

日本語のコード入力

神阪博通

三重大学教育学部物理教室

現在の日本語ワード・プロセッサは、プロセッサ本来の、用語・構文分析とそれに基づく文章構成の方向にはない。「文字入力」の段階で足をとられている。日本語も(中国語も)、音素レベルでの正書法を持つべきである(例えばローマ字)。これは、かな漢字交り文の否定ではない。コード入力と対応した正書法が必要ということである。この正書法は単なる「表音表記」ではない。意味の分化を十分に担いうる、「表意的」表音表記である。和語、音・訓、音・字素、音・概念、「漢語」諸コードを「自然」につなぐ表記法である。本稿では、この表記法の規則作りを試みた。この表記法からの入力には5指コードが深く関わる。5指点字も展望できる。

THE INPUT OF JAPANESE TEXTS BY CODE METHOD

Hiromiti KOUZAKA

MIE UNIVERSITY

1515 KAMIHAMA-Cho TU-City MIE-Ken JAPAN

The Japanese texts are written by hieroglyphic style partly. It seems that the input of those texts to computer is very difficult. But this difficulty is mere appearance. If appropriate spellings are formed to those hieroglyphic characters, the same procedure goes on like the case of non-hieroglyphic style. Spelling means code for word. The problem is the formation of human spelling. At this stage, Japanese, also Chinese has an advantage of newly making them in this era of computer. Some rules to make Japanese text code are proposed. These rules are related to 5-finger code and 5-finger braille points.

日本語の一般のワード・プロセッサは世界に例を見ない入力方式をとっている。入力する文字の特定化に、文字系列内での位置の相関を用い、相関の及ばないところを辞書機能で補う方式である。辞書の読みとりは視覚に依る。この方式は、一部では、手が施せないかに見えた日本語の機械入力を果した点で「画期的」と評価されている。ワープロ使用人口は大変に多くなった。しかし、この「かな漢字変換方式」は、用字・用語法を保守的に固めてしまう。また辞書検索は、範囲が限られているとは言え、必要のない文字を確認するのであり、心理的には「雑音」である。文章の書きおろしでは、特に問題になる。ワープロは「完成」したとは言いがたい。

この雑音を避けるためには、より広い文字関連・相関・文脈を用いることになるが、これは、用字・用語の保守性を一層固めることでもある。かな漢字変換が最終的に目指すのは、かな・アルファベット入力による完全な文章記述であり、これは音声入力という「理想」にもつながる。音声入力は、元来がキーボードになじんでいない日本人にとっては、脱キーボードの大変に「良い」目標とされてもいる。

言表・文章は、新しい語系列を構成することで、意味を開拓し、あるいは意味・イメージを分化・確定していく。ここに言語の活力また成長がある。文字列相関・意味関連に頼る入力は、完成すればするほど言語の活性を奪うという矛盾を持っている。また「理想」ともされている音声入力は、それをそのまま文字表記化したのでは、概して貧弱な内容であることも見ておくべきである(原稿読みあげは別)。音声表現の価値は、こうした表記化とは全く別のところにある。現行ワープロは、定形文章にはなじんでいる。この点では実用性も大いにあるが、基本的には、完成し得ないものと見なくてはならない。ここで問題である。表意文字を用いる以上、こうした事態は避けられないのかどうか、である。

現実には、日本語のコード入力、つまり個々の文字にしかるべきコードを付ける努力は一貫して様々な形で行われている。かな漢字変換は日本語入力の唯一の方法ではない。現に漢字他相当数の世界中の文字には JIS コードが付いている。このコードの存在が「かな漢字変換」を可能ならしめてもいるのである。ただし、JIS コードはヒューマンな記号として考えられたものではない。問題は、個々の文字に対して、ヒューマン(自然)なコードをどう付けるかである。

