

# VISIBLE LANGUAGE

---

*The quarterly concerned with all that is involved in our being literate*

Volume XVIII Number 3 Summer 1984 ISSN 0022-2224

## *Special Issue*

### ASPECTS OF THE JAPANESE WRITING SYSTEM

Guest Editor: Chris Seeley

- 213 Introduction  
Chris Seeley
- 219 Aspects of the Graphological Structure of Japanese  
A. E. Backhouse
- 229 The Adoption of Punctuation in Japanese Script  
Nannette Twine
- 238 Japanese Orthography in the Computer Age  
J. Marshall Unger
- 254 Japanese Braille  
J. Marshall Unger
- 267 The Japanese Script since 1900  
Chris Seeley
- 303 The Authors

Visible Language, Volume XVIII, Number 3, Summer 1984. Published quarterly in January, April, July, and October. Postmaster: send address changes to Visible Language, Box 1972 CMA, Cleveland, OH 44106. Copyright 1984 by Visible Language. Second-class postage paid at Cleveland, Ohio, and at additional mailing offices.

Merald E. Wrolstad, Ph.D., Editor and Publisher  
P.O. Box 1972 CMA, Cleveland, OH 44106

### **Advisory Board**

Colin Banks, Banks and Miles, London  
Naomi S. Baron, Brown University  
Fernand Baudin, Bonlez par Grez-Doiceau, Belgium  
George Bauer, University of Southern California  
Pieter Brattinga, Form Meditation International, Amsterdam  
Gunnlaugur SE Breim, London  
Dennis Fisher, Aberdeen Proving Ground , Maryland  
I.J. Gelb, Oriental Institute, University of Chicago  
James Hartley, University of Keele, England  
Albert Kapr, Hochschule für Grafik und Buchkunst, Leipzig  
Paul Kolers, University of Toronto  
C.L. Lehman, Tigard School District, Oregon  
Dominic Massaro, University of California, Santa Cruz  
Peter Mosenthal, Syracuse University  
Alexander Nesbitt, Newport, Rhode Island  
Thomas Ockerse, Rhode Island School of Design  
P. David Pearson, University of Minnesota  
Sharon H. Poggenpohl, Middletown, Pennsylvania  
Marvin A. Powell, Jr., Northern Illinois University  
Wayne Shebilske, University of Virginia  
Jack W. Stauffacher, The Greenwood Press, San Francisco  
William C. Stokoe, Jr., Galladet College, Washington  
Michael Twyman, University of Reading  
Richard Venezky, University of Delaware  
Peter Wason, University of London  
Dirk Wendt, Christian-Albrechts-Universität, Kiel  
Hermann Zapf, Darmstadt, Germany

Authorization to photocopy items for internal or personal use, or the internal or personal use of specific clients, is granted by Visible Language for libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that the base fee of \$1.00 per article, plus \$0.10 per page is paid directly to CCC, 21 Congress Street, MA 01970. 0022-2224/84 \$1.00 + .10.

## General Information

*Visible Language* is concerned with research and ideas that help define the unique role and properties of written language. It is a basic premise of the Journal that writing/reading form an autonomous system of language expression which must be defined and developed on its own terms.

Published quarterly since 1967, *Visible Language* has no formal organizational affiliation. All communications should be addressed to

Visible Language Telephone 216/421-7340  
Box 1972 CMA  
Cleveland, OH 44106 USA

### Subscription Rates.

	<i>One Year</i>	<i>Two Years</i>	<i>Three Years</i>
Individual subscription	\$15.00	\$28.00	\$39.00
Institutional subscription	25.00	47.00	66.00

Foreign subscribers: add \$2.00 for postage to each year's subscription. All orders must be prepaid. To be honored free of charge, claims for missing issues must be made immediately upon receipt of the next published issue.

**Back Copies and Reprints.** A limited quantity of all back numbers is available at a per issue cost of \$4.00 to individuals and \$5.00 to institutions. A booklet listing the contents of all past Journal issues is available on request. Individual reprints are not available.

**Manuscripts.** Manuscripts, inquiries about research articles, and other contributions to the Journal should be addressed to the Editor. A guide for the organization, preparation, and submission of manuscripts is available. Three copies of a manuscript should be accompanied by an abstract typed on a separate sheet of paper.

**Letters to the Editor.** Comments on articles and letters that have appeared in the Journal are welcome and should be addressed to the Editor. The Editor will also relay to the author questions or comments on any article. Your response—and the author's comment in reply— will not be published without your permission and your approval of any editing.

**Advertising.** Detailed information is available upon request.

The calligraphy on the cover is the title of the special issue in a normal combination of kanji, hiragana, and katakana. It was written by Chizuko Takaoka of Cleveland. Japanese characters interspersed throughout the text are by Hiroko Aikawa.

## Aspects of the Japanese Writing System

---

**Edited by Chris Seeley**

Department of Asian Languages  
University of Canterbury, New Zealand

In the late 1940s the Japanese writing system underwent a series of reforms. Despite these reforms, which were successful in bringing about a substantial degree of simplification, the writing system for modern Japanese still retains the dubious distinction of being the most complicated in use in the modern world.

The intricate nature of the system derives partly from the large number of relatively intricate written signs which a logogram script such as kanji (Chinese characters) inevitably requires, partly from the fact that kanji in Japanese regularly have more than one way of being read (how a kanji is to be read in any given case is determined on the basis of context), and partly from the pluralistic nature of the writing system (kanji together with two Japanese syllabaries). The situation is exacerbated by kanji usage for names, which traditionally frequently involve special readings or obscure kanji. In recent years legislation has been passed which simplifies kanji usage for given names of new-born children, but difficulties are still posed by family names and place names, not to mention the given names of deceased persons or individuals of the older generation.

In this special issue five aspects of the Japanese writing system are discussed. Anthony Backhouse (Monash University) discusses certain general features of the writing system in a way which should be of interest to specialist and non-specialist alike. Nanette Twine (Griffith University) depicts how, under Western influence, punctuation was adopted in Japanese texts in the late nineteenth and early twentieth century. James Unger (University of Hawaii at Manoa) has contributed two articles: one dealing with the problem of computer input in Japanese, the other describing a specialised variety of written Japanese—Japanese braille. The article by Chris Seeley (University of Canterbury) deals with modifications to the Japanese script since 1900. The guest editor of this special issue is

much indebted to Jim Unger for his energetic correspondence on several of these topics.

Down to the present, little has been written in the West on the Japanese writing system other than basic descriptions such as are found in Japanese language textbooks for Westerners. It is hoped, therefore, that the present issue of *Visible Language* will serve to focus attention on and stimulate interest in this general topic.

For the guidance of general readers who are not familiar with the Japanese language or its writing system, the five articles in this issue are preceded by a very brief sketch of the history of writing in Japan, followed by a short description of the basic features of modern Japanese writing, plus one or two observations on public opinion in Japan on the subject of the writing system.

A small number of frequently occurring romanized Japanese words (e.g., kanji, kana) are as a rule not italicized here, nor are the names of major Japanese organizations and institutions.

For further information on the Japanese language and its writing system, see R.A. Miller, *The Japanese Language* (University of Chicago Press, 1967); for the writing system alone, see the Introduction to W. Hadamitzky & M. Spahn, *Kanji and Kana: A Handbook and Dictionary of the Japanese Writing System* (Charles E. Tuttle & Co., 1981); for script reform in late nineteenth-century Japan, see N. Twine, "Toward Simplicity: Script Reform Movements in the Meiji Period".

## Historical Context

Originally the Japanese had no writing system of their own, and so at an early stage in relations with China (ca. 400 AD) the system used by the Chinese was borrowed. The Chinese script is based on the morphemic principle; that is, each 'character' (Jap.: kanji) represents in writing a unit of the language such as *ren* "person", *da* "big", and *shui* "water". The Chinese and Japanese languages are completely different in structure, and so although the Chinese script was adequate as a means of representing the Chinese language in writing, it was ill-suited when it came to representing Japanese, a language with a wide variety of inflectional categories.

Because of the difficulties involved, it appears that initially at least the Chinese script was employed in Japan to write only Chinese, not Japanese. At that earliest stage the task of reading and writing was entrusted to scribes who were immigrants to Japan from the Korean peninsula and China or the descendants of such persons. As time passed, gradually a small number of Japanese acquired the literate skills. According to the

*Nihon shoki* (720), native Japanese scholars were already taking an active part in the compilation of at least some documents by the late seventh century. Even from the earliest period, when writing in Chinese style in Japan there naturally arose the need to represent Japanese proper names in texts in some way. This was done by utilising a technique already in use in China, namely, using kanji not for their meaning, but for their sound value—i.e., as phonograms. Thus, for example, the kanji 伊 and 太 (*i* and *t'ai* respectively in Middle Chinese, the literary variety of Chinese which formed a standard dialect during the seventh century) were borrowed to represent the Japanese proper noun *Ita-* (incomplete name to be found on a Japanese sword inscription, ca. sixth century). From about 700 onwards this method was sometimes employed to write not just isolated names but continuous sequences of Japanese. Use of kanji in this way was, however, rather cumbersome, as a relatively complicated sign was required to represent each syllable of Japanese. In addition to this, there was a lack of prestige attaching to the Japanese written style, and for some time the Japanese continued to write largely either in Chinese style, or in a hybrid style which showed the linguistic influence of both Chinese and Japanese.

At the outset Chinese texts are thought to have been read in Japan as Chinese. As small numbers of native Japanese began to read and write, they tended to accompany the Chinese rendering by a Japanese translation, rearranging the order of characters as appropriate and supplying additional elements required by the conventions of Japanese grammar. At first the process of rearranging a Chinese text was done mentally, but from about 800 Japanese Buddhist priests began to add kanji in their phonogram function in order to serve as a memory aid to help with the Japanese rendering. In the oldest surviving texts marked in this way the phonograms are unabbreviated ('full') kanji, but very soon abbreviated forms were being used also. Two basic methods were used for abbreviation: cursivizing and isolating techniques (the isolating technique is so called here because graphs of this type were created by isolating and using just one part of a 'parent' graph which was more complicated in shape). The following are a few examples of abbreviated phonograms taken from a manuscript of the *Jōjitsuron* (a Chinese Buddhist text) annotated in Japan in 828:

isolating phonogram	syllable represented
尹 (derived from 伊)	<i>i</i>
ム (derived from 牟)	<i>mu</i>
才 (derived from 才)	<i>o</i>

cursivized phonogram	syllable represented
ぬ (derived from 奴 )	nu
に (derived from 為 )	wi
え (derived from 江 )	ye

The *Jōjitsuron* is annotated with both isolating and cursivized phonograms. Before long, however, the former (forerunners of signs in the modern kata-kana syllabary) predominated in the glosses to Chinese Buddhist texts, while the latter (forerunners of the hiragana syllabary) became increasingly cursivized and simpler in shape (a development which took place mainly in the realm of secular texts), and came to be used quite extensively for writing narrative prose (*monogatari*), private letters, *waka* poetry, and other informal texts in the Japanese style.

The development of the two kana syllabaries represents a very important milestone in the history of writing in Japan. However, for the majority of texts of an official or scholarly nature, the Chinese or hybrid styles continued to be preferred (cf. the use of Latin for scholarly writings in mediaeval Europe).

### Modern Japanese Writing

Early in the first part of the twentieth century a written style based on modern colloquial Japanese came into widespread use, replacing earlier archaic styles. The normal orthography for modern Japanese texts is a combination of kanji and kana.

Kanji are related to the morphemes of Japanese by a system of 'readings'. Readings are of two types: *on* and *kun*. A *kun* reading represents a Japanese morpheme corresponding to the meaning of a particular kanji and which has come to be associated with that kanji, for instance 東 "east" *higashi*, 馬 "horse" *uma*.

The *on* reading of a kanji is based on its pronunciation in Chinese. The *on* readings of the above two kanji are *tō* and *ba* respectively, and these occur in *Tōkyō* 東京 "Tokyo" (lit. "eastern capital") and *bariki* 馬力 "horse-power" (the second kanji used in writing *bariki* has the meaning "strength"). Words of this type, which involve morphemes of Chinese origin, tend to be written in kanji rather than kana. Kanji are also frequently employed to write certain common native Japanese nouns, and verb and adjective stems.

Kana (Japanese syllabic signs) are of two types: *katakana* and *hiragana*. *Katakana*, which are angular in appearance, are employed for writing loanwords (other than those of Chinese origin), and have several other specific uses. *Hiragana*, which have a flowing visual aspect, constitute the more widely-used syllabary, and are employed to write inflectional

endings, grammatical function-words (particles), some nouns, etc. Since late 1946 a system of kana usage based on the pronunciation in modern standard Japanese has been in use; this replaced an earlier unwieldy system based on historical principles.

The way in which kanji and kana are combined in writing modern Japanese is perhaps best illustrated by and actual example such as the following: 弟さんはいつその自動車を買いましてか

*Otōto-san wa itsu sono jidōsha o kaimashita ka* ("When did your younger brother buy that new car?). Kanji are used here to represent the Sino-Japanese noun *jidōsha* "car", the (native Japanese) noun *otōto* "younger brother", and for the stems of *atarashii* "new" and *kau* "buy". Kana are used for: the polite suffix *-san*, *itsu* "when?", the particles *o* and *ka*, the demonstrative *sono* "that", and the endings *-shii* and *-mashita*. The above Japanese sentence would typically be written in the way just described, but there is some fluctuation in actual usage (see the article by Anthony Backhouse.)

From late 1946 the number of kanji in general use was considerably restricted. As a result of this simplification, certain Sino-Japanese words which were traditionally written with very uncommon kanji were in some cases rewritten with homophonous kanji having a higher frequency of use. In cases where no suitable replacement kanji were available, such items were either subsequently written in kana, or tended to drop out of common use.

The ensuing decline in the use of certain less common Sino-Japanese words has led some writers and other intellectuals to regard the postwar orthographic reforms as an unfortunate set of moves leading to a deterioration in the general standard of writing and learning. Representative of this group is the Japan Literati Association, which is in favour of at least about 3000 kanji being adopted for general use instead of the present list of 1945 kanji (see the article by Chris Seeley).

Certain other individuals and groups, on the other hand, regard the writing system as being too complicated still, and would like to see a reduction in the number of kanji in everyday use. Generally speaking, school teachers, who are confronted every day with pupils who are severely pressed to assimilate the prescribed curriculum, tend towards this view.

Given these opposing viewpoints, it would be impossible to modify the Japanese writing system in such a way as to please all parties. There does, though, appear to exist a fairly widespread feeling that the number of kanji in general use should not exceed 2000, and in this sense the current kanji list (the *Jōyō kanji* List) may be said to reflect the mood of a substantial proportion of the Japanese people.

The current kanji list contains almost one hundred more kanji than its predecessor which was in use for 35 years from 1946. While this represents only a small overall increase, shortly after World War II the Japanese considered *reducing* the number of kanji in common use to about 1000. The question of how this recent development is to be interpreted is a difficult one, but it may perhaps be appropriate to regard this new trend as being equivalent to a vote of confidence by the Japanese of the 1980s in their own writing system, complicated though it is.

# Aspects of the Graphological Structure of Japanese

---

**A. E. Backhouse**

Monash University, Australia

Certain important general features of the graphological structure of modern Japanese are reviewed. The Japanese writing system is characterized in particular by its plurality of scripts and by the distinctive nature of the kanji (character) script. Three major structural repercussions of these properties in the normal orthography are surveyed: firstly, the differential function of the various scripts in distinguishing grammatical from lexical items and, within the lexicon, in marking membership of different lexical strata; secondly, the integrative role of the kanji script in motivating, by way of graphological linkage, the learned Sino-Japanese vocabulary in terms of the neutral native vocabulary; and finally, the differential role played by kanji in two important types of homophony.

1. The aim of this article is to describe some of the major structural features of the modern Japanese written language as they are manifested at the level of graphology. As the work of Vachek has long since made clear, a written language is properly viewed as an entity in its own right, structured and patterned in ways which are at least partly independent of the corresponding spoken language.<sup>1</sup> Our central concern is with the place occupied in this patterning by the graphological resources of the language: how does the Japanese writing system serve to reflect, and in some cases create, structure in the written language?

2. The main features of the Japanese writing system which will concern us are, firstly, its plurality of scripts, and secondly, the nature of the kanji (character) script. For the benefit of readers unfamiliar with the system, these features may be briefly summarized.

Written Japanese employs three basic scripts, one morphographic (kanji) and two phonographic (hiragana and katakana). To these have been added, in more recent times, Arabic numerals (used in horizontal writing in place of kanji numerals) and Roman capitals (cf. SL 'steam locomotive', OL 'office girl', 3DK 'having three rooms plus a dining room-cum-kitchen', etc.<sup>2</sup>). As in Chinese,<sup>3</sup> characters in Japanese basically constitute a morphemic script, the major difference being that in Japanese a given character commonly represents several morphemes rather than a single

morpheme. The prototypical situation may be illustrated by the kanji 川, corresponding to the pair of morphemes *sen* 'river' and *kawa* 'river'. This kanji is typical in that (a) it represents two morphemes, (b) etymologically, both Sino-Japanese (*sen*) and native (*kawa*) morphemes are involved, and (c) the morphemes are synonymous. However, many kanji deviate along one or more dimensions of this pattern.

3. Before looking at more particular structural features, it is worth pointing out certain general consequences for the Japanese written language which follow from the nature of the writing system as outlined above.

Firstly, the presence of a plurality of scripts makes for a potential flexibility of orthography on a scale which is inconceivable in the case of more familiar writing systems. Thus, while written Japanese normally combines the three basic scripts referred to, the language can theoretically be written purely in kana (but see 4.3 below); indeed it is so written in, for example, certain types of children's literature (hiragana), in telegrams (katakana), etc. Aside from such extreme cases this kind of flexibility is quite generally exploited in various degrees in different written genres, and the stylistic implications of this variation deserve closer study than they have received. In discussing the system here we shall appeal to the notion of 'normal' orthographic practice as set out in Government recommendations and reflected in dictionaries based thereon, but the underlying variation should not be forgotten. Note that some such variation is found even in such highly standardized areas as newspapers: a sample based on one year's editions (1966) of the *Asahi*, *Mainichi*, and *Yomiuri* newspapers showed orthographic variation in over 14% of words, and almost two-thirds of this total involved alternation between kanji and one of the kana scripts.<sup>4</sup>

Secondly, the coexistence of different scripts has led to the not uncommon practice of using two scripts in tandem, one acting as a gloss or commentary on the other. (This immediately adds a further dimension to the variation referred to above: as well as being written in one script or another, a word may also be written in two scripts at once.) The central case here is the use of one of the kana scripts (usually hiragana) alongside individual kanji to provide additional identification of the morpheme in difficult or ambiguous cases, a practice which prompted Sansom's well-known characterization of Japanese writing.<sup>5</sup> But the facility is exploited beyond this central case. For example, hiragana (as the script learned earliest) is used to gloss katakana in children's literature; and more importantly, the superimposition of kana on kanji is extended from the basic function of specifying the correct 'existing' reading to that of

appending a 'new' reading, thereby effecting a kind of amalgam, in the written medium, of two distinct linguistic items.<sup>6</sup>

Finally, from a more diachronic perspective, it is worth commenting briefly here on the highly open-ended nature of the Japanese writing system. Borrowing and change operate in this area of language as in others, and clearly no writing system can be described as fully closed in this sense. However the nature of the Japanese system appears to render it more open than most. In a broad sense this is reflected in the incorporation, in relatively recent times, of Arabic numerals and Roman capitals as new borrowed inventories in the system. More narrowly, the kanji script itself is open-ended in that it permits the construction of new graphemes on the basis of the structural principles inherent in the script. Many such graphemes have, of course, been coined in Japan over the centuries, and several are established elements of the normal modern orthography. A further recent illustration of this open-endedness will be mentioned in 4.1 below.

**4.** Under the sub-headings in this section we propose to review some of the more prominent structural features of modern written Japanese as they involve graphology. We must emphasize again that, as indicated above, our discussion and illustrations will in general be based on 'normal' orthographic practice as set out in recent Government recommendations.<sup>7</sup>

**4.1 Stratification.** Our concern here is with the differential function of the scripts. The Japanese writing system offers particularly clear opportunities for marking structural divisions, and in modern usage the three basic scripts serve to distinguish, firstly, grammatical and lexical elements and, secondly, different strata within the lexicon.

The first general principle is that grammatical elements are written in hiragana. This applies most clearly to core grammatical elements such as enclitics and inflectional affixes but extends well beyond these to take in most items belonging to closed systems. Examples include spatial deictics (*kore* 'this', *sore* 'that', *are* 'that', etc.), auxiliary verbs (*miru* 'try doing', *oku* 'do for future use', etc.), and conjunctions (*shikashi* 'but', etc.). Inconsistencies are certainly present (for example, *shikashi* 'but' is normally written in hiragana, but *tadashi* 'however' is optionally written in kanji) and the borderline between grammar and lexicon is in any case indistinct, but there can be no doubt about the strength of this principle in modern written Japanese.<sup>8</sup> Note, too, that it is grammatical elements which provide the most extreme exceptions to the regular sound/kana correspondence rules: the topic particle *wa* (written *ha*), the case particles *e* (written *he*) and *o* (written with a unique grapheme), and the auxiliary element *yuu* (written *iu*) all constitute unique irregularities.<sup>9</sup>

Turning to the lexicon, the second general principle is that kanji are used to write neutral vocabulary; lexical items falling outside the sphere of kanji are written in the kana scripts, and are thereby graphologically branded as belonging to marked lexical strata.<sup>10</sup> It is the lexicon, therefore, that serves as the locus for the full interplay of the basic scripts.

Thus, firstly, Sino-Japanese items and the neutral core of native Japanese items are in principle written in kanji.<sup>11</sup> (In actual usage some neutral native items show a preference for hiragana: the clearest examples turn out to lie close to the borderline with grammar, or else to be members of homophonous sets (cf, 4.3).) Outside the sphere of kanji stand expressive and mimetic items on the one hand, and Western loanwords on the other. Expressive items typically encode the emotional or affective attitude of the speaker and are colloquial in style: some examples are *butsu* 'clout', *chippoke* 'small', *hebo* 'bungling', *chachi* 'flimsy'.<sup>12</sup> Mimetic items (*giongitaigo*) are a clearly defined morphological group of words denoting sounds, movements, sensations, and the like: cf. *pokin* 'with a snap', *bettari* '(adhere) stickily', *jirojiro* '(stare) fixedly', *gyorogyoro* '(look) gogglingly', etc. Both expressives and mimetics are, etymologically, virtually all native Japanese, but structurally they are clearly distinguishable from the neutral native vocabulary and share a good many features in common.<sup>13</sup> Graphologically, they are written in kana: in general, the norm prescribes katakana for mimetics denoting sounds (i.e., onomatopes) and hiragana for other mimetics and for expressives, but in current usage katakana appears to be gaining ground.<sup>14</sup> Western loanwords are also written in katakana, but it is important to note that they do not fall exactly together with onomatopes in this respect: the katakana script used with Westernisms in the modern language includes various supplementary conventions, and this supplemented script thereby functions as a graphological marker of this particular stratum.<sup>15</sup> The conventions concerned, which consist essentially in an extended repertory and range of use of diacritics, may be seen as another recent example of the open-endedness of the Japanese writing system noted in 3 above.

