

OS/omicon 仮名漢字変換システム第2版

3N-5

におけるモードレス変換の評価

雀ヶ野史子, 丸山芳男, 天野純一, 早川栄一, 並木美太郎, 高橋延匡

(東京農工大学)

1. はじめに

従来の仮名漢字変換において文章を入力する際には文字種の境界でモード切替えが必要であり、思考の中断が避けられない。そこで、我々は OS/omicon 仮名漢字変換システム第2版[1]（以下、本システムとする）において、モード切替え操作不要の入力方式であるモードレス変換を実現した。しかし、実際にモードレス変換が文章作成に適しているのかなど、その入力方式の有効性をまだ明らかにしていなかった。

そこで、筆者はモードレス変換と従来のモード切替え操作を必要とする入力方式とで比較を行なうことによって、モードレス変換という入力方式の定量的な評価を行った。

2. モードレス変換の概要

仮名漢字変換を使用している状態で英字を入力する場合、一般に入力モードの切替え操作が必要となる。これに対し、モードレス変換とは、英字の入力を変換処理の枠組みの中に取り込み、英字の入力時にもモード切替え操作を必要としない入力方式である。具体例を図1に示す。

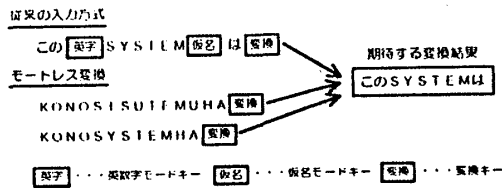


図1 変換結果「このSYSTEMは」を得るために必要なユーザの操作

3. 評価実験

モードレス変換という入力方式の有効性を検証するために、次に述べる評価実験を行なった。

3.1 実験方針

評価実験の方針は次のとおりである。

(1) 未知語を含むデータは測定の対象外とする

An Evaluation of the Modeless Input Method in the Kana to Kanji Transliteration System Version 2 on OS/omicon
Fumiko SUZUMEGANO, Yoshio MARUYAMA, Jun-ichi AMANO, Eiichi HAYAKAWA, Mitarou NAMIKI and Nobumasa TAKAHASHI

Tokyo University of Agriculture and Technology

入力文中に未知語が含まれていた場合、期待する変換結果を得ることができない。未知語の問題は、辞書の問題などが関係するが、本実験の目的は入力方式の評価であるので、測定の対象外とする。ただし、英単語の未知語については、モードレス変換が変換対が辞書に登録されていることを前提とする方式であるので、今回はユーザ辞書に登録を行なった状態で実験を行なう。

(2) 学習機能は用いない

学習機能が実験結果にどのような影響を与えるか、まだ不明である。本実験は比較実験であるので、初期値を同じくするため学習機能は用いない。

3.2 評価尺度

モードレス変換の有効性を「使い勝手のよさ」という尺度で評価する。「使い勝手のよさ」は次式で定義する正変換率と平均変換キー操作数からなる変換精度によって表す。

$$\text{正変換率} = \frac{\text{期待する結果が1回の変換で得られた数}}{\text{変換対象文字列数}} \times 100 [\%]$$

$$\text{平均変換キー操作数} = \frac{\text{期待する結果を得るまでの変換キー操作数}}{\text{変換対象文字列数}} \quad [\text{回}]$$

変換キー操作数とは、変換結果を得るために必要なキー操作数のことで、具体的には変換回数と文節移動回数である。従来の入力方式の場合には、さらにモード切替え操作回数に加わる。

3.3 評価方法

本システムにおいて整備された評価環境[2]を用いて、モードレス変換と従来の入力方式とで、変換精度を測定し、比較を行なう。本実験では、我々のように計算機を頻りに用いるユーザに一般的なローマ字仮名変換による仮名入力を対象とし、モード切替えについては、英字の入力時の前後に

だけ起きるものとする。実験作業の流れを次に示す。

- (1) 実験用にベンチマークテキストを用意する。
本実験では、理工系文書[3]の一部(テキスト①、6,332文字、英単語数189語、英字含有率20.1%)を用いた。変換を行なう文字列の単位は、句読点までの長さで連文節変換を行なった。さらにこのテキストからさらに英単語を含む変換対象文字列だけを抜き出したテキスト②(2,371文字、英単語数189語、英字含有率53.7%)を用意した。
- (2) 用意したテキスト①②を用いて、それぞれについてモードレス変換と従来の入力方式で、変換精度の測定を行なった。
- (3) モードレス変換と従来の入力方式とで、実験によって得られた変換精度の比較を行なった。

3.4 実験結果

実験結果を表2に示す。この結果を踏まえて、次の考察を行なった。

- (1) 正変換率はテキスト①の場合で5.6%、②の場合で40.2%もモードレス変換の方が上回っている。これは、従来の入力方式では英単語を含む文字列の場合、必ずモード切替えが必要となり、またモード切替えの前に文字列の確定を行わなければならないので、1回の変換では変換が不可能である。しかし、モードレス変換ではこのような文字列に対しても、モード切替え操作や結果の確定の操作が必要ないので、高い正変換率を得ることができた。
- (2) 平均変換キー操作数はテキスト①で約0.4回、②で約0.6回モードレス変換の方が多い。これは、モードレス変換では入力文字列のままローマ字仮名変換を行なったものとの2種類で結果を探索しているため変換候補が増え、その分変換キー操作数が増えたものと考えられる。また、従来の入力方式では、モード切替えの前に文字列の確定を行なうため、変換を行なう文字列(文節)が短い、モードレス変換ではモード切替え操作が必要ない分、文節が長く、文節移動回数も多くなる。

これらの結果から、モードレス変換は従来の入力方式と比較して、正変換率はよいが、平均変換キー操作数は若干多くなることがわかった。平均変換キー操作数が多くなるのは、1回の変換で期待する結果が得られなかった場合に必要となる変換キー操作数が、従来の入力方式よりも多いからであり、これがモードレス変換の問題点でもある。しかし、本実験では学習機能を用いていない。本システムでは、学習を行なうことにより、最大で43.7%もの正変換率の向上を得ることができる[4]。正変換率が43.7%向上したと仮定した場合の平均変換キー操作数の見積りを表2に示す。この場合、モードレス変換の方が変換キー操作数は少なくなると予想される。

表1 モードレス変換と従来の入力方式の変換精度

	モードレス変換		従来の入力方式	
	①	②	①	②
正変換率 [%]	34.1	55.6	28.5	15.4
平均変換キー操作数 [回]	5.03	6.30	4.59	5.68
入力テキスト	①	②	①	②

表2 学習を用いた場合の平均変換キー操作数

入力テキスト	モードレス変換	従来の入力方式
テキスト①	2.36 [回]	2.69 [回]
テキスト②	1.07 [回]	3.76 [回]

4. おわりに

本稿では、モードレス変換という入力方式について、従来の入力方式と比較することによって、その評価を行なった。今後は学習機能を用いた場合での実験を行ない、モードレス変換における変換キー操作数の減少を検討する必要がある。

参考文献

- [1]下村他：OS/omicon 仮名漢字変換システム第2版の設計思想，情報処理学会第42回全国大会5Q-1，1991
- [2]天野他：OS/omicon 仮名漢字変換システムの評価環境とその整備，情報処理学会第50回全国大会3N-6，1995
- [3]I.G.Lewis 著：最新ソフトウェアエンジニアリング-CASEの考え方と実践，1993
- [4]酒井貴子：仮名漢字変換における変換手法と変換精度についての比較実験，情報処理学会第46回全国大会5L-8，1993