

拝啓,国語審議会の皆さん

まずは、日本語をつかう者の一人として、今回のパブリックコメントのような機会をつくられたことに、共感の拍手を送りたいと思います。寄せられた意見がどのように活用されるのか、まだ慎重に見守る必要があるとしても、ごく一部の限られた人間だけで、日本語をつかう者すべてに影響をあたえるような重要な施策が決定されるべきでないことは明白で、このようなオープンで賢明な皆さんの姿勢は、これからもぜひ継続、そして拡大されるべきと考えます。

さて、私はこの1年間、フリーライターとして文字コードの周辺を取材してきました[*1]。コンピューターで文字を使うための仕組みの一種である文字コードを調べることで、あまりに身近すぎて、かえって姿が見えづらくなっている日本語の現在、そしてひょっとしたら未来の姿まで、うまく見出すことができるのではないか、そのように考えたのです。

残念ながら、いくら取材してもそのようなものは見つけられず、見えたと思った背中もすぐに霧の彼方に隠れてしまう、そのような失望をあじわう毎日です。

[*1]……手前味噌ながら、インターネット上での私の原稿の掲載場所のURLを挙げておきます。もしよろしければご笑覧下さい。

小形克宏の「文字の海,ビットの舟」——文字コードが私たちに問いかけるもの

<http://www.watch.impress.co.jp/internet/www/column/ogata/>

それはともかく、そうやって文字コードを追いかけている者の一人として、今回皆さんの手によって発表された『表外漢字字体表(案)』(以下、表外字体案)を、おもにコンピューターの実装という視点から検討してみたいと考えました。どうかお付き合い下さい。

●ワープロで打ち出す文字だけを考えていいの？

日々をパソコンのディスプレイと向きあうことで過ごしている私のような者にとって、皆さんがお書きになった表外字体案を読んで、おもわず首を傾げてしまったのは、たとえば以下のような箇所でした。

しかし、ワープロ等の急速な普及によって、表外漢字が簡単に打ち出せるようになり、常用漢字制定時の予想を超えて、表外漢字の使用が日常化した。

(p.2 前文 表外漢字の字体に関する基本的な認識)

(前略) この問題は、(1)一般の書籍類や教科書などに用いられている「鷗」「瀆」がワー

プロ等から打ち出せないこと、(2)仮に「鷗」と「鷗」の両字体を打ち出すことができたとしても、どちらかの字体を標準と考えるかのルールがないこと、の2点にまとめられる。(p.2 前文 表外漢字の字体に関する基本的な認識)

ここでいわれる〈ワープロ等〉は、常に〈打ち出す〉、すなわち印刷機能とセットで語られています。となると私には、ワープロ専用機のことを念頭におかれているかのように読めるのですが、如何でしょう。もしそうだとすれば、ワープロ専用機はパソコンに押されて、すでに販売の中止を決定したメーカーすらあることをご存知でしょうか[*2]。

[*2]……朝日新聞2000年8月21日夕刊の記事『ワープロは今 パソコン攻勢、東芝も撤退 愛好家ら復活運動』などを参照のこと。

本当に表外字体案をまとめる際にワープロ専用機が念頭におかれていたとすれば、それは実態からかけはなれた過去の姿にすぎず、そのような場所からの思考は、やはり実態からかけはなれたものになってしまうでしょう。

なぜならば、文字の入力と印刷に特化した単機能コンピューターであるワープロ専用機では、たしかにJIS文字コードの異体字問題があらわになるのは入力や印刷の時だけかもしれません。しかしごく普通の、大多数をしめるパソコン・ユーザーにとってJIS文字コードの問題点に直面するのは、メールやウェブページなどインターネット上で情報交換をするときだからです。しかし、表外字体案では“インターネット”という言葉も“メール”という言葉も登場しません。そう、表外字体案が文字を語る際にゴッソリと欠落してしまっているのは、こうした“情報交換”という視点であるように私には思えます。

●“印刷文字”とは、どんな文字？

つぎに私がいただいた疑問は、表外字体案はどの範囲まで適用されるのだろうかということでした。

表外漢字字体表は、法令、公用文書、新聞、雑誌、放送等、一般の社会生活において、表外漢字を使用する場合の、字体のよりどころを、印刷文字（情報機器の画面上で使用される文字や字幕で使用される文字などのうち、印刷文字に準じて考えることのできる文字を含む。）を対象として示すものである。

(p.5 2 表外漢字字体表の性格 (1)表外漢字字体表の適用範囲)

そもそも、ここで書かれている“印刷文字”という概念は、いったい何を言い当てようとしているのでしょうか？

表外字体案を読むと、1,022文字の選定には、凸版印刷と読売新聞社などが印刷に使用した文字の頻度調査を規準にしたとあります (p.10 表外漢字字体表に掲げた漢字(表外漢字)の選定方法について)。してみると、この“印刷文字”とは、大規模な商業目的の印刷につかわれる文

字のことのようには思えます。

しかし、一方で表外字体案は〈一般の書籍や教科書などに用いられる〉字体が〈ワープロ等から打ち出せない〉(p.2 前出引用箇所) ことを問題にしています。ワープロを対象にしているのなら、それは商業目的の印刷のような大規模でプロフェッショナルな用途というよりも、個人が仕事に使ったり、あるいはごくプライベートな生活の中で使う文字のようにも思えます。

自分のパソコンの使い方を振り返っても思うのですが、私たちはかつて紙に鉛筆で文書を書いていたのと同じような感覚で、現在パソコンを使ってないでしょうか？ だとすれば、これは手書きの代替手段でもあると言えます。その手書きの代わりにつかう字体を、商業目的の印刷を対象とした頻度調査で決定できる根拠は、いったいどこにあるのでしょうか？