The differential functions of the scripts outlined here are summarized in the following table:

GRAMMAR		Hiragana
	Neutral	Kanji
LEXICON	Expressives, Mimetics	Hiragana/Katakana
	Westernisms	Extended katakana

**4.2 Native/Sino-Japanese homography.** In what follows we transfer our attention from the overall system to certain structurally significant aspects of the kanji script. Here we are concerned with an integrative function of that script.

As noted in 4.1 above, in general kanji are used to represent both Sino-Japanese (SJ) items and items of the neutral native vocabulary (NJ). To repeat the example given in 2, the kanji 川 represents both the SJ morpheme *sen* 'river' and the NJ morpheme *kawa* 'river'. (From the viewpoint of the kanji concerned, the associated SJ and NJ morphemes are said to constitute its *on*-reading and *kun*-reading respectively: in these terms, *sen* is the *on*-reading and *kawa* the *kun*-reading of 川.)

The fact that kanji are used in this way in Japanese means that in general SJ and NJ items, though formally unrelated in the spoken language, are graphologically linked in written Japanese. And since the SJ and NJ items linked through a given kanji are as a rule closely related in meaning, the bond is normally also a semantic one. A particularly significant consequence of this situation is that in Japanese the learned vocabulary, which is composed largely of SJ items, tends to be semantically transparent, graphologically, vis-a-vis the ordinary native vocabulary.

This point has been made on several occasions by Suzuki, who has emphasized the advantage Japanese enjoys in this regard in comparison with a language like English, where the learned vocabulary based on Latin and Greek is formally unrelated to the basic vocabulary in both the spoken and written mediums.<sup>16</sup> To repeat one of Suzuki's examples, the SJ term *tantoo* 短頭 'brachycephalic' is graphologically linked to the NJ words *mijikai* 短かい 'short' and *atama* 頭 'head', whereas there is no formal link amongst the corresponding English words.<sup>17</sup>

This is clearly a highly important feature of the graphological structure of Japanese, and it provides a strong argument in favour of the particular nature of the Japanese writing system.<sup>18</sup> Not that it applies throughout the whole modern kanji script: as already noted, the semantic transparency referred to clearly depends on (a) a kanji having both an *on*- and a *kun*-reading, and (b) the SJ and NJ morphemes being synonymous or at least closely related in meaning. Of the currently prescribed 1,945 *jooyookanji*, 737 are assigned only *on*-readings and 40 only *kun*-readings, leaving 1,168 kanji (60% of the total) to provide graphological SJ/NJ linkage.<sup>19</sup> Semantic relatedness is a difficult property to quantify: however, clear exceptions are not easy to find, and synonymy or near-synonymy is undoubtedly the rule.

**4.3 Two kinds of homophony.** It can easily be seen that the two main features of the Japanese writing system—viz. the presence of several scripts and the nature of the kanji script—both serve to magnify the potential of the language for homophony (in the sense of ‘same sound, different writing’). In the first case like-sounding items may be distinguished in writing by being represented in different scripts; in the second, a morphographic script will by its very nature distinguish distinguish morphemes independently of their phonological shape.

Clearly, examples of inter-script homophony may arise out of the graphological stratification discussed in 4.1: thus, *soko* **そこ** ‘that place’ (grammar) vs. *soko* **底** ‘bottom (of sea, etc.)’ (lexicon), *oku* **おく** ‘do for future use’ (auxiliary verb: grammar) vs. *oku* **置く** ‘put’ (main verb: lexicon); and, within the lexicon, *erai* **偉い** ‘great’ (neutral NJ) vs. *erai* **えらい** ‘terrible’ (expressive), *sooseiji* **双生児** ‘twin’ (SJ) vs. *sooseiji* **ソーセージ** ‘sausage’ (Western loan), etc. Such cases, though structurally important, are in practice relatively infrequent. By contrast, kanji homophony is rife in Japanese, and it is this differential function of the kanji script that will occupy us here.

The fact that kanji are used to represent both SJ and neutral NJ items makes for three potential types of homophony in terms of these lexical groups, namely SJ, NJ, and SJ/NJ homophony. SJ/NJ homophones are in fact relatively rare, due to the different phonological characteristics of the two groups, but homophony within each group is both common and structurally important.

SJ homophones are extremely numerous, reflecting the highly restricted number (less than 300) of phonological shapes of SJ morphemes. Discussions of this topic tend to concentrate on the morphemic level, pointing out the many different characters used to represent a given SJ morphemic shape.<sup>20</sup> Since SJ words are canonically dimorphemic, the magnitude of the phenomenon at word level is clearly reduced, but it nevertheless remains considerable. Thus, a typical smaller Japanese-English dictionary lists 8 homophonous SJ words pronounced *koosei*, 10 pronounced *kootei*, 10 pronounced *kooki*, and 11 pronounced *kikoo*,<sup>21</sup> and these figures would rise further with the inclusion of more specialized vocabulary. Were it not for the graphological differentiation among such items made possible by the morphemic script, this would represent an intolerable level of formal identity. The role of the kanji script in supporting the characteristic high proportion of SJ items in the written language is quite obvious, and an all-kana orthography (cf. 3 above) faces important limitations in this area.

Unlike SJ homophones, the great majority (and structurally the most important group) of NJ homophones consist of near-synonyms. Here are some examples: *abura* 油 'oil', 脂 'fat'; *kawa* 皮 'skin', 革 'leather'; *hayai* 早い 'early', 速い 'fast'; *atsui* 熱い 'hot (to the touch)', 暑い 'hot (of ambient temperature etc.)'; *semeru* 攻める 'attack', 責める 'criticize'; *naku* 泣く 'cry(= weep)', 鳴く 'cry (of animals)'. In these cases words which, in an analysis of spoken Japanese alone, would be regarded as single items with a wide range of meaning (e.g., *abura* 'fat, whether liquid or solid', etc.) are in the written language differentiated into two formal items each with a narrower meaning. That is to say, meanings are formally distinguished among NJ items in writing which are not so distinguished in speech, and to this extent the kanji script has effected a semantic re-structuring, along Chinese lines, of the native vocabulary.

These distinctions, of course, stem from the NJ/SJ graphological link discussed in 4.2, and they are echoed and reinforced in SJ items: thus, the distinction made in *abura* 油 'oil' vs. 脂 'fat' reflects that seen in such SJ words as *shokubutsuyu* 植物油 'vegetable oil' (with 油) vs. *shiboo* 脂肪 'fat (in meat, on body, etc.)' (with 脂). The distinctions, however, are not always as clear as in the examples illustrated. The NJ adjective *katai* 'hard, firm, stiff', for example, is assigned three kanji in the norm, viz. 固, 堅 and 石硬,<sup>22</sup> and the illustrative examples in the relevant official publication include the phrases *katai ishi* 'hard stone' (with 石硬) but *katai zaimoku* 'hard timber' (with 堅).<sup>23</sup> The semantic distinction in these particular examples is difficult to pinpoint, and the division here is not clarified in SJ items: SJ terms such as *koodo* 硬度 'hardness' and *kooshitsu* 硬質 'hard' (with 硬) denote hardness as a physical quality of substances in general, including wood. The various spelling dictionaries (*yoojiyogoji-ten*) on the market generally attempt to supplement the illustrative phrases provided in the Government directive with semantic explanations. Takeda (1975:83), for example, has the following notes under the entry for *katai* on the use of the two kanji:

石硬 : Mainly refers to the property of not bending or denting when force is applied (to stone, metal, etc.)

堅 : Mainly refers to the solidity of substance of objects

Apart from the indication on collocability, however, the distinction remains far from clear, and in area such as this the writing system is, in the last resort, indeterminate: which kanji, for example, is to be used for *katai*

in describing the hardness of an object composed of both metal and wood?

As we have indicated, this type of NJ homophony is the natural accompaniment of the NJ/SJ homography discussed in 4.2. As we saw there, the graphological integration of these two sections of the vocabulary has the advantage of creating transparency in Sino-Japanese words. On the other side of the coin, however, the incorporation through kanji of SJ semantic distinctions into the native vocabulary leads on occasion to undecidable choices for the writer. This is a point of instability in the system, and the general theoretical and practical difficulties of these NJ homophones have often been pointed out. Miyajima, for example, sees the recommendations of the norm in this whole area as running counter to the principle of 'one word, one spelling' which he discerns as gaining ground in actual usage. He produces many actual examples from literature which indicate that the recommended distinctions are commonly ignored in practice, a single basic kanji being used throughout at the expense of others of more restricted distribution.<sup>24</sup> In addition, as noted in 4.1, there is a tendency in current usage towards writing NJ homophones in hiragana.<sup>25</sup> The extent of such variation and the precise nature of individual cases require further study, but clearly any wholesale trend here towards 'one word, one spelling', particularly in the direction of hiragana, would serve to weaken the NJ/SJ link and have important consequences for the graphological structure of Japanese in general.

5. This concludes our survey of some of the major structural features of written Japanese as they involve the writing system. Starting from the two outstanding characteristics of the system, namely its mixture of scripts and the nature of kanji, we have attempted, hopefully in terms accessible also to the non-specialist in Japanese, to bring together some of the structural links and divisions in which it plays a central role.

Though we have had occasion to mention significant contrasts with actual usage, our discussion has in general been in terms of the 'norm'. As we have emphasized in 3, a good deal of work remains to be done on the stylistic options afforded by the flexibility of the system, but of course any such work must take the norm, or at least some 'unmarked' usage fairly near it, as its starting point.

My thanks are due to Harold Rowe for his comments on an earlier version of this article.

1. On the autonomous status of the written norm, cf. Vachek, 1939; 1945-1949; 1973.
2. In speech these words are pronounced *esueru*, *ooeru*, and *sandiikee*, respectively. For cited Japanese forms we follow the Hepburn system of romanization, except that we write long vowels double.
3. Cf. Kratochvil, 1968: 146-162 on the nature of the Chinese writing system.
4. See Kokuritsu Kokugo Kenkyuujō, 1983: 33-36.
5. 'One hesitates for an epithet to describe a system of writing which is so complex that it needs the aid of another system to explain it.' (Sansom, 1928: 44)
6. Some examples and discussion can be found in Martin, 1972: 95-97.
7. On the orthographic form of individual words we follow Takeda (1975), which incorporates Government recommendations up to that date. Our examples have not been affected by subsequent directives.
8. Cf. Miyajima's discussion (1977: 209-217) on this point. Miyajima is mainly concerned to point out inconsistencies around the borderline in actual usage, as well as the influence of other factors such as frequency of use in favouring kana. But he does not deny the status of the general principle.
9. Graphological differentiation of grammar and lexicon is, of course, a familiar feature of written language: compare the well-known avoidance of two-letter lexical words in English reflected in pairs of grammatical and lexical homophones such as *in/inn*, *be/bee*, *by/buy*, etc. (cf. Venezky, 1970: 57; Vachek, 1973: 54; Bolinger, 1975: 475). Vachek comments (1973: 54, fn. 26) that 'it appears perfectly natural that the words of the former category (i.e., our lexical elements: AEB) ... were to symbolize their greater semantic and communicative weight by a more impressive graphemic extent of their written forms'. In the case of Japanese, it appears no less natural that this distinction be basically symbolized by the opposition between the morphographic and phonographic scripts.
10. By 'lexical items falling outside the sphere of kanji' we refer to lexical groups which are systematically excluded from kanji representation: as specified below, these comprise mimetics, expressives, and Western loans. Individual members of other groups may lack kanji as a result of, for example, the official reductions in the recommended kanji inventory, but these sporadic exceptions are not our concern here.
11. In the case of inflected native items it is, of course, the stem which is represented in kanji, inflectional (i.e., grammatical) endings being written in hiragana.
12. On expressives in Japanese, see Backhouse, 1983.
13. Cf. Backhouse, 1983: 69-70.
14. For some actual examples of variation see Kokuritsu Kokugo Kenkyuujō, 1983: 45-46.
15. As with grammatical vs. lexical differentiation (cf. note 9), graphological marking within the lexicon is generally attested. Cf. Vachek's remarks concerning English in e.g. 1971: 215.
16. See Suzuki, 1978a, especially 24-27. See also Suzuki, 1975, especially 182-191, for a discussion in English.
17. Suzuki, 1975: 190.
18. Miyajima (1977), who sees no practical objection to writing all NJ items in kana, appears to overlook the importance of this NJ/SJ link.
19. Figures from M. Nomura, 'Jooyokanji no onkun', *Keiryookokugogaku* 13, 1981, as summarized in Miyajima (1982: 251).

20. Thus Suzuki (1978b: 173) mentions that, even in a smaller dictionary, the SJ morpheme shape *koo* is represented by 75 different kanji.

21. Yamada, 1968. Some of these SJ 'homophones' are in fact distinguished actually, but this is irrelevant to the problem of written differentiation at issue here: prosodic differences are regularly ignored in phonographic writing systems, and all these words would be formally identical in kana.

22. We ignore the additional kanji 糞 here, which is mainly restricted to representing *-gatai* as a suffix.

23. See Kokugoshingikai, 1972.

24. See Miyajima, 1977: 204-209. As indicated above (cf. note 18), Miyajima concentrates here on problems of writing NJ items and does not discuss the related issue of NJ/SJ homography.

25. See the examples and discussion in Kokuritsu Kokugo Kenyuujo, 1983: 43-44. Attested examples of 'incorrect' kanji usage are also given, on 53-54.

## References

- Backhouse, A.E. 1983. 'The expressive stratum in modern Japanese', *Gengo Kenkyuu* 83, 61-78.
- Bolinger, D. 1975. *Aspects of Language* (second edition). New York: Harcourt Brace Jovanovich.
- Kokugoshingikai. 1972. '*Ijidookun' no Kanji no Yoohoo*.
- Kokuritsu Kokugo Kenyuujo. 1983. *Gendai Hyooki no Yure*. Tokyo: Kokuritsu Kokugo Kenyuujo.
- Kratochvil, P. 1968. *The Chinese Language Today*. London: Hutchinson.
- Martin, S.E. 1972. 'Nonalphabetic writing systems: some observations', in J.F. Kavanagh and I.G. Mattingly (eds.), *Language by Ear and by Eye*, 81-102. Cambridge, Mass.: MIT Press.
- Miyajima, T. 1977. 'Wago no kanji hyooki', in S. Suzuki (ed.), *Kokugo Kokuji Mondai no Riron*, 187-217. Tokyo: Mugi Shoboo.
- Miyajima, T., et al. (eds.). 1982. *Zusetsu Nihongo*. Tokyo: Kadokawa Shoten.
- Sansom, G.B. 1928. *An Historical Grammar of Japanese*. Oxford: Clarendon Press.
- Suzuki, T. 1975. 'On the twofold phonetic realization of basic concepts: In defence of Chinese characters in Japanese', in F.C.C. Peng (ed.), *Language in Japanese Society*, 175-192. Tokyo: University of Tokyo Press.
- Suzuki, T. 1978a. 'Gairaigo o megutte', in T. Suzuki, *Kotoba no Ningengaku*, 9-27. Tokyo: Shinchoosha.
- Suzuki, T. 1978b. 'Kanji no amari shirarete inai tokusei ni tsuite', in T. Suzuki, *Kotoba no Ningengaku*, 167-186. Tokyo: Shinchoosha.
- Takeda, M. (ed.). 1975. *Gendai Yooji Yoogo Jiten*. Tokyo: Meiji Shoin.
- Vachek, J. 1939. 'Zum Problem der geschriebenen Sprache', *Travaux du Cercle Linguistique de Prague* 8, 94-104.
- Vachek, J. 1945-1949. 'Some remarks on writing and phonetic transcription', *Acta Linguistica* 5, 86-93.
- Vachek, J. 1971. Review of Venezky 1970, *Language* 47, 212-216.
- Vachek, J. 1973. *Written Language*. The Hague: Mouton.
- Venezky, R.L. 1970. *The Structure of English Orthography*. The Hague: Mouton.
- Yamada, K. (ed.). 1968. *The New Crown Japanese-English Dictionary* (revised edition). Tokyo: Sanseido.

# The Adoption of Punctuation in Japanese Script

---

Nanette Twine

Griffith University, Australia

Exposure to Western writings in the nineteenth century had a beneficial effect on the spatial organization of Japanese texts, which had hitherto been either sketchily punctuated or not punctuated at all. The use of such devices as word spacing, paragraphing, commas, and full stops was advocated first in essays by scholars of the West and applied in a few school textbooks in the early Meiji period (1868-1912). The real impetus for the adoption of European-style punctuation, however, came from the novelists who produced the various schools of modern fiction which began to appear from the mid-1880s. They experimented with the whole spectrum of European devices in conjunction with a new colloquial style, though not all proved appropriate to Japanese. Through their efforts the practice of punctuating texts became well established and later spread outside the realm of fiction to other areas of written Japanese.

Written Japanese in the mid-nineteenth century encompassed a great variety of written styles and forms of orthography. Script was determined by style, and style by the nature of the text. There were a number of literary styles,<sup>1</sup> none of which bore more than a passing resemblance to the spoken language.

The rift between the spoken language and written texts effectively limited the range of readers who had access to the information presented in writing. The education of commoners had been quite widespread before the Meiji Restoration in 1868, but the degree of literacy attained thereby was usually only sufficient for carrying on the business of everyday life. To master the classical forms required years of study; lower-class children, whose physical labour contributed substantially to their family's welfare, had no leisure to devote to such pursuits. In effect, therefore, written Japanese remained almost exclusively the province of the upper class, a situation which further reinforced feudal class distinctions. It was by no means an efficient medium of communication. Chinese style in particular was so difficult that the method of instruction in that subject in *samurai* schools often consisted of mere chanting of passages learned by heart from the Chinese classics. As the newspapers which began to circulate in 1872 were controlled and written by upper-class

scholars who perpetuated the use of pure and hybrid Chinese styles,<sup>2</sup> townsmen, women and children were effectively cut off from the main source of information until popular newspapers written in a simple colloquial style began to appear after 1874. At the other end of the stylistic spectrum, classical Japanese was too long-winded, vague, and elegantly archaic in its expression to be fitted for dealing with modern situations.

These difficulties were compounded by the vast number of Chinese characters then in use in Japan. It was a favourite ploy of scholars wishing to display their erudition to pad out the text of their discourse with unnecessarily complex characters in order to impart a more educated appearance and tone to their prose.<sup>3</sup> To complicate matters still further, there was an often total lack of punctuation to show phrasing within the sentence itself. No spaces separated words, which rendered the mainly hiragana classical Japanese style particularly difficult. To extract the sense of the passage, the reader had to recognise sentence finals and other grammatical signals indicating function. It is not uncommon to find page after page in texts of this period almost totally without punctuation or punctuated only by commas and full stops. Where these last were in evidence, there does not appear to have been any rigorously adhered to convention governing their use. Examination of nineteenth-century government documents, scholarly texts,<sup>4</sup> and novels shows that in some cases the symbols , and ◦ were used to indicate clause and sentence boundaries respectively; in others, the symbol ◦ fulfilled both functions; sometimes commas were used for both, varying the size from that ordinarily used today to others so small as to be barely visible. Often, long sentences were broken up with commas but contained no full stops. Documents written in pure Chinese style or epistolary style<sup>5</sup> were in many cases not punctuated at all. Paragraphing was often non-existent, or at least rudimentary, which meant that readers of Chinese style texts in particular were faced with solid blocks of characters unrelieved by visual variation. While it may be argued that no more than basic clause, sentence, and paragraph separation are needed in official documents, the same cannot be said of the novel, where dialogue and narrative need separation and where effective punctuation gives life to the text. Figure 1 is an extract from *Seiyōdōchū hizakurige* (*By shanks' pony through the west*, 1870-76) by Kanagaki Robun (1829-1894). Its only punctuation is quotation marks to indicate dialogue embedded in the body of the narrative; this is true of the entire work with the exception of English words rendered in katakana which are separated from each other by full stops.<sup>6</sup>

In this field, as in many others, the need for reform was highlighted by exposure to the West. Those who came in contact with Western prose

支那の上海ハヤンハ共を出帆して英領なる香港まで船路四日にて着すべし乍麼香港ハ支那の東南の方にある孤嶋なり長サ五里巾三里岩山のみにて草木少く平地なし元來支那の領地なりしが近世英吉利領となりしより英人追々住居を移し交易場を開き寺院を建立學問所を設けて人の數も次第に増し繁昌の港となれり且説大腹屋廣藏が博覽會へ赴く船ハ上海を出帆して直地に香港に着せしかバ旅宿を設けて彼処に舍るに例の弥次郎北八等ハ同行の中にて取分て睦まじき通次郎と三個等しく別間にござり弥「ライ／＼通さん上海で北八が大ぼろを出したので逗留中氣をもんでしまつて面白くもおかしくもなかつたから今夜ハえんぎ直しに何処へかしけこんで浮れやうじやアねへか通「そいつハ奇妙だが北さんの珍事でおいらアこりたからめつたな処へハ御めんだヨ北「それだつておめへがぜんてへ仕出かした事だから仕方がねへ通「ヲヤなぜおいらが出かしたことだへ北「夫でもおめへの言にやア英吉利の婦人ハ多淫で女の方から持かけるといつたからアノ牛屋ハ英吉利人だから定めし内にゐる女も英人だらうと思つて當込んだ処がまんざらでもねへ様子だからチヨ／＼の引つかへしではや幕を切らうとしたところが彼婦人が早替りで敵役の邪魔人バタ／＼の立まはりからふり落しの山幕をむかふに見て一人り道行ハ此ところ大不出来／＼と繪本の小がきに書かれるのだしかし鳴物が啼びちやるめらで世界が支那で迷子が日本人だけ新奇しいから日本へ帰つたら河竹にこの脚色をはなして一幕著てもらふのだア

Figure 1. Kanagaki Robun, "Bankoku Kōkai: Seiyōdōchū Hizakurige," *Meiji Bungaku Zenshū*, v. 1 (Tokyo: Chikuma Shobō, 1966), p. 25.

could not help noticing the clarity and versatility of its colloquial style, the conciseness and simplicity of its alphabet, and the ease which punctuation imparted to the deciphering of its content. As the Meiji period got under way and it became increasingly clear that written Japanese was not an efficient vehicle for disseminating the flood of new information pouring into the country, voices calling for script and style reform began to make themselves heard.

