

漢字階層コード方式

石川 皓 勇 (トッパン・ムーン)

1. 概要

高頻度漢字をカナ鍵盤四シフトに配置して直接に選択し、これらとの熟語・類義関係でその他の漢字を指定し、漢字カナまじり文中でカナ表記される頻度に応じてカナ配列した鍵盤でカナを打鍵する方式について説明する。

はじめに、日本語入力方式を道具としてのタイプライター操作の洗練として捉え、日本語文中における同音語の諸形態の頻度を集計分析してカナ漢字変換方式の限界を説明し、漢字テレタイプから2ストローク方式に到る記憶コード方式の能率と柔軟性の発展をたどり、漢字を体系的に促えた本方式の必然性を説明する。

次に操作法とハードウェアの概要を説明し、システム構成の特徴として、少数の鍵盤内漢字との熟語・類義関係により多数の漢字を選択し、漢字とひらがなの字種指定のための打鍵が不要であり、漢字カナまじり文中での頻度に応じた配列によりカナ入力が容易になることを説明し、漢字コードの一部とカナ鍵盤の構成例を紹介する。

2. 先行技術との比較

2.1 道具としての日本語入力装置

改文タイプライター式の鍵盤で漢字を含む日本語文を印字しようという思想の歴史は古く、実に大正11年出願の特許第60142号(杉原甚三)に遡るが、一般に実用化したのは昭和54年であり、キーボード、プリンタ、ディスプレイ、

電子辞書の複合体を、マイコンコンピュータで制御するワード・プロセッサーとしてであった。

日本語ワードプロセッサの実用化の成功は、上記の複合体としての成功であり、ディスプレイ、電子辞書、マイコンコンピュータの発達に要する所が大きく、カナ・漢字変換方式の成功度については、あかち書き、同音語の選択、辞書構成、漢字検索などの問題点について、深刻な疑問が提出されている。(例えば最近の文献には、野村雅昭「漢字の現在」、『言語生活』誌、1983年6月号P.52~P.59)

日本語の入力方式については、日本語と人間工学の要素が重要なので、上記のハードウェアの進歩とは別の観点、即ち、道具としてのタイプライターの有効性の観点から、従来技術の要約を試みる。

それは諸技術の複合化よりも作業の洗練の歴史であり、道具としての技術を能率的に抽出する仕方の探索の歴史である。考え方の類比としては、日本語文の印刷にルビが使われなくなった過程(漢字の学習・使用の普及)、自転車の安全性の追求の歴史(増速伝動手段の階層化)、複葉飛行機が低翼単葉機になってスピードが達成された歴史(厚い翼型という飛行手段の流線型化・最適化)などが想起される。

これらの技術が出現した当時、どれも奇妙な、特殊な印象を与えながら、内に秘めた発展性により、徐々に現実に適応して支配的な方式になった。

2.2 カナ漢字変換のタッチ・二次選択

カナ(ローマ字)漢字変換による日本語文の入力では、打鍵動作のほか、同音語の選択が必要であり、初心者が約50時間の訓練の後、一語の選択に約4.0秒を要し、カナ漢字変換辞書に該当単語がない場合の漢字訂正作業(訂正箇所指定および漢字検索情報の入力)には、漢字1文字当り平均21.2秒を要している。(文献2)

「新明解国語辞典」所収の58,431語のうち、その36.4%にあたる21,270語は同音語である(文献3)が、使用頻度(延べ語数)を数えたものに国立国語研究所報告38「電子計算機による新聞の語彙調査(II)」のP.239~P.304の「同音短単位表」があり、昭和41年の新聞約1年分の全紙面より抽出した延べ約94万語の中に含まれる同音語の使用度数の一覧表である。これをキー操作の難易度を考慮して類別集計したのが右の表であり、合計は下記の構成比になる。

標本度数	940,533	100%
同音語合計	455,158	48.4%
うち A. 同音異義語	191,529	20.4%
B. 個有名詞	21,498	2.3%
C. 漢数字	24,839	2.6%
D. 同語異表記	80,955	8.6%
E. 一字カナ語	136,337	14.5%

} 25.3%
} 23.1%

Eは助詞など、Dは用言の文字ばかり(例、有、有り、あり)であり、同音語から除外するためには、カナが漢字かの字種指定のキー(変換キー)の打鍵が必要であり、文節指定モードで操作を簡単にすれば、文体が固くなる。

A、B、Cについて、接続関係、意味カテゴリ、使用頻度による自動判別が実施されているが、いずれも部分的効果にとどまり、しかも瘦字辞書への登録語数を増して、なるべく長い単位の熟語をいにより選択を完全にしようとする、同音語を増加するというティレンマがある。