おそらく表外字体案は、こういう用途については簡易慣用字体を使えと言っているように読めます。しかしこの簡易慣用字体も選定の基準は、やはり同じ凸版印刷と読売新聞社などのデータであり、これら商業目的の印刷で使われている字体を個人の文字生活でつかうべき理由については、まったく語っていません。

本来ここで表外漢字の指針を決めようとする者に求められているのは、どのような字体が、どのような目的／場所で、どのようにして使われており、今後はどのようにあるべきかという、現実裏打ちされた確固とした指針をしめすことではないでしょうか。

たとえば、私たちは文字を書く目的や内容によって漢字の使用頻度が変わるのではないのでしょうか。そういう基本的なところから“文字生活”の姿が浮かびあがるのではないか、私にはそのように思えるのです。

実はそういった広範囲な調査を、私は表外字体案にもとめたかったのですが、残念ながら公表された案は(それじたいは大規模な調査だったとはいえ)、わずかに限られた分野での数種の頻度調査ですべてを決めようとしているようです。

誤解していただきたくないのですが、たとえばパーソナルな使用が主である情報機器では、筆写字体風のものを使うべきなどと言おうとしているわけではありません。こうした個人的な使用目的の字体と、大規模な商業目的の印刷における字体を、まったく同じものとして扱える理由がわからないのです。

個人から商業目的の印刷までにわたる網羅的な頻度調査ができなかったのならば、ぐっと表外字体案の適用範囲を限定するのも一案です。つまり“印刷文字”とは、個人ではなく大規模な商業目的の印刷においてつかわれる文字とするのです。それならば筋がとおるように思えます。

表外字体案では、『情報機器との関係』として〈将来的に文字コードの見直しがある場合、表外漢字字体表の趣旨が生かせる形での改訂が望まれる〉(p.9) としており、JIS 文字コードの変更をうながしています。しかし私が調査したかぎりでは、そのような変更をした場合、けっして小さくない混乱がおこりそうです。

文字コードの改訂など求めず、商業目的の印刷にだけ通用する限定的な技術に対して表外漢字字体表を適用することが、じつは現状では一番混乱がすくないスムーズな方法のように私は思います。

たとえば、プロフェッショナルなDTP市場におけるデファクト標準であるアドビシステムズの『CID フォント字体切り替え』という技術があります。これは、対応したアプリケーション・ソフトからある操作をすることで、選択した文字の字体を任意のものに切り換えられるというものです。

アドビシステムズは、この技術に対応するソフトを今後拡充していく方針ですし、アップルコンピュータも、これにOSレベルでの対応を発表しています。そして重要なことは、商業目的の印刷を舞台としたDTP市場において、この2社の組み合わせは圧倒的なシェアをもっているのです。

知人の調査によれば、表外字体案のうち、この技術で出すことができないのは3文字だけです(後述)。つまり、この3文字をのぞけば、わざわざ文字コードの改訂などしなくても、たった今、現実に印刷標準字体を使うことが可能になるのです。となれば、きわめて現実的な選択肢であると言えるのではないのでしょうか。

●包括的とは言えない『デザインの違い』

私が文字コードの改訂を求めてほしくない理由は他にもあります。それはJISの文字コード規格が表外字体案に対応したときに、混乱がおこると予想されるからなのですが、それを述べる前に、もうすこし別の角度から表外字体案の問題点を述べたいと思います。

まずあげられるのは、『表外漢字における字体の違いとデザインの違い』(p.34 以下、デザインの違い)が、あまり包括的とは言えないということでしょう。

たとえば、JIS X 0208/0213では『包摂規準連番1』として区別しない、“王”“壬(にん/ノ+土)”“壬(てい/ノ+土)”(図1)の3つの部分字体ですが、表外字体案では、この扱いが不徹底です。

『デザインの違い』では、【逞】(p.37, 4-A-(2)の9番目)が“王”と“壬(てい)”の包摂(図2)、【挺】(p.37, 4-A-(6)の3番目)が“壬(にん)”と“壬(てい)”の包摂(図3)と考えられます。

さて、それでは“王”と“壬(にん)”は包摂されるのでしょうか、それとも区別されるのでしょうか?“王” \leftrightarrow “壬(てい)”, “壬(にん)” \leftrightarrow “壬(てい)”を、別々に分類してあることから、おそらくは包摂しないのだろうという解釈も可能です。しかし、これはあくまで“解釈”にすぎず異論の余地を残してしまいます。

これにより、たとえば表外字体案460番の【閏】は、印刷標準字体として門構えに“王”の字体で収録されていますが、はたして門構えに“壬(にん)”も許容されるかどうか不明になっています。



図1(JIS X 0208:1997規格票p.14より)

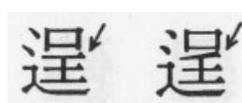


図2(表外字体案 p.37より)

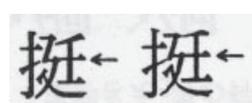


図3(表外字体案 p.37より)

ちなみに、幕末期から1946年にわたる23の明朝体活字の字形が一覧できる『明朝体活字字形一覧』（'99年 文化庁）によれば、この門構えに“王”と“壬（にん）”のふたつは併存しています（図4）。また、道光版康熙字典と大漢和辞典では両方の字体とも収録しています。

年	書	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1833年	康熙字典	1829年	1848年	1869年	1877年	1887年	1887年	1892年	1894年	1903年	1906年	1913年	1913年	1913年	1913年	1914年	1914年	1916年	1916年	1920年	1934年	1935年	1946年	1988年	
	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨
	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨	閨