The first attempt at breaking Japanese prose into less formidable blocks took the form of word spacing. With the exception of the authors of Dutch and English language textbooks, the first to write phonetic script in such a fashion was Shimizu Usaburō (1829-1910), a scholar of Dutch, Russian, French, and German who had published *Engirishi kotoba* (*The English language*), an English conversation manual, in 1860. Shimizu believed that Chinese characters should be replaced entirely with the hiragana script. His main concern was with the rapid dissemination of new knowledge; he felt that this could be greatly facilitated by simplifying script as well as style. The chief problem in using the kana script was its lack of any visual

あんたらしいとはぎりしやのあんたらくすすなはちすみ  
といふことはよりいでたり。これはすみのうまれだち  
なるやまいろものにて、すみねとわづかのみつねを  
ふくむ。いろくろく、ほろく<sup>ろく</sup>とかけてヤ、かねのご  
ときひかりあり。よりてかどやきすみ<sup>ろく</sup>またひかりすみ  
ともいふ。しづかにもえてけぶりなし。よりてまた  
めくらすみともいふ。やましはこれをつねのいしずみ  
のきくさ<sup>きく</sup>をたねとしてなりたつかみわざのあま  
りのやけがらとす。さてこれをいしずみがら  
のごとくたきもちふ。これはいるらんとにおほく、また  
えぎりすすことらんどならびにえうろばのひがし  
のいしずみばらにもいづる。あめりかにはいとお  
ほくいづるなり。

このあんたらしいとはくろくひかりてもろく、も  
えがたく、けぶりなく、わづかにあをきほのほあるの  
み。さりながらひのいきほひはいとつよくしてはくす  
のけなし。ゆゑにじようきしや<sup>きし</sup>などのかまには  
もつともよろし。しろきはひをのこすものをよしとす。  
卯三郎はかねてより假名のみを用ひて書を著し、西洋文を譯  
するにも漢字を假ること無らんとの説あり。即ち此譯文の如  
きも其一例なり。

文三が二階を降りて、ソツトお勢の部屋の障  
子を開けるその途端に、今迄机に頬杖をついて  
何事か物思ひをしてゐたお勢が、吃驚した面相  
をして些し飛上ツて居住居を直ほした。顔に手  
の痕の赤く残ツてゐる所を観ると、久敷頬杖を  
ついてゐたものと見える。

「お邪魔ぢや有りませんか。」

「イ、エ。」

「それぢやア。」

ト云ひ乍ら文三は部屋に這入ツて坐に着いて

「昨夜は大に失敬しました。」

「私こそ。」

「實に面目が無い、貴儀の前をも憚らずして……」

……今朝その事で慈母さんに小言を聞きました。

アハ、ハ、ハ。」

「さう、オホ、ハ、ハ。」

ト無理に押出したやうな笑ひ、何となく冷淡い。

今朝のお勢とは全で他人のやうで。

ヨこはまのかいがんどほりへとりたてたヤそけうくわい  
のれいはいだうが、できあがつて、こんげつとうかにひら  
きになり、ボるべつき、タむそん、フらをん、なぞが  
はるくにて、きやうをとなへおしへのわけをのべ  
ました。

Figure 2. Shimizu Usaburō, “Kishūsekitan Kantei no Setsu,” *Chūgai Shinbun*, no. 29, 26/8/1869. *Meiji Bunka Zenshū*, v. 4, 3rd ed. (Tokyo: Nihon Hyōronsha, 1968), p. 435.

method of signalling word boundaries. Shimizu demonstrated that this could be simply overcome by spacing words in the Western manner in “Kishūsan sekitan kantei no setsu” (“Expert Opinion on Coal from Kishū”, 1869), and later in *Monowari no hashigo* (*The ladder of physics*, 1874) a three volume translation of a German science primer. Figure 2 is an example from the former. Even postpositions indicating the syntactic function of the preceding word are separated from their head nouns.

Another who adopted the device of spacing words in kana prose was Watanabe Shūjirō (1855-1945), who had studied English and lived in Germany in 1870-71 before returning home to become eventually a renowned historian. He gave an example in “Nihonbun o seitei suru hōhō” (“Rules for writing Japanese”, 1875) of how he thought kana should be written (Figure 3). Watanabe suggested the use of katakana (in a hiragana text) to indicate capital letters in personal and place names, and the insertion of commas and full stops. The Japanese should adopt European grammar as a model, he said; they should emulate its best features, such as punctuation, paragraphing, parenthesis, dashes, and quotation marks. Unlike Shimizu, Watanabe was content to preach and did not practise his theories in his own prose, preferring to write in the traditional mode to which he had been educated. In the very early years of the Meiji period, few of the scholars who espoused script and style reform actually departed from custom in their own essays, perhaps fearing the thrust of their argument would be lost in the scorn sure to be heaped upon any such departures by the intelligentsia at large.

One of the first areas in which these new ideas were put to the test was, of course, education. The first use of punctuation outside translations of Western literature was in a primary reader, *E-iri chie no wa* (*An illustrated puzzle ring*, 1872), written by Dutch studies scholar and enlightenment writer Furukawa Masao, who used full stops and commas. These symbols were later used in *Shōgaku kyōjusho* (*Primary textbook*), published by the Education Ministry in 1873, and in Tanaka Yoshikado's *Shōgaku Nihon*

Figure 3. Watanabe Shūjirō, *Nihonbun o Seitei Suru Hōhō* (1875). Quoted in Yamamoto Masahide, *Kindai Buntai Hasei no Shiteki Kenkyū* (Tokyo: Iwanami Shoten, 1965), p. 132.

Figure 4. Futabatei Shimei, “Ukigumo,” *Gendai Nihon Bungaku Zenshū*, v.1 (Tokyo: Chikuma Shobō, 1956), p. 200.

*bunten* (Japanese grammar for primary schools, 1874-75). Tanaka (1841-1879), a *yōgakusha*,<sup>7</sup> was appointed head of the Education Ministry's editorial office for normal school textbooks in 1872. His book was a conscious imitation of English grammars, hence his adoption of the punctuation he found therein. *Shōgaku Nihon bunten* was written in a variety of Japanese characterised by heavy use of Sino-Japanese items which were conventionally associated with the rendering of Chinese-style texts into Japanese; texts of this type were normally left unpunctuated.

The following decade saw several instances of the advocacy and practical application of punctuation. Miyake Yonekichi (1860-1929), secretary to the Kana Club whose members advocated replacing Chinese characters altogether with hiragana, called in "Bun no kakikata ni tsuite" ("On writing", 1884) for the use of commas and full stops. Mozume Takami (1847-1928), a teacher in the Literature Department of Tokyo University and an active member of the Kana Club, published an essay in March 1886 entitled "Genbun'itchi" ("Colloquial style"), remarkable as the first scholarly discourse which combined both the theory and practice of style reform. While no actual reference was made in the text to the need for punctuation, commas and full stops were liberally used to clarify its meaning. That same year Suematsu Kenchō (1855-1920), a politician and scholar who had studied for eight years in England, suggested in "Nihon bunshō ron" ("Written Japanese: a discussion") that those who advocated replacing Chinese characters altogether with hiragana should use word spacing, capitals, and Western-style punctuation to improve their texts. His views were echoed by Nishi Tomisada, physicist and leading member of the *Dai Nippon kyōiku kai*,<sup>8</sup> in "Nihon futsūbun no zento" ("The outlook for written Japanese", 1888). In "Nihon futsūbun dō" ("The future of written Japanese", 1887) Shinpo Iwatsugu, a member of the Kinkōdō publishing firm who had earlier published the first colloquial-style primary textbook, suggested the use of punctuation as part of an over-all progression towards a simple colloquial style, claiming that under the current system what stops were used at all were simply inserted after a certain set number of characters regardless of meaning. Shinpo experimented with applying full stops and commas to the *Kojiki* (*Record of ancient matters*, eighth century) and other Japanese classics, and went on to illustrate his theory further in a primary school reader, *Nihon tokuhon* (*Japanese reader*, 1886). Two other school texts experimenting with punctuation at this time were Nakagawa Kenjirō's<sup>9</sup> *Rika tokuhon* (*Science reader*, 1886) and Nishi Tomisada's *Yōgaku tokuhon* (*Children's reader*, 1887), both of which used commas, full stops, paragraph indentations, and large characters to begin each sentence. Nishi also used capitals for names.

Thus, thanks to the urging of *yōgakusha* who realized that the success of

colloquial style in Western countries was due in part to the convenient division of writing into manageable segments, the practice of punctuating texts slowly gained a foothold in Japan. Much more influential in its eventual spread than the early scholars, however, were the novelists who began to use the colloquial style as a necessary adjunct to the new realistic fiction which appeared in the mid-1880s. First among these was Futabatei Shimei (1864-1909), author of *Ukigumo* (*The drifting cloud*, 1887-89), the first novel which attempted both a realistic portrayal of the problems of an ordinary modern individual and the use of the colloquial style in narrative as well as dialogue passages. Futabatei went so far as to attempt in his colloquial translations of Russian literature<sup>10</sup> to reproduce the exact punctuation of the original.

In translating foreign literature, one runs the risk of spoiling the original if one concentrates only on meaning. I myself was convinced of the necessity to grasp and reproduce the rhythm of the original text; I therefore did not dispense arbitrarily with even a single comma or period. If there were three commas and one period in the original, then I put the same into my translation in an effort to convey its tone. When I first began doing a translation, in particular, I expended considerable effort on the matter of form, even going so far as to use the same number of words as the original in an effort to reproduce the rhythm with absolute fidelity; but this did not go at all as I hoped, and there were times when I could not meet my own standards.<sup>11</sup>

The text of *Ukigumo* itself, which appeared in three volumes, was sparsely punctuated in volume one, but better punctuated in the later volumes, with dialogue separated from narrative on separate lines as well as by quotation marks (Figure 4). Futabatei's prose in the first volume showed frequent traces still of traditional literary styles, and no doubt this affected his attempts at punctuation too. When he became more proficient at using colloquial style in volumes two and three, which he patterned after the prose of Dostoevsky and Goncharov, his punctuation also improved. His diary for 1888, *Kuchiba shū: hitokagome* (*A collection of dead leaves: first basket*), is written in literary style and punctuated by only a few commas, lacking even full stops. By contrast, "Yo ga hansei no zange" ("Confessions of half a lifetime"), an autobiographical essay published twenty years later in 1908, is replete with paragraph indentations, commas, full stops, dashes, lines of dots to suggest reflection, quotation marks, and exclamation marks.

Another who incorporated European punctuation into his novels was Yamada Bimyō (1868-1910), who began writing colloquial-style works at about the same time as Futabatei. His first published attempt at the new style was *Chōkai shōsetsu tengu* (*The conceited demon: a satirical novel*), an unfinished work published in 1886 in "Garakuta bunko" ("Trash library"), the journal of the Kenyūsha group.<sup>12</sup> Its extremely long and

sparsely punctuated sentences were reminiscent of those of Tokugawa period light fiction, but Yamada did introduce certain forms of punctuation which had not been seen in a Japanese work before, separating dialogue and narrative with the symbol = and using the device . . . . . to suggest admiration or lingering memories. In later works he adopted with enthusiasm all forms of European punctuation, including abbreviation, dashes, exclamation marks, question marks, and quotation marks, as well as the basic comma and full stop.

The use of punctuation in Japanese fiction accelerated from then on in direct proportion to the adoption of the colloquial style by Meiji period authors. Because the one major drawback to the colloquial style was its wordiness by comparison with the much more concise Chinese style, it was essential that it be controlled and shaped by orthographic devices, and authors continued to hammer out a system of punctuation appropriate to Japanese as a necessary corollary to their successful use of *gen-bun'itchi*.<sup>13</sup> The text of *Wagahai wa neko de aru (I am a cat, 1905)*, by Natsume Sōseki (1867-1916), perhaps the most famous author of this period, is studded with question marks, quotation marks, dashes, and lines of dots to indicate fragmented speech, in addition to the standard commas and full stops. Paragraphs are indented. In some sections dialogue and narrative are separated with dialogue beginning on a new line; in others Sōseki reverts to the older practice of leaving dialogue in the body of the narrative, separated only by quotation marks. *Hakai (The broken commandment, 1906)*, by Shimazaki Tōson (1872-1943), a prominent writer of the Naturalist school, is similarly embellished, with the added refinement that Tōson used full stops as well as quotation marks to end sections of dialogue whereas Sōseki did not, relying instead on the closing quotation mark alone to indicate the end of an utterance. By then the practice of punctuation was well established, and later writers continued to use it both as an indication of the sense of their text and as a means of stamping their own individual styles upon the written language

Not all European devices were adopted. Inverted commas as used to indicate quotation in the West were not appropriate in Japan, where the addition of " next to the top right hand corner of certain of the phonetic symbols denoted a change from unvoiced to voiced consonant. In their place were used ㄱ ㄴ. Nor was the use of capitals more than a passing fad among those who advocated replacing Chinese characters entirely with phonetic script; while it may have been possible in an all-phonetic text, it would have made typesetting a nightmare in the script situation which actually existed. Word spacing was not generally adopted; it is found today only in certain elementary textbooks, children's fairytales meant to be used as first readers, and some elementary Japanese language textbooks for foreigners.<sup>14</sup>

As the use of the colloquial style gradually spread to areas of written Japanese outside fiction, punctuation naturally accompanied it, until both were well established in education and journalism by 1925. Official documents proved more resistant to change; with the appearance in 1946 of the draft of the revised Japanese constitution written entirely in colloquial style, however, it was not long before these last areas of entrenched traditionalism were also overcome.

1. For a discussion of styles, see Satō Kiyoji (ed.), *Kokugogaku knekyū jiten* (Tokyo: Meiji Shoin, 1977), pp.353-368.

2. Pure Chinese style was a style in which kanji are arranged in accordance with the syntactic order of literary Chinese; Hybrid Chinese style was a form of writing which, while following the syntactic order of literary Chinese to some extent, also shows the influence of Japanese syntax.

3. Fukuzawa Yukichi, "Fukuzawa zenshū chogen", *Fukuzawa Yukichi zenshū*, v. 1 (Tokyo: Iwanami Shoten, 1958), p. 6.

4. For example, those found in the *Dai Nihon komonjo* and *Meiji bunka zenshū* collections.

5. A style which developed about the twelfth century as an offshoot of Hybrid Chinese. It was known as *sōrōbun* because of its characteristic use of the verb *sōrō*.

6. My example is taken from the third volume; the two forewords to this volume offer an interesting example of the range of punctuation practices then extant. One is un-annotated pure Chinese style, totally innocent of punctuation; the other is a variety of Hybrid Chinese in which full stops have been used not only to indicate clause and sentence boundaries but also subjects, topics, objects, and so on, as well as replacing commas in lists.

7. Scholar of things Western.

8. An influential education group with several thousand members.

9. 1850-1928. Meiji period educator.

10. Notably "Aibiki" (1888), a translation of Turgenev's "The Rendezvous"; and Meguriai" (1888-89), Turgenev's "Three Meetings".

11. Futabatei Shimei, "Yo ga honyaku no hyōjun" ("My standards in translation", 1906), *Meiji bungaku zenshū*, v.17 (Tokyo: Chikuma Shobō, 1971), p.108.

12. A literary group who favoured a return to the styles of the late eighteenth and early nineteenth centuries.

13. 言文一致, colloquial style.

14. And, of course, in romanized Japanese.

# Japanese Orthography in the Computer Age

---

**J. Marshall Unger**

University of Hawaii

Modern Japanese writing makes use of a large inventory of Chinese characters. Computers that can output such characters and represent them internally may cost more than computers that need only handle standard alphanumeric data, but there is no difficulty designing and manufacturing them. On the other hand, although many ingenious systems for input have been devised, none comes close to matching the efficiency of touch typing. Why is this, and is there hope for a breakthrough in the years ahead? This paper attempts to answer both these questions by examining the linguistic and technical factors which are responsible for the intractability of the input problem, and the social factors which lend it a sense of urgency. Special care is taken to separate those aspects of the problem which are genuinely cultural in nature from those which are not. The conclusion is not encouraging: an efficient general-purpose input system is unlikely. The Japanese might be better off directing some of the capital they are now investing in the quest for artificial intelligence towards less glamorous pursuits such as fostering public acceptance of romanization for use in computer applications that demand fast, accurate, easily modified input. The political resolve for such an approach, however, seems to be lacking.

## 1. Why is Japanese Script a Problem?

Uenohara Michiyuki, senior vice president of Nippon Electric Company (NEC, one of Japan's leading computer manufacturers), once summarized the problem of handling Japanese-language data on computers as follows:

People using European languages can engage in data-exchange and dialogue with machinery through typing with almost the speed of conversation. But this is impossible in the Japanese language. If oral input becomes possible, the [sic] handicap will be totally eliminated. Because of the phonetic simplicity of the Japanese language relative to European languages, oral input will give the Japanese an advantage, reversing the present situation. The world is made in such a way that advantages and disadvantages are always offsetting relationships.

The ultimate mission of technology is to make up for disadvantages of human beings and society and bring them progress. If it ruins advantages that human beings have, it does not deserve to be called technology.

A local culture is something that has been developed through the long history of the region and human life itself. It should not be altered because of technology. Rather technology must be altered to fit the local culture. If the

Japanese language is abolished for the sake of convenience of usage of computers, the Japanese will be deprived of their identity. (Gregory & Etori 1981:J40.)

There are four major claims in this statement — one valid and three invalid. Although not all Japanese would subscribe to them, they give a feel for the consensus of opinion in Japan regarding the orthography/computer problem.

The valid claim, implicit in the first paragraph, is that input is the only real problem of data processing using Japanese script. The output of Japanese text on printers or television-like hardware, and the internal manipulation of data representing the elements of Japanese orthography are merely engineering problems. Existing methods suffice, and future methods will be better, cheaper, or both. The hard part is devising a suitable code for the machine-internal representation of the orthographic elements, and encoding raw Japanese-language data so that machines can use it.

The three invalid claims are all explicit. The first is perhaps the most obvious. Nothing “is impossible in the Japanese language” because the Japanese language is not Japanese writing. Even if the *current writing system* were “abolished” (which hardly anyone is urging), the Japanese language would remain. The obvious technical problem of improving the speed at which one can input Japanese orthography must not be confused with the more fundamental problem of deciding, in each application, exactly what sort of data one needs to input and manipulate.<sup>1</sup>

Just as the Japanese language (as distinct from Japanese writing) is no barrier to using Japanese-language data on computers, so too is its alleged “phonetic simplicity” of no benefit. As already remarked, the chief difficulty in using Japanese data on computers is inputting it. When this involves transforming phonemic representations of Japanese utterances into traditional orthography, using a microphone rather than a keyset as the means of input not only does not make the process any easier, it actually slows it down by introducing extraneous (non-phonemic) information which must eventually be filtered from the signal.

Finally, to speak of technology as the mere handservant of culture is a distortion: the two are inextricably related.<sup>2</sup> Changes in an unrelated field can suddenly wipe out the usefulness of a seemingly indispensable tool (e.g., the integrated circuit killed the slide rule) or breathe life into an abandoned one (e.g., the Arab oil embargo resurrected the New England wood-burning stove). The alphabet is not merely a cultural property; in a larger view of human history, it is a technological breakthrough. The question is not whether Japanese script can be handled *at all* on computers—no one ever doubted that—but whether it can be handled in a

way comparable in cost and efficiency to what would suffice if alphabetic writing were used instead.

## 2. Linguistic Factors

So far I have said nothing about the specific features of the Japanese writing system which make it unwieldy on computers. A comprehensive exposition of the writing system would require an article all its own,<sup>3</sup> especially if it were to be of value to readers unacquainted with spoken as well as written Japanese. But it is possible to get a feeling for the complexities of Japanese script without knowing too much about its "fine structure" because almost all of them arise from a single cause: the use of kanji.

Kanji are the Chinese characters which are intermixed with syllabic letters called kana in modern Japanese writing. Kanji are completely characterized by their relationships with one another, their symbolic values, and their shapes. These properties are illustrated within the ellipse in Figure 1. The various disadvantages that result from using kanji rather than phonemic notation to represent Japanese-language data on computers are categorized with respect to these three properties in Figure 2. Keep in mind that we are not concerned here with applications in which kanji are treated as data in their own right; in such "print shop" situations, one is merely interested in creating and formatting displays—the semantic content of the displays is immaterial. In computer applications, on the other hand, one usually wants the machine-internal data to be, to some extent, intelligible as language. Thus, only the last three of the following nine problem areas are related to the typographic properties of kanji.

1. Practical kanji inventories must be large. Most of the problems discussed below have their roots in this fundamental requirement. The Japan Industrial Standard (JIS), for example, calls for two sets (high and low frequency) of roughly three-thousand kanji each. Newspapers and book publishers maintain fonts of thousands more since the JIS sets do not include many rarely used and variant kanji.
2. In principle, any Chinese character can be pressed into service in writing Japanese. Even when there is a "standard" inventory of kanji available, provision must usually be made for *ad hoc* additions. The lack of a true standard, based on linguistic principles, complicates software design.
3. Alphabetic order is, of course, arbitrary; but the alphabet is short, and the process of alphabetizing distinct words is strictly mechanical. There is no similar simple algorithm for ordering words written in kanji. Since sorting and retrieving data from sorted lists is a basic task in

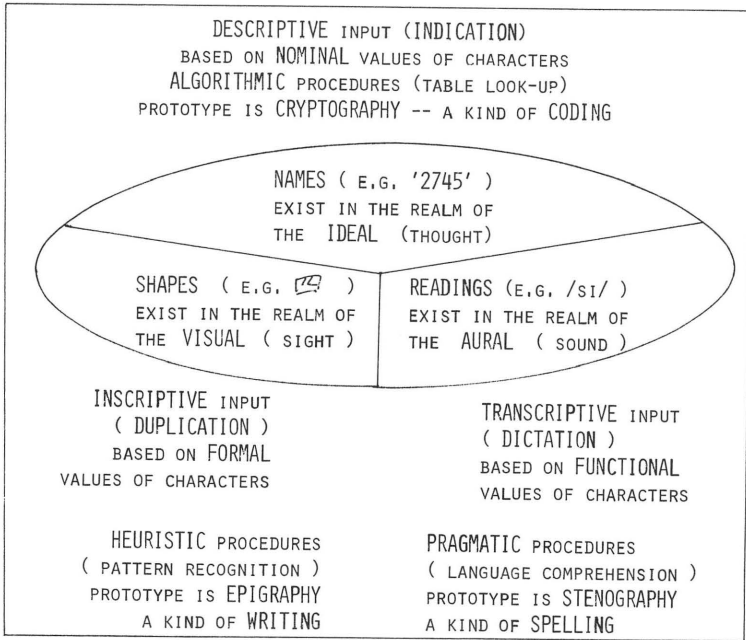


Figure 1.

Figure 2.

USING KANJI TO REPRESENT JAPANESE-LANGUAGE DATA  
 ON COMPUTERS IS

	WASTEFUL	IMPRECISE	DIFFICULT
BECAUSE			
SETS OF KANJI ARE	(1) LARGE	(2) OPEN	(3) ILL-ORDERED
THE READINGS OF KANJI ARE	(4) REDUNDANT	(5) AMBIGUOUS	(6) ARTIFICIAL
KANJI SHAPES ARE	(7) COMPLEX	(8) ABSTRACT	(9) HOMOGENOUS

countless computer applications, the lack of a collating sequence poses an obvious difficulty.

4. The system of "readings" which relate kanji to morphemes of Japanese is highly inefficient because the vast majority of kanji occur very infrequently in ordinary texts.<sup>4</sup> Personal and place names contain hundreds which are otherwise not used at all.

5. Kanji which do occur frequently typically stand for two or more morphemes. Even in context it is not always possible to know with complete certainty the exact words represented by certain sequences of kanji. Proper nouns are notorious in this regard. Further ambiguities arise when kana are used after kanji to represent the endings of inflected words (verbs, adjectives, etc.). Although there are prescribed "spelling rules" for such cases, considerable variation is found in practice.

6. Some strings of syllables ("readings") are associated with only one kanji, but most are not. Dozens of kanji can, in the right context, substitute for the same phonemic strings. This has led to a plethora of Sino-Japanese homonyms, and the use of different kanji to distinguish different senses of the same word. In addition, many kanji usages are optional: kana may be used instead for stylistic reasons.

7. Kanji shapes are wasteful of system resources relative to standard alphanumeric characters because they require higher resolution for displays of the same quality and use up much more memory per character. Falling hardware costs are narrowing the gap between systems that can and cannot output kanji, but that gap will never be reduced to zero.