言語・文字の歴史から見れば、コード入力は言語の本質に沿った自然なものであり、十分に可能である。ただ、文字表記を「自然成長」にまかせ

ず、あらためて整理しなおさなくてはならない。ヒューマンな記号系列・コードとは何か、これはキーボードの構造にも関わる。

1. 音素記号表記

全世界で用いられている言語・文字は多様を極めるがその音声・音節は、これはまた極めて限られた数の「音素」に還元されてしまう。人間の発声器官・機構の共通性に基くものではあるが、著しい事実である。どの民族も 30 程度の音素記号があれば、音声・音節の記述に不足はない。この点では音素記号系であるアルファベット、アルファベットの記号系は、人間にとって大変に普遍性の高い記号系である。

この「アルファベット」が文章表記の上でどう位置づけられているかは、各民族、それぞれの歴史があるが、文章が一面音声レベルで読まれるものでもある以上、何らかの音素記号水準での分析が、どの民族においてもなされている。文章表記での「常用」の文字体系とは別に、「アルファベット・音素記号系で記述する文章」は、表意文字民族も含め、どの民族にもあると言わなくてはならない。問題は、常用文字を何にするかが、自然成長のままに放置されていることである。

音素記号の数が 30 程度であることには、人間内部の情報処理の仕方の特徴と思われるが、この点の解明はともかく、音素記号の普遍性から言えば、どの民族も、音素記号に依る文章表記に、それ相当の大きな評価を与えるべきである。この場合、音素記号表記を単に「表音」と見るのは当たらない。表音表記民族にとってもスペリングは表意記号でもある（スペリングは書いて発音はできない症例、また黙読の場合がある）。表意表記民族が、音素記号表記を仮りのものとみて、「読む」にたえる音素記号表記の文章を持つとしないのは間違いである。漢字等の表意文字はこの音素記号表示の確立を基礎とした上で、はじめて安定した使用・運用のできる、より高次の記号系と見るべきである。こう見れば、この高次の機能を、日中、朝と言わず、全世界で利用する展望も出てくる。表音表記諸民族は、自分達のスペリングが「意味」に対するコードだとは必ずしも自覚してはいなかった（きっかけがなかったとも言える）。表音表記民族はあらためて、「文字」つまり「意味」のコード化にせまられていることになる。漢字は「字」とともに、意味を担う「単語」でもある。日本人、中国人（多分に朝鮮人も）にとって必要なことは、この高次記号系である漢字を特定化する「つづり：スペリング」を作り、このつづりによる文章表記が自ずと高次記号系のコード入力でもあるようにすることである。そうした音素記号による文章表記法を確立することである。「効率のいい」入力コードを直接求めるのはヒューマンなアプローチではない。

日本人にとって「ローマ字書き」か「かな漢字交り文」かは、選択の問題ではなく両方必要なのである（どの民族にも固有語の音素記号系の文章表記を確立する課題がある）。

音素記号の数が 30 程度であることは、この記号系にヒューマンな性格を与えている。つまり、つづり構成上の「適度」な冗長度である。音言語自体に既にあるものではあるが、この冗長性は「人間らしい記号運用の誤まり」を許容し、言語の学習、また発展・深化を可能ならしめているのである。ヒューマンさと冗長性は、記号系に関わるヒューマン・インターフェースの問題である。JIS コードはなぜヒューマンな入力コードでないのか、漢字がなぜ難しいと言われるのか、はっきりしない発音でもなぜ通じるのか、他人のつづりの間違いをなぜ指摘できるのか？

高次記号系の果すべき機能は、今までが表意・表音いずれの民族も、両記号系を未分化に扱ってきたので、「視覚的確認の容易さ」あるいは「音声を経ない意味関連の把握」など様々に言われながらも基礎データは少ない（特異な神経障害の症例はあるが）。冗長度と合せ、あらためて研究すべき課題である。

