67-14	崇	崇
1-19-75	葛	葛	1-27-13	鮫	鮫	1-36-42	鎚	鎚	1-42-46	蔑	蔑	1-67-62	竈	竈
1-19-83	鞆	鞆	1-27-33	餐	餐	1-36-52	辻	辻	1-42-51	篇	篇	1-68-7	筵	筵
1-19-88	釜	釜	1-28-61	杓	杓	1-36-82	挺	挺	1-42-58	媿	媿	1-68-32	篝	篝
1-20-45	翰	翰	1-28-62	灼	灼	1-37-2	鄭	鄭	1-42-60	鞭	鞭	1-71-7	臄	臄
1-20-69	翫	翫	1-29-22	酋	酋	1-37-7	擢	擢	1-42-89	庖	庖	1-71-59	艘	艘
1-21-11	徽	徽	1-29-61	楯	楯	1-37-14	溺	溺	1-43-9	蓬	蓬	1-71-74	芒	芒
1-21-32	祇	祇	1-29-82	薯	薯	1-37-38	兎	兎	1-43-80	鱒	鱒	1-73-42	虔	虔
1-21-66	汲	汲	1-29-83	藹	藹	1-37-40	堵	堵	1-43-88	迄	迄	1-73-71	蜃	蜃
1-21-68	灸	灸	1-30-5	哨	哨	1-37-43	屠	屠	1-44-57	儲	儲	1-74-4	蠅	蠅
1-21-72	笈	笈	1-30-68	鞞	鞞	1-37-50	賭	賭	1-44-63	餅	餅	1-75-35	訝	訝
1-22-10	卿	卿	1-30-83	杖	杖	1-38-52	漚	漚	1-44-66	粿	粿	1-80-43	靄	靄
1-22-34	饗	饗	1-31-10	蝕	蝕	1-38-59	遁	遁	1-44-76	爺	爺	1-80-55	鞞	鞞
1-22-47	僅	僅	1-31-54	訊	訊	1-38-70	謎	謎	1-44-90	鎧	鎧	1-81-57	騙	騙
1-22-84	喰	喰	1-31-64	逗	逗	1-38-71	灘	灘	1-44-92	愈	愈	1-82-77	鴉	鴉
1-22-91	櫛	櫛	1-32-2	摺	摺	1-38-74	檣	檣	1-45-18	猷	猷			
1-22-93	屑	屑	1-32-81	撰	撰	1-39-9	襴	襴	1-46-90	漣	漣			

※字形はリュウミンProだが一部筆者が修正した文字もあり、正確性は保証の限りではない。

すし、新しいMSフォントはいずれ字形の切り替えができるようになるはずです。

とはいえ、どのような環境で入力されたかによって同じデータでも字形が違ってくるとい点には変わりありません。たとえば、入稿されたテキストデータが

JIS X 0213：2000の字形なのか2004の字形なのかを後工程側で判断する術はないのです。

このことは、DTP現場で極めて重大な問題を引き起こすことになるでしょう。環境が異なるだけで文字化けが高い確率で発生してしまうのです。しかも、

トラブルを防ぐ方法は、入力側にどの文字を想定してデータを作ったのかを確認する、くらいしかありません。

Windows Vistaが登場すれば必然的に起きるこの問題について、DTP業界全体として今のうちに対応を考えておく必要があるでしょう。

OpenType フォントの使い方

田村信幸

OpenType フォントのメリットとは？

DTPの世界にOpenTypeフォントが登場して数年が経ちました。新しい環境への移行に極めて慎重なDTP業界において、新しいフォントが普及するにはかなり時間が必要だというのはCIDフォントのケースを思い起こしても明らかですが、OpenTypeフォントの場合、プラットフォームの変更とも密接に関係しているだけに、事態はさらに複雑です。

Mac OS XやInDesignに完全に移行してしまったユーザーだとOpenTypeの使用率もかなり高くなりますが、Mac OS 8/9やQuarkXPress 3.3/4.xなど旧来のシステムに頼らざるを得ない現場ではOpenTypeへの移行は依然として進んでいないというのが現状でしょう。ただ、旧来のシステム自体いつまで使えるかわからないという状況では、OpenTypeは当然使わないから関係ない、などと安閑

