

## 7C-5 漢字字画による日本語漢字入力補助方法に基づくシステムの評価実験

張小剛\* 高建斌\* 高橋勇\*\* 黒岩丈介\*\* 小高知宏\*\* 小倉久和\*\*

\* 福井大学工学研究科 \*\* 福井大学工学部

## 1 はじめに

現在、キーボードを用いた日本語入力方式には、主にカナ文字入力とローマ字入力がある [1]。文字の読み方によるこの二種類の入力方式は、日本語の読み方が分かる人にとっては操作は簡単であるが、日本語の読み方、特に日本語の漢字の読み方が分からない外国人日本語初学者にとっては大変難しい操作である。そこで、我々はこの問題を解決する方法の 1 つとして、日本語の読み方によらない日本語漢字字画入力補助方法 JCSIM (Japanese Character Strokes Input Method) 法を提案した [2]。本報告では、JCSIM 法が実際に外国人留学生に対し有効であることを示すために行った評価実験の結果を報告する予定である。

## 2 JCSIM 方法とシステムの概要

## 2.1 JCSIM 入力方法

JCSIM 法では、全ての漢字字画を表 1 に示すような 6 種類に分ける。そして、これらの 6 種類の字画に対し表 1 のように 6 つのキーに対応させた。つまり、本手法では、漢字を入力する際には、漢字を入力字画に分解し、漢字の書き順に従って一つずつ字画と対応するキーをタイプすることで漢字を入力することとなる。実際の入力方法は字画数によって次の 3 つの場合に分けることができる：

1. 5 画未満：漢字の書き順により、実際の字画をすべて入力してから、最後にスペースを入力する。
2. 5 画：漢字の書き順により、実際の字画をすべて入力する。
3. 5 画以上：漢字の書き順により、最初の 4 画を入力してから、最後の 1 画を入力する。

同じ字画コードに多数の漢字を持つ場合は、すべての候補に番号をつけて入力者に表示する。入力者は番号で入力したい漢字を選ぶ。

Evaluation experiments of the system based on the Japanese character input auxiliary method by character strokes  
Xiaogang Zhang\* Jianbin Gao\* Izamu Takahashi\*\* Takesuke Kuroiwa\*\* Tomohiro Odaka\*\* Hisakazu Ogura\*\*  
\*Graduate School of Engineering, Fukui University  
\*\*Faculty of Engineering, Fukui University

表 1: 漢字字画と入力キーの対応

漢字字画	一	丨	ノ	丶	丨	その他
キー	S	D	F	J	K	L

## 2.2 システム概要

実装したシステムは「漢字字画コード生成」、「漢字検索処理」、「字画コード定義辞書」、「多漢字出力と多漢字出力入力者選択」及び「単漢字出力」からなる。まず、システムは、作成した「Emacs とのインタフェース」を介して、入力者が入力したキー列を受け取り、「漢字字画コード生成」部で漢字字画コードを生成する。次に、「漢字検索処理」部は、「字画コード定義辞書」を用いて、入力者が入力したい漢字候補を検出する。「単漢字出力」あるいは「多漢字出力と多漢字出力入力者選択」は、その結果を「Emacs のインタフェース」とを介してユーザに提示する。

システムの「字画コード定義辞書」には、JIS 第一水準 2,965 漢字と JIS 第二水準 3,390 漢字を登録しているため、通常の日本語文章の漢字入力は、本システムで入力できる。

## 3 実験と評価

JKIMS の有効性を確認するために、「評価に用いたテキストに外国人留学生の読みの分からない漢字がどの程度あるか」、及び「外国人留学生が JKIMS を組み込んだローマ字入力システムを使用することにより入力時間をどの程度軽減できるか」を検証した。

入力者は、外国人留学生日本語初学者である (表 2)。入力者 A~C は日本在留期間と日本語学習時間が短く、日本語レベルが初級程度 (日本語能力試験 3 級以下) である。入力者 D~F は日本在留期間と日本語学習時間が長く、日本語レベルが中級程度 (日本語能力試験 3 級以上 2 級以下) である。そして、A~F の順に、日本語学習時間が長く、日本語レベルが高くなっている。

被験者である外国人留学生らは、論文や発表原稿などの専門的文章を入力する機会が多い。そこで、本実験では、専門雑誌から漢字率が中程度 [3] の文章を選んで実験対象とした。3 つの文はそれぞれ違う分野である「情報処理」、「生物化学」及び「電子電気」に関する文である。実験用文に関する詳細を表 3 に示す。表 3 の「新出漢字数」とは複数に出現した漢字を 1 つとしてカウントした数値である。実験においては、同じ実験用文中に出現する読めない漢字を 2 回以上入力する