図4『明朝体活字字形一覧』（p.561）より

●“いわゆる康熙字典体”とは、具体的にどんなもの？

しかし、これよりももっと影響力がおおきいと思われるのは、表外字体案に収録された1,022文字と、収録されなかった表外字一般の閨係が、きわめて曖昧であることです。表外字体案のなかで数少ない例としてふれられているのは、以下の箇所です。

なお、この字体表は、芸術その他の各種専門分野や個々人の漢字使用まで及ぶものではなく、従来の文献などの固有名詞に用いられている字体にまで及ぶものでもない。（p.5 2表外漢字字体表の性格(1)表外漢字字体表の適用範囲）

なお、表外漢字字体表に示されていない表外漢字の字体については、基本的に印刷文字としては「いわゆる康熙字典体」によることを原則と考える。

（p.6 2表外漢字字体表の性格(2)対象とする表外漢字の選定について）

たとえば、表外字体案が選定した1,022文字以外の文字をコンピューターでつかう際に、どうすればよいのでしょうか？ 1,022文字以外の文字のフォントを変更してよいのでしょうか、それとも変更しなくてもよいのでしょうか。この抽象的な記述からは分かりません。

同様に、『デザインの違い』が、表外漢字一般にも適用されるのか、それともこの表外字体案の1,022文字だけを対象にするのかについても不明です。ただ1カ所だけ、以下のよう記述はあります。

また、この字体表に掲げられていない表外漢字についても、この字体表に掲げられていない表外漢字についても、現に印刷文字として「亓/衤/飠」の字形を用いているものについては、3部首許容を適用してもよい。（p.13 字体表の見方 5）

この記述によれば、3部首だけは、表外字体案だけでなく明確に表外字一般に及ぶようです。しかし、これ以外の字体差や字形差は、どうなのでしょう？ 一部分だけが明確なだけにかえって分からなくなっています。こうして、表外字体案を読んだうえで表外字一般について起こしうる行動が、すくなくとも以下の4つになってしまいます。

- ①これからは表外字体案1,022文字以外を、〈芸術その他の各種専門分野や個々人の漢字使用〉以外には使わないことにする。
- ②これからは、“いわゆる康熙字典体”を表外字体案の1,022文字以外でもつかうために、『デザインの違い』のうち“デザイン差とされていない”違いを字体表から類推して表外字一般に当てはめ、字体をあらためる。
- ③これからは、“いわゆる康熙字典体”を表外字体案の1,022文字以外でもつかうために、『デザインの違い』のうち“デザイン差とされていない”違いでも無視し、より“いわゆる康熙字典体”と考えられる字体にあらためる。
- ④表外字体案1,022文字以外は無関係であるから、従来通りとして一切手を加えない。

ここであげた②と③の違いは、②が『デザインの違い』を表外字体案の1,022文字以外においても尊重しようとするのに対して、③が1,022文字以外には適用しないという違いです。前述したとおり、『デザインの違い』の適用範囲が表外字体案にかぎるのかどうか不明であるので、③のような解釈も十分に成り立ちうるわけです。

このようにして、表外字体案における最大の不明点が浮上します。それは、

“いわゆる康熙字典体”とは何なのか、具体的に明示されていない

ということです。表外字体案では、これについて以下のように述べてています。

印刷標準字体には、「明治以来、活字字体として最も普通に用いられてきた印刷文字字体であって、かつ、現在においても常用漢字の字体に準じた略字体以上に頻度高く用いられている印刷文字字体」及び「明治以来、活字字体として、康熙字典における正字体と同程度か、それ以上に用いられてきた略字体や俗字体などで、現在も、康熙字典に掲げる字体そのものではないが、康熙字典を典拠として作られてきた明治以来の活字字体（以下「いわゆる康熙字典体」という。）につながるものである。

(p.3 1 表外漢字の字体に関する基本的な認識 (2)表外漢字字体表作成に当たっての基本的な考え方)

なんとも晦渋で分かりづらい、悪文と言ってもさしつかえない表現ですが、ここでは印刷標準字体・簡易慣用字体と“いわゆる康熙字典体”の関係について、前者が後者にくつながらるもの〉という曖昧な表現しかしておらず、イコールのものとしては書いていません。印刷標準字体と簡易慣用字体は具体的に字体が明らかにされていますから自明ですが、ではこれらとも違うらしい、“いわゆる康熙字典体”とは、いったい如何なるものでしょう？ 不明である以上、当然利用者は恣意的な、それぞれの解釈にしたがって“いわゆる康熙字典体”を考え出して使うことにならざるをえないのです。

こうして、〈実態を混乱させないことを最優先に考えた〉(p.4, 8～9行目)はずの表外字体案が、結果として正反対の結果を招くものになりかねなくなるわけです。もし答申にいたっても上記の点について明確化されない場合、表外字体案の目論見とはうらはらに、かえって字体が乱立して大混乱になることでしょう。これは、絶対に絶対に避けねばならない事態です。

この表外字体案がもとめるものの曖昧さは、JIS文字コード改訂の際の対応にも影をおとすと思われます。それも、残念ながらけっして小さくない影をです。では具体的に、JISの文字コードにそって検討してみましょう。

●表外字体案の字体は、JIS X 0213ですべて表現可能

JISの文字コード規格は、現在4つあります。

- (a)JIS X 0208 (一番普及している“JIS漢字コード”)
- (b)JIS X 0213 (0208の拡張版)
- (c)JIS X 0212 (補助漢字)
- (d)JIS X 0221 (= ISO/IEC 10646 ≡ Unicode)

ここでは、今後おそらく改訂はないと思われる(c)の補助漢字をのぞいた、(a)(b)(d)の3つについて検討したいと思います。ただし(c)は(d)の部分集合として収録されていますので、けっして無関係ではありません。