8. Many kanji have variant shapes; recent reforms (different in China and Japan) have added hundreds more. These are not matters of font style but variations in the number, size, relative location, and scale of the graphic elements which are used in combination to form the vast majority of kanji images (*jita*).<sup>5</sup> Thus, graphically different kanji may have the same symbolic value. In order to identify allographs, the best one can do is try heuristic or brute-force methods of the kind used in chess-playing programs.

9. Each of the basic graphic elements that combine to form kanji images have considerable visual redundancy, but kanji often differ only in one of these elements. Certainly 'i' and 'j' look similar; but when one is dealing with thousands of symbols, optical scanning begins to get intolerably difficult as such "near misses" multiply. And, of course, handwriting makes matters much worse.

Although the nine characteristics of kanji just enumerated present difficulties for kanji use on computers, many resulted from or allowed for the successful use of kanji in Japan for centuries.<sup>6</sup> People have little

difficulty, for example, identifying traditional variant forms of kanji with their mental norms because, unlike machines, they perceive kanji as complete *Gestalten* situated in context. They easily accommodate themselves to high levels of ambiguity and artificiality in kanji/morpheme relationships far beyond what existing input programs can tolerate. They have devised ingenious tricks for sorting and retrieving Japanese data written in kanji because they take in *signe* and *signifié* holistically.<sup>7</sup> Thus, the kanji problem has nothing to do with the cognitive difficulties of learning and using kanji to write Japanese. Whatever the difficulties, people obviously cope with them. The problem is precisely that machines don't. Our tendency to associate literacy with culture makes us forget that a writing system is first and foremost a technology. Once this fact is recalled, it becomes clear that what we have in the case of Japanese data processing is a conflict between an ancient technique for reducing the Japanese language to writing and a modern technique for reducing writing to numbers.

It is important to stress this point for two reasons: first, because there are people (like Uenohara) who mistakenly think the conflict is cultural rather than technological; and second, because there *are* intercultural pressures against kanji use which must not be confused with the very specific problem of handling Japanese orthography on computers. For example, the older practice of making up Sino-Japanese compounds such as *jidōsha* 'automobile' and *denwa* 'telephone' has given way to deluge of Japanese-English that makes Franglais seem like a summer shower. Kabashima Tadao (1981:178ff.) has pointed out an important reason for this: part of the post-war orthography reform involved banning many *kun-yomi* or native glosses for kanji. Thus the percentage of kanji that had only *on-yomi* (i.e., could represent only Sino-Japanese morphemes) rose dramatically, and Japanese school-children were suddenly deprived of the mnemonics that had aided past generations of learners. For them a Sino-Japanese neologism is more often than not a troublesome piece of pedantry rather than a meaningful part of their inner symbolic landscape.

Indeed, kanji exist on a semiotic level congruent with neither the phonological nor the semantic structure of the Japanese language. When Japanese say that writing Japanese without kanji is inconceivable, they are referring, I believe, to this stratum of extra meaning. It is non-linguistic but it is not an illusion; one must not ignore its existence even if some Japanese (e.g., Suzuki 1977) wrongly insist that it is part of the synchronic structure of the Japanese language. In fact, the phenomenon of non-linguistic meaning that is experienced when one reads Chinese or Japanese is, I think, a hitherto overlooked piece of evidence against the Platonist/rationalist perspective now favored by many linguists. This is the

epistemological view which underlies rigorously formalized theories of grammar, much work on machine translation, and above all artificial intelligence research. Ironically, as we shall see below, the future of the kanji input problem is tied closely to just such applied research.

### 3. Technical Problems

As remarked at the beginning, the nine points summarized in Figure 2 are merely inconveniences to the extent they pertain to output and machine-internal calculation; but to the extent they relate to input and the problem of devising a suitable code for kanji, they are serious obstacles to progress.<sup>8</sup> Liu (1983:415) estimates that in China alone "nearly 400 schemes" for character encoding have been proposed, of which "several dozens" have actually been implemented. Many of these coding schemes may differ only slightly from one another, but distinctive systems developed in Japan, Taiwan, Hong Kong, the US, and other countries should also be taken into account; clearly, an enormous amount of work is going into input and encoding of characters.

Figure 1 gives a typological summary of Chinese-character input systems. The inclusion of kana in Japanese texts makes the input of Japanese different in some respects from the input of Chinese; but the methods available for both languages are essentially the same, and all are based on these three basic types in one way or another.

Inscriptive input, in which the user actually "draws" characters on a digitizer or similar device, is the least developed of the three because pattern recognition is such a hard computational problem, and because even a perfect system could not proceed faster than ordinary handwriting. Research in this area is now focused primarily on optical character recognition (OCR). Its goal is to produce machines that can scan printed documents and transfer them character-by-character to electronic media with minimum user intervention.

Transcriptive input is now the "method of choice" in Japan. The user types strings of kana or roman letters and instructs the computer to convert portions of them into kanji. Although some programs use frequency statistics or elementary syntactic analysis to narrow down the number of alternative kanji, the user must often choose among them. This means that, if a text is being copied, the user must momentarily look away from it. A maximum rate of about 50-60 characters per minute can be achieved this way, though there are claims for higher rates when ergonomically designed keysets are used in conjunction with this kind of input. As already mentioned, using voice input rather than typing for transcriptive input does nothing to ameliorate its inherent shortcomings.<sup>9</sup>

Descriptive input has more variants than the other two methods. There

are four subtypes depending on whether the codes for the kanji are mnemonic or arbitrary, and whether the hardware implementation uses a small or a large keyset. Mnemonic codes may make reference to kanji shapes and/or readings. Large keysets (both "tablets" and "multi-shift" keyboards) almost invariably use some sort of mnemonic arrangement. Characters may be grouped according to relative individual frequency, co-occurrence in pairs, shared graphic components (radicals), common readings, and so forth. On small keysets (standard or specially designed keyboards with roughly the same number of keys as a typewriter), the individual keys almost always stand for primitive elements (phonemes and syllables, strokes and radicals) into which kanji readings or shapes can be analyzed. The idea is that the user should be able to derive the code for any available kanji by applying a simple set of rules; this presumably makes the system easy to learn and efficient to use. In practice, however, the reported results for small-keyset systems are 120 characters per minute or less.

Transcriptive and many descriptive input methods are really two-stage processes. During stage one, the user inputs raw character codes. Stage two involves disambiguating kanji that share the same code. Novices generally interrupt stage-one work and attend to stage-two work whenever an ambiguity arises. More experienced users often try to type in a whole paragraph or page "blind" and go back to alter incorrect kanji. But the twofold nature of the process cannot be eliminated entirely; as long as the codes are mnemonic, they will be ambiguous to some extent. Stated differently, kanji will always cost more to process because they involve extra-linguistic information that cannot be automatically predicted from raw speech or an abbreviated form of writing.

The bifurcation of the input process, however, is not the only way to handle this extra cost. Another approach is to use completely arbitrary codes, thereby sacrificing "ease of learning" for the benefits that come with true touch typing.<sup>10</sup> In English beginning typists are expected to be able to handle 55 words per minute (five strokes = one word; each error = -1 word). Experienced typists often break 100 words per minute, and the Dvorak Simplified Keyboard puts such rates within the grasp of people of only average ability. Assuming two-stroke codes for each high-frequency character, with a smaller number of three-stroke codes for uncommon kanji, we would therefore look for a rate of two-hundred or more characters per minute for Japanese. As already stated, existing input techniques don't come close to this; moreover, those that require constant user intervention (stage-two work) cause excessive fatigue, muscular discomfort, and eyestrain. The extra labor necessitated by kanji cannot be eliminated; but undesirable side-effects can be avoided by shifting this

extra effort to the initial skill-acquisition phase of user activity and away from *ad hoc* disambiguation of characters.

Yamada Hisao and his colleagues at the University of Tokyo are well-aware of this trade-off. They are concerned with the problems faced by office-workers who input Japanese text five to six hours a day, five or more days a week. With such typists in mind, they have devised a keyboard arrangement in which the most frequent 2,304 (= 48 squared) kanji, kana, and punctuation marks are represented by two-stroke sequences assigned solely to optimize input speed (Hiraga et al. 1980). The keys have no names or meanings in this system; a map of the kanji and kana in relation to key pairs appears to be totally patternless. Intuitively, such a scheme should be hard to learn; but experimental results suggest quite the opposite. Yamada's hypothesis is, in fact, that no mnemonic coding scheme can be optimal. Any reference to the shape or function of kanji actually slows down input because it interferes with those subliminal kinesthetic skills that make touch typing viable. To use a popular metaphor, mnemonic codes access so-called "left-hemisphere" cognitive functions whereas touch typing depends on concentrated "right-hemisphere" processing.<sup>11</sup>

This brings us to the boundary between technology and economics. The hardware and software needed to implement Yamada's touch typing input system is all "old technology," well understood, widely available, and cheap. By contrast, the transcriptive kana-to-kanji and roman-to-kanji conversion systems being pushed by the major computer manufacturers in Japan right now require extra memory for look-up dictionaries, elaborate software, and sometimes non-standard keysets as well. Thus, if a touch typing standard is to be adopted, it will have to be adopted soon; otherwise manufacturers and their customers will have too much invested in research and development costs, outmoded machines, and user training to give up transcriptive input.<sup>12</sup> On the other hand, if a touch typing standard is not adopted, Japanese offices, laboratories, and factories that cling to kanji will, it seems, have to process information about half as fast as the rest of the industrialized world, and pay a premium for the privilege of doing so. Finally, there is the possibility that a touch typing standard will be established, but that it will be found impractical for use except by the office-workers for whom it was originally conceived.

Under the circumstances, the only way out would seem to be rescue by a *deus ex machina*. And that is literally what appears to be on the horizon: Japanese government and industry are now collaborating on a project aimed at producing the world's first genuine artificial intelligence (AI) computers by 1990. These "supercomputers" will allegedly be capable of creative problem-solving, natural-language translation, general pattern

recognition, and other seemingly miraculous feats. According to some experts (e.g., Feigenbaum & McCorduck 1983), it is only a matter of time before such machines are built; others have questioned whether such machines are possible even in principle (Dreyfus 1979) or, if possible, whether they are always socially desirable (Weizenbaum 1976). In view of this controversy—and at this stage, I think it fair to say that the doubters have the preponderance of hard evidence on their side—it is natural to ask: to what extent is the ultimate solution of the kanji input problem dependent on the outcome of the fifth generation project?

I had a chance recently to pose this question informally to several Japanese experts attending the 1983 International Conference on Text Processing with a Large Character Set in Tokyo. Responses were of three kinds: the majority felt existing input techniques are essentially adequate and merely need to be refined; they see the outcome of the fifth generation project as unrelated. Most of the rest held that AI machines would be needed for a truly general and elegant solution of the kanji problem. A few, however, went further; in particular, Motooka Tōru of the University of Tokyo, chief planner of the fifth generation project and (significantly) the conference chairman, opined that there would still be difficulties with input *even after* AI machines become available. This is probably the most prudent view for two reasons. First, the majority consisted mostly of engineers, many of whom were personally involved in the production of systems on the market. Second, a certain amount of “not invented here” thinking (abetted by the Japanese penchant for professional factionalism) was in evidence. For example, it was members of the majority who most strongly criticized the goal of voice input, one of the items high on the agenda of the fifth generation project.

#### 4. Social Aspects

At any rate, if there is a lack of unanimity among experts on the technical connection between the fifth generation project and the kanji problem, the historical evidence is clear enough. A recent “think-tank” report summarizes the situation this way:

There is no dispute among industry participants that Japan lacks the software capability needed to fuel its computer sales. . . . The reason for this lag stems from two interrelated sources: (1) the lack of emphasis on national projects that needed high-level software capabilities, and (2) the low esteem of software engineers/programmers as a profession. For example, 20 years ago the United States launched its National Aeronautics and Space Administration (NASA) space programs, which acted as a catalyst for developing software systems. Conversely, . . . Japan had no comparable government-sponsored program and did not establish software priorities until the late 1970's. In part a reflection of the national focus (or the lack of it) and in part a reflection of cultural values, the

Japanese education system, as a result, did not prepare the talent it needed to develop software. Although Japan has some 50 scientists and engineers per 10,000 in the labor force, less than 6 percent of them become software engineers or programmers. In actual numbers, Japan has only 7 percent of the number of software professionals in the United States. (Ohmae 1983:70-71.)

It goes without saying that the fifth generation project is exactly the kind of national endeavor that Japan neglected in the past. And the “cultural values” that contributed to that neglect are, of course, tied to kanji-based literacy.

Japanese study the reading and writing of their native language through nine years of compulsory education and three years of high school; those going on to university continue to study kanji, though on a less systematic basis. Anything that calls into question the value of their long years of effort in acquiring conventional literacy makes them feel uneasy, and they do not want to spend more time learning new ways of writing Japanese, however simple or useful they may be. On the other hand, they dislike having to work in a computerized world in which the flood of English-language data is growing daily. Given these conflicting emotions, it is easy to see how the fifth generation *deus ex machina* can grow from a dry intellectual hypothesis into an object of faith.

Of course, the Japanese are not the first to become infatuated with the dream of a disembodied Platonic intelligence. Already in 1976 Weizenbaum identified an American “artificial intelligentsia” that showed all the symptoms of chronic orthodoxy. One of its leading exponents, Edward A. Feigenbaum of Stanford University, has wasted no time in championing Japan’s drive towards AI. Unlike Vogel 1979 (*Japan As Number One*) and Ouchi 1981 (*Theory Z*), bestsellers which obviously inspired it, *The Fifth Generation* (Feigenbaum & McCorduck 1983) contains hardly any solid information about its namesake; rather, it is an extended polemic in support of the authors’ personal views on where AI research should be going. The message is unabashedly sensationalist: the Japanese are going to build genuine AI machines while the complacent, skeptical West slumbers; when it awakes, it will be too late—the Japanese will be in control of the key to the economic power of the future. The following passage illustrates the smug confidence in AI theory that informs the argument:

It’s a distinct pleasure to report that while the Japanese have put a lot of planning and thought into their Fifth Generation project, they’ve spent no time at all in those arid little debates so beloved by Western intellectuals, debates centered on the question whether a machine really can be said to think. They regard our obsession with that topic the same way we regard their eating raw fish—an odd, puzzling, but harmless cultural quirk. Instead, their debates are about the best way to design an intelligent machine, truly a new generation, the engine that will produce the new wealth of nations. (1983:17.)

I can only hope that the authors' quaint ideas about *sashimi* put readers on guard: the absence of "those arid little debates" at which Feigenbaum and McCorduck sneer may be a cause for rejoicing in some circles; but a more plausible interpretation is that Japan is eager to rush in where even IBM fears to tread at least in part because of intellectual shallowness.

Feigenbaum and McCorduck are so wrapped up in their self-congratulatory envy of the "forty samurai" spearheading the fifth generation project that they overlook its linguistic and sociopolitical aspects entirely. They discuss "Japanese Computer Science Education—an Achilles Heel?" (1983:144-147) without suspecting that the inadequacies they perceive are not confined just to computer science. "Another myth," they say, "goes something like this: *The Japanese might be able to build terrific cars, stereos, cameras, and baseball mits, but computers are different. The Japanese can't produce software. But it's not the fault of their intelligence; it's a limitation of their language.*" (1983: 141; emphasis in the original.) The only people who think that way seem to be Japanese (like Uenohara) who habitually confuse writing in kanji with thinking in their own language. In "Some Views from the Companies" (1983:108-110) Feigenbaum and McCorduck document the doubt and hostility with which some of Japan's hi-tech corporations view the national commitment to the fifth generation project, but they totally misinterpret this important fact. Companies vie fiercely for bright college graduates in Japan; the competition for computer science majors is particularly intense. Since many managers believe (as we have seen) that consumers will accept the latest crop of machines and programs, they are loathe to loan out their "whiz kids" to a project that, at best, will produce a prototype by 1990.

One must be careful, however, not to go to the other extreme and ignore the cultural aspects of computerization. The widespread belief that kanji are indispensable for writing Japanese should not be ignored even though it is rarely more than the result of misunderstanding or prejudice. To someone accustomed to reading mixed kanji/kana script, sentences without kanji do have a phantom-like, insubstantial quality about them; they seem somehow naked without the clothing of extra-linguistic meaning which kanji provide. As remarked at the end of Section 2, this intuitive feeling that the linguistic meaning of a written Japanese sentence needs to be completed by information unrelated to actual speech is one of those everyday (and therefore easily overlooked) cognitive experiences found in all cultures that cast doubt on the reductionist epistemology on which AI is based; yet ironically it is the very thing that makes so many Japanese want computers that can handle kanji. Without kanji, most Japanese feel they have no way to show others that they are educated and possess the knowledge which entitles them to social acceptance. Indeed,

education, knowledge, and kanji are often just one big blur in the Japanese mind.<sup>13</sup> Unless and until there is a perception that the leaders of Japanese society are prepared to foster a transition to the non-use of kanji, it is doubtful that many Japanese will be willing to take the risks involved in beginning that transition individually.

## 5. Conclusion

Thus, in looking ahead to the proliferation of computer power throughout Japanese society, our attention should, I think, be focused on government ministries rather than corporate laboratories. The bottleneck of kanji input simply cannot be widened. The obvious move is to use standard alphanumeric characters in all applications that do not require Japanese script, including real-time communications, creative programming, and so forth. Since it is unlikely that, even in the twenty-first century, the bulk of one's life will be spent in front of a computer terminal, the establishment of romanization as a special-purpose alternative mode of writing would probably not mean the end of traditional script. Unfortunately, it seems that many of the key "players" in Japan today are unwilling to countenance any compromise whatsoever.

1. Suzuki Takao (1977) contends that, at least in the case of Japanese, orthography is an integral part of what we should properly call language. It is obvious, however, that even if we were to grant some special psycholinguistic status to Japanese orthography, it would be of no relevance to machines in which all information is reduced to sequences of binary digits (bits).

2. Burke (1978) presents dozens of historical cases that illustrate this point. For example: "The change in the weather which struck northern Europe like a sledgehammer in the twelfth and thirteenth centuries provided urgent need for more efficient heating. The chimney answered the need, and in doing so had the most profound effect on the economic and cultural life of the continent." (1978:291)

3. For example, the introductory material in Hadamitzky & Spahn (1981)—a handbook for students starting to learn modern Japanese reading—runs to nearly seventy pages.

4. Between July 15 and July 21, 1971, a sample of 4,252 sentences on the Kyōdō News Agency wire were studied for frequency. Of the 1,001,554 character tokens, 462,209 or 46.1% were kanji; however, these kanji accounted for 2,258 or 86.8% of the 2,601 character types tallied (see Hayashi 1982:206).

5. According to Hayashi (1982:216) kanji composed of simpler graphic elements (*kaiji moji*) and *keisei moji* accounted for 1,624 or 87.8% of the 1,850 *tōyō kanji*; the 95 new kanji in the *jōyō kanji* list are almost all of these two types.

6. Of course, "success" must here be understood in a relative sense. Tōdō (1982:170-175) comments on illiteracy in pre-war Japan and argues that the post-war orthography reform was as important for Japan's recovery as the land reform and new constitution. Miller (1982:186) questions the reliability of official

literacy statistics, claiming that no country could have 94.7% literacy simply because of blindness and other impairments.

7. These are precisely the aspects of human intelligence which Dreyfus (1979) argues are not formalizable and therefore permanently beyond the grasp of digital computers.

8. I hasten to add, however, that even if input is left out of consideration, the extra costs of kanji output and manipulation are substantial relative to the alternative of using standard alphanumerics.

9. Confusion on this point is not confined to Japan. Panko (1983) is an example of an "in depth" article by an American who, though fairly knowledgeable about the linguistic facts of Japanese life, mistakenly seizes on the keyboard as the nub of the kanji input problem.

10. Yamada (1980) summarizes the development of touch typing in the West and its implications for Japan. Touch typing is much more than a mere convenience; it is a remarkable breakthrough in its own right. "Consider a typical secretarial task, the retyping of a document . . . . [I]t closely resembles a series of choice-reaction-time tasks, in which a subject is presented with a single visual stimulus from a set of two or more stimuli after being instructed to rapidly press a particular button for each of the possible stimuli. Under optimal conditions . . . the average latency, or delay between the presentation of the stimulus and the pressing of a button, is approximately 250 milliseconds. The paradox of typing is that a latency of 250 milliseconds yields a typing rate of 48 words per minute . . . . Yet speeds of twice that rate are fairly common." (Salthouse 1984:128.)

11. In recent lectures Yamada gave two pieces of evidence which suggest that touch typing, once mastered, is strictly "right-hemisphere" (non-linguistic). One is the observation that many experienced American secretaries can accurately copy-type manuscripts while carrying on conversations. The other is the fact that many of these same secretaries prefer to place the manuscript to their left; this contradicts the practice of many typing schools, but does fit the "right-hemisphere" interpretation. I can think of two similar pieces of circumstantial evidence. First, it is hard to sightread the lyrics of a song and its piano accompaniment at the same time; however, once either the music or the words have been learnt, it is easy to concentrate on the other, singing and playing at the same time. Second, skillful magicians never perform the same trick twice for the same spectators unless they can vary the method. In one classic effect, a spectator makes some simple moves with half a deck of cards, imitating what he sees the performer doing with the other half. This makes it impossible for the spectator to see through the trick because "the mind cannot think of two things at once while executing a manual operation which involves the use of the eyes . . . ." (Hugard & Crimmins 1961:74-75.)

12. In the West the Dvorak Simplified Keyboard has been available for nearly fifty years, but similar factors have kept it from replacing the grossly inefficient "qwerty" arrangement in common use.

13. Miller (1982) discusses this and related Japanese sociolinguistic attitudes at length. He attributes them to a "sustaining myth," a kind of obsession with the Japanese language, that has emerged in the post-war period as a psychosocial compensation for the misery and humiliation of the defeat. Although I agree with many of Miller's observations in a general way, I cannot accept mass psychological trauma as a valid explanation. (Indeed, Miller's arguments are so extreme and his tone so scathing that one suspects the "myth" is just an extended metaphor designed to allow the author to attack Japanese linguists he personally dislikes

under the guise of establishing a serious theory.) In my opinion, the attitudes which Miller catalogues are the result of the abolition of elite higher schools by the Allied Occupation, and the accompanying changes in public secondary education. A full discussion of this hypothesis is beyond the scope of this paper, but I think it important to mention one aspect of it here. Research on the *level* of literacy among the *great majority* of Japanese prior to 1945 is almost entirely lacking (Richard Rubinger, personal communication). Although compulsory egalitarian elementary education was instituted early in the Meiji Period, other evidence (such as *Tōdō's* remarks referred to in note 6) suggests that a thorough command of kanji of the sort now required of all Japanese high-school students was not expected of most Japanese before 1945. If this is so, then the blurring of kanji, education, and knowledge is a relatively recent phenomenon, and would account for the kind of attitudes which Miller attributes to his "myth."

