「電子行政における文字コードと外字について」に反論する

安岡孝一*

貴誌『行政＆情報システム』Vol.49 (2013 年 10 月号) に掲載された榎並利博氏の「電子行政における文字コードと外字について」という記事を、拝読いたしました。私個人は、第 3・第 4 水準漢字を含む JIS X 0213 という規格に 15 年近く携わっており、その流れで、第 1・第 2 水準漢字を含む JIS X 0208 という規格の改正もおこなっているのですが、その私から見ると、榎並氏の本記事は第 1〜第 4 水準漢字に関して、あまりに大きな誤解を含んでいます。他の読者に対する注意喚起の意味も込めて、この文章をお送りする次第です。

榎並氏の本記事において、もっとも論旨がおかしいと考えられるのは、以下の部分です (pp.64-65)。

> 我が国における「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を示すもの」として、文部科学省文化庁から常用漢字表が公開されている。この字種は 2,136 字であり、表外漢字字体表を合わせても、せいぜい 3,000 字である。すなわち、一般の社会生活においては約 3,000 字があれば十分であるとされている。
>
> 結論として、現在日本語を扱うことができるすべてのコンピュータに搭載されている JIS 第 1 水準と JIS 第 2 水準の範囲内だけで社会的には十分であり、それ以上漢字を増やしていくことは、漢字に対する誤認識や無理解を助長するだけであり、社会的利益にはならない。

「一般の社会生活においては約 3,000 字があれば十分」という榎並氏の主張が、仮に正しいと仮定したとしても、その「3,000 字」が第 1・第 2 水準漢字に納まっているとは限りません。実際、常用漢字表 2,136 字[1]の中でも、叱、塡、剝、頬、の 4 字は第 3 水準漢字です[2]。ましてや、表外漢字字体表 1,022+22 字[3]ともなると、鷗や鄧や沪など 42 字が 第 3 水準漢字です。少なくとも、榎並氏がここで主張している「3,000 字」は、第 1・第 2 水準漢字から現実にハミ出ています。

また、社会生活に使う「3,000 字」は、人それぞれで異なっています。たとえば、富山市には冨居という地名があるので、そこに住む人たちの「3,000 字」には「冨」と「富」が含まれているでしょう。あるいは、笠岡市には大宜という地名がある[4]ので、そこに住む人たちの「3,000 字」には「宜」と「宜」が含まれているかもしれません。そのように、人それぞれで「3,000 字」が異なっているのですから、日本中の人々の「3,000 字」を集めたならば、もちろん 3,000 字や 6,000 字に納まることなど無くて、当然、もっと大きな文字集合が必要となります。ところが、榎並氏の提案にしたがうと、「冨」と「富」は使ってもよいが、「宜」は使ってはダメで、でも「宜」は使ってよい、というワケの分からない結果となってしまいます。

そもそも「宜」は、本来は第 2 水準漢字に収録されるべきであったところ、1974 年ないし 1975 年に起こった作業上のミスで、収録から漏れてしまった漢字なのです。『国土行政区画総覧』の漢字は全て収録する、というのが、第 1・第 2 水準漢字の設計方針であり[5]、笠岡市大宜も、当時すで

---

*京都大学人文科学研究所附属東アジア人文情報学研究センター

1) 常用漢字表, 平成 22 年内閣告示第 2 号 (2010 年 11 月 30 日).

2) 安岡孝一: 新常用漢字表が迫る Unicode 移行、「シフト JIS」では対応不可能, 日経 ITpro (2009 年 12 月 11 日).

3) 表外漢字字体表, 国語審議会答申 (2000 年 12 月 8 日).

4) 安岡孝一, 安岡素子: ぽすたるガイド '97 に見る辞書にない漢字, 情報処理学会研究報告, Vol.97, No.108『人文科学とコンピュータ研究報告 No.36』 (1997 年 11 月 15 日), pp.1-6.

5) 笹原宏之: 「JIS X 0208」における音義未詳字に対する原典による同定—『標準コード用漢字表 (試案)』と『国土行政区画総覧』—, 国語学研究と資料, 第 20 号 (1996 年 12 月), pp.1-17.

に『国土行政区画総覧』に載っていたからです。第1・第2水準漢字の開発は、その大半が手書きの紙カードを用いておこなわれたものなので、ミスが発生してしまうのは仕方のないことだったでしょう。

「亘」が第1・第2水準漢字から漏れてしまっていることは、かなり早くから指摘されていました。日本IBMは、1979年8月発表のIBM日本語文字セット[6]に、「亘」を追加収録しました。榎並氏が所属する富士通は、1987年8月発表のJEF拡張漢字(第3版)[7]に、「亘」を追加収録しました。1988年6月発表の日本印刷産業連合会要望拡張漢字表5,843字は、やはり「亘」を含んでおり、JIS X 0212という「補助漢字」規格[8]になりました。

これらの指摘に対し、私を含むJIS X 0213の開発委員は、1995年から2004年にかけて膨大な作業をおこない、過去にミスで漏れてしまった漢字を、優先的に第3・第4水準漢字に収録しました[9]。もちろん「亘」も、第3水準漢字に収録しました。また、この作業中に、子の名づけに用いることのできる漢字(いわゆる人名用漢字)の一部が、第1・第2水準漢字から漏れていることが判明し、それらを第3水準漢字に収録した結果、現在は人名用漢字861字[10]のうち、105字が第3水準漢字となっています。第3・第4水準漢字は、本来、第1・第2水準漢字に含まれるべきであった漢字を、サルベージする過程において開発されたものであり、現代社会においては、第1〜第4水準漢字は一体運用されるべきものなのです。

したがって、榎並氏の「JIS第1水準とJIS第2水準の範囲内だけで社会的には十分」という結論は、一般には全く首肯しがたいものである、ということが言えます。第1・第2水準漢字だけでは、もはや常用漢字も人名用漢字も十分には書けず、現代社会においては不十分だということです。行政の場において使う漢字を減らしたい、という榎並氏の考えは理解できなくもありませんが、しかし、その結論は正しくない、と断言いたします。

大亘納涼大会にて(2009年8月8日撮影)[11]

[6]小形克宏: Windows外字とその互換性をめぐる争い, 歴博, No.180 (2013年9月), pp.7-11.

[7]富士通株式会社: FACOM JEF文字コード索引辞書, 第3版 (1987年8月).

[8]田嶋一夫: JIS漢字補助集合案の設定と今後の課題, 情報処理学会研究報告, Vol.89, No.13『情報学基礎研究会報告No.12』(1989年2月20日), 12-1.

[9]豊島正之: 文字の符号化—新JIS漢字第3・第4水準の開発から見た—, 京都大学大型計算機センター第64回研究セミナー報告 (2000年3月24日), pp.3-18.

[10]戸籍法施行規則の一部を改正する省令, 平成22年法務省令第40号 (2010年11月30日).

[11]http://www.city.kasaoka.okayama.jp/site/kabutogani/kabuni7.html