日本語の音素入力コードは、高次記号系を紛れなく呼び出すと共に、それ自体、音素記号文章表記として、日本語発達の現段階にふさわしい、安定した構造を持たなくてはならない。ここに日本人は、既にアルファベット記述の長い歴史を持つ、いわゆる表音表記民族には望むべくもない、今からつづりを定める、という大きな「利点」を持つのである。「つづり」に歴史の「あか」は、全く付いていない。

日本語の「つづり」には、同音語識別のための付加コード(補コード)を付ける。この付加コードは「読み」の上では無声(サイレント)であるが、一定の「意味」を分担する。無声(付加)部分は、つづりの特徴から無声であることが明示されるようにする。これは、教育上、一般に日本語学習上不可欠である。

以下に示す「つづり」は、一応の試みとしてみていただきたい。和語コード、熟語コード、音(訓、字素、概念)コードよりなる。同じ高次系の文字(漢字)に複数のコードがつくが、これは漢字使用の実態のもたらすものである。コードの分ち書きは、教育・学習の場合を除き、原則として行わない。現在の「かな漢字交り文」と同じである。句読点は、文体に属する問題である。

2. 和語コード

日本への漢語の到来は、和語の発達を妨げた面がある。和語の意味の拡大・分化はあまり自律的には行われず、漢語・漢字の力を借用したために、和語は未分化なままに残り、和語の「同音語」は非常に多いことになった。到来・導入当時の文化格差から言ってやむを得ないことではあったが、以来、「同じ和語」に多数の漢語をあてて細かな表現をするようになった。この格差は漢字を本字とみる意識を生み、「かな」の工夫の後も一点一点に意味があるかのような考えはつづいている。しかし、和語は死んだのではなく、漢語にはない諸語の統括能力を持つことにもなっている。和語コ

ードは、この統括性と漢語による意味の分化を表現するつづりである。

一応考えられるコード規則(つづり構成規則)を示したい。

新しいローマ字表記法「かな等価式ローマ字表記法」を定める。この方式でのローマ字列は、かな漢字交り文のかなの部分とは、ほぼ 1-1 に対応する。漢字部分へは、和語コード、熟語コード、音訓・音字素・音概念コード他が、それぞれ固有のつづりで対応する。以下、諸コードを通して漢字を用いる場合(またそのときに限り)、大文字で始める。従って、文頭大文字とはしない。この表記法は、「表意的」表音表記法である。

shi, chi, fu は用いず, si, ti, hu とする。子音を重ねて促音を表すことはせず、「っ」に対応する「c」を用いる。軟音表示のための「ゃ, ゅ, ょ」は直接コード化せず、ローマ字つづり kya, sya, tya, … を, きゃ, しゃ, ちゃ, … のコードとする。カタカナにはカタカナ母音として W を援用する。W は漢字母音のうちの、かな母音と同じものを形式上区別すべく導入したものであり(a, i, …)。外来語に対しては, c, f, v も使う。ツア, ファ, ウァ等である。ツ, ウ, フには cu, vx, fx 等を用いる。「ん, ン, ー, 「, 」, 「.」」は、ローマ字表記上は n, N, 母音 2 重表記, 「, 」, 「.」とするが、入力コード上は, nn, nq, ll, j, f である。カタカナ促音記号は「q」にする。

助詞の「は, へ, を」は、そのまま「ha, he, wo」とする。

コードは先ず、和語の語幹を採る。語幹は、文法変化をしない部分である。動詞語幹のほとんどは子音で終るが、これを、すべての動詞語幹が子音で終るように整理する。つまり、現行ローマ字記法で、原形が、母音が 2 個以上重なって終るものについては、最後の u を w に換える。r-変等、語幹に変化の及ぶ不規則変化動詞の語幹は「変らないもの」とみなす。この語幹の先に必要な補コード(付加コード)を付けてコードを完成させる。文章は、その先さらに、語幹からの「かな使い」を伸ばして記述する。動詞語幹の後の文法変化では先ず母音ないし促音記号 c が続くが、そのうちの u は、ローマ字つづりから省く。動詞から派生した名詞は、動詞の一変化形として扱う。補コードの優先順位は、先ず音、ついで字素、概念である。漢字の音・音節、字素コード、概念コードのつづりは、後節で表して示す。