見出し	A. 同音異義語	B. 個有名詞	C. 漢数字	D. 同語異表記	E. 一字カナ語	合計
あ	2,619	1,031	0	6,044	4	9,698
い	6,959	1,176	5,576	5,018	1,307	20,036
う	1,220	79	0	1,825	743	3,867
え	4,297	235	0	762	47	5,341
お	2,081	951	0	3,280	1,406	7,718
か	7,943	1,039	0	14,708	12,803	36,493
き	10,139	198	337	15,445	399	12,618
く	2,195	106	957	2,499	2	5,759
け	5,629	459	0	153	0	6,241
こ	18,524	482	40	1,065	495	20,607
さ	5,819	849	3,743	972	1,341	12,724
し	32,334	1,110	1,244	1,050	6,636	42,374
す	379	306	0	1,593	634	2,912
せ	7,555	153	357	673	399	9,137
そ	2,422	21	0	1,468	4	3,915
た	11,614	1,016	0	2,379	12,322	27,331
ち	4,888	163	0	316	7	5,174
つ	2,978	81	0	1,926	3	4,788
て	10,413	130	0	1,705	19,556	31,804
と	9,659	2,973	0	1,648	10,182	24,462
な	4,473	881	0	5,044	2,377	12,775
に	630	2,472	5,359	1,246	17,040	26,747
ぬ	171	3	0	2,285	157	2,616
の	210	122	0	865	32,590	33,787
は	4,737	191	1,412	1,675	14,072	22,087
ひ	3,341	629	90	1,275	2	5,337
ふ	5,136	733	21	826	1	6,717
へ	5,618	406	0	825	0	6,849
ま	2,758	604	3,237	4,924	3	11,326
み	1,193	1,041	12	457	407	3,110
む	1,911	202	0	1,891	11	4,015
も	840	99	0	2,367	0	3,306
や	1,636	614	0	508	823	3,581
ゆ	1,089	228	0	342	0	1,659
よ	2,485	383	757	3,210	146	6,981
び	5,248	152	1,697	1,294	404	8,795
わ	586	179	0	1,692	14	2,471
合計	191,529	21,498	24,839	80,955	136,337	455,158

表1. 同音語の類別度数

上記の分類で、問題のある例として

あこ (2893)	れき (678)	れんじ (84)	
或 2	歴 (677)	連合 (82)	}D
有 546	層 1	聯合 (2)	
有 13	ヒト (184)	ヒリ (163)	
あこ 2332-D	一 (133)	一人 (144)	}D
	ヒト (51)	ヒリ (19)	

細部では分類する人によって相違が生じ得るが、数字が大きく相違する事は無い。

欧文タイプライター式の鍵盤により漢字を入力して、表示装置により選択する方法は、比較的早くから知られて居り、特公昭29-5809 (リン・ユータン) は同一部首の漢字を、特公昭41-18187 (新保光正) は同音の漢字または熟語を表示する機械式の装置を開示している。

電子辞書による同音熟語の逐次表示によっても、同音語選択か入力速度を大きく制約する事は、上述のデータにより明白であり、読み方による漢字入力法の宿命であろう。

2.3 大型鍵盤による直接選択

和文タイプライターにおける全文字配列、漢字テレタイプにおける多段シフト配列も、古くから実用されているが、前者は多数の漢字の中から必要な漢字を目視により検索するので速度が遅く、後者は鍵盤が大型になるので、文字キーを選択するのに腕を振り廻すことになり、やはり速度が制約される。なおし、漢字テレタイプは終つては、特定用途に高頻度に出現する文字を手元に近く配置する事により、腕の動きを少なくして、やや速度を向上させ得る事が知られている。

2.4 2ストローク・タッチ法

欧文タイプライター式の鍵盤を用いてタッチ法により打鍵し、しかも同音異字の表示選択工程を不要とした方法は、比較的最近になって発表され、先駆的な特公昭40-9312 (後述する) 以外は、特公昭50-33369 (工業技術院、木沢誠)、特公昭50-33368 (谷村新興、小川注連男)、特公昭50-35453 (川上星、川上義) などに開示され、後二者は、2ストローク方式として実用されている。

