国立国語研究所学術情報リポジトリ

〈受賞紹介〉 文字コードの標準化：ISO/IEC 10646の開発

メタデータ	言語: Japanese 出版者: 国立国語研究所 公開日: 2015-10-30 キーワード: 作成者: 高田, 智和 メールアドレス: 所属:
URL	https://doi.org/10.15084/00000685

大学共同利用機関法人人間文化研究機構
国立国語研究所
National Institute for Japanese Language and Linguistics

国語研プロジェクトレビュー　NINJAL Project Review
No. 6 pp. 43–45 (October 2011)

〈受賞紹介〉

情報処理学会情報規格調査会では，情報処理に関する国際・国内規格の審議やそれに関わる調査研究を行っています。標準化貢献賞は，所属委員会委員として顕著な貢献のあった方々の中から選ばれる賞です。高田氏は，文字情報の整理，グリフ開発，文字符号標準化に関わる貢献が認められ，2010 年度に受賞しました。

文字コードの標準化
——ISO/IEC 10646 の開発——

高田　智和
国立国語研究所　理論・構造研究系　准教授

ISO/IEC 10646 は，世界中の文字・記号を情報機器で扱うための国際文字コード規格です。1993 年に最初の版 (ISO/IEC 10646-1: 1993) が制定され，以降，追補・改定を重ね，2011 年 3 月には最新版 (ISO/IEC 10646: 2011) が制定されました。追補・改定のたびに収録文字を増やし，現在では 10 万字を超える巨大文字セットになっています。

ISO/IEC 10646 の大部分を占めているのは漢字です。CJK 統合漢字と呼ばれる漢字領域は（図 1），最新版の規格では，拡張 C 領域，拡張 D 領域を収録するに至っています。拡張 C 領域，拡張 D 領域ともに，日本から提案された文字も含まれています。

図 1　CJK 統合漢字（部分）

CJK 統合漢字拡張 C 領域には，日本製漢字（いわゆる「国字」）が提案されました。標準化の審議を経て，日本から提案された「国字」は，最終的に 367 字が ISO/IEC 10646 に収録されました。

この 367 字の中には，日常生活では，ほとんど見たこともないような文字が含まれています。例えば，偏「哥」，旁「舞」につくる「𠮷 (2A8A6)」は，「契情お国𠮷妓（けいせいおくにかぶき）」（享保 15［1730］年刊，浮世草子），「酒迎御馳走𠮷妓（さかむかえごちそうかぶき）」

（明和 8［1771］年初演，歌舞伎脚本）のように，江戸時代の文芸作品の題目に用いられます。「哥舞」の二合字（二文字を組み合わせて作った文字）と考えられる文字です。

「𠮷」は，江戸文学を研究したり，図書館が目録を作成したりするときには，現代でも十分に使われる可能性のある文字です。例えば，国立国会図書館の蔵書検索では，「契情お国哥舞妓」と一文字ではなく二文字「哥舞」で表示されます（図 2）。「𠮷」が標準的な文字コードで表現できないために，「哥舞」で代用しているとみられます。所蔵原本（請求記号 196-161）の内題（目録題）では，一文字の「𠮷」を確認することができます（図 3）。

請求記号	196-161
タイトル	契情お国哥舞妓 5巻
責任表示	其磧
責任表示	自笑

図 2　NDL-OPAC の検索結果画面（一部）

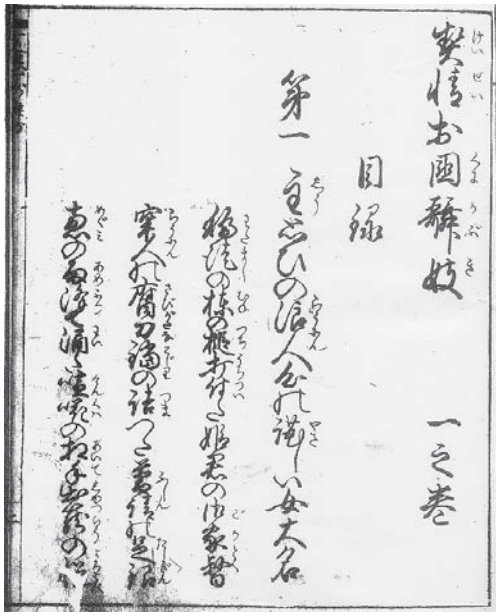

図 3　国立国会図書館蔵『契情お国哥舞妓』巻一内題

また，CJK 統合漢字拡張 D 領域には，各国・地域で緊急に必要とされる漢字が提案されました。日本からは，行政の情報処理で使われる漢字の中で，特に緊急度が高いものを選んで提案しました。このときには，汎用電子情報交換環境整備プログラム（平成 14 ～ 20 年度，国立国語研究所・情報処理学会・日本規格協会）の成果を利用して文字の選定を行い，最終的に 107 字が ISO/IEC 10646 に収録されました。

拡張 D 領域に収録された 107 字は，いずれも地名や人名に使われる文字です。例えば，偏「金」，旁「当」につくる「鍓 (2B7F0)」は，北海道弟子屈町の地名「鍓別（とうべつ）」に使われます。「鍓」は「鐺」の略体です。

2008 年 9 月に弟子屈町を訪れた折，街路の表示には，「鍓」「鐺」の双方が使われているのを見ました（図 4，図 5）。しかし，公式の字名表記は「鍓別」であるため，住民票や公図などの公文書では「鍓」を用いる必要が生じます。そのため役場では，パソコン用の外字（標準的に用意されている文字以外で，ユーザが特別に追加した文字）を作成して，日々の業務を行っています（図 6）。住民の転出などにより，弟子屈町以外の自治体と公文書を電子的にやりとりする場合に不便が生じると，困っている現状を役場の方はうったえておられました。

図 4　バス停留所

図 5　河川標示

図 6　弟子屈町役場パソコン外字

ISO/IEC 10646 に追加収録されたからといって，追加された文字がすぐに実装されるわけではありません。しかし，国際規格として標準化されたことにより，パソコンや業務システムで標準的に扱え，情報交換ができるような環境の実現に，一歩近づいたと言えるでしょう。

今後の文字コード標準化は，実務の現場のニーズを把握しつつ，さらに進められていくものと思われます。

参 照 文 献

高田智和（2001）「合成字」『北海道大学大学院文学研究科研究論集』1: 89–107.
高田智和（2009）「行政用文字の調査研究における文字同定―辞書同定と辞書非掲載字に対する文献資料・非文献資料による同定―」『日本語科学』25: 131–141. 東京：国書刊行会.
高田智和・井手順子・虎岩千賀子（2008）「行政用文字の調査研究―汎用電子情報交換環境整備プログラム―」『日本語科学』23: 95–110. 東京：国書刊行会.

高田　智和（たかだ・ともかず）
国立国語研究所理論・構造研究系准教授。博士（文学）（北海道大学）。国立国語研究所研究開発部門研究員を経て，2009 年 10 月より現職。
主な著書・論文：『例解辞典』（新版監修，ぎょうせい，2010），「漢字処理と『大字典』」（『訓点語と訓点資料』109，2002），「漢字字体規範データベース」（共著，『日本語の研究』1(4)，2005）.
受賞：標準化貢献賞（日本規格協会，2007）.
社会活動：計量国語学会理事，情報処理学会人文科学とコンピュータ研究会幹事，日本言語学会事務局委員.