

# 仮名漢字変換システムの現状

2 Q-3

芝野耕司 大場康成 摂待清明 小堀博之 武田千春

東京国際大学 商学部 経営情報学科

## 1. はじめに

(1) 仮名漢字変換辞書交換形式の規格化原案作成、  
(2) 仮名漢字変換に関するアプリケーションプログラムのインタフェース (API) 規格仕様の検討、以上2項目の原案作成のため日本規格協会電子文書処理システム標準化調査研究委員会のWG3で行った仮名漢字変換システムについてのアンケート結果を報告する。(これは、各企業 [約40社] のアンケート回答に基づく。)

## 2. 回答者の立場

回答者のプロフィールは、エンドユーザ16社、システム従属/独立系仮名漢字変換システム開発者がそれぞれ13、12社、システム従属/独立系仮名漢字変換辞書開発者がそれぞれ10、9社となっている。

また、プラットフォームは、パソコン23社、ワークステーション19社、ワープロ16社、不特定(各種プラットフォーム全般)8社、次いでミニコン、メインフレーム各5社の割合であった。

## 3. 辞書交換形式に必要な情報について

### 3.1 読み情報

— 読みとして許される字種 [アンケート結果]

- ひらがな 28 (全角 15・半角 4)  
 カタカナ 25 (全角 7・半角 10)  
 英数字 24 (全角 8・半角 6)

読みとしては、平仮名にすることが常識的には考えられるが、アンケート結果は、最大限に緩くするか、またはなくすという意見が強かった。また、「ん(ン)」で始まる字についても、それを認める方がよいという意見があった。

— 読みとしての文字列最大長 [アンケート結果]

- 制限なし 19  
 制限あり 18

制限ありと制限なしの二つに同じ割合で分かれた。

制限ありとする理由として、(1) 読みとして長いものを入力する必要はない、(2) 常識的な範囲で制限すべきというのが挙げられる。最も多い意見は12~32文字に制限するといった意見であったが、回答範囲は12~250文字の範囲であった。

制限なしとする理由としては、ユーザーの立場に立ち可変長などのデータ構造を考えて、(1) 最大長はエ

ントリー形式に依存しており一概には決められず、

(2) システム移行の場合、最大長を越えた場合は、無視することで解決できるなどが挙げられた。

これらの結果は、仮名漢字変換を入力短縮手段として用いることもあるという利用実態を反映したものである。

### 3.2 漢字(表記)情報

— 漢字として許される字種 [アンケート結果]

- JIS X 0208 全般 17  
 JIS X 0212 全般 9  
 JIS X 0208 の漢字部分 (16区~84区) 5  
 JIS X 0212 の漢字部分 (16区~77区) 2  
 JIS X 0201 の英数字 9  
 JIS X 0201 の半角カナ 6  
 文字列であれば制限無し 16

漢字として許される字種についてJIS X 0208 全般、文字列であれば制限なしがそれぞれ半数という回答結果になった。8~15区、85区以降が他のパソコンのシステムと交換できないとか、システム移行中にそのシステムで許される漢字としての字種以外の物を含む場合はその辞書項目を無視する事で解決できるなどの理由から全体的に漢字として許される字種に関して制限を行わないのが望ましいようである。

— 漢字としての文字列最大長 [アンケート結果]

- 制限なし 18  
 制限あり 18

読みとほぼ同じ結果であった。

### 3.3 品詞情報

— 必要な品詞情報 [アンケート結果]

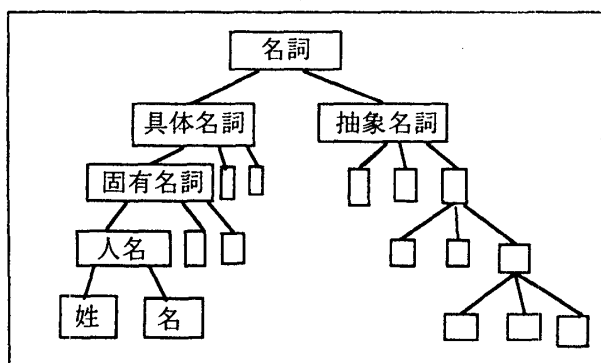
- 自立語 14     付属語 8     名詞 35  
 サ変名詞 (サ変動詞・名詞型) 33  
 固有名詞 29     5段動詞 24  
 上一段動詞 21     下一段動詞 20     サ変動詞 23     カ変動詞 20  
 形容動詞 26     形容詞 26     連体詞 19  
 数詞 21     代名詞 19     接続詞 21  
 感動詞 19     助動詞 13     助詞 16  
 接頭語 26     接尾語 26     副詞 26