コード終了記号として x を用いる。一連のローマ字列が、字列が伸びるたびに別の単語に変る頻度はかなり高く、これらを特定するためである。

かな漢字交り文では、すべての和語が漢字となるわけではない。漢字にしたい文字を大文字で始めるのは「区別」としては妥当であるが、他方、この大文字化は、ローマ字書き日本文に、日本文正文としての見易さを与えるのにも使える。この場合は、本当に漢字にする和語には別の記号が必要であり、それには w を利用すればよい。カタカナとの区別は「大文字化」でついている。この大文字化は、漢字・漢語を、日本語に「同化」させる 1 つの方法である(漢字だから目立つのではない)。

記述例を示す。ローマ字つづり，そのコード上の扱い，かな漢字交り文の3通りで示す。

HayasoꝑkuHasir.

H·a·y·a·s·oꝑ·k·u·H·a·s·i·r·f

速く走る。

HayasoꝑkuKurrajitekudasai.

H·a·y·a·s·oꝑ·k·u·K·u·r·r·a·i·i·t·e·k·u·d·a·s·a·i·f

早く来てください。

kokowoToortuꝑcteMigides.

k·o·k·o·w·o·T·o·o·r·t·uꝑ·c·t·e·M·i·g·i·d·e·s·f

ここを通過して右です。

Hayou H·a·y·oꝑ 葉 Hasi H·a·s·i 齒 Hayax H·a·y·a·x 統

Warin W·a·r·iꝑ 輪 Watax W·a·t·a·x 綿

Watasto W·a·t·a·s·t·ꝑ 渡す

Watastoi W·a·t·a·s·t·ꝑ·i 渡し

Watasi W·a·t·a·s·i 私

例文に見る通り，これらが「読める」のは，漢字の音のつづりが直ぐ見てとれることが前提である。この点では，ローマ字表記上も，wꝑばかりでなく，他の漢字母音にも「ひげ」を付ける方がはっきりする。ouꝑ, okꝑ, aiꝑ等である。文章書きおろしのコード入力では問題にならない。

動詞の不規則変化および促音・音便については，つづり上は不規則とせず，読み・発音を不規則とする。

「des. d·e·s·f です。」は特例。

文章の入力は，このローマ字列を逐一追っていくものではない。漢字母音は1単位，1ストロークで処理すべきものである。逐次ローマ字入力では，ストローク数も多く，コード上の紛れも多々生じる。日，中ともに音節が基本的に「子音・母音」の構造であることを利用した，1音節2ストロークのコード入力となるようにプログラム(ソフト)を組む。後述の5指コードも含め，キーは子音と母音の二重の機能を持つ。このソフトで必要とされるキーの総数は，日，中ともに，シフト・キーを含め，12あるいは22にできる。5指コード系では5+1キーである。なお，コード文字を用いた記述は，即コード入力となる。

音で常用され，一般性のある訓がない文字については，名詞の場合，コード終了符を付け音をそのままコードとする。Myakx 脈 Batx 罰 Noux 農 Tenx 点 Senx 線 Menx 面 Zix 字 Bunx 文 Tyoux 腸 Aix 愛 Reix 礼…。動詞では，やはり第1字大文字化の上でs, zを付けて，動詞原形とする(文語的)。Enz 演ず Kous 抗す Kouz 講ず Henz 変ず Tais 対す Syouz 生ず…。これらのコードは，漢語コードである。