## References

- Burke, James. 1978. *Connections*. Boston, Toronto: Little, Brown & Co.
- Dreyfus, Hubert L. 1979. *What Computers Can't Do: The Limits of Artificial Intelligence*. Revised edition. New York: Harper Colophon.
- Feigenbaum, Edward A., and Pamela McCorduck. 1983. *The Fifth Generation: Artificial Intelligence and Japan's Computer Challenge to the World*. Reading, Massachusetts: Addison-Wesley.
- Gregory, Gene Adrian, and Eiori Akio. 1981. "Japanese Technology Today: The Electronic Revolution Continues." Advertising supplement. *Scientific American*, October. 97-144.
- Hadamitzky, Wolfgang, and Mark Spahn. 1981. *Kanji & Kana: A Handbook and Dictionary of the Japanese Writing System*. Rutland, Vermont: Charles E. Tuttle & Co.
- Hayashi Oki, et al., comps. 1982. *Zusetsu Nihongo: Gurafu de miru kotoba no sugata* [Japanese in diagrams: The form of language seen through graphs]. Kadokawa shōjiten 9. Kadokawa shoten.
- Hiraga Yuzuru, Ono Yoshihiko, and Yamada Hisao. 1980. "An Assignment of Key-Codes for a Japanese Character Keyboard." University of Tokyo Department of Information Science Technical Report 80-12 (July). In *Proceedings of the 8th International Conference on Computational Linguistics (Tōkyō)*, 249-256.
- Hugard, Jean, and John J. Crimmins, Jr., eds. 1961. *Encyclopedia of Card Tricks*. London: Faber & Faber.
- Kabashima Tadao. 1981. *Nihongo wa dō kawaru ka* [How will the Japanese language change?]. Iwanami shinsho 145. Iwanami shoten.
- Liu Yongquan. 1983. "Language Engineering in China." In *Proceedings of the ICTP '83*, 412-418.
- Miller, Roy Andrew. 1983. *Japan's Modern Myth: The Language and Beyond*. New York & Tōkyō: Weatherhill.
- Ohmae Kenichi, ed. 1983. *Japan: Obstacles and Opportunities, (A Binational Perspective for U.S. Decision-makers Prepared by McKinsey & Company, Inc. for the United States-Japan Trade Study Group Tōkyō: President Inc.*
- Ouchi, William G. 1981. *Theory Z: How American Business Can Meet the Japanese Challenge*. Reading, Massachusetts: Addison-Wesley.
- Panko, Raymond R. 1983. "Kanji Keyboard Chaos." *Computerworld*, November 28.
- Salthouse, Timothy A. 1984. "The Skill of Typing." *Scientific American*, February, 128-135.
- Suzuki Takao. 1977. "Writing Is Not Language, or Is It?" *Journal of Pragmatics*, 1:4 (December) 407-420.
- Tōdō Akiyasu. 1982. *Kanji no kako to mirai* [The past and future of Chinese characters]. Iwanami shinsho 205. Iwanami shoten.
- Vogel, Ezra F. 1979. *Japan As Number One: Lessons for Americans*. Cambridge, Massachusetts & London: Harvard University Press.
- Weizenbaum, Joseph. 1976. *Computer Power and Human Reason: From Judgment to Calculation*. San Francisco: W. H. Freeman.
- Yamada Hisao. 1980. "A Historical Study of Typewriters and Typing Methods from the Position of Planning Japanese Parallels." *Journal Information Processing (Tōkyō)* 2:5.175-202.

# Japanese Braille

---

**J. Marshall Unger**  
University of Hawaii

Braille is perhaps the only area of Japanese linguistic life in which serious attention was paid to the problem of word and phrase delimiters before the advent of computers. Japanese orthography does not use spaces in the Western manner, but braille texts must if they are to be intelligible. The first part of this paper describes the fundamentals of Japanese braille, and outlines the spacing rules now in general use. Unlike English braille—in which cells correspond to letters of the alphabet, punctuation marks, or special contracted forms—Japanese braille cells are associated with elements of the syllabic script called kana. This leaves no room for contractions, although it does result in some savings in space. Are these savings superior to what could be achieved in a roman-based Japanese braille system? The second part of this paper answers that question in the negative.

This paper has two purposes: to provide a brief description of the Japanese braille writing system (*tenji*); and to point out the relevance of Japanese braille for the computer treatment of the Japanese language. In particular, braille is one of the very few areas of Japanese linguistic life in which the problem of word spacing (*wakachigaki*) has received serious attention. Also, the Japanese approach to braille transcription raises some interesting questions about the compressibility of Japanese texts.

Krebs (1974:1) defines braille succinctly as “a system of touch reading for the blind which utilizes a cell of 6 dots, three high and two wide, from which 63 different characters can be formed by placing one or more dots in specific positions or combinations within the cell. Each dot within the cell is designated by a position number. Dots 1 & 4 form the top pair, dots 2 & 5 the middle pair, and dots 3 & 6 the bottom pair. The embossed dots . . . are numbered 1-2-3 on the left and 4-5-6 on the right.”<sup>1</sup> The following paragraph from Loomis 1936:14 adds some historical perspective:

Before the adoption of Braille, there were countless systems invented for the use of the blind. These various systems represented to a large extent the work and ideas of sighted people. For many years everyone thought that the raised alphabet should bear a resemblance to our alphabet. Boston Line Letter represents this idea. New York Point and American Braille were both point systems and were based on frequency of occurrence, the letters most commonly

used having the fewest number of dots. To a sighted person it seemed a very good and sensible idea. But time and experience proved that the signs containing the least number of dots were not the signs most easily recognized by the blind reader, so frequency of occurrence was abandoned. Of all the systems devised in the years gone by, and there were many of them, only two remain in use today. One is Braille, which is used in all countries, and our own Braille alphabet is to-day the same as Louis Braille devised it in 1829. The other system is Moon, which is read largely by the foreign born and those who lose their sight late in life and cannot master the complicated Braille system. These two systems alone survive, and it is most significant that both these systems were devised by blind men.

To bring this statement up to date, we should note that Moon script is no longer in active use, and that the alphabet of Louis Braille (1809-1852) differed slightly from the one in use today. The original alphabet omitted 'w' since it was not used in French at that time (Krebs 1974:2). More importantly, accented letters, diacritics, non-roman alphabets (e.g., Cyrillic, Greek, Hebrew), and non-standard punctuation (e.g., Spanish inverted interrogation and exclamation points) require special treatment (see Krebs 1974:102-112 [Appendix B]); but they are all built on the original foundation laid by Braille. By contrast, the Japanese braille system, created in 1890 by Ishikawa Kuraji (1859-1945), is radically different from the roman-based braille system of the West.<sup>2</sup> (Special extended codes for music and mathematics [e.g., Nemeth code] have been devised, but are beyond the scope of this paper.)

Braille—unlike, for example, Gregg shorthand—is a transliterative system.<sup>3</sup> Although, as we shall see, it includes many non-alphabetic symbols, English braille does not attempt to treat the language phonemically. Ishikawa's braille, on the other hand, is based on the so-called 50 syllables (*gojūon*) which form the basis of kana orthography. In that sense, Ishikawa's system is also transliterative; but the connection between the structure of the kana writing system and the morphophonemics of the Japanese language is much closer and deeper than the connection between the roman alphabet and the sound system of English, as any Japanese who has had to struggle with English spelling can testify.

Because all forms of braille are transliterative, special care must be taken to keep the physical length of texts within reasonable limits. Experienced braille transcribers of English have told me that a normal page of ink-print prose is roughly equivalent to three braille-print pages. In order to keep transcribed texts short, a system of 189 contractions, word signs, and short-form words has been developed to supplement letter-by-letter transcription.<sup>4</sup> According to Krebs (1974:2),

Through the use of contractions, the finger in coursing across one full line of writing, encounters almost twice the number of words as would be present if full spelling were employed. For example, the sentence 'You have more knowledge.' requires but 9 spaces when contracted, rather than 25 spaces when uncontracted. Due to the advantages derived from the use of contractions, the 'average' braille reader can attain a speed in reading to himself equal to that of a person with sight when reading aloud.

This system of contractions is possible in part because the 26 letters of the alphabet leave 38 cells (including space, ' ') available for other functions. Thus 3-4-5-6 serves as a number sign, allowing the cells for the letters 'a', 'b', . . . 'j' to double for the digits '1', '2', . . . '0'. The same cell (3-4-5-6) also represents the sequence 'ble' when not initial in a word. This illustrates the principle of using mutually exclusive environments (complementary distribution) to distinguish among symbolic functions and thereby increase the power of the system of contractions. Another example is the cell 1-5-6, which stands for the common English digraph 'wh' when part of a word; the word 'which' when in isolation; 'where' when preceded by a dot 5 cell; and 'whose' when preceded by a dot 4-5 cell. Similarly, dot 2-3 is the whole-word contraction for 'be', the digraph 'bb' when word-medial, and ';' when word final. In order to prevent confusion and guarantee consistency when hyphenating, there are numerous rules concerning preferences among contractions when conflicts occur, omission of spaces in certain phrases, and so forth.

In Japanese, since almost all the cells are required from the start for the kana and their associated diacritics, such contractions are impossible. In English, there are many frequently occurring words of more than one or two letters: representing these high-frequency words with a one- or two-cell abbreviation results in a reliable rate of compression because the highest frequency words in any language will show up at roughly the same rate, on average, in any text regardless of content. (Actually, the English contractions deviate from this principle slightly. The word 'blind', for instance, has a contraction, presumably because it is very common in texts for the blind.) Japanese is compressed more homogeneously: each braille cell corresponds to a kana and hence (usually) to more than one phoneme or roman letter. Figures 1 and 2 graphically demonstrate this difference between Japanese and English braille. Another way to describe the situation is to say that in English braille the basic cells are relatively more redundant than in Japanese braille. Such formal redundancy can be seen as a desirable property of individual cells with respect to the rest—desirable because it lessens the chances of misspelling or misreading. In Japanese braille this redundancy is exchanged for overall compression, a desirable global property of texts.<sup>5</sup> An interesting theoretical question is

Figure 1. Letter assignments in English Braille.

Dots	none	4	5	4-5	6	4-6	5-6	4-5-6
none	⠠							
1	a	c	e	d				
2		i		j				
1-2	b	f	h	g				
3								w
1-3	k	m	o	n	u	x	z	y
2-3		s		t				
1-2-3	l	p	r	q	v			

Figure 2. Kana assignments in Japanese braille.

Dots	none	4	5	4-5	6	4-6	5-6	4-5-6
none	⠠	-y-	”	”-y-	P	P-y-		*
1	a	u	ra	ru	ka	ku	sa	su
2	Q	o	:	ro	-w-	ko	”-w-	so
1-2	i	e	ri	re	ki	ke	si	se
3	wa	ya	wo	yo		yu	N	
1-3	na	mu	ta	tu	ha	hu	ma	mu
2-3	wi	no	we	to		ho		mo
1-2-3	ni	ne	ti	te	hi	he	mi	me

*Legend*

- ⠠ space
- y- *yōon*; inserts -y- in following syllable. Before ‘te’, indicates innovative pronunciation [ti]
- ” *dakuten*; changes k-to g-, s- to z- &c.
- P *handakuten*; changes h- to p-. Before ‘te’, indicates innovative pronunciation [tyu]
- \* before ‘te’, indicates innovative pronunciation [dyu]
- Q *sokuon*; doubles following k-, s-, t- &c.
- :
- ” *chōon*; indicates vowel lengthening
- w- *yōon*; inserts -w- in following syllable. Before ‘to’, indicates innovative pronunciation [tu]; before ‘ha’, ‘hi’ & c., indicates [fa], [fi] & c.; before ‘ta’, ‘ti’ & c., indicates [tsa], [tsi] & c.
- N *hatsuon*; syllabic nasal
- N.B. *dakuten* and *handakuten* can be combined with *yōon*; dots 2-5-6 (”-w-) before ‘wa’, ‘wi’, ‘u’, ‘we’, ‘wo’ indicate the innovative syllables [va], [vi], [vu], [ve], [vo].

whether the Japanese alternative to contractions, which capitalizes on the phonological structure of the Japanese language at the cost of drastically reducing the redundancies among the different basic cell shapes, is in some sense optimal. More concretely, how would it compare with a braille system based on romanized rather than all-kana strings?

There is a second reason why a system of contractions for Japanese would, at present, be impossible to implement. In Japanese orthography there is no tradition of using spaces to separate words. The Grade 2 braille contractions are made possible to a large extent because the same signs can serve different functions in different positional contexts. Without spaces in Japanese texts, this strategy could not be applied. As a matter of fact, Japanese braille does use spaces to separate words; as remarked at the beginning of this paper, this is perhaps its most interesting feature for those in fields unrelated to the problems of the blind. The task of teaching Japanese Braille to space properly occupies a position comparable to the problem of teaching U.S. braille to use the contractions correctly. Of course, this complicates the question of optimal compression for Japanese texts since some sort of spacing is evidently necessary: surely Japanese braille would not bother with spaces at all if blind Japanese readers could do without them.

The primary purpose of the small dictionary NMFK 1982 is to clarify the proper use of spaces in braille transcription, although it also discusses other aspects of Japanese braille usage. In standard orthography or *kanji kanmajiribun*, in which semantically pregnant morphemes are generally written with Chinese characters (kanji) while inflectional endings, particles, and other high-frequency low-content morphemes are written in kana, there is little if any need for the visual reinforcement of word separators; but in braille, such separators become crucial. All-kana text written without spaces is called *betagaki* to distinguish it from *wakachigaki* (*wakatsu* 'to separate, split'). Because of a lack of unanimity among Japanese experts as to the proper unit for analyzing sentences, and because there is room for disagreement when different people try to apply the same method of analysis, the parsing of *betagaki* strings has become an outstanding problem of Japanese computer research. According to Yoshida and Tanaka (1981:9), there are five different ways in which all-kana text can be input on computers:

- 1a. *betagaki hōshiki*—no spaces.
- 2a. *kanjibu shitei hōshiki*—parentheses surrounding those substrings which are to be converted into kanji.
- 3a. *bunsetsu wakachigaki*—spaces separating syntactically defined phrases.

4a. *jiritsugo/fuzokugo wakachigaki*—spaces separating so-called ‘independent’ and ‘dependent’ words.

5a. *tango wakachigaki*—spaces separating all lexical items.

As an example, they write the sentence “There are some remarkable things in technological development” in five different ways which are reproduced here in romanized transliteration:<sup>6</sup>

1b. *gizyutunohattennihamemasii monogaaru.*

2b. *(gizyutu)no(hatten)niha(meza)masii monogaaru.*

3b. *gizyutuno hattenniha mezamasii monoga aru.*

4b. *gizyutu no hatten niha mezamasii mono ga aru.*

5b. *gizyutu no hatten ni ha mezamasii mono ga aru.*

Where does the *wakachigaki* of Japanese braille practice fall on this spectrum of all-kana input techniques? The following translation of the *wakachigaki* rules in NMFK 1982:18-20, supplemented by references to the examples in TTB 1981:36-47, provides the basis for an assessment:

1.1 Break before independent words (*jiritsugo*). [Rule 1.5 makes it clear that this means all words other than particles (*joshi*) and inflecting suffixes (*jodōshi*). The example *Hananoyōni kireidatta* ‘It was as pretty as a flower’ in TTB 1981:35 shows that *yōni* and *datta* are considered *jodōshi*.]

1.2 As a rule, break before formal nouns (*keishiki meishi*) except when they serve as components of compounds or affixes. [*Keishiki meishi* refers to nouns such as *koto* ‘thing’ when used with a weak sense as the heads of noun phrases. According to this rule, one should write *kau kotoga dekiru* ‘can buy’ but *kaimono* ‘shopping’, not \**kai mono*. TTB 1981:36 calls attention to the fact that the same phrase can be treated differently depending on context; e.g., *sono ue* is correct when the meaning is literally ‘its top’, but *sonoue* ‘moreover’ should not have a space.]

1.3 Even in cases in which a formal noun gives rise to a sound change, break before it; however, if the final sound of the preceding word is altered, the break is optional. [E.g., following TTB 1981:37, *donna monda* < *donna monoda* ‘What sort of thing is it?’, not \**donnamonda*, but either *antoki* or *an toki* < *ano toki* ‘that time’.]

1.4 Break before auxiliary (*hojo*) verbs and adjectives since they are independent words. But if part of an auxiliary verb or adjective is contracted with the end of the preceding word, run them together. [Thus, *itte iru* ‘has gone’ shortens to *itteru*, not \**itte ru*.]

1.5 Run dependent words together with independent and other dependent words. [Dependent words (*fuzokugo*) are particles (*joshi*) and inflecting suffixes (*jodōshi*). Besides *jiritsugo* and *fuzokugo*, there is a small class of words called *setsuji* or ‘affixes’; hence the “other”.]

1.6 [But] break before particles and inflecting suffixes which immediately follow roman letters or foreign-language quote signs. [E.g., from TTB: 1981:35, *CM dattamitai* 'It seems that it was a CM (= Commercial Message)', not \**CMdattamitai*.]

1.7 Break before dependent words that follow commas even when not transcribing commas. [A later rule of punctuation on NMFK 1982:21 says, "When using commas in the middle of sentences, make no break before the comma and take one space after it. When, however, commas are used in ink-print in lieu of *wakachigaki*, to break up multidigit numbers or in a similar manner, ignore them."]

2.1 Do not break up short compounds that have become single lexical items. [E.g., TTB 1981:40, *otokonoko* 'boy', not \**otokono ko*, literally 'male child'.]

2.2 As a rule, when compounds are enlarged with prefixes or suffixes, do not break them up; however, divide the compound if the meanings of the constituents are independent and the pronunciation allows. [E.g., according to TTB 1981:41, *hankakumei* 'counterrevolution', not \**han kakumei*, but *han shakaiteki* 'antisocial'. Apparently, the presence of an accent-phrase juncture in normal speech plays a role here since in both cases the *han* 'anti-' is the same prefix.]

2.3 As a rule, break up long compounds if there are two or more parts which can be taken as having independent meanings. [E.g., TTB 1981:41, *Nihonkoku kenpō* 'the Japanese constitution'.]

2.4 Even if a compound is composed of words which could [otherwise] be independent, do not divide it if doing so results in the loss of the original meaning. [E.g., TTB 1981:42, *keizaigakusha* 'economist', not \**keizai gakusha*, literally 'economy' + 'scholar'.]

2.5 Even if a compound is composed of two words which could [otherwise] be independent, do not divide it if the initial consonant of the second word has been voiced. [E.g., *udedokei* 'wristwatch' < *ude* 'arm' + *tokei* 'clock'.]

2.6 Words written with three or more kanji in which the kanji are all on a par with one another should either be written all together or divided character by character depending on the pronunciation and meaning. In four-character compounds, when the first and second pairs could be independent words, one may break between the second and third characters. [Thus, TTB 1981:42 shows *shunkashūtō* 'spring/summer/autumn/winter', the Sino-Japanese pronunciation for the four-character compound indicating the four seasons, but *haru natsu aki fuyu*, the Japanese rendition of the same four-character compound. On the other hand, *tōzai nanboku* 'east/west/south/north', the four directions, has only one space since *tōzai* and *nanboku* are compounds in their own right.]

2.7 Run together compound names used in botany, zoology, and other sciences in which the compounding is strong, or insert hyphens between the constituent parts.

2.8 Break between numbers and their units. [In Japanese, this includes two important cases besides the '10 meters' sort of phrase common in English. One is *san gatu* 'March' (literally, 'Month 3'), *jūku niti* 'the 19th' (literally, 'Day 19') and other dates, including names of years; the other is 'numeral-plus-counter' constructions such as *yon mai* 'four sheets (of paper etc.)', *ni hai* 'two cups (of coffee, etc.)'.]

2.9 In the case of compound verbs formed by a four-or-more character compound plus *suru*, break before the *suru* if you put breaks in the body of the compound.

2.10 Conjunctions and compound adverbs are, as a rule, undivided; but they may be divided where there are breaks in pronunciation.

2.11 Do not divide reduplicated words of up to four syllables. For six syllables or more, divide the word unless this results in a change of meaning. [This appears to be overly cautious; it is hard to think of a reduplicated word of six or more syllables which would change in meaning even if broken.]

3.1 Divide surnames from personal names. [E.g., *Tanaka Tarō*.] Chinese and Korean names of three or fewer kanji may, however, be run together. [E.g., *Mōtakutō* 'Mao Zedong'.] And in other foreign names, hyphens may be inserted as necessary.

3.2 Divide titles following names if they could be independent words; otherwise, run the name and title together or insert a hyphen. [E.g., *Yamada sensei* 'Professor Yamada' but *Homeinishi* 'Ayatollah Komeini'.] But to make it easier to recognize surnames, break before the [common] suffixes *san*, *sama*, *kun*, *dono*, *shi*, and *uji*.

3.3 Run together nicknames, affectionate names, etc.

3.4 Run together titles when they are attached to common nouns.

3.5 Break between prefecture and city, etc., in place names; but between place names and their designations, use a hyphen or make no break. [E.g., *Tōkyōto* or *Tōkyō-to* 'metropolitan Tokyo', not \**Tōkyō to*; but *Tōkyōto Fuchūshi* 'Tokyo metropolitan area, Fuchu City'.]

3.6 Run together the old names for the provinces. [E.g., *Yamatonokuni* 'Yamato province' which, were it not for this rule, would be *Yamatono kuni* 'province of Yamato'.]

3.7 Names of rivers ending in *-kawa* ['river'], and of mountains ending in *-san* or *-yama* should be divided if longer than two characters; otherwise, they should be run together. [Presumably, in view of 2.5, this

rule does not apply to names such as *Shinagawa* or *Hieizan*; however, the use of kanji in the text leaves some room for doubt.]

3.8 Names of countries, products, groups, corporations, etc., should be divided where the meaning permits.

4.1 Division of classical Japanese should conform to modern usage.

4.2 Division of dialects should conform to the standard usage.

Although the word *bunsetsu* 'syntactic phrase' does not occur anywhere in these rules, their net effect is to prescribe a system of *bunsetsu wakachigaki*.<sup>7</sup> Significantly, certain extra breaks are called for even when the *bunsetsu* criterion does not require them. A similar though more complex set of rules for word divisions has been adopted for bibliographic purposes outside Japan (Library of Congress 1982), suggesting that these extra breaks are desirable for reasons which go beyond the immediate exigencies of braille. At any rate, when considering the question of text compression, we can assume that the rules for Japanese braille spacing are also suitable for spacing romanized Japanese texts as well.<sup>8</sup>

What, in fact, is the compression achieved in Japanese braille by using a kana-based system rather than a roman-based system? To answer this question, we can use the statistics reported in Hayashi et al. 1982:319. This is a synopsis of a 1960 study of Japanese digrams (two-kana sequences) based on 40,000 kana of running text culled from a wide variety of popular and professional literature. The units for this survey were 102 syllables of following types: the Q and N syllables (*sokuon* and *hatsuon*) shown in Figure 2; 33 syllables with intrasyllabic glides (*yōon*, e.g., 'hya', 'byu', 'pyo'); 5 syllables with initial /p/ (*handakuon*, e.g., 'pa'); 18 syllables with voiced obstruent initials (*dakuon*, e.g., 'ba'); and 44 other syllables (*seion*, e.g., 'ha', 'ya', 'a').<sup>9</sup> Syllables of the type 'bya' and 'pya' are counted as *yōon*, not *dakuon* or *handakuon*. In Japanese braille, there are an additional three *seion* ('wi', 'we', 'wo'), two *dakuon* ('di', 'du'), and three *yōon* ('dya', 'dyu', 'dyo'). Except for 'wo', which is used exclusively to write the accusative marker, these and the 'innovative' syllables found in some recent foreign loanwords occur very seldom. In the survey, long /a/, /i/, /u/, and /e/ are counted as double vowels; long /o/ is treated as 'o' + 'u'. The braille rules for long vowels are different, but always call for one cell (sometimes a 'long mark', sometimes the single-vowel syllable corresponding to the vowel of the preceding). Also, the postpositional particles pronounced /wa/ and /e/ but written with the kana for 'ha' and 'he' respectively are counted in the survey and written in Japanese braille as they are pronounced. Thus, the reported frequencies of syllables in the survey can be assumed to reflect the frequencies of syllables in transcribed Japanese braille texts.