これらはいずれも使用する漢字すべてについて、それぞれのコードを記憶しなければならないのが困難な点であるが、この訓練により高速度が得られるのが長所である。自分が常時使用する漢字のコードの記憶は鮮明であり敏速に入力できる。記憶されたコードは使用頻度順に、また、意味カテゴリー別に頭脳の中で整理されるので、各自に最適なコード体系が頭脳の中に生成される。望ましい単純化の極限として、Tコード (東大、山田尚勇) が実験されている。全文字配列や、漢字テレタイプなどの多段シフト配列が、大型の鍵盤における文字配列によって打鍵動作が制約されるのに対して、暗記コードによれば、各自に最適なコード体系が脳中に生成されてタッチ打鍵の技能が習得される。(文献4)

2.5 漢字階層コード方式

日本語に常用される約2000字の漢字のうち、使用頻度の上位200字をいし400字は、一字で語として用いられる意味の明瞭なものが多く、また、同音熟語の成分として多くの熟語を形成するので、同音異義語の主たる要素ともなっている。

上位200字で漢字の延べ使用度数全

体の52.0% (雑誌) ないし58.1% (新聞) に達し (文献5)、その他の多数の漢字は、これらとの熟語または類義関係で使用されることが多い。

従って、これらの使用頻度の高い漢字を指定すれば、その他の多数の漢字は、読み方を指定するだけで、これらの漢字との熟語・類義関係により、同音異字の中から特定の漢字を指定し得る。これは、同音の漢字を電話で確認する方法と同じである。例えば「義」は義理のギであり、「藤」は藤の花のフジである。(文献6)

この方法によれば2ストローク法のように多数の漢字に個有のコードを記憶する必要が無く、使用頻度の高い漢字の配列さえ覚えれば、その他の漢字は、熟語および類義関係により、読み方の第一音節を指定することが出来る。漢字の音読みは、第一音節が、イ、ウ、キ、ク、チ、ツ、ン、に限られるので、指定上有効ではなく、省略することが出来る。

熟語を同音異字の特定のちめに使用した装置は、既に特公昭40-93/12 (芝電発、園印成) に開示されているが、これは所定の二文字熟語の読み方を仮名鍵盤で打鍵し、更に前字キーまたは後字キーを打鍵して一字を特定するものである。現代の電子メモリにより、使用頻度の高い少数の漢字との熟語および類義関係に於て、その下位区分として使用頻度の低い多数の漢字を階層的に記憶させておけば、漢字のみでなく熟語をも第一音節のみの短縮打鍵により能率的に選択できる。

この方法を効果的に実施するためには、改文タイプライター式の鍵盤に、使用頻度の高い漢字を直接に配列して、

シフトキーの併用により選択することが必要であり、改文タイプライターの鍵盤に配列されるものは、アルファベット、仮名などの表音文字に限るという先入見から脱却することが、先ず必要である。次に、漢字テレタイプの多段シフト配列のように漢字をグループ化して並列的に配列するだけでなく、上記のように熟語および類義関係により階層メモリに構成して、漢字および熟語を選択することが必要である。

2.6 キー選択の順列

アルファベットは26字にすぎないが、実際に打鍵操作により印字される文字連続の数は多数のものが必要であり、これは日本語における常用される漢字の字種の数とあまり大きな相違ではなく、また、これらの文字連続が、それぞれ特定の意味を持つていて、文脈展開の要素となっていることも漢字の場合と同じである。(例えば、a, the, of, ab-, ad-, -ed-, -er などには意味を持つていて。) 決定的な相違は、すべてを26字のアルファベットで表記するので、25種類以下の順列の逐次的な反復しか有り得ず、必ずしも系統的な打鍵操作になることである。

更に、26字のアルファベットのうちでも、母音を印字する文字は、a, e, i, o, u, y など数個の字種に限られ、音節は CV C、CV、VC (たかしCは子音、Vは母音) などの型式で構成されるので、約20種の子音文字と数種の母音文字が文中で交互に出現することが多い。従って、子音から母音への文字連続には、数種類の順列しか存在し得ない。また子音同志の文字連続も、例えばCに連続する文字は ch-, cl-, cr-, ce, c, l, ll, z に限られるように、実際には数個の順列

のうちの一個を選択打鍵すれば良いので、この事実が欧文タイプライターのタッチ法による迅速な打鍵を容易なものにしている。

従って、同時に選択判断すべき順列の数を数個に限定すれば、特殊なコード等の暗記に依らずに、欧文における迅速な容易な打鍵法を日本語のタイプライターに採用できる。

3. システム構成

3.1 鍵盤内漢字

高頻度漢字192字をカナ鍵盤4シフトに配置、文字キー1個、シフトキーの順序で打鍵して選択する。例えば、「け」の1は「決」であり、同2は「戦」、同3は「生」、同4は「死」である。