としてもいられないはずですが。

DTP環境の対応度はひとまず置いて、OpenTypeフォントのメリットをあらためて考えてみましょう。Adobeやフォントメーカーはさまざまなメリットを声高に叫びますが、本当にOpenTypeフォントにはメリットがあるのでしょうか。

OpenTypeフォントのメリットとしてよく挙げられるのが「文字数の多さ（異体字）」「組版機能」「クロスプラットフォーム」などです。このうち、クロスプラットフォームというのはWindowsとMacintoshの環境でデータをお互いにやり取りするケースでのみメリットとなるものであり、誰にでも当てはまるものではないのでここでは一応除外します。

また、イタリックや上付き・下付きなどの指定をフォント内蔵の機能を使って行うという組版機能について考えてみると、確かに、この機能があれば便利なこともあるでしょう。ただし、通常は、こういった機能は使わないという仕事のほ

うが多いようにも思います。

結局、OpenTypeフォントのメリットといっても、普通の仕事で実感するメリットというのは文字数の多さくらいなのかもしれません。逆に言うと、文字数の多さをいかに活用できるかがOpenTypeフォントのメリットを享受できるかどうかのカギになるわけです。

GSUB テーブルの役割

OpenTypeフォントを使うと、異体字の変換や合字などいろいろな処理が可能ですが、これはOpenTypeフォントの機能や特性を記述したOTL (OpenType Layout) テーブルという格納部分に含まれている「GSUBテーブル」を利用することで実現されています。GSUBとはGlyph Substitution (字形置き換え) の意味で、異体字変換、縦横用文字やルビ文字の切り替えなど、文字の置換を指定するためのものです。

GSUBにはさらに細かな特性を指定するテーブルがあり、ここでタグを使ったさまざまな指定ができます。たとえば、ユニコード番号が同じ2つの文字 (字形) があるとして、1つには「jp78」、もう一つには「jp83」というタグが付けられているとします。アプリケーションがこれらのタグに対応していれば、jp78の文字からjp83の文字への置換を自動的に行うことも可能なのです。

InDesignの字形パレットを使うと、GSUBテーブルで付加されたタグの一端を確認することが出来ます。字形パレ

OpenType フォントの特徴

特 徴	詳 細
クロスプラットフォーム	データフォークにデータを格納するためMacintoshでもWindowsでもまったく同じデータ
組版機能を持つ	文字属性の自動指定や、合字、分数、文字詰めなどがフォントの機能で可能
ユニコード対応	OpenType フォントにはUTF-16のcmapテーブル (グリフ参照表) が内包されており、標準でユニコードに対応している
文字数が多い	OpenType フォントで事実上標準の文字セットとなっているAdobe-Japan1-4で15,444字、Adobe-Japan1-5では20,317字もの文字が用意されている
異体字変換が可能	タグによって異体字がグループ化されており、親字から変換することができる
ダイナミックダウンロード	プリンタ側にフォントがなくてもフォントデータをそのまま出力することができる



InDesignの字形パレットで文字の異体字を確認してみると、異体字の中には文字コード番号以外に「JIS78字形」などと表示されるものがある。これはOpenTypeフォントに含まれるタグに基づいている。

トで「選択された文字の異体字を表示」し、ポインタを異体字一覧の中の文字に置くと、その文字のCID (GID)、ユニコード、Shift JISの各番号がポップアップで表示されますが、GSUBテーブルのタグが付加されている文字にはそのタグも表示されます。

たとえば「辺」という文字を選択して字形パレットで異体字を表示させると小塚Proフォントで23文字もの異体字が現れます。一つずつポインタを当ててみると、ポップアップ表示の一番下のところに「JIS78字形 (jp78)」とか「全ての異体字.5 (aalt)」といった表示がある文字があります。これは、それぞれの文字にこういったタグが指定されているということを意味しています。