3. 漢字の音の「音節」

前節では既に漢字の「音節」を用いた。日本語の中で用いられている漢字の音を、実際の発音とは別に、ローマ字表記上1音節化したものが「音節」である。中国では、かつて今も漢字1字は1音節で発音されているが、日本に入ってきた音の一部は2音節で発音されるようになった。しかし、かつて1音節であった規則性は残っている。表記上「1音節」の形に整理することは可能であり、これは既述の和語コード、後節で扱う熟語コード、音(訓, 字素, 概念)コードすべてに用いる。

漢字の「音節」は日本語の拡大50音図の中に位置づける。漢字の母音自体、大変に規則的な構造を持っている。日本語で用いられる母音を示す。

かな母音	W	a	i	u	e	o
単(短)母音	W ₇	a ₇	i ₇	u ₇	e ₇	o ₇
二重(長)母音		ai	ui	uu	ei	ou
~k(k末)		ak	ik	uk	ek	ok
~n(n末)		an	in	un	en	on
~t(t末)		at	it	ut	et	ot

W と W₇ は音価は同じであるが、つづり・コードの上では区別する。

現代中国語(普通話)の母音の構造は、下記のように整理できる。音韻学の立場からではなく、Y, W も子音とみて、つづりの上で必要とされる母音である。中国語入力でのキー構成を考えている。3個所、同一キーに異なる母音をわりふっている(シフトではない)。

a	i	u	e	o
ai	ui·ü	ua·ue·üe	ei	ou
ia	iu	ao	ie	uo
an	in	un	en	uan
ang	ing	iang·uang	eng	ong·iong
iao·uai	ian			

一般に軟音の表示には、軟子音あるいは軟母音の2通りの方法がある。かな等価式ローマ字の立場では、漢字音節に対しては軟母音の立場をとる。

ya, yak yui yu, yuu yuk yun yut yq, you yok

の11個である。中国語の表記法(pinyin)では音の硬軟は表に出ない。日本語の拡大50音図は、この計41個の母音と子音の組合せ表である。日本語のローマ字表記は、漢字音節の明示がなければ生きてこない。

日本語「音節」表 (拡大 50 音図)

a	ka	ga	sa	za	ta	da	na	ha	ma	ra	wa	ba	pa
a ₇	ka ₇	ga ₇	sa ₇	za ₇	ta ₇	da ₇	na ₇	ha ₇	ma ₇	ra ₇	wa ₇	ba ₇	pa ₇
ai	kai	gai	sai	zai	tai	dai	nai	hai	mai	rai	wai	bai	pai
ak	kak	gak	sak		tak	dak		hak	mak	rak	wak	bak	pak
an	kan	gan	san	zan	tan	dan	nan	han	man	ran	wan	ban	pan
at	kat	gat	sat	zat	tat	dat	nat	hat	mat	rat		bat	pat
ya ₇			sa ₇	za ₇	ta ₇								
yak	kyak	gyak	syak	zyak	tyak		nyak	hyak	myak	ryak		byak	pyak
i	ki	gi	si	zi	ti	di	ni	hi	mi	ri	wi	bi	pi
i ₇	ki ₇	gi ₇	si ₇	zi ₇	ti ₇		ni ₇	hi ₇	mi ₇	ri ₇		bi ₇	pi ₇
(ui)			sui	zui	tui					ru			
ik	kik		sik	zik	tik		nik	hik		rik		bik	pik
in	kin	gin	sin	zin	tin		nin	hin	min	rin		bin	pin
it	kit		sit	zit	tit		nit	hit	mit	rit			pit
yui													
u	ku	gu	su	zu	tu	du	nu	hu	mu	ru		bu	pu
u ₇	ku ₇	gu ₇	su ₇	zu ₇			nu ₇	hu ₇	mu ₇	ru ₇		bu ₇	pu ₇
uu	kuu	guu	suu		tuu			huu					
(uk)								huk					puk
un	kun	gun	sun					hun				bun	pun
ut								hut				but	put
yu ₇			syu ₇	zyu ₇	tyu ₇								
yuu	kyuu	gyuu	syuu	zyuu	tyuu		nyuu			ryuu		byuu	
(yuk)			syuk	zyuk									
(yun)			syun	zyun									
(yut)			syut	zyut									
e	ke	ge	se	ze	te	de	ne	he	me	re	we	be	pe
e ₇	ke ₇	ge ₇	se ₇	ze ₇									
ei	kei	gei	sei	zei	tei	dei	nei	hei	mei	rei		bei	pei
ek		gek	sek	zek	tek	dek		hek		rek		bek	pek
en	ken	gen	sen	zen	ten	den	nen	hen	men	ren		ben	pen
et	ket	get	set	zet	tet	det	net		met	ret		bet	
o	ko	go	so	zo	to	do	no	ho	mo	ro	wo	bo	po
o ₇	ko ₇	go ₇	so ₇		to ₇	do ₇		ho ₇	mo ₇	ro ₇		bo ₇	po ₇
ou	kou	gou	sou	zou	tou	dou	nou	hou	mou	rou		bou	pou
ok	kok	gok	sok	zok	tok	dok		hok	mok	rok		bok	pok
on	kon	gon	son	zon	ton	don		hon	mon	ron		bon	pon
ot	kot		sot		tot			hot	mot			bot	
yo ₇	kyo ₇	gyo ₇	so ₇	zo ₇	to ₇		nyo ₇			ryo ₇			
you	kyou	gyou	syou	zyou	tyou		nyou	hyou	myou	ryou		hyou	pyou
yok	kyok	gyok	syok	zyok	tyok					ryok			