シフトキー1、2、3、4を、スペースバーの位置(サ1回でシフトバー

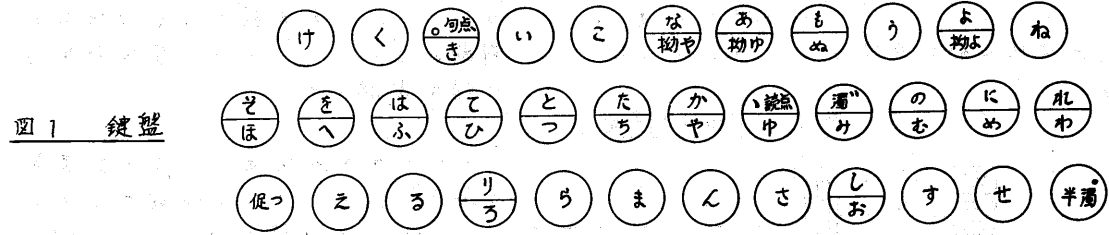


図1 鍵盤

図2 文字配列

各キー内の数字は
上段...カナの使用頻度%
下段...漢字の使用頻度%(鍵盤外漢字を含む)
(句点)

け 1.09 決生 戦死	く 1.09 苦性 租苦	い 4.06 一合 二計 三数 四行 五行	こ 1.30 国軍 民家	な 成城 全 中	あ 私同 戦太 中	は 物同 中 中 中 中	う 1.41 運進 取撃	よ 用要 利 利 利	ね 0.10 0.10			
とほ	をへ	はふ	てひ	とつ	たち	かや	い	濁	のむ	にめ	ル巾	
促つ	え	る	り	ろ	う	ま	ん	さ	し	す	せ	半濁

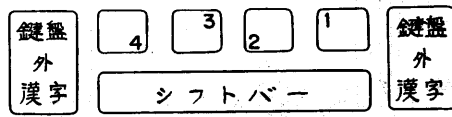
の上)に設けて、左右いおれの手でも操作可能にし、文字キー、シフトキーの左右手交互打鍵を可能にする。

文字キー上の4字の漢字は句を形成し、各段の意味特徴を、上段(上)主体、中段(下)形式的認識、下段動作物体に限定し、各指の意味特徴を、人(1)指から起、承、転、結(小指)に限定して、選択を容易にする。各キーの句の頭文字が、カナキーと読み方を合わせる。

鍵盤内漢字は、配列を暗記する必要がある。

3.2 鍵盤外漢字

鍵盤内漢字の下位区分として、鍵盤外漢字の熟語メモリを熟語および類義の漢字により構成し、鍵盤内漢字の打鍵に鍵盤外漢字のキー音節を追加打鍵して指定する。例えば鍵盤内漢字「決」(「け」のシフト1)には、次の鍵盤外



漢字が付属していて、付記したカナの通りオノ音節を追加打鍵すると同時に鍵盤外漢字キー（「四」の最下段の左右端にある）を打鍵して指定する。

カ キ サ シ ソ シ ヲ ハ
可 許 裁 審 訴 訟 断 判

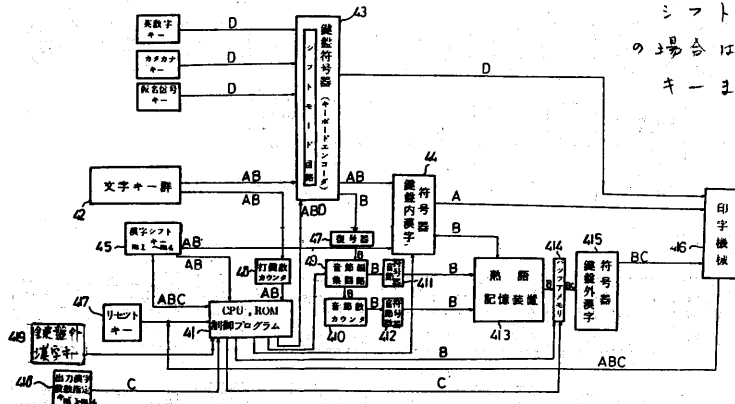
訴-訟とハイフンで結んだのは、訟が訴に付属する下位区分であることを示し、これにより「審」の「シ」と区別される。

3.3 仮名

本方式に於ては、漢字の配列とカナキーの配列がほぼ独立しているため、漢字カナまじり文中で付属語、活用語尾等のためにカナ表記されている頻度に応じてカナキーを配列している。