以上の結果から各品詞の重要度を問うよりも、各品詞の接続・流通情報に重要性があるようだ。

この接続・流通関係の定義には品詞レベルからさらにプロファンドな分類をする必要があるようだ。

(例) 接頭・接尾語 → 接続条件により分類  
 五段動詞 → 活用行により分類

固有名詞 → 人名、地名などに分類



人名の姓の品詞を場合などを例にとると、図のようになる。このように階層を追って表現することにより、その文字がどのような品詞のカテゴリーに含まれ、また表現されるのかの手掛となる。

### 3.4 まとめ

読み情報にしても漢字(表記)情報にしても、出来るだけ制限を緩め、より広い範囲で使える規格化をすべきであるとのことである。

品詞情報については、その品詞の定義というよりも品詞と品詞のつながりに重点がおかれ、そのために品詞の階層的分類が必要となる。

## 4 仮名漢字変換関連APIの標準化に関する技術的可能性について

### 4.1 標準化の対象

[アンケート結果]

- A1. アプリケーションと仮名漢字変換 (FEP/BEP) とのインタフェース 33
- A2. アプリケーションとシステムとのインタフェース 6
- A3. 仮名漢字変換 (FEP/BEP) とシステムのインタフェース 18
- A4. 仮名漢字変換に関するヒューマンインタフェース 17

アプリケーションと仮名漢字変換 (FEP/BEP) とのインタフェース、仮名漢字変換 (FEP/BEP) とシステムのインタフェース、仮名漢字変換に関するヒューマンインタフェースの順で標準化されるべきという回答が得られた。その他として、機能キーの特定、半角の標準化という要求が得られた。

### 4.2 標準化の範囲

[アンケート結果]

- B1. 実行環境 (OS環境など) ごとの仮名漢字変換そのものの実装インタフェース標準化とそれに伴うユーザ・アプリケーションインタフェースの標準化 17
- B2. 実行環境 (OS環境など) に因らない一般的な仮名漢字変換機能とユーザ・アプリケーションインタフェースの標準化 23

[1] 実行環境 (OS環境など) に因らない一般的な仮名漢字変換機能とユーザ・アプリケーションインタフェースの標準化と [2] 実行環境 (OS環境など) ごとの仮名漢字変換そのものの実装インタフェース標準化とそれに伴うユーザ・アプリケーションインタフェースの標準化が回答を二分したが、[1] に対し、メモリ管理、イベント制御の面などで不可能なのではないかという意見も出された。

### 4.3 標準化のアプリケーションモデル

[アンケート結果]

- C1. キー入力を仮名漢字変換より先に取り必要もなければ、変換中の文字表示も仮名漢字変換に任すアプリケーション 21
- C2. キー入力は仮名漢字変換に先んじて取る必要があるが、変換や変換中の文字表示は仮名漢字変換に委ねるタイプのアプリケーション 14
- C3. キー入力も変換候補表示も自前で行い、単に仮名漢字変換機能を背後で利用するアプリケーション 14

キー入力を仮名漢字変換よりも先に取り必要もなければ、変換中の文字表示も仮名漢字変換に任すアプリケーションがよいという回答が多かった。またこれはアプリケーションによって要求が異なることからレイヤーに分けて考えた方がよいという意見が出された。

### 4.4 標準化の機能候補

仮名漢字変換に関するアプリケーションからの制御機能について [アンケート結果]

1. オン/オフ制御 27
  2. 仮名漢字変換機能の利用 24
  3. 変換表示に関する画面制御 (表示位置、大きさ、フォント、表示属性等) 21
  4. 操作及び制御 (各種辞書設定・解除、新語登録・削除等) 21
  5. キーボード制御 (入力モード等の制御) 20
- の順で機能候補が挙げられた。

### 4.5 まとめ

仮名漢字変換辞書同様、広い対象での標準化が望まれている。

標準化の範囲では実行環境に因らず、複数のアプリケーションモデルをサポートし国際化機能を取り込むことが要請されている。

標準化の候補についても、幅広い回答を得たが、それらの要求の重要度はこれから検討していく必要がある。

## 5. 参考文献

日本規格協会電子文書処理システム標準化調査研究委員会/WG3 アンケート