まずJIS X 0208ですが、表外字体案の1,022文字であげられている印刷標準字体と簡易慣用字体は、739番の1字体をのぞいて、すべて包摂規準と『6.6.4 過去の規格との互換性を維持するための包摂規準』(以下、互換規準)の範囲内におさまります。ただ、このうち1005番としてあげられている印刷標準字体そのものの表現は不可能ですが、表外字体案では1005番の備考として、JIS X 0208の区点71-37で例示されているのと同じ字体をあげ、これと印刷標準字体のちがいをデザイン差として区別しないことがしるされています。

もっとも、JIS X 0208には収録されていない739番もJIS X 0213には面区点1-92-80にあります。また1005番の印刷標準字体そのものも面区点1-91-26にあります。(表1)

◆表1 表外字体案のうち、JIS X 0213だけで表現できる文字

表外字体案	JIS X 0213
鄧	鄧
739	1-92-80
藹	藹
1005	1-91-26[*1]

*3……ただし表外字体案では、JIS X 0208の【藹】71-37との違いは字体差ではないとされているので、正確には“表外字体案のうち、JIS X 0213だけで表現できる文字”には入らない。

したがって範囲を JIS X 0213 にひろげれば、表外字体案はいずれかの符号位置に包摂されますので、これに対応するのにわざわざ符号位置を変えるなどといった非互換な変更をする可能性はないと思われます。つまり、JIS X 0208 も JIS X 0213 も例示字体だけを、包摂規準の範囲内で変更して表外字体案に対応することになるのでしょう。

●“互換規準”の29文字のゆくえは？

さて、JIS X 0208にかぎって言えば、おおきな問題は以下のふたつに絞ることができると思います。

- ①表現不可能な表外字体案 739 番を追加するか否か。
- ②互換規準の 29 文字の対応をどうするか。

①については、前述のとおり JIS X 0208 の拡張である JIS X 0213 に 739 番があることから、わざわざ JIS X 0208 を変更してまで追加することもないでしょう。

一方②です。これについては少し説明がいきりそうです。そもそもの発端は'78年に制定された JIS X 0208 が、'83年に最初の改訂をうけた時にさかのぼります。この時おおくの非互換な変更がされ、それから長い間日本の文字コードには混乱がつづきました。その'83年の非互換な変更とは以下の3つです。

(i) 4 文字の例示字体が簡略化された字体に変更され、元の字体の文字がそのまま新たな符号位置をあたえられた。

(ii) 22 組の簡略字体と旧字体（つまり計 44 文字）が、第 1 水準と第 2 水準のあいだで符号位置を入れ換えられた。

(iii) 250 文字[*4]が字体・字形を変更された。

[*4]……『日本語情報処理』（Ken Lunde ソフトバンク '95年刊）379ページの『相違点4』による数え方。文字数は資料によってことなる。

まだパソコンが普及していない'83年当時の時代意識のなかで、'81年制定の常用漢字表とそれともなう人名用漢字別表の改正に対応しようとしたのが、この結果だといわれています。しかしこの'83年変更（以下、83JIS）は、のちにおおきな禍根を残しました。

この時よりも前からコンピューターを作っていたメーカーは、あまりにも多くの変更がされていて、従来の'78年版（以下、78JIS）を搭載した製品との間で互換性がとれなくなるとして対応しない一方で、'83年以降に作りはじめたメーカーは、83JISの方を搭載し、結果として両者のパソコンの間で文字化けが発生してしまったのです。

前者のメーカーの代表的なものに NEC があり、後者の代表にエプソン、アップルコンピュータがあります。つまり'80年代当時、圧倒的なシェアをもっていた NEC の MS-DOS

マシン PC-9801 と、その互換機エプソンの PC-386 シリーズの間で文字が化けました。

この混乱を収束しようとしたのが'97年におこなわれた JIS X 0208 の3回目の改正 (JIS X 0208:1997) です。この版では、符号位置と例示字体については一切の変更をおこなわず (i) と(ii)はこの時点で対処の方法なしということでしょう)、そのかわり“包摂規準”を明文化することにより、JIS X 0208の中では(iii)で変更される前と後の字体・字形を一定の枠の内におさめ、78JIS と 83JIS を同じ文字として区別しないことにしました。

千変万化の漢字の特性をよく知悉した、おそろしくクレバーな対応策だと思います。包摂規準が明示された結果、それまで曖昧だったデザイン差、字形差、字体差が明確になり、JIS X 0208 の解釈・運用が簡単になったことも特筆されます。

しかし、それでも 29文字が包摂規準の枠内におさまらないものとして残ってしまいました。これをどうにかしようとしたのが件の『互換規準』です。

この互換規準では、変更された83JISの例示字体を(A)とし、これとは別に元の78JISの例示字体を(B)として2つのセットにわけ、(A)セットの字体を規格票の例示字体とします(つまり(A)がデフォルト)。そのうえで、(A)と(B)のどちらを使っても規格に適合する、ただしどちらを使ったか明示せよ、としたのです。

この規定により、ようやく'83年以降の版と'78年版との互換性が維持できるようになるということで『過去の規格との互換性を維持するための包摂規準』と名づけられたわけです。

この互換規準の29文字は、変更の前後で特に字体の変化がおおきかったため、一部では“JISのウソ字”とまで言う人まであらわれました。この29文字は従来からあった字体をJISが収録しただけで、別にJISで創作したわけではないのですから、彼らの責任にするのも可哀想ですが、結果としてJIS文字コードを広く世の中に知らしめる結果になったのは皮肉なことでした。