By using the graph in Hayashi 1982 it is easy to calculate that for every thousand syllables of running Japanese text, 250 will be expressible with one letter in romanization, 718 with two letters, and the remaining 32 with three letters.<sup>10</sup> Thus, 1,782 letters will on average be needed to transcribe 1,000 syllables. On the other hand, 142 of every thousand syllables will require two Japanese braille cells; the remaining 858 will require only one.<sup>11</sup> Therefore, 1,142 braille cells on average will suffice to transcribe the same text.<sup>12</sup> Thus the compression ratio of roman over kana is 1.560. If, instead of comparing roman with kana, we were to compare it with a hypothetical roman-based braille that makes use of contractions, what would be the result?

Let's assume that contractions can be formulated for Japanese which are at least as efficient as the Grade 2 English contractions at reducing sentence length. Even if Krebs' estimate of "almost twice the number of words" is interpreted conservatively, this would still be better than the current kana-based system. Of course, it is one thing to *assume* that such a set of contractions exist for Japanese and quite another to produce it. On the other hand, since Japanese syntax is agglutinative, the same strings of syllables occur over and over at the ends of phrases and clauses; a set of Japanese contractions might actually be more efficient at shortening texts than the English contractions. A rough calculation using KKK 1970:281ff., a table of word frequencies in newspapers based on extensive random sampling, supports this view. The definition of 'word' in the context of KKK 1970 is closer to the linguistic concept of morpheme than to the 'word' defined by the rules of Japanese braille spacing; but for our purposes, these differences can be ignored. If the first N most frequently occurring words of Japanese were, in a roman-based Japanese braille, replaced by single-cell contractions, how would this affect the ratio of cells to syllables? KKK gives the cumulative number of the N most frequent words per mil (i.e., per thousand); the average number of syllables per word is just over two (Yoshida & Tanaka 1981:318). Thus every high-frequency word replaced by a contraction removes two syllables and adds a single cell; the number of cells per thousand syllables, C, can be calculated from the cumulative number of high-frequency words per mil, w, for the first N words in the list with the formula  $C = 1782 (1 - 2w/1000) + w$ , which makes use of our earlier finding that 1,782 cells would be needed on average to transcribe a thousand syllables without contractions. Sample results follow:

N	W	C
100	343.815	900.46
125	359.669	859.81
150	372.393	827.18

Grade 2 braille has about 190 contractions, but, of course, many of these are two cells long; in this respect, our C are too low. On the other hand, our calculation excludes all those strings of two or more syllables within lower frequency words that would also be contracted; in this respect, the C are too high. These two effects tend to cancel each other out. Clearly, our hypothetical roman-based Japanese braille with contractions would outperform the current kana-based system. Of course, there are many cultural, practical, and historical factors which favor the continuation of the current Japanese braille system. But, interestingly, compression efficiency does not seem to be one of them. Kana are, at least in this respect, not optimal.

1. Those familiar with computers may find it convenient to think of dots 1 through 3 and 4 through 6 as the three bits of an octal digit; thus each braille cell is a six-bit byte with a value in the range 000 to 077 (i.e., 0 to 63). E.g., 'm' (dots 1-3-4) is 051 (= 41); 'j' (dots 2-4-5) is 023 (= 19); cf. Figure 1. This is one way to number the cells so that they can be ordered.

2. The Kokugogaku jiten (1955) does not say whether Ishikawa himself was blind, but the outline of his career suggests that he was not. (Incidentally, the biography of Ishikawa and other 'short entries' [*shōkō*] have been omitted from the new [1982] edition of the Kokugogaku jiten.)

3. In discussing the problems that arose in unifying British and American Grade 2 braille in the '30's, due largely to American disapproval of the British practice of contracting sequences of letters even if they crossed syllable boundaries, Loomis (1936:6) comments that, "In the effort to protect the reader from anything that might be considered illiterate or confusing, one thing was always permitted and never questioned: *the use of the contraction irrespective of its sound or pronunciation.*" The emphasis is Loomis's. For our purposes, however, the interesting part is what comes before the colon. Statements like this tend to suggest that braille is transliterative rather than phonological as a result of preconceptions held by sighted educators; but this is probably not the case. Although the principles of phonology were not clearly articulated in the early 19th century, when the words 'letter' and 'sound' were often used interchangeably, the fact remains that, in French, rules of *liaison* and *enchaînement* would have made a strictly phonemic treatment quite difficult; and, in view of dialectal and personal variations, it is obviously much easier to train transcribers to transliterate according to a fixed standard than to phonemicize (consider, for example, English words like 'either'). Finally, one must take the social and historical setting into account: Braille's object was to put the blind on as nearly an equal footing with the sighted as possible. The major obstacle to be overcome was, as Loomis points out, resistance to the idea of characters that bore no visual relationship to the shapes of the letters of the alphabet.

4. Grade 2, with 189 contractions and short-form words, is now standard in the English-speaking world. Grade 1 is simply uncontracted (straight transliterative) braille. A Grade 1½ was standard in the U.S. before the compromise discussed by Loomis: it used only 44 contractions, and avoided contractions across syllable

boundaries. Grade 3 "is an extension of Grade 2, by using additional contractions and short-form words, and by the use of the outlining (the omission of vowels). Grade 3 contains more than 500 contracted forms and is used mainly by individuals for their personal convenience" Krebs 1974:109). Krebs also mentions braille shorthand "for use by blind stenographers."

5. Considerations of this kind obviously played some role in Braille's original formulation of signs for the letters of the alphabet. Only one letter, 'a', is represented by a one-dot cell, viz. dot 1. Presumably, this is because it is relatively difficult to sense the relative location of a single dot within its frame of reference unless one can depend on certain kinds of cells being adjacent to provide supporting context. Likewise, it is obviously difficult to sense the difference between a 1 & 3 and a 4 & 6 when flanked by spaces. Thus, not all cells are equal: the whole theoretical range of 64 distinct cells cannot in practice be used with complete freedom. (This seems to be what Loomis has in mind when she criticizes those systems in which high-frequency letters were associated with those cells which contain the fewest dots.)

6. I have altered examples 3b through 5b since the sentences given in Tanaka & Yoshida do not conform to the standard definitions and might be misprints.

7. This point is made explicitly in TTB 1981:34. Another (but less detailed) presentation of the *wakachigaki* rules is to be found in Honma et al. 1983:64-68. Apparently, there are some differences among the various agencies that serve the blind in Japan, particularly between Kantō (Eastern) and Kansai (Western) organizations.

8. Honma et al. (1983:91f). do not mention automated transformation of *betagaki* into *wakachigaki* in their brief discussion of computer-aided braille printing. I have seen brochures for U.S. computer systems designed to carry out the analogous task of transforming standard English input into Grade 2 braille (i.e., with the correct contractions). These systems are reported to be quite expensive for this age of cheap personal computers (in the \$20,000 range), but this is no doubt due to low demand, not the complexity of software. In Japan, on the other hand, it would seem that the problem is finding a program that works at all.

9. I am using the word 'syllable' as an equivalent for the Japanese term *onsetsu*, which is sometimes translated as 'mora' in linguistic studies.

10. The 250 single-letter syllables are the five *agyō* 'single vowel' syllables (180), *sokuon* 'Q' (23), and *hatsuon* 'N' (47); the 32 trigrams are the *yōon*.

11. The 142 two-cell syllables are *dakuon* (107), five *handakuon* (3), and the *yōon* (32).

12. Notice that even kana-oriented braille requires more cells than actual syllables; this is due to the way in which *dakuon*, *handakuon* and *yōon* syllables are handled.

## References

- Hayashi Ōki et al., eds. 1982. *Zusetsu Nihongo: gurafu de miru kotoba no sugata* (Japanese in diagrams: the form of words as seen in graphs) = *Kadokawa shōjiten* (Kadokawa handbook series), 9. Tokyo: Kadokawa shoten.
- Honma Ichio, Iwahashi Hideyuki, Tanaka Nobuo, eds. 1983. *Tenji to rōdoku e no shōtai* (An invitation to braille and talking books). Tokyo: Fukumura shuppan.
- Kokugo Gakkai (Society for the Study of the Japanese Language). 1955. *Kokugogaku jiten* (Dictionary of Japanese language studies). Tokyo: Tōkyōdō shuppan.
- Kokuritsu Kokugo Kenkyūjo (National Japanese Language Research Institute) = KKK. 1970. *Denshi keisanki ni yoru shinbun no goi chōsa* (Computer survey of newspaper vocabulary) = *Kokuritsu Kokugo Kenkyūjo Hōkoku* (KKK Report), 37. Tokyo: Kokuritsu Kokugo Kenkyūjo.
- Krebs, Bernard J. 1974. *Transcriber's Guide to English Braille*. New York: The Jewish Guild for the Blind.
- Library of Congress. 1982. *Manual of Romanization, Capitalization, Punctuation, and Word Division for Japanese*. Washington D.C.
- Loomis, Madeleine Seymour. 1936. *Sequence and Syllabication*, Monograph Number 3, New York Institute for the Education of the Blind Series.
- Nihon Mōjin Fukushi Kenkyūkai (Japan Society for Research on the Welfare of the Blind) = NMFK, comps. 1982. *Hyōjun tenji hyōki jiten* (Standard dictionary for braille transcription). Tokyo: published by the compilers.
- Nihon Mōjin Shakai-Fukushi-Shisetsu Kyōgikai, Tenji Toshokan Bukai (Japan Coordination Council of Society of Welfare Institutions for the Blind, Braille Library Subcommittee) = TTB, eds. 1981. *Tenji no tebiki, nyūmon-hen* (Braille handbook, introduction). Tokyo: published by the editors.
- Yoshida Shō & Tanaka Takemi. 1981. "Kana kanji henkan ni yoru Nihongo nyūryoku shisutemu ni kansuru chōsa" ("A survey of Japanese input systems based on kana-kanji translation"), pp. 8-13 in *Kana kanji henkan o chūshin to shita Nihongo nyūryoku shisutemu no kaiatsu kenkyū hōkokusho* (Research report on the development of Japanese input systems built around kana-to-kanji conversion). *Kyūshū daigaku*.

# The Japanese Script since 1900

---

Chris Seeley

University of Canterbury, New Zealand

The postwar *Tōyō kanji* reforms were successful in helping very substantially to simplify the writing of ordinary, non-specialist Japanese texts by restricting the number of kanji in use, their readings, and their form (*jitai*). In 1981 the *Tōyō kanji* List of 1946 was replaced by the more extensive *Jōyō kanji* List. The compilation of the new list over a period of eight years, and its formal adoption, were the subject of much debate. This article gives an account of the new list in terms of its development through two draft stages to the final version, examines its contents and nature, and provides a comparison with the *Tōyō kanji* List. By way of introduction, a summary is given of changes in the Japanese script from 1900 onwards.

## 1 Script Reform: Incipient Attempts

In the late nineteenth and early twentieth century the written language was the topic of much controversy and debate in Japan. Various proposals for script reform were put forward prior to 1900, most of which fall into one or the other of three categories: (1) use of kana only, (2) use of romanisation only, (3) continuation of the existing writing system, but with restrictions on the number of kanji employed.<sup>1</sup> For the last of these measures in particular there appears to have been substantial support; not only did it have the support of those who saw restrictions on use of kanji as the ultimate goal,<sup>2</sup> but also it was favoured by certain other scholars who, though they wanted to see the conventional orthography replaced by a phonogram script in due course, considered that a direct transition to such a script would be too abrupt and drastic to have much chance of success.<sup>3</sup>

In response to the mood of the times, in 1900 the Monbushō (Education Ministry) issued new regulations aimed at simplifying the teaching of written Japanese at the primary level. The regulations involved the following: the setting-up of a limited number of kana signs as standard,<sup>4</sup> restricting the number of different *kanji* taught in primary schools to about 1200,<sup>5</sup> and finally, changing the complicated and cumbersome system of kana usage for Sino-Japanese (SJ) items from a historical system to one based on pronunciation.<sup>6</sup>

The first two of the above measures seem to have been generally welcomed, but the third was not. The main features of the rules for SJ kana usage were as follows: (1) use of the length mark in representing the long vowels *ō* and *ū* (e.g., うんとうー, べんきょー); (2) adoption of small-sized や・ゆ・よ in writing syllables such as *kyū* and *ryō* (きゅ and りょ respectively), and small-sized つ in writing a cluster of two like consonants (e.g., がっこう *gakkō*); (3) abolition of the following orthographic distinctions— か:くわ、じ:ぢ、す:づ (thereafter written as かじす respectively).<sup>7</sup>

These modifications to SJ kana usage meant that the writing of SJ items in kana was greatly simplified at the primary level—the very level where historical kana usage presented the greatest obstacle, being prominent there because of necessity only a limited number of kanji could be used. The new kana usage meant, for example, that SJ *ō*, which was traditionally written in one of six different ways, came to be written in all cases as おー.<sup>8</sup>

Since the Monbushō's new rules for SJ kana usage clearly resulted in considerable simplification, what were the objections to them? One source of difficulty was that the new kana usage related only to the writing of *kango* (SJ items), while native Japanese words and elements (NJ) continued to be written in historical kana usage, e.g., (てふ)ませう (not \*たり)ましょー). This distinction in the way of using kana to write SJ on the one hand and NJ on the other proved to be a problem for primary school children, at least those at the more junior levels, since they were not yet able to appreciate the difference between the two layers. Another problem was that this type of kana usage was employed only in the primary schools, the historical variety being used for the SJ layer in junior high schools and beyond. There was, in other words, a lack of consistency between the forms of SJ kana usage employed by different groups within society. In addition, the new kana usage was criticised by some for use of the length mark (traditionally used only in conjunction with katakana) in combination with hiragana, and for the writing of *ē* as えい and not えー

Aware of such difficulties as these, the Monbushō requested the Kokugo chōsa iinkai (Japanese Language Investigative Committee)<sup>9</sup> to look into the whole question of kana usage for both SJ and NJ. Meanwhile, the movement opposing the new kana usage gained momentum, with men such as Mori Ōgai arguing that the historical kana usage represented the essential form of the Japanese language and so should be preserved in perpetuity. As a result of the controversy which arose concerning the precise form which kana usage should take, in September 1908 the

Monbushō rescinded the regulations of 1900 relating to kana signs, kana usage, and kanji.

In this way, a temporary victory was won in retaining the script unchanged. There were, though, several factors at work within society which actively favoured simplification of the script. One such factor was a desire to shorten the length of time needed to complete an education: compared with the West, the Japanese education system entailed 2-3 years' extra study to complete a university degree, and this was seen as a desirable area for reform.<sup>10</sup> After considering how to ameliorate the situation without bringing about a decline in academic standards, Hoshina Kōichi, a language scholar who had studied under Ueda Kazutoshi (1867-1937: one of the first Japanese to study modern Western linguistic science), came to the conclusion that one possible answer lay in simplification of the script. The first aspect of the script to be taken up by Hoshina was the form (*jiti* 字体) of individual kanji.<sup>11</sup> A number of scholars brought together by Hoshina worked on the problem of *jitai*, and as a result of this a list of kanji entitled *Kanji seirian* (Proposed modifications to kanji) was published in 1919. This list, which contained about 2600 kanji, gave one recommended form of each kanji for use in both printed and handwritten texts. Though not officially adopted, the *Kanji seirian* is nevertheless of significance as an early attempt by the Monbushō to produce some sort of standard list of kanji.

During the first decades of the twentieth century, moves in the direction of simplifying the script were also being made by the newspaper companies. From early in the Meiji period, newspapers had been written in a difficult style which employed a very large number of kanji—a feature which meant that they were intelligible to only a relatively small readership. The first step towards simplification was taken in 1888, when the *Ōsaka Asahi shinbun* began adding small-sized kana to the side of kanji (so as to indicate the pronunciation of the latter) in all articles.<sup>12</sup> In 1900 Hara Takashi, chief editor (and later company president) of the *Ōsaka Mainichi shinbun*, wrote a series of articles in that newspaper in which he argued in favour of reducing the number of kanji used.<sup>13</sup>

The trend towards adoption of a less complicated orthography and more colloquial written style received further impetus after Hara Takashi rose to the position of Prime Minister in 1918.<sup>14</sup> In addition to this, with the principles of democracy taking root in Japan round about that time, there were calls from various quarters for the “democratization” of newspapers so as to make them accessible to everyone rather than to just a privileged few. This proposal was one which was attractive to the newspaper companies, for one thing because processes such as printing and proof reading were easier using a small set of kanji than a large one. On March

21, 1921, a number of Tokyo and Osaka newspapers carried a joint statement entitled *Kanji seigen o teishō su* (Advocation of restriction on the number of kanji), which announced the proposed dropping of particularly difficult kanji from newspapers and use of katakana in their place.<sup>15</sup>

In the same year that the above joint statement appeared, a new body, called the Rinji kokugo chōsakai (Interim Committee on The National Language), was established.<sup>16</sup> Three topics were decided on for early investigation by the committee, viz. a kanji survey (with a view to reducing the number of kanji in use), revision of kana usage, and modification of the colloquial written style.<sup>17</sup> As a result of deliberations, in May 1923 the Interim Committee produced the *Jōyō kanjihyō* (List of kanji for general use), a list containing the kanji which were considered necessary for writing ordinary Japanese texts.<sup>18</sup> The *Jōyō kanjihyō* was welcomed by the major newspaper companies, which on August 5, 1923, announced their intention to base kanji use in newspapers on the list as of September 1 of that year.<sup>19</sup> Unfortunately, September 1, 1923, turned out to be the day of the Great Kanto Earthquake, and so as a result of the ensuing chaos and disruption the proposed kanji restrictions had to be postponed. Two years later a number of Tokyo newspaper companies held discussions and then put out another joint statement,<sup>20</sup> reaffirming their intention to restrict the number of kanji used in newspapers, and proposing to follow the same principle as far as possible in newspaper advertisements also.<sup>21</sup> It was decided to employ a modern version of the 1923 *Jōyō kanjihyō* consisting of 2108 kanji, and not to use kanji outside the list other than in proper nouns, the texts of Imperial rescripts and of laws, and unavoidable quotations. The amended list of 1925 did not, however, have any greatly beneficial effect. The reason for this was that particularly difficult kanji were simply replaced by kana which were identified as substitute signs by the addition of a dot or dots, e. g. 疑 *しんじゆん* for 狐疑 *しんじゆん* *kogi shunjun* "hesitancy, vacillation", or 苛 *かれんちうきう* for 苛 *かれんちうきう* *karen chūkyū* "extortion". Rewriting certain *kango* in this way had the advantage of reducing the overall number of kanji used, but in some cases only at the cost of semantic clarity. To help overcome this difficulty, therefore, between 1926 and 1928 the Interim Committee compiled its *Kango seirian* (Proposed modifications to *kango*) in which a number of complicated kanji were replaced by simpler ones. The Interim Committee also set itself the task at about that time of making modifications to the *Jōyō kanjihyō* of 1923, and in 1931 made public proposed revisions resulting in a kanji list of 1856 signs.<sup>22</sup> However, in September 1931 there was the outbreak of the Manchuria Incident and this made it difficult to keep within the limits of the new kanji

list for several reasons: firstly, there was the need to write a large number of Chinese personal and place names, many of which involved uncommon kanji; secondly, the military, who reported on the Manchuria Incident, favoured a written style which was characterized by heavy use of *kango* and difficult kanji.

The Manchuria Incident, with the ensuing mood of conservatism and nationalism which it evoked, also created an obstacle to the revision of kana usage. In 1931, at the same time as the kanji revision proposals, the Interim Committee put forward a proposal to modify an earlier projected pronunciation-based kana usage (proposed in 1924).<sup>23</sup> Subsequently—in June of that year—the Monbushō determined to adopt the modified kana usage in primary school textbooks from the following year. However, public opinion on the issue was so divided that, in the end, the Monbushō abandoned its proposed course of action.<sup>24</sup>

The Interim Committee on The National Language, had, as its name suggests, originally been set up only as a temporary body, and so in 1934 a new organisation called the Kokugo shingikai (Deliberative Council on The National Language) was formed. The following year, the Education Minister (Masuda Genji) requested the Deliberative Council to inquire into the following four topics: “control of the national language” (*kokugo no tōsei*, a kanji survey, revision of kana usage, and improvement of written style. The first of these topics to be taken up by the Council was the kanji survey, but a new and major obstacle was presented at about this time by the emergence of a view of kanji which was linked to the preservation of the *kokutai* “national policy”.<sup>25</sup> The new intellectual climate that prevailed can be illustrated by the case of the Education Minister Hirao Hachisaburō. Originally in favour of abolishing kanji, Hirao was subsequently forced to draw back from this view in both the Upper and Lower House after his critics maintained that to abolish kanji would mean changing the form of Imperial rescripts, an act which, they argued, would be disrespectful towards the Emperor and would have a negative effect on the Japanese spirit because writing and thought were, in their view, inseparable.<sup>26</sup>

In this intellectual environment the Deliberative Council clearly had to exercise great circumspection in its deliberations on kanji reform. In consequence its *Kanji jitai seirian* (Proposed modifications to the form of kanji) of 1938 contained two categories of kanji — one category for use in school textbooks and for general use, the other for special texts such as Imperial rescripts. Thus, the kanji 關 辭 變 繼, for example, were to be replaced in special texts by their unsimplified, more formal equivalents 關 辭 變 繼 respectively.<sup>27</sup> Opposition to the list from conservative elements meant that the *Kanji jitai seirian* was not put into effect at

that time. It is, though, considered to have had substantial influence on the form of kanji appearing in the *Tōyō kanji jitaihyō* (List of forms for the kanji for current use) adopted in 1949.<sup>28</sup>

The *Kanji jitai seirian* addressed itself primarily to the question of the recommended forms of kanji. Shortly before this list was completed, the Deliberative Council set about the task of compiling another list which was to contain the kanji needed for writing ordinary texts. Completed in June 1942, the new list was given the title *Hyōjun kanjihyō* (List of standard kanji). The purpose of this list as explained in the Introductory Remarks (*Hanrei*) was as follows: "In recent times kanji have been used in Japan without limit, at considerable inconvenience to life in society, and so by modifying and controlling kanji usage this List shews the standard for kanji to be used in Government offices and society in general."<sup>29</sup>

With regard to kanji not appearing in the list, the introductory remarks state: "Kanji which are not in this list should, with the exception of those used for writing proper nouns, in principle be written in kana."<sup>30</sup>

The *Hyōjun kanjihyō* consisted of a total of 2528 kanji—considerably more than other lists of its type. Kanji were divided into three categories, viz. those of high frequency (*jōyō kanji*; 1134) those of lower frequency (*jun jōyō kanji*; 1320), and special kanji occurring in the text of the Imperial Constitution, Imperial rescripts, and the like (*tokubetsu kanji*; 74). In his report of June 17, 1942, to the Education Minister (Hashida Kunihiko), Deliberative Council Chairman Minami Hiroshi explained *inter alia* how it was envisaged that kanji of the three categories would be assimilated into the education system.<sup>31</sup> It was expected firstly that school pupils would be taught to read and write accurately all 1134 of the high frequency kanji; with regard to kanji in the other two categories, the ability to read these was to be regarded as sufficient.