同じグループの異体字には必ずタグやタグの番号が与えられています。たとえば「aalt」というタグは異体字全般を意味しており、InDesignは通し番号によって同じグループの異体字を管理しています (異体字をInDesignタグ付きテキストに書き出すと「aalt¥.6」といったように指定される)。これによって異体字を正確に特定できるようになっているわけです。

OpenTypeフォントの文字 (字形) には、aaltやjp78、jp83以外にも trad (旧字体) や expt (エキスパート字形) といった字形に関するタグが付けられていま



InDesignの「文字スタイル」および「段落スタイル」には、字形を指定する「異体字」欄が用意されている。ここで字形を指定することで、文字をそれぞれ指定していく手間を省くことも可能になる。

す。これらのタグを指定すれば字形パレットでグループごとに一覧表示することができます。

さらに、字形パレットのメニューやスタイルで字形タグを指定すれば、テキスト全体を旧字体やJIS78字形など特定の字形へ一気に変換するといった処理も可能です (ただしJIS78字形といっても完全にJIS78に準拠しているとは限らない)。

高まる字形指定の重要性

これまで、OpenTypeフォントの字形変換は、クライアントから特定の異体字を指定されるなどの必要に迫られた場合に限って利用されることが多かったのではないのでしょうか。あくまでも例外的な処理というのが一般的な捉え方だったように思います。

しかし、まもなく字形変換の機能に大きな注目が集まるはずです。というのも、あと数ヶ月でリリースされるマイクロソフトの新OS「Windows Vista」に搭載される標準フォントでは、表外漢字に印刷標準字形を採用した改正JIS X 0213に従い、百数十文字の字形が変更されることになっています。

フォントの字形そのものが変更されるので、DTPレベルでどうこうできる問題ではありませんが、DTP現場ではで

きるだけトラブルのないように対策を考えなければなりません。

たとえば、元の前稿テキストがXPで入力されたものであれば、入力者は従来の字形のつもりで入力したとみなすことができます。また、Vistaで入力したテキストは、入力者が改正JIS X 0213準拠の字形であることを確認しながら入力したと考えられるでしょう。DTP現場でテキストを支給された場合、原則として入力者の想定した字形でデータを作るべきということになります (実際にはクライアントの意向を改めてきちんと確認しなければならなくなるはず)。

InDesignにテキストを取り込んだ場合、OpenTypeフォントを始めとする現行のフォントは、そのままですべて従来の字形になってしまいますから、改正JIS X 0213準拠の字形を使いたい場合は字形パレットで字形を変更しなければなりません。

字形パレットのメニューで一括変換できれば便利ですが、パレットメニューにある字形セットでは改正JIS X 0213を全てカバーすることはできません。結局今のところ一つずつ字形を変換していくしかないようです。ただし、トラブル防止や効率向上を考えると、字形を簡単にチェックするプログラムなどを用意するといった対策を今のうちに考えておいたほうがよいかもしれません。

バックナンバーのご案内

いんぷおるむマガジン別冊は、これまで Vol.1 ～ Vol.3 が刊行されています。

Vol.1 Professional DTP 誌 2005 年 10 月号～ 12 月号に掲載された記事に未掲載の記事を加え、「Professional DTP ダイジェスト 2005/10-12」というタイトルで発行。DTP 現場の失敗防止法と印刷文化史を掲載。

Vol.2 「いんぷおるむマガジン」に掲載した記事を元に、DTP 制作の基本や最新技術の記事をまとめたもの。

Vol.3 「いんぷおるむマガジン」掲載の記事を元に、印刷と伝統文化に関する記事をまとめたもの。

執筆者プロフィール

田村信幸

1985 年上智大学仏文科卒。印刷会社に勤務し、デジタル化および DTP の起ち上げを行う。

1997 年より Professional DTP 編集部にて Professional DTP および Windows DTP 誌の執筆・編集を担当。

現在インフォルムにて校正ならびに記事の執筆を行っている。

制作(編集・レイアウト・デザイン) いんぷおるむマガジン編集部

発行 株式会社インフォルム

発行日 2006 年 11 月 11 日

質の高い DTP と印刷を提供する
株式会社インフォルム
www.informe.co.jp