さて、話はようやく表外字体案にもどります。ここでも、この互換規準の29文字は強く意識されています。冒頭『(1)従来の漢字施策と表外字体問題』で〈ワープロ等から打ち出せない〉として取り上げられている、“鷗”“禱”“瀆”などの旧字体は、すべてこの29文字のうちですし、ですから〈打ち出せない〉旧字体は、すべてJIS X 0208でいう(B)のセットの内です。

そんなわけですから、当然29文字はみんな表外字体案に選定されています。その印刷標準字体はすべて78JISの字体とする一方で、うち12文字は83JISの字体を簡易慣用字体として認めています。

ところで、JIS X 0213では、78JISの(B)の字体をそのまま収録しています。つまりJIS X 0213を使えば、(A)と(B)の両方が使用可能になるわけです。これらの関係を表にすると、以下の表2のようにまとめられます。

◆表2

JIS X 0208:1997の『互換規準』に指定された文字と表外字体表の関係

(JIS X 0208では包摂,JIS X 0213で新たに収録された文字。《29文字》)

表外字体案		JIS X 0213 (78JIS)	JIS X 0208 (83JIS)
印刷標準	簡易慣用		
啞	啞	啞	啞
1		1-15-08	16-02
焰	-	焰	焰
60		1-87-49	17-75
鷗	鷗	鷗	鷗
75		1-94-69	18-10
搨	-	搨	搨
123		1-84-89	36-47
麴	麴	麴	麴
184		1-94-79	25-77
俠	-	俠	俠
213		1-14-26	22-02
頰	-	頰	頰
222		1-93-90	43-43
軀	-	軀	軀
234		1-92-42	22-77
齷	齷	齷	齷
275		1-94-74	24-20
嚙	嚙	嚙	嚙
336		1-15-26	19-90
攢	-	攢	攢
375		1-85-06	58-25
繡	繡	繡	繡
452		1-90-22	29-11
蔣	蔣	蔣	蔣
485		1-91-22	30-53
醬	醬	醬	醬
493		1-92-89	30-63
蟬	-	蟬	蟬
570		1-91-66	32-70

表外字体案		JIS X 0213 (78JIS)	JIS X 0208 (83JIS)
印刷標準	簡易慣用		
搔	搔	搔	搔
592		1-84-86	33-63
驪	-	驪	驪
628		1-94-20	34-45
箎	-	箎	箎
653		1-89-73	35-29
填	-	填	填
715		1-15-56	37-22
顛	-	顛	顛
717		1-94-03	37-31
禱	禱	禱	禱
745		1-89-35	37-88
瀆	-	瀆	澆
748		1-87-29	38-34
囊	-	囊	囊
764		1-15-32	39-25
潑	-	潑	澆
792		1-87-09	40-14
醜	-	醜	醜
793		1-92-90	40-16
麵	麵	麵	麵
920		1-94-80	44-45
萊	-	萊	菜
951		1-91-06	45-73
屢	-	屢	屢
981		1-47-64	28-40
蠟	蠟	蠟	蠟
1008		1-91-71	47-25

※表中“78JIS”の字体は『Illustrator』の『CID フォント字体切り替え』による“78JIS”を、

“83JIS”の字体は同“標準”の字体を、平成明朝 W3 によって再現した。以下の表も同じ。

これら 29 文字を JIS X 0208 でどう対応するかですが、私は理論上はノータッチになると理解しています。先日あるセミナーに出席した折り、講演者が 78JIS の字体に例示字体を変更するのではないかという推測をしていました。しかし、この見方はあたらないと思います。

原則として JIS X 0208:1997 では(A)と(B)は別の文字です。あくまで(A)を符号化しており、(B)の方はいわば例外措置と考えるべきです。となれば互換規準を廃止して、例示字体を(A)から(B)に変えるのは“文字を変える”わけですから、これは非互換な変更になってしまいます。

さらに JIS X 0208 と一体で運用するのが前提である JIS X 0213 ではすでに(B)を収録済みですから、そこに JIS X 0208 が(B)を例示字体にすれば整合性がとれなくなりなります。こうしたことから、この 29 文字の例示字体変更はありえないと思います。つまり、これらを表外字体案の印刷標準字体で使いたければ、JIS X 0213 を使って、ということです。

● JIS X 0208のうち、73文字の字体変更が必要

つぎに、JIS X 0208 で例示字体を変更すると思われる字体をあげます。便宜上、表を 2 つに分けます。(1)が 83JIS で字体・字形変更をした文字の中で今回も変更しなければならないと思われる文字(表 3)、それから(2)が 83JIS の時には変更されなかった文字の中で今回変更すると思われる文字(表 4)です。

表 3 の文字は、もし JIS X 0208 が表外字体案に対応すると、83JIS について、実に 3 度目の変更となります。

もっとも、例示字体が変更されたとして、これはあくまで包摂規準の範囲内であり、別にフォントの字形を印刷標準字体に変えないとしても、規格に不適合とはならないということは知っておくべきでしょう。

◆表3

JIS X 0208が表外字体案に対応すると字体変更が必要な文字①

(83JISで変更され、表外字体案で78JISの字体が採られた文字)《66文字》

JIS X 0208	表外字体案 印刷標準字体
溢	溢
16-78	26
鰯	鰯
16-83	27
淫	淫
16-92	31
冤	冤
49-45	55
襖	襖
18-08	71
晦	晦
19-02	109
葛	葛
19-75	132
翰	翰
20-45	153
翫	翫
20-69	160
徽	徽
21-11	177
厩	厩
17-25	201
卿	卿
22-10	217
僅	僅
22-47	229
繫	繫
23-50	256
倦	倦
23-81	265
捲	捲
23-94	268
諺	諺
24-33	279