The *Hyōjun kanjihyō* as compiled by the Deliberative Council was circulated by the Monbushō to other Government departments for them to comment upon. As a result of this, a number of requests were received for further kanji to be added, and in December 1942 the Monbushō published an amended *Hyōjun kanjihyō* containing 2669 kanji—141 more than in the list as first compiled. In addition to increasing the number of kanji, the *Monbushō* also abolished the division of kanji into three categories. Both these moves diminished the value of the *Hyōjun kanjihyō* as a list to help reduce the number of kanji in daily use.<sup>32</sup>

In June 1942 the Deliberative Council also presented its *Shin jion kanazukaihyō* (New kana usage list for Sino-Japanese). The kana usage proposed in this list was based on pronunciation of SJ elements in the standard language, and was in essence similar to that outlined in earlier proposals of 1924 and 1931.<sup>33</sup> Like its predecessors, however, the New

Kana Usage List failed to be adopted because of the controversy which developed between the 'traditionalists' on the one hand and the 'phoneticians' on the other.<sup>34</sup>

In this way, a thorough-going reform of the Japanese writing system was not carried out in the first part of the 20th century, despite calls and proposals from various quarters. As a result of this failure to reform, Japan was left heir to an intricate and complex writing system which, though it shewed respect for historical principles, often did so only at the expense of clarity of meaning, and sometimes this had disastrous consequences. This point is best illustrated by the case of the Japanese Army. In the late 1930's the shortage of suitable recruits meant that the Army was obliged to take on recruits of an educational standard that became lower and lower, to the point where eventually some of them were unable to read certain of the kanji used to write the names of weapons and related terms. This, it is reported, was the cause of a number of accidents involving weapons.<sup>35</sup> Faced with this situation, the Army Ministry decided to simplify the use of kanji and kana for weapon nomenclature and related texts. To this end, in February 1940 it issued a directive entitled *Heiki meisshō oyobi yōgo no kan'ika ni kansuru kitei* (Regulations relating to the simplification of weapon nomenclature and terms), restricting the number of kanji for weapon names and terms to 1235. This was followed in March 1941 by a directive to the effect that a pronunciation-based kana usage be adopted for weapon names and related texts; this latter directive, *Heiki ni kansuru kanazukai yōryō* (Summary of kana usage relating to weapons), followed the kana usage set out in the Interim Committee's proposals of 1931.

## 2. The Tōyō Kanji Era

Although in the earlier part of the twentieth century there were already various scholars and other prominent figures who were in favour of simplification of the script, until the end of the Pacific War such reform was frustrated by more conservative elements. After the war defeat of 1945, however, the situation changed dramatically, and there sprang up overwhelming support for major orthographic revision.<sup>36</sup> The first postwar meeting of the Deliberative Council was held in November 1945, and at this meeting the then Education Minister emphasized that the solution of problems relating to the Japanese language and script lay at the basis of all the other reforms which were necessary for the rebuilding of Japan. The Monbushō was requested by the Allied Forces High Command to limit the number of kanji used in textbooks to about 1500,<sup>37</sup> and so from December 1945 deliberations went ahead on the basis of making additions and deletions to the 1134 kanji designated as *jōyō kanji* in the *Hyōjun kanjihyō* (List of standard kanji) of 1942.

In March 1946 an American educational delegation visited Japan with the purpose of studying the Japanese educational system and then advising the Allied High Command and the Monbushō. With regard to language reform, the delegation recommended in its report the eventual adoption of romanization.<sup>38</sup>

In April and May 1946 the Deliberative Council discussed in plenary sessions a proposed kanji list of 1295 signs, *Jōyō kanjihō* (List of kanji for general use), that had been compiled by a committee of Council members under the chairmanship of Yanada Kyūjirō.<sup>39</sup> The list was, however, considered too expensive for use in elementary education, yet too restrictive for general use. In consequence, a new kanji committee was set up under the chairmanship of Yamamoto Yūzō, a writer who had shown considerable enthusiasm for simplification of the script. At the same time, a kana usage committee was set up in order to put forward proposals for reform in that area also.<sup>40</sup> The new kanji list compiled at this time, which profited by experience and included all the kanji found in the text of the new Constitution, was the *Tōyō kanjihyō* (List of kanji for current use), consisting of 1850 signs. The new list was based on the list of 1295 kanji proposed earlier in 1946, but supplemented with additional kanji which were selected after soliciting the opinion of Government offices, newspaper companies, and the like.<sup>41</sup> After being approved by the Deliberative Council at a plenary session on November 5, the List of Kanji for Current Use (hereafter normally referred to as 'the TK List') was announced and promulgated by the Cabinet just eleven days later, on November 16, 1946. That date also saw the promulgation by Cabinet of the *Gendai kanazukai* (Modern kana usage), the rules for which are based on pronunciation in the modern standard language.<sup>42</sup>

The TK List and modern kana usage together represent a major simplification to the Japanese script compared with before the War, but it is interesting to note that both are in fact based on earlier proposals — the former on the *jōyō kanji* in the List of Standard Kanji of 1942, the latter on the New Kana Usage List for Sino-Japanese (also of 1942).<sup>43</sup> Thus, although various of the earlier attempts at script reform had been rejected at the time, they did provide a basis for the reforms of 1946.

The new orthography of 1946 was adopted in newspapers, magazines, and official texts of all kinds almost immediately and was put into effect in school textbooks from April of the following year. Although the number of kanji in the TK List represented a substantial reduction compared with, for example, the List of Standard Kanji of 1942, it was found that it was still not possible to teach school pupils to satisfactorily read and write all of them within the nine years of compulsory education. The situation was remedied by compiling the *Tōyō kanji beppyō* (Separate list of kanji for

current use; referred to below as the “Separate TK List”). Submitted to the Education Minister in September 1947 by the Deliberative Council, the Separate TK List, which set out 881 TK List kanji, was approved and promulgated by the Cabinet in February 1948.<sup>44</sup>

From the very outset the TK List had the desired effect of reducing very substantially the number of different kanji in general use. There remained, though, several areas in which further regulation was required. Firstly, while the TK List contained all the different kanji for general use, it did not give any guidance concerning *on* and *kun* readings; in other words, even obscure readings were permitted, provided that the kanji themselves were in the TK List. Clearly this ran counter to the underlying aim of the TK List, and so in order to rectify this state of affairs an official list of *on* and *kun* readings was announced by Cabinet in January 1948, *Tōyō kanji onkunhyō* (List of *on* and *kun* readings for the kanji in current use; referred to below as the “TK *Onkun* List”). After this list came into effect, many older and literary *kun* readings ceased to have the official seal of approval (e.g., *tagui* for 類, *nottoru* for 貝), as did most irregular *kun* readings for kanji combinations (e.g., *momiji* for 紅葉, *tabako* for 煙草). *On* readings were also reduced in number; in the case of 請, for example, the 1948 TK *Onkun* List contains the *on* readings *sei* and *shin* (the latter being noted as a relatively uncommon reading), but not the reading *shō*; similarly, the reading *jin* is given for 仁, but not *nin*.

In addition to *on* and *kun* readings, another main area in which a degree of standardization was desirable was the form of individual kanji. Several earlier attempts had been made to deal with this problem (see above), but these had not advanced beyond a draft stage, and in any case the question of the form of kanji is a very intricate area, and not something for which a solution could be found overnight. For a substantial number of kanji there was commonly some difference between (a) the printed form and (b) the handwritten form - e.g., (a) 衣, (b) 衣; (a) 玄, (b) 玄. For some kanji there were variant forms which were used more or less interchangeably—e.g., 叙 ~ 叙 ~ 叙, 宝 ~ 寶 ~ 寶, 場 ~ 土場. In addition, simplification of forms was sometimes achieved only at the expense of structural principles of the kanji script; 突 *tsuku* “to thrust, pierce”, for example, is a kanji of the semantic compound type, consisting of the two elements 宀 “hole, cave” and 犬 “dog”, but this etymology is obscured in the slightly abbreviated form 突.

In order to reduce the inconvenience and problems caused in such areas as education and the printing trade by the form of kanji, in July 1947 the Deliberative Council set up a special committee, the members of

which represented the world of education, printing, newspapers and Government offices. As the title of the committee, *Katsuji jitai seiri ni kansuru kyōgikai* (Committee on the printed form of kanji), suggests, it was concerned primarily with the standardization of printed forms, but the draft proposals which it presented to the Deliberative Council in October 1947 did nevertheless attempt to reduce the gap which existed between printed forms and their handwritten equivalents.<sup>45</sup> After due consideration, which included the setting up of a further committee,<sup>46</sup> the Deliberative Council submitted its *Tōyō kanji jitaihyō* (List of forms for the kanji for current use, referred to below as the "TK Forms List") to the Education Minister in June 1948. The list was then announced and promulgated by Cabinet in April of the following year.

As a result of the implementation of the TK List, the reading and writing of Japanese texts was greatly simplified, but in one area of everyday life—the use of kanji to write given names—the new, simplified orthography was generally felt to impose an excessive restriction: in the past, parents had been free to choose virtually any kanji to represent in writing the given names of their offspring, but after December 1947 they were limited to the 1850 kanji in the TK List, the only alternative being to write given names in *hiragana* or *katakana*.<sup>47</sup> So as to give parents a little more flexibility in choosing kanji for given names, in May 1951 the Deliberative Council submitted a written recommendation to the Attorney General and the Education Minister to the effect that 92 specified kanji be permitted for use in given names in addition to kanji in the TK List.<sup>48</sup>

The 92 kanji set out in the Council's recommendation were approved by the Cabinet, and announced and promulgated later the same month in the form of the *Jinmeiyō kanji beppyō* (Separate list of kanji for use in given names; referred to below as the "TK Name Kanji List").<sup>49</sup>

The term *tōyō kanji* has been rendered in this article as "Kanji for current use"; it could equally well be translated as "Kanji for temporary use". Implicit in this term, in other words, there was the suggestion that the TK List was to be modified as necessary at some time in the future. From 1946 onwards, newspaper companies proposed to the Deliberative Council in the light of experience that a number of kanji be added to the TK List. The Kanji Section of the Council, after considering this matter, recommended in March 1954 that 28 new kanji be added to the TK List, and that 28 other kanji already in the list (but of very low frequency) be dropped. In plenary session (March 15, 1954), however, the Deliberative Council was of the general view that the time was not yet ripe for the TK List to be modified in this way. Consequently, instead of proposing to the Education Minister that the TK List be amended, on this occasion the Council simply received the Kanji Section's recommendations, and made

them public.<sup>50</sup> This meant that from that point on there was some divergence, albeit relatively minor, between kanji usage in newspapers and that in school textbooks and a variety of other texts of a general nature.

During the late 1940's, major reforms in the usage of kanji and kana were carried out as described above. One transitional area between the two still awaited reform—or rather standardization—at that time, namely *okurigana* (kana added following kanji in order to shew inflectional or other endings). *Okurigana* usage falls into one of two broad categories: the 'explicit' type, which often employs kana to represent not just the changing inflectional ending of verbs and adjectives, but also part of the stem (e.g., 行<sup>な</sup>は<sup>る</sup>; *okonau*, 異<sup>な</sup>は<sup>る</sup> *kotonaru*) and the 'implicit' type, which tends to use as few *kana* as possible (e.g., 行<sup>ろ</sup>, 異<sup>る</sup> ). *Okurigana* of the former type tended to be favoured in school textbooks, while the latter type was common in other texts.

Faced with this situation, the Deliberative Council's Orthography Section sought to standardize *okurigana* usage. To this end, it drew up 26 rules for usage, which tended towards the 'explicit' type. These rules, presented by the Deliberative Council as a written recommendation to the Education Minister in November 1958, were approved and promulgated by the Cabinet in July 1959.<sup>52</sup>

By this time all the main areas of the modern Japanese writing system had undergone revision, this often being extensive. The work of reform, though, was by no means complete: some problems remained from the pre-War period, while others had arisen as a result of the new orthographic reforms themselves. The need for continuing revision in this area was brought to the attention of the Education Minister (Nadao Kōkichi) in a report submitted by the Deliberative Council in October 1963 entitled *Kokugo no kaizen ni tsuite* (Reform of the national language).<sup>53</sup> Several years later (in June 1966), Education Minister Nakamura Umeyoshi responded to this report by requesting the Deliberative Council to re-examine the areas which had been specified in its report of 1963, namely: the TK List, together with the Separate TK List (policy regarding selection of kanji therein; scrutiny of kanji in the lists; *on* and *kun* readings; the form of kanji); *okurigana*; modern kana usage; and, finally, any other related points.<sup>54</sup>

The first areas to be given further consideration by the Deliberative Council were *on* and *kun* readings, and *okurigana*.<sup>55</sup> The new *Tōyō kanji onkunhyō* of June 1973, which represented the Council's recommendations, was more liberal than the TK *Onkun* List of 1948, containing approximately 800 additional such readings,<sup>56</sup> and also giving official approval for a group of 106 irregular but long-established *kun* readings for

certain kanji combinations such as 今日 *kyō* "today", 梅雨 *tsuyu* "rainy season", and 息子 *musuko* "son". A less restrictive approach was also evident in the *Okurigana no tsukekata* (Method of adding *okurigana*) announced by Cabinet in June 1973, which allowed more flexibility of usage than the equivalent rules of 1959. The more liberal nature of the 1973 TK *Onkun* List, and the new *okurigana* rules announced at the same time, reflects a change in the collective opinion of the Deliberative Council which was due to a re-organisation of that body in the early nineteen-sixties. The less restrictive attitude of the re-organised Council was, as we shall see below, also brought to bear on the TK List and its eventual successor.

### 3. Beginnings of the New Kanji List

The Deliberative Council took up anew the question of the TK List in November 1972. Two years later, a preliminary report on the Council's basic policy with regard to kanji was completed. The basic policy, the purpose of which was to provide guidelines for future deliberations, included the following points: (1) a kanji list was necessary, but it should not be restrictive in nature like the TK List; (2) kanji for the list were to be selected for use in the ordinary life of society, i.e. for use in official texts, newspapers, magazines, broadcasting, and the like; such kanji were not intended to encompass usage in specialist fields such as science and technology of arts, or the form of writing of individual persons; (3) the field of application of the kanji list was *modern Japanese writing*.<sup>57</sup> The question of whether the kanji list should serve only as a guide (*meyasu*)—the status which had been accorded to the 1973 *Oonkun* List—was deferred for later consideration, as also was the question of how many kanji should be included in the list (though with regard to this latter point, it was understood that there should not be any dramatic departure from the number of kanji in the TK List at that time).<sup>58</sup>

In January 1974 a Kanji Committee was set up by the Deliberative Council. The Committee agreed that it would follow the basic policy on kanji that had been formulated, and resolved to refer to as wide a range of relevant material as possible (studies of kanji frequency in present-day newspapers and magazines, selection criteria for kanji in earlier kanji lists, and material which focussed on the question of how certain *kango* originally written with kanji not included in the TK List were dealt with after the TK List came into effect). The Committee also resolved that it would not be excessively influenced by the TK List in compiling a new kanji list, but that it would nevertheless still treat it as an important source material. During the course of its investigations, the Kanji List Committee drew up a series of points to be borne in mind when considering the

selection of kanji for the new list. These points were as follows: (1) frequency of use; (2) function; (3) the question of kanji for proper nouns, and how to deal with these; (4) the need to attempt a typological classification of kanji in terms of semantic areas—topography, natural phenomena, flora and fauna, the body, kinship terms, and so on; (5) the part of speech being represented; (6) usage in terms of vocabulary—single kanji each of which represents a *kango*, e.g., 胃 *i* “stomach”, 塾 *juku* “private coaching school”; kanji for monosyllabic NJ words, e.g., 蚊 *ka* “mosquito”, 藻 *mo* “seaweed”; different kanji which are of identical or similar meaning, and which share the same *on* or *kun* reading, e.g., 回廻 *kai* “go round”, 膚肌 *hada* “skin”; kanji used to represent foreign loanwords; e.g., 缶 *kan* “can, tin”, 頁 *pēji* “page”; (7) kanji structure—degree of complexity (complicated kanji such as 鬱鬱 *utsu* “gloom, depression”), and whether a kanji occurs as a constituent element in other, more complex kanji 皿 *sara* “dish”, 屯 *ton* “station, barracks”).<sup>59</sup> Consideration was also given by a special subcommittee to the form of kanji in the new list, but it was decided that the forms in the existing TK Forms List should be followed in principle.<sup>60</sup>

The progress made by the Deliberative Council in regard to a new kanji list in the two years from November 1972 to November 1974 (the period corresponding to the 11th Session of the Council)<sup>61</sup> was made public in the form of the *Shingi keika hōkoku* (Report on proceedings). The Report was distributed widely, and responses received from a number of interested groups. One such group was the Japan Newspaper Association Editorial Committee, which in March 1976 submitted a reply to the effect that it would like great care to be exercised in the selection of kanji, and that it would like the total number of kanji in the new list to be kept below 2000; the Editorial Committee also submitted a list of 79 kanji in the TK List as candidates for omission from the putative new list, and 16 kanji as candidates to be added.

In the latter part of 1976, it is reported, the Kanji List Committee set about its task with great vigour,<sup>62</sup> and in the following January the *Shin kanjihyō shian* (New kanji list draft) was drawn up, and submitted to the Education Minister (Kaibu Toshiki).

The New Kanji List (Draft) contained 1900 kanji, and so in terms of number of signs did not constitute any major departure from the TK List. However, the very fact that any difference at all was proposed in this respect was of considerable significance, marking a new and more liberal attitude towards use of kanji. This same liberality was also reflected in the fact that it was also proposed that the new List should serve as a guide

(*meyasu*) for kanji use rather than as a rigidly prescriptive list. The nature of the New Kanji List (Draft) is made clear in the following excerpt from the Preamble to the List: "The New Kanji List contains kanji which are efficient and which also share a high degree of common usage, for use in the general life of society—laws and ordinances, official texts, newspapers, magazines, broadcasting, etc. The list aims to provide a guide in the general life of society for writing in an easily-intelligible and clear manner."<sup>63</sup>

The 1900 kanji in the New Kanji List (Draft) consist of all the TK List kanji except for the following 33, which are of low or very low frequency:

芋謁歐効堪蕪嗣遊嫡遙痘陪畔濫且但虞脹勺畝匆侯爵帥朕錘銑薪奴翁隸婆嚇  
together with the following 83 additional kanji:

猿凹渦靴稼涯垣殼瀉褐岳頑挾襟隅溪蚩嫌洪溝昆崎皿傘肢遮蛇酌汁塾尚宵  
繩唇基搢杉齊逝仙栓挿槽藻馱濯挑眺釣塚漬亨泥棟洞凸把霸漠肌鉢披扉猫頻  
瓶霧塀泡倅褒朴僕堀磨岬妄厄癩悠羅竜戾。

The above list of additional kanji includes 24 of the 28 recommended in 1954 by the Deliberative Council's Kanji Section for adoption at a future date. Similarly, 15 of the 33 kanji omitted from the New Kanji List (Draft) were among those recommended for deletion earlier in 1954.

The TK List of 1946 had consisted of a bare list of kanji, which were set out according to the order of the 214 radicals; this List was then supplemented by the TK *Onkun* List (January 1948) and the TK Forms List (April 1949). The New Kanji List (Draft) is more convenient in this respect in that individual kanji are set out in the approved form, together with the recommended *on* and *kun* readings, in the one list.

At the stage of the New Kanji List (Draft), the following points were noted as needing to be dealt with at a later date: (1) the question of kanji for use in proper nouns, especially the 120 kanji designated for the writing of given names only; (2) the question of how to accommodate kanji in the new List into the education curriculum; (3) various specific problems relating to the form of kanji.<sup>64</sup>

Once the New Kanji List (Draft) had been made public, the Cultural Agency (*Bunkachō*) set about the task of assessing public opinion on that topic. In June and July 1977 special conferences were held at five centres (Sendai, Tokyo, Osaka, Hiroshima, Fukuoka) to explain and discuss the List; opinions were sought from all government departments, from public bodies, and those engaged in publishing and printing, education, the arts, and language; in addition, a public opinion survey was conducted by the Information Bureau of the Prime Minister's Secretariat. From such enquiry, the following points emerged: there was majority support for the proposed

status of a new kanji list as a guide (though at the same time there was also a feeling of concern among many lest the guide status of the list be ignored, and kanji outside the list be used indiscriminately); regarding the number of kanji in the List, there was general agreement over the total of 1900 (some felt that the total should not be increased beyond the 1850 of the TK List, but certain others considered that an increase would be in order provided that it did not go beyond the 2000 level); as for individual kanji, there was substantial support (over 20%) for nine kanji (芋 畔 奴 但 婆 侯 薪 且 翁) which had been dropped from the New Kanji List (Draft) to be retained, and more than 10% support for the three kanji (瘵 瘡 霸) to be deleted. There was, in other words, overall consensus of agreement concerning the new list, with differences of opinion essentially over points of detail.

In due course, after considering reaction to the New Kanji List (Draft), the Deliberative Council resolved the following policy in regard to the proposed final form of the list: (1) particular attention would be paid to opinions which related to laws and ordinances, official texts, and newspapers; close attention would also be given to opinions obtained from public opinion surveys; (2) kanji which had been omitted from the New Kanji List (Draft) would be retained if they had hitherto been widely used in laws and ordinances, official texts, or newspapers, and if there had also been a strong request for retention; (3) further kanji would not be added to the list without substantial cause; (4) no kanji in the New Kanji List (Draft) would be deleted without substantial cause.<sup>65</sup>

With regard to the form of kanji, a special subcommittee of the Deliberative Council's Kanji List Committee determined that: (1) the matter of the modification of the form of kanji *not* in the proposed list be handled with the greatest care;<sup>66</sup> (2) kanji in the new list should be shown using a variety of Ming type;<sup>67</sup> (3) there was a need to shew by means of examples and explanation that small differences in movable type design did not come within the scope of the list; (4) there was a need to shew that there were, in some cases, small differences between kanji in Ming type and their handwritten equivalents.<sup>68</sup>

On the basis of the above guidelines, a new list called the *Jōyō kanjihyōan* (Proposed list of kanji for general use) was compiled, being completed in March 1979. This list was very similar in nature to the New Kanji List (Draft), but there were certain overall differences. The most immediately noticeable difference was in the number of kanji: The New Kanji List (Draft) contained 1900 signs, whereas this latest list contained all these together with a further 26 signs—14 TK List signs re-adopted after being dropped from the New Kanji (Draft) 芋 謁 歐 勃 堪 蕪 嗣 遊 嫡 通 痘 陪 畔 濫

and the following 12 kanji: 拐喝矯棧褒曹棚偵搭屯抹粹 .<sup>69</sup> Another feature of the March 1979 list was the addition of a section entitled *Jitai ni tsuite no kaisetsu* (Explanation concerning the form of kanji) which set out examples of very minor variation in the form of one and the same kanji in Ming type which represented nothing more than a difference in type design. This section also gives examples of minor differences to be found in some cases between Ming type kanji and their handwritten equivalents (see below for examples from the JK List of both these types of minor variation).