4. 熟語コード

熟語内の漢字は、和語、漢語とも、そのつづりだけで、ほとんどが特定される。熟語内の文字は、単独であるときよりは意味が限定されているからである。同音語の競合を避けて来た「歴史」もある。

熟語、特に漢字熟語については、その語の第何字目を指定する「1字呼出し法」が成立する。単独では先ず用いられない文字の呼出しには有効である。

和語の熟語については、第2番目以降の漢字にやはり大文字化規則を用いるのかどうか、つづりの字面の上で一考の余地がある。w 化の方法も現実的である。和漢混成熟語は和語コード的に扱う。

熟語コードは、一応コード的側面を持つてはいるが、基本的には用語登録である。どの範囲の熟語・用語を用意すべきかは、分野また個人的にかなり違ってくる。この点では同音熟語の区別に用いる補コード構成には、あまり厳密な一般則はたてられない。便宜性優先でよい。俗手ともいふべき、化学 Kagakba・Kagakch、科学 Kagaksc、なども認められてよい。一応の補コード規則として、第1字の音(訓、字素、概念)コード・漢語コード(次節)、これが区別に無効であれば第2字で比較、などが考えられる。

同音熟語の区別にわずらわしい補コード規則を作るよりは、同音語の多い熟語は熟語としての熟成度が低いことも見て、登録からは外し、次節で述べる1字呼出しコードの活用を計る方がよい。

熟語 Zyukgo、分野 Bunya、概念 Gainen 重要 Zyuuyou
妥 1Dakyou 協 2Dakyou
呼出し YobiDasi, Yobidasi 刺身 SasiMi, Sasimi
普通 Hutuuan 不通 Hutuar
重箱 Zyuubako, Zyuubako 仕事 SiGoto, Sigoto
度合い DoAi, Doai 都落ち MiyakoOti
名古屋 NaGoYa, Naggya 浅草 AsaKusa, Asakusa

文章入力の手速・軽快さは、この熟語登録で大幅に上がるが、それは妥当な登録数の場合である。過度の登録は、その同音語判別の補コードが足かせになる。他の諸コードとのバランスが肝心であるが、このバランスは個人個人の文体でもある。

5. 音訓、音字素、音概念コード

日本へ入って来た文字：漢字は音を伴っていた。文字だけの借用ではない。「外国語」そのものの導入である。それだけの文化格差があった。このために、音から、日本語の一部として生きている漢語：漢字を呼出す道もあることになるが、音だけでは漢字はほとんど特定しない。「音節」の後に、訓による補コードを付ける「音訓コード」が成立する。一応考えられる、同音内の漢字を区別する補コード規則を示す。