JIS X 0208	表外字体案 印刷標準字体
巷	巷
25-11	305
薩	薩
27-07	369
櫛	櫛
22-91	417
酋	酋
29-22	441
薯	薯
29-82	464
藪	藪
29-83	465
哨	哨
30-05	470
秤	秤
39-73	471
蛸	蛸
34-93	481
摺	摺
32-02	483
鞞	鞞
30-68	487
鏞	鏞
27-12	539
鯖	鯖
27-10	541
漚	漚
38-52	542
屑	屑
22-93	550
詮	詮
33-07	562
煽	煽
32-90	565

JIS X 0208	表外字体案 印刷標準字体
撰	撰
32-81	567
瘦	瘦
33-73	596
噌	噌
33-25	598
甌	甌
25-89	602
樽	樽
17-29	615
樽	樽
35-14	616
鱒	鱒
43-80	617
灘	灘
38-71	655
箸	箸
40-04	674
儲	儲
44-57	675
瀦	瀦
35-85	676
抄	抄
36-29	688
鄭	鄭
37-02	703
擢	擢
37-07	709
溺	溺
37-14	710
堵	堵
37-40	727
屠	屠
37-43	728

JIS X 0208	表外字体案 印刷標準字体
賭	賭
37-50	729
剥	剥
39-77	782
謬	謬
41-21	828
屏	屏
54-02	832
瀕	瀕
41-46	835
蔽	蔽
42-35	861
瞥	瞥
42-45	868
庖	庖
42-89	884
鞞	鞞
19-83	890
粿	粿
44-66	924
愈	愈
44-92	931
檣	檣
38-74	935
猷	猷
45-18	936
煉	煉
46-91	987
榔	榔
47-17	1003

◆表4

JIS X 0208が表外字体案に対応することで字体変更が必要な文字②
 (83JISで変更されなかったものの中で、表外字体案により変更が必要な文字)《7文字》

JIS X 0208	表外字体案 印刷標準字体
嘘	嘘 _{..}
17-19	206
饗	饗 _{..}
22-34	225
妍	妍 _{..}
53-11	264
歎	歎 _{..}
35-23	651
叛	叛 _{..}
40-32	800
并	并 _{..}
54-85	857
簾	簾 _{..}
46-92	990

マークの説明

#……Macintosh版『Illustrator』(アドビシステムズ)ver.5.5J以降に搭載されている『CIDフォント字体切り換え』の“Expert”を選択すると表現可能な字体。

##……『CIDフォント字体切り換え』を使用しても印刷標準字体を表現できない文字。ちなみに、JIS X 0213特有の字体である表1はともかくとして、表2と3の78JISの字体は“78JIS”，そして簡易慣用字体は標準の字体で再現可能あるから、JIS X 0208収録の文字で表外字体案で表現できないのは、この表4にあげた3文字だけということになる(直井靖氏のご教示による)。

●おおきな混乱を生む“いわゆる康熙字典体”

さて、ここまでは JIS X 0208 の表外字体案への対応として順当であると考えられるところですが、おそらくは、JIS としての対応は、ここまでにとどまるでしょう。

しかし、問題なのはフォント・ベンダーの“解釈”です。さきに表外字体案でく表外漢字字体表に示されていない表外漢字の字体については、基本的に印刷文字としては「いわゆる康熙字典体」によることを原則[*5]と書かれている一方で、肝心の“いわゆる康熙字典体”なるものの具体像が示されていないために、表外字体案以外の表外字一般にたいする行動として、4つも選択肢ができてしまうと述べました。以下にその4つを再掲します。

[*5]……p.6/2 表外漢字字体表の性格(2)対象とする表外漢字の選定について

①これからは表外字体案1,022文字以外を、く芸術その他の各種専門分野や個々人の漢

字使用〉以外には使わないことにする。

②これからは，“いわゆる康熙字典体”を表外字体案の1,022文字以外でもつかうために、『デザインの違い』のうち“デザイン差とされていない”違いを字体表から類推して表外字一般に当てはめ、字体をあらためる。

③これからは，“いわゆる康熙字典体”を表外字体案の1,022文字以外でもつかうために、『デザインの違い』のうち“デザイン差とされていない”違いでも無視し、より“いわゆる康熙字典体”と考えられる字体にあらためる。

④表外字体案1,022文字以外は無関係であるから、従来通りとして一切手を加えない。

このうち、問題は②と③です。不明である以上、当然利用者は恣意的な、それぞれの解釈にしたがって“いわゆる康熙字典体”を考え出して使うことにならざるをえないわけです。表外字体案のような“標準”に、曖昧さがあるてはいけないと考えます。ベンダーによって解釈の差が出るようなことがあれば、それは混乱につながります。

たとえば、以下の表の文字などは、表外字体案には収録されていない文字ですが、83JISで比較的はっきりとした字体変更がされている文字の一例です。こういった文字は、まっさきに“いわゆる康熙字典体”を類推してあてはめる候補にあげられるのではないのでしょうか。

◆表5

JIS X 0208のうち，“いわゆる康熙字典体”を表外字体案以外の文字に類推して当てはめようとした場合、真っ先に影響が予想される文字の一例

(83JISで字形変更された全250文字から、表外字体案に収録されていない文字を取り出し、これから(1)常用漢字とその異字、(2)人名用漢字[許容字体表をふくむ]とその異体字(3)表外字体案の異体字、(4)表外字体案のデザイン差によって包摂される文字、(5)『明朝体活字字形一覧』で78JISの字体が優勢でない文字をはずしたもの)《13文字》

規格	78JIS	83JIS
字体	廠	廠
区点位置	30-19	
字体	鱈	鱈
区点位置	35-13	
字体	唳	唳
区点位置	51-26	
字体	捩	捩
区点位置	57-64	
字体	柺	柺
区点位置	59-42	
字体	湮	湮
区点位置	62-48	
字体	甄	甄
区点位置	65-11	