オノ図に於て、キーを上下に分割してあるものは上段（例「と」）がシフトなし、下段（例「ほ」）がシフトつきキーである。「ほ」のキーを押したまま、または指を離した後に、シフトバーを叩く。

図3 ブロック構成



指定系統、Bは鍵盤外漢字の指定系統である。通常の鍵盤符号器43（バッファメモリ付）の他に漢字シフト信号△を追加する鍵盤内漢字符号器44、熟語記憶装置40を介する鍵盤外漢字符号器45が機能する。音節編集回路は濁音、拗音等の入力マトリクス回路である。CPU41の機能は信号の分岐指令である。詳細は文献1の公報を参照。

3.4 操作法

- を文字キー1個打鍵
- △を漢字シフトキー（シフトバーの上の□□□□の1個）打鍵
- ◎を文字キーを押したまま鍵盤外漢字キーを打鍵
- △を文字キーを押したまま漢字シフトキー打鍵とすると、
- で平仮名例、け、を指定し、
- △で鍵盤内漢字例、決
- △◎で鍵盤外漢字例、可
- △◎で鍵盤内漢字と付属の鍵盤外漢字例、決裁、可決
- △◎◎で同一の鍵盤内漢字に付属する鍵盤外漢字を2字例、訴訟、判断、許可を指定する。

以上の打鍵によって生ずる信号の連続は、平仮名の打鍵がある場合のみ、△と△の間の○が2個以上になるので、これを検出して平仮名と判断させることにより、字種指定の打鍵を省略できる。

シフトつきの仮名に所屬する漢字の場合は、シフトバーと漢字シフトキーまたは鍵盤外漢字キーに親指をまねかかせて打鍵する。
例、本（「作」のシフト1）
基（「き」のシフトバーと鍵盤外漢字キー）

3.5 ハードウェア

左の図3は、ブロック構成に、Aは鍵盤内漢字の

4. 最適化とコード

4.1 カナ鍵盤

日本語の通常の表記である「漢字かなまじり文」に於ては、名詞等の自立語は主として漢字で表記され、助詞、助動詞、活用語尾等の付属語は仮名で表記される。これに対して、従来のカナ鍵盤に於ては、漢字で表記される自立語を含めて全文をカナ入力しているため、カナに過重な負担を強いる結果になっていないだろうか。カナ入力をカナで表記される語、即ち付属語だけに限れば、英文タイプライターと同様に三段鍵盤でシフトの回数が僅かなカナ鍵盤が得られないだろうか。

この疑問に答えるためには、漢字かなまじり文中における仮名の頻度を調べ、全文をカナ入力する場合の頻度と比較するのが妥当である。前者の資料として朝日新聞東京本社(1950年)の「活字使用度数調査熟語使用度数調査」より、漢字かなまじり文中のひらがなの頻度を求め、後者の資料として今泉国晴の「日本語の digram の相対頻度とその特性」(雑誌「心理学評論」1960年4月号、85頁～100頁)より、全文カナ入力の場合の頻度を求めたものを下記する。比較しやすいようにそれぞれの頻度順に併記する。(文献7)

朝日			今泉		
1	濁点	11.04%	1	濁点	9.32%
2	の	7.78	2	い	5.80
3	て	6.78	3	か	5.42
4	か	5.25	4	し	5.33
5	に	4.95	5	う	4.67
6	は	4.81	6	た	4.57
7	た	4.47	7	ん	4.15
8	と	4.32	8	で	3.89
9	る	4.13	9	と	3.74

朝日			今泉		
10	を	4.09%	10	の	3.55%
11	い	3.92	11	一、 ^{促音}	3.21
12	し	3.60	12	お、を	2.82
13	な	2.64	13	こ	2.53
14	つ、 ^{促音}	2.49	14	き	2.50
15	ら	1.89	15	に	2.50
16	れ	1.89	16	く	2.47
17	す	1.79	17	わ、 ^{助詞は}	2.47
18	こ	1.56	18	よ、 ^{拗音}	2.45
19	も	1.50	19	な	2.05
20	り	1.44	20	る	1.88
21	う	1.41	21	あ	1.58
22	あ	1.26	22	ら	1.40
23	ま	1.23	23	も	1.37
24	さ	1.20	24	け	1.34
25	よ、 ^{拗音}	1.18	25	ゆ、 ^{拗音}	1.29
26	く	1.09	26	さ	1.25
27	け	1.08	27	す	1.24
28	き	0.96	28	ち	1.21
29	そ	0.85	29	れ	1.21
30	わ	0.82	30	せ	1.18
31	め	0.76	31	ま	1.17
32	ん	0.74	32	り	1.14
33	せ	0.70	33	と	1.12
34	へ	0.67	34	は	1.01
35	お	0.66	35	え	0.98
36	え、 ^②	0.64	36	や、 ^{拗音}	0.97
37	ち	0.61	37	ひ	0.89
38	ろ	0.52	38	み	0.74
39	み	0.49	39	ふ	0.69
40	ひ	0.48	40	め	0.60
41	ふ	0.39	41	ほ	0.56
42	や、 ^{拗音}	0.32	42	ろ	0.55
43	を	0.31	43	ぬ	0.36
44	ほ	0.29	44	へ	0.30
45	ぬ	0.29	45	を	0.29
46	ね	0.18	46	ぬ	0.06
47	ゆ、 ^{拗音}	0.16	47	半濁点	0.18
48	ぬ	0.14			
49	ゑ、	0.09			
50	半濁点	0.03			