表 2: 入力者データ

入力者	専門	日本への在留期間(月)	日本語学習時間(時間)
A	情報処理	6	約 300
B	経済経営	6	約 350
C	日本文学	6	約 500
D	機械	12	約 1,000
E	材料開発	18	約 1,200
F	材料開発	18	約 1,300

表 3: 実験用文データ

文番号	総字数	漢字数	新出漢字数	仮名数	漢字率
F1	336	159	88	177	47%
F2	234	86	66	148	37%
F3	338	137	86	200	41%

場合は、前の同じ漢字をコピーしてよいこととした。

実験は日本語レベルに応じて入力者を二つのグループに分けて実施した。日本語レベルが初級の入力者 A, B, C は、JKIMS を組み込んだローマ字入力システムだけを用いて各実験用文を入力する。実験では、読みの分からない漢字を入力する際には、かならず JKIMS を使うこととした。そして、各入力者が各実験用文を入力する際に、JKIMS で入力した漢字数を記録した。

日本語レベルが中級の入力者 D, E, F に対しては、まず JKIMS を組み込んだローマ字入力システムで文 F1 から F3 までを入力する。読みの分からない漢字を入力するには、JKIMS を使い、JKIMS で入力した漢字数と入力時間を記録する。その後 10 日間以上おいてから、JKIMS を組み込んでいないローマ字入力システムで文 F1 から F3 までを入力する。この際、読みの分からない漢字を入力するには、外国人留学生が一般によく使う「紙製の辞書から漢字の読みを調べ、漢字を入力する」という方法により入力することとした。そして、入力時間を記録した。

各入力者が JKIMS を組み込んだローマ字入力システムで各実験用文を入力する際に、JKIMS で入力した漢字数と、新出漢字についての JKIMS の利用率及び各入力者の平均利用率を表 4 に示す。日本語レベルが中級の入力者 D, E, F における、JKIMS を利用することによる入力時間軽減率を表 5 に示す。表 5 の「辞書時間」は読みの分からない漢字を辞書で引く方法で入力にかかった時間であり、「JKIMS 時間」は JKIMS 使用時の入力時間である。表 5 の表記時間の単位はすべて「分」である。

表 4: JKIMS の利用率

入力者	実験用文	提案方法で入力した漢字数	新出漢字での利用率	入力者の平均利用率
A	F1	65	74%	63%
	F2	44	67%	
	F3	41	48%	
B	F1	45	51%	51%
	F2	40	61%	
	F3	38	44%	
C	F1	24	27%	21%
	F2	14	21%	
	F3	13	15%	
D	F1	9	10%	9.6%
	F2	5	7.6%	
	F3	9	10%	
E	F1	6	6.8%	5.8%
	F2	3	4.5%	
	F3	5	5.8%	
F	F1	5	5.7%	3.8%
	F2	2	3.0%	
	F3	2	2.3%	

表 5: JCSIM による入力時間軽減率

入力者	実験用文	辞書時間	JCSIM 時間	軽減時間	軽減率	総軽減時間	平均軽減率
A	F1	47	32	15	32%	40	31%
	F2	29	21	8	28%		
	F3	51	34	17	33%		
B	F1	89	80	9	10%	20	8.5%
	F2	67	62	5	7.5%		
	F3	80	74	6	7.5%		
C	F1	52	45	7	13%	15	13%
	F2	26	22	4	15%		
	F3	36	32	4	11%		

## 4 まとめ

評価実験により、JCSIM が使いやすい漢字入力補助方法であり、日本語レベルが中級である外国人留学生の日本語入力においても有用であることが示された。そして、JCSIM は日本語レベルの低い入力者に有用性が高いことが分かった。

今回は漢字圏からの留学生だけを対象として実験を行ったため、非漢字圏からの留学生に対する有用性の確認は行っていないので、今後の課題としてこれらの留学生を対する実験を行う予定である。

## 参考文献

- [1] 尾関雅則, “第 32 回全国大会にあたって—日本語入力方式について”, 情報処理, Vol.27, No.5, pp493-497 (1986)
- [2] 張小剛, 高建斌, 高橋勇, 黒岩丈介, 小高知広, 小倉久和, “違う読みで入力された日本語単語の読みを表示するシステムの提案”, 情報科学技術フォーラム講演論文集 FIT2003 E-007(2003.9).
- [3] 国立教育政策研究所: <http://www.nier.go.jp/homepage/jouhou/cooperation/saka28.htm>