規格	78JIS	83JIS
字体	綰	綰
区点位置	69-25	
字体	綰	綰
区点位置	69-38	
字体	葱	葱
区点位置	72-27	
字体	蔗	蔗
区点位置	72-84	
字体	蟒	蟒
区点位置	74-29	
字体	鬪	鬪
区点位置	82-13	

ここであげた文字はあくまでも一例です。そもそも私が自分の作業を容易にするために、83JISで変更した250文字を出発点にして絞っていった訳ですから、変更されなかった数千もの文字に“いわゆる康熙字典体”を類推適用される可能性を考えねばなりません。そうなれば、もっともっとうこうした変更候補の文字の数は増えるでしょう。

こうした混乱をまねかないためにも、私は表外字体表の適用範囲を厳密にそれ自身に限定することを盛り込むよう提案したいと思います。つまり、上記の選択肢でいえば、(4)です。

●Unicodeに致命傷をあたえる表外字体案

最後に、冒頭あげた JIS 文字コードのうち、(d)JIS X 0221 (= ISO/IEC 10646 ≡ Unicode) を取り上げましょう。実は、これが一番深刻なのです。

◆表6

JIS X 0208が表外字体案に対応すると、JIS X 0221との変換表作成に問題が出そうな文字

表外字体案		JIS X 0208	JIS X 0221				
印刷標準	簡易慣用		J-0	J-1	C-B	C-T	K
繫	-	繫	繫 0-2350 7E4B	繫 1-525B	★	★	★
250		23-50	7E0B				
表外字体案		JIS X 0208	JIS X 0221				
印刷標準	簡易慣用		J-0	J-1	C-B	C-T	K
瘦	瘦	瘦	瘦 0-3373 75E9	瘦 1-4587	★	★	★
590		33-73	7B20				
表外字体案		JIS X 0208	JIS X 0221				
印刷標準	簡易慣用		J-0	C-T	C-T	K	
妍	-	妍	妍 0-5311	妍 1-418D	妍 E-1211	★	
204		53-11	59B0		59F8		
表外字体案		JIS X 0208	JIS X 0221				
印刷標準	簡易慣用		J-0	C-T	C-T	K	
剥	-	剥	剥 0-3077	剥 E-1507	剥 1-5277	★	
782		39-77	52B5		525D		
表外字体案		JIS X 0208	JIS X 0221				
印刷標準	簡易慣用		J-0	C-T	C-T	K	
屏	屏	屏	屏 0-5402	屏 1-4B7B	屏 E-2135	★	
832		54-02	5C4F		5C5B		
表外字体案		JIS X 0208	JIS X 0221				
印刷標準	簡易慣用		J-0	C-T	C-T	K	
并	并	并	并 0-5485	并 1-3904	并 E-08B7	★	
857		54-85	5E70		5E77		

☆……規格票には記載があるが、意図を明確にするため字形を掲載しない文字
JIS X 0221 欄の中の記号は以下の規格であることをしめす

J-0…JIS X 0208

J-1…JIS X 0212(補助漢字)

C-G…中華人民共和国の繁体字の規格等

C-T…TCA-CNS-11643(中華民国の規格)

K …KS C 5601(大韓民国の規格)

解説しましょう。本来が各国の文字コードの集合体である JIS X 0221 (= ISO/IEC 10646 ≡ Unicode) では、制定当初から収録されている規格を“原規格”といいます。日本の文字コード規格としては JIS X 0208 と JIS X 0212 (補助漢字) が、これにあたります。また JIS X 0221 では、漢字について、同じ字体として考えられるものは一つの符号位置に統合しています。これを“統合漢字”といいます。

しかし、実際に JIS X 0221 を実装する際には、以前からある文字コードとの間で変換して使うために、これらとの間で交互に変換しても、最初の符号位置が変わらないことが求められます。これを“ラウンド・トリップ・コンバージョン”といいます。“ラウンド・トリップ・チケット”は往復切符のことですから、往復旅行が保証されるというほどの意味でしょうか。

さて、これを確保するために、JIS X 0221 では、“原規格分離漢字の取扱い規則”（詳細は JIS X 0221 の解説 p.883 以降を参照）とよばれるルールをもうけています。これは、本来は JIS X 0221 自身の統合規則によって統合される字体がある場合でも、必ず分離され、双方の字体とも違う符号位置に収録することが保証されるということです。これにより、例えば JIS X 0208 と JIS X 0212 については、JIS X 0221 との間で相互に変換を繰り返しても、元の符号位置のままで、しかも JIS X 0208 と JIS X 0212 を区別して JIS X 0221 を使用できることが保証されるわけです。

これは各国の複数の規格から成り立つ JIS X 0221 を、従来からあるシステムと互換性を保ちながら使用するためのいわば“生活の知恵”なのですが、では例えば JIS X 0208 の例示字体が変更されたらどうなるか、というのが、この表 6 なのです。

JIS X 0208 では区点 23-50 の【繫】ですが、これを表外字体案のとおり例示字体を変更すると、補助漢字の字体と衝突してしまいます。しかも“原規格分離漢字の取扱い規則”により、符号位置を統合することはできませんから、同じ字体が2つの符号位置をもつという、文字コードでは一番忌むべき“重複符号化”をまねいてしまいます。

JIS X 0208 から JIS X 0221 に【繫】を変換する際、JIS X 0208 出身の符号位置である 7E4B に変換されればよいものの、もしも補助漢字出身の 7E6B に変換されてしまえば、再度 JIS X 0208 に変換する際、たとえば補助漢字を扱えないシフト JIS を使っていれば文字化けになってしまいます。他に【瘦】も同じケースです。こうなると、規格としての JIS X 0221 は致命傷を負ったこととなります。