表2
カナの頻度比較

(朝日、漢字かなまじり文中のひらがな、今泉、全文カナ)

上記のデータを比較すれば、以下の事が明らかである。

1. 漢字かなまじり文中のひらがな(朝日)では下記の通り上位12位の頻度が65.1%に達し、全文カナ入力(今棠)の56.5%よりも8.6%高い。これは最も打鍵しやすいホームポジションの段(第2段)へ配置し得る文字の頻度が8.6%高い事を意味する。

	朝日	今棠
1位~12位の計	65.14%	56.47%
13位~22位の計	17.78	21.83
23位~32位の計	9.91	12.40
33位以下の計	6.97	9.30
全体の合計	99.80%	100.00%

2. 三段のカナ鍵盤に於てシフト付きの打鍵は頻度が低い事が望ましいので、上記の33位以下の文字が対象になるが、その頻度は漢字かなまじり文中のひらがな(朝日)は6.97%、全文カナ入力(今棠)は9.30%であり、前者の方が頻度が低い。

3. 漢字かなまじり文中のひらがな(朝日)の第1位~12位の文字は、助詞等に用いられるが、このうち、の、て、に、は、る、を、の6字は、全文カナ入力の場合と頻度が大きく異なり、特に、に、は、る、の3字は、全文カナ入力(今棠)の第13位以下になっている。従って、全文カナ入力用の鍵盤では、これらの文字全部をホームポジションの段に収容できず、その上の段等をも使用している。の、に、は、を、等の助詞は、日本文に基本的なものであり、これらがホームポジションの段だけで打鍵できない事は、特に初心者にとって、全文カナ入力のカナ鍵盤が使えない原因になっている。

上記の三項目を考慮に入れて構成し

た鍵盤配列の一例を図1に示した。シフトバーの打鍵9.5%が必要なのと、これを含めて100%に計算し直すと、漢字かなまじり文中の仮名を入力するのには概要下記の通りの頻度配合になる。

シフトなしの打鍵(句英、読英を除く)	
ホームポジションの段	51.4%
上段および下段の計	31.2%
シフト付きの打鍵(全段)	8.7%
シフトバーの打鍵	8.7%
合計(但し句英、読英以外)	100%

各段の配合(シフトバー、句英、読英を除く)

	シフトなし	シフト付き
上段	15.22%	2.22%
中段	56.23%	6.00%
下段	18.95%	1.18%
合計	90.40%	9.40%

各指の配合(シフトバー、句英、読英を除く)

左手 48.85%				右手 50.95%			
小	中	人	小	中	薬	小	
9.18	10.42	10.18	19.07	11.57	12.58	11.29	10.51

文字連続(文献8による新聞テキストの数字は頻度%)

両手交互打ち		片手連続打ち
から 2.548%	より 0.876%	いる 4.314%
ある 3.370	れり 0.830	こと 2.103
ない 2.010	なる 0.817	もの 1.618
いう 2.378	する 0.793	まで 1.017
この 2.123	あり 0.765	つい 0.821
では 1.674	あつ 0.654	です 0.770
その 1.534	それ 0.634	たか 0.745
ます 1.534	よる 0.606	でも 0.733
には 1.499	ても 0.602	ほか 0.717
なつ 1.373	たら 0.587	にも 0.710
ため 1.264	なく 0.517	また 0.704
これ 1.058		たが 0.577
など 1.050		いう 0.567
れた 1.048		とろ 0.553
れた 1.048		では 0.537
れて 0.895	合計26語 26.1%	これ 0.536
		合計16語 18.0%