以下で用いる記号 C, W は、子音、かな母音を示す。

(1) 訓(和語)の、かな等価式ローマ字表記の第 1・2 桁 (CW, WC, WW)。

(2) 前項で区別されなければ、さらに第 2・3 桁、以下必要に応じて桁をずらす。語幹を使いきった場合は、語幹の最後の桁と x で補コードとする。この桁ずらしは、既に確定している補コードとの競合がないようにするが、第 1・2 桁が異なっておれば、双方排除でなく、いずれか一方を残す(そういうプログラムにする)。どれを残すかでは、使用頻度が一応の尺度。この場合、コード構成の一意性が失われるが、試行学習方式の入力アルゴリズムをとれば、この非一意性は表には出てこない。

桁ずらしは第 3・4 桁までで止どめ、以下の字素・概念コードへ移るのも有力な方法である(本論では採用していないが)。

(3) 訓で決着がつかない場合は、漢字の字素(大略は部首)のコードに移る。字素コードの表を示す。字素コード構成規則は、それ自身も一字をなし得る字素についてはその音、音が同じであれば和語の頭 2 桁、が原則である。ここでの和語は部首の通称。このコードの第 1 桁は大文字にする。字面はよくないが、音訓、熟語コードとの混コードを避けるためである。部首に対する「音」の採用は、異和感もあるが、中国語も含めた立場では、音字素コードは、漢字コードの「本流」と思われる。

表中、同じコードが異なる字素にわりふられているが、問題は同音内の区別であり支障はない。

(4) 字素まで行っても適当なコードが得られない場合は、以下の概念コードを用いる。

zb : 自然(有機, 無機とも)・動植物(この諸器官)・鉱物・・・・

dm : 人工(造営)物・道具・建物・・・・

bn : 文化・文字・文章・出版・情報・技芸・・・・

cp : 社会(人間)関係・政治・経済・身分・親族・・・・

cq : 抽象概念・数・数量・尺度・論理・幾何・・・・

ともかく、これらの項目のどれに「関連」するかのファジィな判断をする。字素コードも、実質的には、部首が意符である点で、拡大運用が考えられる。Syok 食いもの・食事関連、等である(本では採用していない)。

音のまま日本語に定着した漢語については、既に和語化したものとみて、一応和語コードの節で扱った。しかし、本節の(3)、(4)項の扱いもできる。どちらが妥当かは判断しづらい。両コード併用も可である。

日本語の中の漢語については、単にコード化という以上の問題がある。文字ばかりでなく、漢語的表言(言いまわし)・造語が日本語の中に占める重みは大変に大きい。熟語ならざる「熟語」も多々現われる。音で定着した漢語も含め、「漢語コード」として扱う。