別のケースもあります。例えば JIS X 0221 の 598D にある【妍】ですが、もう1つよく似た字体が 59F8 としても割り当てられているのです。これは中華民国の“原規格”(C-T 欄)

が、微妙に異なる字体を2つ収録しており、前述の“原規格分離漢字の取扱い規則”により統合されず、そのまま2つの符号位置をあたえられたものです。JIS X 0208の字体は、前者の598Dに割り当てられています。

ところが、ややこしいことに、表外字体案の印刷標準字体は、後者の59F8と同じなのです。まさかJIS X 0208で例示字体を後者の字体に変更しても、JIS X 0221が符号位置を変える非互換変更をすることは考えられません。しかし、基本的に一定のルールにより同じと見なされる字体が1つの欄(符号位置)に統合されていたのに、表外字体案への対応によって、“字体は別の符号位置と同じなのに、違う字体の符号位置に割り当てられている”という混乱が発生します。同様のケースとして【剥】【屏】【并】があげられます。

こうなると、JIS X 0208からJIS X 0221に変換する際、“原規格の符号位置に注目した結果”と“例示字体に注目した結果”で、変換先の符号位置の解釈が2つに分かれてしまうこととなります。これまた大混乱です。

つまり、表外字体案はこの表6の文字によって、実装がすすみつつあるJIS X 0221(Unicode)に避けようのない致命傷をあたえるものになるのです。こういった事態を、はたして国語審議会の皆さんは認識されているのでしょうか？

●むすび

さて、結論です。以上みてきたように表外字体案がそのまま答申され、これにJIS文字コードが対応した場合、少なからぬ混乱がおこると思われまます。そうならぬためには、以下のような修整が必要と考えまます。

- (a)表外字体表の適用範囲は、字体表に収録された文字だけであることを明記する。表外字を使うにあたっては、なるべく表外字体表にある文字を使うよう推奨する一方で、これ以外の表外字一般には、表外字体表の効力は及ばないとする。
- (b)『デザインの違い』が、より包括的になるよう修整する。
- (c)明確な表現を本旨とし、あいまいな解釈をうむ“いわゆる康熙字典体”に関する記述を一切削除する。規範とされるべきは字体表にある字体、それに『デザインの違い』の2つだけであり、これらについての必要最低限の説明が付属すれば十分に運用は可能なはずだ。
- (d)情報機器に搭載される表外字の字体については、表外字体表の趣旨が活かされることを希望する一方、その具体的な実現手段としてJIS文字コードの改訂はもとめない。

以上、よろしくご検討ください。

敬具

2000年11月1日

小形克宏拝

※本稿の執筆にあたり、直井靖氏、豊島正之氏より貴重な示唆をいただいた。また、表6として掲載したリストは、小池和夫氏がメーリングリスト『Font-G』(<http://www.l-h.co.jp/lhcontents/maillinglist.html>)に発表したものにもとづく。各氏のご教示に深く感謝いたします。もちろん本稿の文責は私一人に属し、もしもこの原稿に誤りがあっても、すべて私の責任に帰する。

● PDF 版公開のためのノート

このPDFファイルは、『小形克宏の「文字の海、ビットの舟」』の特別編7～9「国語審議会への手紙」(全3回)として書かれたものを、再編集して作成したものである。原文のタイトルおよび掲載日、URLは以下の通りだ。

特別編7 国語審議会への手紙(上) 表外漢字字体表案に欠けている視点とは？(2000年10月28日)
<http://www.watch.impress.co.jp/internet/www/column/ogata/special7.htm>

特別編8 国語審議会への手紙(中) 表外字体案は私たちに何を求めているのか？(2000年10月31日)
<http://www.watch.impress.co.jp/internet/www/column/ogata/special8.htm>

特別編9 国語審議会への手紙(下) 表外字体案への対応がまねくJIS文字コードの混乱(2000年11月7日)
<http://www.watch.impress.co.jp/internet/www/column/ogata/special9.htm>

当初2000年11月1日を期限とする国語審議会のパブリックコメントに応募するため、早めに原稿を完成して読者の皆さんにレビューしていただき、修整した上で応募しようと思論んだのだが、生来の遅筆のゆえ最初の2回分がかろうじて応募期限前に掲載できたのみだった。3回目の原稿の脱稿は11月1日深夜。それ以前に文化庁に電話して、2日の朝早くなら間に合うことを確認して(私はこういうことには聡い)、どうにか間に合わせたのだった。

そのようなわけで、このPDFファイルはまことに遺憾ながら、応募原稿そのものではない。ただし、本文は完全な間違いをのぞいてはそのままである。まず原文と応募原稿の違いをのべる。

①原文の特別編7～9をひとつにまとめ、文字コードの仕組みを説明した部分を中心に、冗長とおもわれる表現をばっさり削除。

②そのうえで、首尾一貫した文章になるよう各部分を細かく調整。

次に、応募原稿と本PDFファイルの違いをのべる。これ以外にも拙い本文に手を入れようと思ったが、それも未練たらしいのであえてしなかった。

①“鷗”と“鷗”等、テキストファイルでは表現できない字体の違いを、『CIDフォント字体切り換え』をつかって再現。

②ウェブで使った表を取り込んだ(作成していただいた編集担当者の永沢茂さん、ありがとうございます)。それにともない若干の語句を修整した。

③p.4の下半分、本来は「“壬(てい)”」と書くべきところを「“壬(にん)”」としていた誤りを修正。

以上、不完全な原稿で応募した自らの至らなさを恥じ入るばかりである。

(2000年11月7日記)