4.2 鍵盤内漢字

鍵盤内漢字の選定と配列は、これに付属する鍵盤外漢字を含めた漢字全体の、体系のキーとなる重要なものであるが、その選定の条件としては、

- (1) 使用頻度
- (2) 構成する同音語の数
- (3) カナ鍵盤とシフトの照合
- (4) 所属する鍵盤外漢字の数
- (5) 各キーの四字句の有意味性
- (6) 各段、各指の有意味性

があげらる。実際は、(1)、(2)、(3)の条件により一次的に構成し、(4)を検証の上、(5)、(6)を修正してゆく。

(1)、(2)、(4)の条件については、国立国語研究所の語彙統計(文献9)を利用できる。(1)については、下記の表4の様式の索引(A3判4枚に収まる)を用意すると便利である。

表4. 音読み別使用順位分布表(一部)

順位	1	201	301	401	501	601	701	801	901	1,001	1,101	1,201	1,301	1,401	1,501	1,601	
音	100	200	300	400	500	600	700	800	900	1,000	1,100	1,200	1,300	1,400	1,500	1,600	1,995
ア										阿							亜
アイ		愛															哀
アウ		悪															握
アツ																圧	
アン		安	案					暗									
イ	意	以	位			伊	医	衣	遺	依	偉	威		胃			椅
		遣				圓	移	為	易				葦				魁
						異		雜									
イウ												域					
イフ				育													
イツ																	逸
イン		引	頁			院	印	飲					隱	陰			蔭
							因										

(使用(頻度)順位は、国研「雑誌90種」による。)

4.3 鍵盤外漢字

鍵盤外漢字は、複数の鍵盤内漢字に所属し得る。例えば「資」は熟語「資本」により「本」(ほ・1)、「資金」により「金」(な・2)、「資産」により「産」(お・4)に所属し得る。しかし、下記のように類義の漢字と共に「産」に所属させる。

イ サ シ ハ 空
遣 財 資 嗣 空

このように鍵盤外漢字1字に対して1つのコードを定めて最少コード(基本コード)とする。

最少コードの第一音節を清音で指定し、一段メモリで配置する。但し、同音異字を混同する恐れのある場合だけ濁音、拗音による指定や、二段配置を併用する。例えば前記した「決」(け1)に付属する

シ ソ シ
審 訴 訟

の場合、「訴訟」を二段配置とする。

最少コードは、オペレーターの漢字の知識を整理すれば、鍵盤内漢字の配置から容易に想起できるように、高頻度熟語と意味カテゴリー(意味特徴)により、図2および3.1項の説明のように配置する。

図2の各キー内の数字のうち下段は国立国語研究所報告22「現代雑誌90種の用語用字、第二分冊、漢字表」による漢字の使用頻度の構成比で、

「雑誌90種」の標本全体	280,094	100%
最少コード表全体(1,817字)	260,004	92.8%
15鍵盤内漢字(192字)	121,704	43.5%
15鍵盤外漢字(1,625字)	138,300	49.4%
{その他13.漢数字三四五、 (六、七、八、九、十、百、千、不徳の計)}	10,864	3.9%
合計		98.7%

各段の配分

上段	25.16% (2,928)	3.54%
中段	26.31%	21.84%
下段	16.10%	9.74%
合計	67.57%	29.12%

各指の配分

左手	52.26%	右手	47.74%
小	9.58	中	26.54
中	17.30	人	13.84
人	18.84	人	13.68
小	13.94	小	6.00
合計	9.81	合計	9.81

(396)

表5 最少コマ表の一部(右手中段の上)