字素コード表

乙(レ)	ot	おつしよう	人(人)	Nin	にんべん	立	Rit	たつへん	彳	Ns	にすい
儿	Nn	にんによう	口	Ku	くちへん	米	Bei	こめへん	彳水	Sz	さんずい
几		きしよう	土	Dq	つちへん	糸	It	いとへん	リ刀	Rt	りっとう
父(父)		すいによう	女	Zyq	おんなへん	缶		ほとぎへん	彳	San	さんづくり
支	En	いんによう	山	San	やまへん	糸(糸)		らいすき	彳	Bat	はんづくり
支	Sa	しによう	巾	Hb	はばへん	耳	Zi	みみへん	父(父)	Bok	ぼくづくり
无		むによう	弓	Kyuu	ゆみへん	舌	Zet	したへん	斤	On	おのづくり
爪		そうによう	彳	Gn	ぎょうにんべん	舟	Syuu	ふねへん	斗	Tq	とます
走	Sou	そうによう	心	Rs	りっしんべん	虫	Tyuu	むしへん	欠	Ket	けんづくり
走	Sin	しんにゆう	手	Te	てへん	衣	/i	ころもへん	欠	Rn	るまた
鬼	Ki	きによう	方	Hou	かたへん	角	Kak	つのへん	皮	Hj	けがわ
表(表)		ばくによう	日	Nit	ひへん	言	Gen	ごんべん	羽	/y	はねづくり
文	Bun	ぶんによう	日		ひらび	豆	Tou	まめへん	韋		ふでづくり
一	Nb	なべふた	月	Get	つきへん	豕	Nk	いのこへん	卩	Oz	おおざと
一	Wk	わかんむり	木	Ki	きへん	豕	Mz	むじなへん	卩(心)	Set	ふしづくり
一	Uk	うかんむり	止	To	とめへん	貝	Bai	かいへん	佳	Ht	ふるとり
穴	An	あなかんむり	歹	Gat	がつへん	足(足)	Sok	あしへん	頁	Oo	おおがい
竹	Tik	たけかんむり	火	Ka	ひへん	身	Sin	みへん			
老(老)	Rou	おいかんむり	片	Hen	かたへん	車	Syq	くるまへん			
一	Kk	くさかんむり	牙	Ga	きばへん	酉	Tr	とりへん	口	Kok	くがまえ
虎	Tk	とらかんむり	牛	Gyuu	うしへん	采		のこめへん	口		まきがまえ
雨	/y	あめかんむり	犬	In	いぬへん	里	Ri	さとへん	口		かんがまえ
麻	Ma	あさかんむり	彳	Km	けものへん	金	Ka	かねへん	口	Ku	はこがまえ
		つめかんむり	王(王)	Syuu	しょうへん	卩	Kz	こざとへん	口		かくしがまえ
			王	/ou	たまへん	革	Kak	かわへん	气		きがまえ
			目	Me	めへん	韋(韋)		なめしがわ	行		ぎょうがまえ
八	Hat	はちがしら	矛	Mu	ほこへん	食	Syok	しょくへん	門	Mon	もんがまえ
五(五)		けいがしら	矢	Ya	やへん	馬	Ba	うまへん	門		たたかいがまえ
六	Hg	はつがしら	石	Sek	いしへん	骨	Kot	ほねへん	勺	Tt	つつみがまえ
七(七)	Ag	あみがしら	示	Si	しめすへん	魚	Gyq	うおへん			
影	Kg	かみがしら	禾	Ng	のぎへん	鳥	Tyou	とりへん	尸	Sk	しかばねかんむり
			羊	You	ひつじへん	鼻	Bi	はなへん	尸	Md	まだれ
						齒	Ha	はへん	尸	Kq	とだれ
皿	Sa	さら							尸	Yd	やまいだれ
子	Ko	こ							尸	Gd	がんだれ
力	Rik										
見	Ken										
又	Yuu										

7. おわりに

この「かな等価式ローマ字表記」による文章記述から、かな漢字交り文への変換は、実際には5指コードによるのが最も簡便であり、かつ本質的なことではあるが、詳論は別とする。日本語に音素水準(ローマ字等)の表記法が確立され、高次記号系の自由さが生きてくることを望みたい。ローマ字表記で、漢字書かないし助詞以外の単語の第1字を大文字化する規則は、現行かな漢字交り文での、ひらかなと漢字のコントラストに対応するものである。この方法は、中国語へも通じる(介詞は小文字で)。

コード入力になじむ、この表記方式の確立で、日本語の「ワー・プロ」も、用語・構文の解析・制御を課題とできる。

- (1) 神阪博通 5指コードおよび5キー送受信系 私信(1988)
- (2) 神阪博通 点字および5指コード文字 私信(1988)
- (3) 神阪博通 5指コードによる日本語入力
情報処理学会 日本語文書処理 10-1 (1987)