鍵盤内	鍵盤外漢字
カ / 会	議 達 司 道 堂 達 社
2 見	達 眺 復 塔 看 每 又
3 重	荷 榎 榎 榎 榎 榎
4 壁	壁 壁 壁 壁 壁 壁
、 / 年	過 去 昨 壯 每 音 翌 齡
2 月	才 戲 伊 英 欧 休 唐 八 巴 朱 蒙 曜 旬 暇
3 日	無 伊 英 欧 休 唐 八 巴 朱 蒙 曜 旬 暇
4 時	刻 暫 隨 閉 閑 閑 閑 閑 閑 閑
、 / 海	荒 瀉 瀉 瀉 瀉 瀉 瀉 瀉 瀉 瀉 瀉 瀉
2 山	浪 灣 越 嶺 嶺 嶺 嶺 嶺 嶺 嶺 嶺 嶺
3 川	石 河 江 次 源 源 源 源 源 源 源 源
4 水	浸 滂 滂 滂 滂 滂 滂 滂 滂 滂 滂 滂
の / 野	球 鼠 鼠 鼠 鼠 鼠 鼠 鼠 鼠 鼠 鼠 鼠
2 火	炒 燗 燗 燗 燗 燗 燗 燗 燗 燗 燗 燗
3 吉	犬 羊 羊 羊 羊 羊 羊 羊 羊 羊 羊 羊
4 木	梅 枝 枝 枝 枝 枝 枝 枝 枝 枝 枝 枝
に / 入	押 介 介 介 介 介 介 介 介 介 介 介
2 加	負 減 減 減 減 減 減 減 減 減 減 減
3 全	完 完 完 完 完 完 完 完 完 完 完 完
4 業	卸 卸 卸 卸 卸 卸 卸 卸 卸 卸 卸 卸
れ / 令	向 勅 勅 勅 勅 勅 勅 勅 勅 勅 勅 勅
2 文	歌 歌 歌 歌 歌 歌 歌 歌 歌 歌 歌 歌
3 言	換 換 換 換 換 換 換 換 換 換 換 換
4 切	碎 碎 碎 碎 碎 碎 碎 碎 碎 碎 碎 碎

表6 漢字頻度集計(国研雑誌90種、12/2)

合計	鍵盤内	鍵盤外漢字
カ	1,077	14 24 18 88 10 40 64 社
2	2,251	150 283 39 20 復 塔 看 7 每 又
3	5,413	1,507 47 荷 榎 榎 榎 榎 榎
4	1,907	415 壁 壁 壁 壁 壁 壁 壁
、	1,726	137 151 25 25 翌 齡
2	8,211	951 134 才 戲 伊 英 欧 休 唐 八 巴 朱 蒙 曜 旬 暇
3	1,193	221 日 無 伊 英 欧 休 唐 八 巴 朱 蒙 曜 旬 暇
4	2,978	130 刻 暫 隨 閉 閑 閑 閑 閑 閑 閑
、	1,216	409 29 瀉 瀉 瀉 瀉 瀉 瀉 瀉 瀉 瀉 瀉 瀉
2	6,827	64 浪 灣 越 嶺 嶺 嶺 嶺 嶺 嶺 嶺 嶺 嶺
3	2,253	970 170 21 18 60 399 172 12 23 14 18
4	1,509	480 317 15 14 16 19 19 19 16 16 16 16
の	827	574 175 鼠 鼠 鼠 鼠 鼠 鼠 鼠 鼠 鼠 鼠 鼠 鼠
2	3,788	727 22 炒 燗 燗 燗 燗 燗 燗 燗 燗 燗 燗
3	1,422	900 犬 羊 羊 羊 羊 羊 羊 羊 羊 羊 羊 羊
4	1,524	441 53 梅 枝 枝 枝 枝 枝 枝 枝 枝 枝 枝 枝
に	1,501	122 152 押 介 介 介 介 介 介 介 介 介 介 介
2	7,264	305 149 負 減 減 減 減 減 減 減 減 減 減 減
3	1,233	527 完 完 完 完 完 完 完 完 完 完 完 完
4	1,187	298 卸 卸 卸 卸 卸 卸 卸 卸 卸 卸 卸 卸
れ	1,091	130 14 向 勅 勅 勅 勅 勅 勅 勅 勅 勅 勅
2	5,326	469 238 84 歌 歌 歌 歌 歌 歌 歌 歌 歌 歌 歌 歌
3	1,217	684 28 28 換 換 換 換 換 換 換 換 換 換 換 換
4	1,170	247 400 碎 碎 碎 碎 碎 碎 碎 碎 碎 碎 碎 碎

5. 参考文献 (本稿1172の照会は Tel 03-878-1424 石川宅へ)

- 特公服58-26574(石川) 特開0857-14935(石川)
- 日本入力研究9-3, 自鳥嘉勇(横須賀通研)
- 望月八百吉「中国語研究学習双書13、中国語と日本語」(光生館, 1974年, 1981年)、文献60 P.132以下。
- 手塚と経産, 1982年1月号, 横山晶一(電研研)

- 国立国語研究所報告56「現代新聞の漢字文献60R245
- 林大地「国語日本語」(角川書店1982年) P.305穴吹義教
- 田中二郎, 山田尚勇「70年代打鍵法による日本語入力法の研究」(東京大学理学部情報科学科, 1978年)
- 国立国語研究所報告37「電子計算機による新聞の語彙調査」P.281
- 国研報告22「現代雑誌九十種の用語用字(2)漢字表」