In the case of the TK List, those kanji which all schoolchildren were to be taught to read and write during the six years of primary education were specified in the Separate TK List; with regard to the new kanji list, however, it was resolved that when this came into effect the Separate TK List should cease to be used, and that the question of how to accommodate the new list into the education system should be the subject of separate enquiry. Also to be given separate consideration were kanji for special use in writing given names.<sup>70</sup>

It had originally been expected that the new kanji list as described above (i.e. the Proposed List of Kanji for General Use) would constitute the Deliberative Council's final recommendations on this topic, but Council Chairman Fukushima Shintarō was in favour of testing public opinion on the new list somewhat further, and so the Proposed List of Kanji for General Use was only presented to Education Minister Naitō Yosaburō as an interim report.<sup>71</sup>

#### 4 List of Kanji for General Use (Jōyō kanji)

After the Proposed List had been made public in March 1979, public opinion on the subject was tested. As in the case of the New Kanji List (Draft), explanatory conferences were held (this time in Sendai, Tokyo, Fukuoka, Gifu, and Okayama), and copies of the Proposed List sent to interested bodies and organisations. As before, there was a measure of concern among some that the concept of the new list as a guide (*meyasu*) was too vague and needed clarification so as not to be interpreted as being tantamount to a complete absence of control on kanji use. With regard to the number of kanji in the Proposed List, there was substantial support for the figure of 1926, but at the same time there were those who chose to differ: the Japan Literati Association, for example, was of the view that the number of kanji in the List was far too small,<sup>72</sup> while the *Kanamoji kai* (lit. "Kana Society")<sup>73</sup> considered that the number should be decreased. Concerning the question of individual kanji to be included in or omitted from the List, there were various views. There was, though, a considerable body of opinion to the effect that those kanji in the TK List which had

been dropped in the New Kanji List (Draft) and not re-adopted in the Proposed List of March 1979 should be reinstated.<sup>74</sup> There was substantial support also for a more detailed section in the Preamble to the proposed kanji list in its final form dealing with the form of kanji. Educational groups, understandably, requested that suitable steps be taken to ensure that an increase in the total number of kanji in the list would not impose any further burden on young pupils.

After ascertaining reaction to the Proposed List, certain modifications were made. Further consideration was given to the term *meyasu* (guide), and it was decided that it would be necessary to emphasize that the use of this word in indicating the status of the new list was not a signal for the dropping of all control on kanji use; accordingly, it was resolved that a note clarifying *meyasu* be added to the Preamble to the kanji list in its final form. With regard to the 19 TK List kanji not included in the Proposed List, it was decided to re-instate all 19, on the grounds that they were familiar through over 30 years of use.

The duly modified new kanji list, the *Jōyō kanjihyō* (List of kanji for general use), was approved at a plenary session of the Deliberative Council held on March 23, 1981, and submitted to Education Minister Tanaka Tatsuo. The List was then approved and promulgated by Cabinet on October 1, 1981.

The List of Kanji for General Use (hereafter normally referred to as “the JK List”), consists of a Preamble, an explanation of the List, the kanji list itself (1945 kanji), and an appendix containing kanji combinations which involved irregular *kun* readings.

The Preamble is comprised of six sections. Section one, *Hajime ni* (Introduction), describes how the Deliberative Council took up the task of revising the TK List. Section two, *Jōyō kanjihyō sakusei no kei* (Circumstances of compilation of the JK List), goes on to explain that while the TK List and related lists had had the beneficial effect of simplifying the writing system, there were certain negative aspects to the reforms also: the restrictive nature of the TK list (and related lists) meant that words involving kanji not included in the List, or *on* and *kun* readings not contained in the TK List of *On* and *Kun* Readings, had either to be written in kana, or a different word used, thus restricting range of expression and leading to unnaturalness of expression: in addition, the TK List kanji included some which were very seldom employed,<sup>75</sup> while on the other hand there were other kanji of high frequency (or potentially high frequency) which were not included. It was to meet the changing needs of the times, therefore, that in 1966 the Education Minister asked the Deliberative Council to consider ways of improving the national language policy (with particular reference to the script). The Deliberative Council had

approached its task by firstly examining all kanji lists which had appeared from the Meiji period onwards, and had then decided to compile a new kanji list which, while it did not involve any abrupt changes compared with the TK List, would facilitate communication and understanding in the general life of society. Section two concludes by re-affirming the effectiveness of using kanji and kana together for writing Japanese, point out the advantages of kanji from the point of view of representing SJ items and expressing the hope that Japanese orthography will become even more clear and attractive in the future.

Sections one and two of the Preamble provide the historical background to the JK List. Section three makes clear the intended application and status of the List. This part of the Preamble, of which a translation follows, is particularly important because of the differences which exist in this respect between the JK List and the TK List.

### **Nature of the JK List**

The JK List contains kanji which are efficient and share a high degree of common usage, for use in the general life of society—laws and ordinances, official texts, newspapers, magazines, broadcasting, etc. The List aims to provide a guide for kanji use for writing in an easily-intelligible and clear manner.

The JK List is to be used in the general life of modern society. It does not attempt to encroach upon kanji use in the various specialist fields in science, technology, and the arts, nor upon the kanji use of individuals; nor does it attempt to negate kanji use in earlier documents. The List does not concern itself either with kanji used in place names, personal names, etc.

Furthermore, “kanji use in the general life of society” refers to usage by persons who have to some extent experienced life in actual society or educational institutions after finishing study in the period of compulsory education.

Since the JK List aims to provide, as mentioned above, a guide to kanji use in the general life of society, it is not a restrictive list which demands that texts be written using only kanji which are given in the List. [Rather], there is scope for suitable reflection in its application, according to circumstances at the time. In cases where kanji seem difficult to read, one method might be to consider using furigana where necessary.

However, in order to facilitate mutual communication and understanding in the general life of society, it is expected that kanji use will conform to this List as far as possible. [Note] The nature of the JK List is not restrictive; rather, it has been made into a guide to kanji use in general life of society. The sense of the word ‘guide’ (*meyasu*) may be supplemented as follows: (1) It is expected that in the general life of society—laws and ordinances, official documents, newspapers, magazines broadcasting etc.—there is no objection to deviation in certain areas from the way kanji are treated in this List, as, for example, in taking decisions on special kanji use on the basis of the List [but] in accordance with actual conditions.

The first paragraph of the above states that the JK List “aims to provide a guide to kanji use” (*kanji shiyō no meyasu to naru koto o mezashita*). For the general public, who had in the postwar period become used to a kanji list the purpose of which was to shew the limits of the different kanji to be used,<sup>76</sup> this was a very pertinent point.

Some degree of difference is also to be seen between the two lists in relation to their range of application. The TK List Preamble states that that list shews the limits of kanji to use “in laws and ordinances, official documents, newspapers, magazines and society in general”. The stated range of the JK List as described in the first paragraph of section three (“the general life of society—laws and ordinances, official documents, newspapers, magazines, broadcasting, etc.”) would appear to be very similar to that of the TK List, but some reduction in the intended range of the new List is clearly indicated in the subsequent paragraph (paragraph two of the above translation).

A further point to be noted here is the suggestion that furigana may be used where necessary, essentially to indicate the readings of kanji not in the JK List. Until shortly after the Pacific War, furigana were used quite extensively, but the orthographic simplification brought about by the TL List reforms meant that from then on this device was far less necessary.<sup>77</sup>

Section three of the Preamble concludes with two notes which attempt to clarify the meaning of *meyasu* (guide), the term used to describe the intended role of the JK List. The notes were added as a result of criticism of the two earlier drafts of the new kanji list to the effect that the word *meyasu* was excessively vague.<sup>78</sup>

Section four, *Jishu to onkun* (The different kanji; *on* and *kun* readings), sets out the main guidelines which were employed in determining the kanji to be included in the List, and their *on* and *kun* readings. The following is the translation.

#### The Different Kanji; *On* and *Kun* Readings.

As mentioned above, the different kanji are shewn in the JK List together with their *on* and *kun* readings also. With regard to *on* and *kun* readings the List of *On* and *Kun* Readings for the Kanji in Current Use (June 1973) has been followed in principle; readings for kanji which have been added [to the JK List] for the first time have been selected in accordance with [the principles for selection of] readings in the June 1973 List.

Selection of kanji and *on* and *kun* readings was determined in a collective manner from the viewpoint of representing words and sentences in writing, on the basis of the actual state of kanji and *on* and *kun* readings used in modern Japanese. The main way of thinking was as follows:

- 1 Kanji of a high degree of frequency or function (especially as elements in compounds) are included. Furthermore, consideration is given also to extent of field of use.

- 2 Kanji which from the viewpoint of expression of concept are hard to understand when written in kana and seem particularly necessary are included, even if they do not possess a very high degree of frequency or function.
- 3 Kanji which are used mainly for proper nouns—place names, personal names, etc.—are not included.
- 4 Kanji for interjections, auxiliary verbs, and particles are not included.
- 5 Kanji for pronouns, adverbs, and conjunctions which are widely-used are included.
- 6 Different kanji which have the same *kun* reading are avoided as far as possible, but kanji [of this type] which can be distinguished in use, or for which there is a strong custom of use, are included.
- 7 In the case of kanji chosen arbitrarily for their sound value (*ateji*) or kanji combinations involving irregular *kun* readings, those which have been used widely and for a long time are included.

Furthermore, taking into account the effect on various areas, all the kanji in the TK List were included.<sup>79</sup>

The above guidelines regarding selection of kanji in the JK List are set out in a somewhat different manner to the corresponding guidelines for the TK List (for instance, frequency of use is specified as a selectional criterion in the case of the JK List, but not in the case of the TK List, even though it was clearly a guiding principle for this latter list also; similarly, kanji used mainly for proper nouns are specified amongst the numbered guidelines above as not being included in the JK List, while in the case of the TK List criteria the same general principle operates, but is given in unnumbered form),<sup>80</sup> and so it is difficult to compare the guidelines for the two lists directly simply by noting differences between the two sets of numbered criteria. One general difference which may be noted is that the selectional criteria for the JK List tend to be couched in more general terms than those for the TK List. This has the advantage that very few—if any—exceptions need to be recognized in relation to the JK List criteria, whereas with the TK List as many as five out of the eight numbered criteria involve exceptions—e.g., “(3) Kanji of a difficult shape are omitted

難形 難 難 are omitted; exceptions: 穢, 廳”.

Allowing for differences of format and arrangement, there are nevertheless certain points of clear divergence between the selectional criteria for kanji in the two Lists. Guidelines (4) of the TK List is as follows: “Kanji possessing a small range of use are omitted— 挨拶 曖昧 (exceptions: 矛盾, 膨脹)”. The JK List makes allowance for certain low-frequency kanji such as those used in writing 矛盾 *mujun* “contradiction” and 膨脹 *bōchō* “expansion” in the form of its second guideline on kanji selection (see above). The eighth guideline for the TK

List states that: “Kanji used mainly only in Government offices are omitted—俸牒傭 (exceptions: 叙轄)”. This principle was set aside in compiling the JK List, which includes kanji with a pronounced bureaucratic flavour such as 屯 *ton* “station, barracks” (as in 駐屯 *chūton* “stationing of troops”) and 拐 *kai* “falsify, kidnap” (as in 誘拐 *yūkai* “kidnapping, abduction”). The JK List guidelines make clear provision for kanji representing widely-used pronouns (e.g., 私 *watakushi* “I” and adverbs (e.g., 殊 *shū* = *koto ni* “particularly”), whereas the TK List guidelines note that words of these two categories should “as far as possible be written in kana”(see above). The treatment of kanji combinations having irregular *kun* readings (e.g., 今日 *kyō* “today”) is a further difference between the two lists (though readings of this type were officially approved prior to 1981 when the revised TK *Onkun* List came into effect in 1973).

The New Kanji List (Draft) of 1977 contained 1900 kanji—a figure which increased to 1926 in the Proposed List of Kanji for General Use. In the JK List itself, the total number of kanji increased further to 1945 due to calls from various groups (especially the Monbushō) for re-instatement of the 19 TK List kanji which had been omitted from the New Kanji List (Draft) and not re-adopted in the interim list of March 1979.

The JK List, which replaced the TK List in October 1981, consists of the 1850 TK List kanji together with the following 95 kanji:

猿凹渦靴稼涯垣袂馮褐缶頑挾襟隅溪蚩嫌洪溝昆崎皿傘肢遮蛇酌汁塾尚宵  
 繩唇甚据杉齊逝仙栓插槽藻馱濯挑眺釣塚漬亨泥棟洞凸把霸漠肌鉢披犀猫頻  
 瓶霧塀泡俸褒朴僕堀磨岬妄厄癒悠羅竟戾 芋謁欣幼堪繭嗣遊嬌遙痘陪畔濫  
 拐喝矯棧壤曹棚偵搭屯抹粹 且但廣脹勺畝匆侯爵帥朕錘銃疥奴翁隸婆嚇

The JK List was compiled to provide a kanji list<sup>81</sup> suited to the needs of writing modern Japanese in the nineteen-eighties and beyond. Just as the TK List had been selective in its choice of kanji from earlier lists in accordance with general changes in society, so too the JK List reflects changes in everyday life which took place in 35 years between compilation of the TK List and JK List. The following are examples of kanji added to the JK List on the basis of such changes: 缶 *kan* “can, tin” (as in 缶詰 *kanzume* “tinned goods”; reflects the rise in popularity of such items since the War); 搭 *tō* “load (a vehicle), ride” (as in 搭乗受付 *tōjō*

uketsuke "(airport) check-in"; reflects the rise of air travel); 凹 *ō* (as in 凹凸レンズ *ōtotsu renzu* "concavo-convex lens"; reflects the greater importance of optical equipment).

Among the new kanji added to the JK List, there are a number which represent everyday words, e.g., 垣 *kaki* "fence, hedge", 塀 *hei* "(outside) wall", 堀 *hori* "ditch, moat", 皿 *sara* "dish". Probably kanji of this type were omitted from the TK List out of the need just after the War to compile a list of well under 2000 kanji.

In considering additions to the new kanji list, the Deliberative Council is reported to have taken into consideration the question of whether a particular kanji was necessary in order to avoid writing certain SJ nouns partly in kanji and partly in kana. An example of this is *sentaku* "washing, laundry": in the TK List era the approved way of writing this word was 洗汰く (the second element *taku* being written in kana because the kanji for *taku* "wash, rinse" is not in the TK List): when the JK List came into effect, 洗濯 *taku* was included and so *sentaku* "washing, laundry" could legitimately be written as 洗濯 in newspapers, magazines, and the like, thus avoiding the aesthetically displeasing visual effect created by kanji/kana SJ hybrid writings (known in Japanese as *mazegaki*, "mixed writing"). The following are other examples of kanji added to the JK List to avoid such hybrid writings (examples of words involving the kanji concerned are added in parentheses): 猿 *en* "monkey" (猿人 *enjin* "pithecanthrope, ape-man"); 拐 *kai* "falsify, kidnap" (誘拐 *yūkai* "kidnapping, abduction"); 喝 *katsu* "scold, become hoarse" (恐喝 *kyōkatsu* "threat, blackmail"); 昆 *kon* "descendants, elder brother" (昆布 *konbu* "kelp"; 昆 is used here for its sound-value only); 涯 *gai* "shore" (生涯 *shōgai* "career"); 霧 *fun* "fog" (霧圍氣 *fun'iki* "atmosphere"); 把 *ha* "bundle, grasp" (把握 *haaku* "grasp").

Suggestions for new kanji for the JK List which were put forward by government departments tended to be regarded favourably by the Deliberative Council. Thus, at the request of the Self Defence Agency, the kanji 屯 *ton* "station, barracks", 曹 *sō* "an official" (as in 陸曹 *rikusō* "sergeant"), and 偵 *tei* "spy" (as in 偵察 *teisatsu* "scouting, reconnaissance") were approved, as were two kanji put forward by the Justice Ministry, viz. 拐 *kai* and 喝 *katsu* (see above for meanings and examples). However, by no means all kanji requested for inclusion by Government departments were approved. The Self Defence Agency had also wanted 哨 *shō* "scout, sentinel" (as in 哨艦 *shōkan* "patrol

ship”) and 曳 *ei* “pull” (as in 曳航 *eikō* “towing”) approved, but neither in fact was;<sup>82</sup> the same fate was also shared by three kanji requested for inclusion by the Construction Ministry: 堰 *seki* “dam, embankment”, 峻 *shun* “high, steep” (as in 峻路 *shunro* “steep road”), 汎 *han* “spread out, wide” (as in 汎濫 *hanran* “flooding”).<sup>83</sup>

The influence of government departments on the new kanji list was not restricted to kanji being added for the first time. 遵 *jun* “follow, obey” (as in 遵守 *junshu* “obey, observe”) and 濫 *ran* “overflow, spread over” (as in 濫用 *ran’yō* “misuse, misappropriation”) were omitted from the New Kanji List (Draft), but subsequently re-instated at the request of the Justice Ministry.<sup>84</sup> Similarly, 遞 *tei* “in turn” (as in 通信 *teishin* “communications”) was re-adopted at the wish of the Ministry of Postal Services. Finally, all 19 of the TK List kanji omitted from the proposed List of Kanji for General Use were re-adopted—a move which was due largely to a request to that effect from the Monbusho.<sup>85</sup>

The fact that all TK List kanji were in the end retained in the JK List is somewhat surprising when one bears in mind the following statements in the two earlier stages of the JK List:

In the TK List there are some signs which are at present hardly used at all, while on the other hand there are some signs which are frequently used in society but not included in the List.—New Kanji List (Draft), Preamble. <sup>86</sup>

Kanji such as 朕 *chin* “(Imperial) we” and 虞 *osore* “fear, anxiety”—very rarely used in everyday texts (but included in the TK List because they are used in the text of the postwar Constitution)—are among the 19 TK List kanji re-adopted in the JK List. In view of their very low frequency of occurrence, one might expect the passage corresponding to the above (translated) passages in the kanji lists of 1977 and 1979 to be omitted from the JK List itself, reflecting the more accommodating stance of the Deliberative Council in relation to TK List kanji of very low frequency at the time of compiling the JK List. Such is not the case, however: the above passage in the 1979 Proposed List (“It has, moreover, also been pointed out....”) is included verbatim in the JK List Preamble.<sup>88</sup> The final incorporation of all 1850 TK List kanji into the JK List is referred to in a vague manner at the end of section four (The different kanji: *on* and *kun* readings) of the JK List Preamble, where it is stated: “Furthermore, taking into account the effect on various areas, all the kanji in the TK List were included.”<sup>89</sup>

As described above, from 1954 onwards there was minor divergence between kanji usage in official documents, school textbooks and so on on the one hand, and that in newspapers on the other, due to the adoption in

newspapers of 28 kanji not in the TK List to replace 28 other kanji already in the List. When the JK List report was submitted to the Education Minister in March 1981, the press agreed that it would follow the JK List when it came into effect, but with minor modifications. These modifications consisted of not using the 11 following uncommon kanji 謁 *etsu* “audience (with a ruler)” (as in 謁見 *ekken* “audience”), 虞 *osore* “fear”, 箇 *ka* (counter), 且 *katsu* “furthermore”, 遵 *jun* “follow, obey”, 但 *tadashi* “however”, 脹 *chō* “swell, expand” (as in 膨脹 *bōchō* “swelling, growth”), 朕 *chin* “(Imperial) we”, 附 *fu* “attach, append” (as in 附錄 *furoku* “supplement”; this and other compounds such as 附着 *fuchaku* “adhesion” are now commonly written with the simpler kanji 付), 又 *mata* “again, furthermore”, 溢 *ran* “overflow, spread over”, and using six kanji not in the JK List, viz.: 龜 *ki* “turtle, tortoise” (as in 龜裂 *kiretsu* “cleft, fissure”), 舷 *gen* “gunwale” (as in 右舷 *ugen* “starboard”), 痕 *kon* “mark, footprint” (as in 血痕 *kekkon* “bloodstain”), 挫 *za* “break, sprain” (as in 挫折 *zasetsu* “setback, frustration”), 哨 *shō* “scout, sentinel”, 狙 *so/nerau* “to aim”. It was also resolved that the press would adopt the additional *on* reading of *ka* for 個 (counter), a kanji for which only the reading *ko* is listed in the JK List and the earlier TK List equivalent. These points of deviation from the JK List in kanji use in newspapers represent a reduction in the degree of divergence from the norm compared with the period when the TK List was in effect.

Section five of the JK List Preamble—entitled *Jitai* (The form of kanji)—explains that kanji in the List are given in a variety of the widely-used Ming type; that the forms are based on the TK Forms List (promulgated in 1949), with the form of new kanji being determined as appropriate in accordance with those in the 1949 List (e.g., 螢 *kei* “firefly—as in 螢光灯 *keikōtō* “fluorescent light”—in the form 螢, on the basis of the upper element in kanji such as 勞 *rō* “labour, toil” and 榮 *ei* “prosperity, glory”); it is also mentioned that the JK List is not concerned with the minor variations in type design that are to be found in different varieties of Ming type, nor with the minor differences which are sometimes to be found between kanji in Ming type and their handwritten equivalents.<sup>90</sup> These last two points are dealt with in an additional section in the Preamble entitled *Jitai ni tsuite no kaisetsu* (Explanation concerning the form of kanji).<sup>91</sup> The Proposed List of 1979 did already include a small amount of guidance on these two points, but this was elaborated on further here in the JK List in response to requests to this effect.<sup>92</sup>

Section six of the Preamble deals with three miscellaneous topics. The first topic is that of how many and which kanji should be taught at school. During the TK List era there was a set list of kanji which were to be taught during the six years of primary education (at first, the 881 kanji of the Separate TK List, but in 1968 a further 115 kanji were added to these), but in the JK List the question of precisely which kanji should be taught at primary school, and during the period of compulsory education as a whole, is left for later deliberation, as the following passage shews:

The JK List, as explained in [the section on] its nature, was compiled as a guide for kanji use in the general life of society, but in school education it is desirable that the teaching of kanji in an appropriate manner, taking into consideration the gist and contents of the JK List.

Furthermore, with regard to guidance in kanji during the period of compulsory education, it is not necessary to teach all the kanji in the JK List at that time; treatment of the List is left for separate appropriate measures in education which take into account circumstances in the teaching of kanji in the past, and which also pay due regard to such matters as the stage of development of child pupils.

The question of how to relate the JK List to the school curriculum was a difficult problem for the Deliberative Council, especially as those involved in the education process tended to be of the view that teaching the 1850 kanji in the TK List was already a very demanding task, given the limited time allocated to that topic in the school timetable.<sup>93</sup> In view of this situation, it is in a sense not surprising that the Deliberative Council should have decided to leave this matter for later consideration by another body.

The next topic to be treated in Section six is the question of special kanji for use exclusively in writing given names. 92 kanji of this type had been approved in 1951 to be used in conjunction with the TK List, and in 1976 a further 28 kanji were approved for this purpose. The first group of 92 kanji was drawn up by the Deliberative Council, and a recommendation made to both the Education Minister and the Attorney General. The second group of 28 kanji, however, was selected by the Civil Affairs Administrative Council (CAAC) (a consultative body which reports to the Justice Minister); this change in procedure was due to the ramifications which changes in name kanji had in relation to the Family Registration Law.

From 1976, then, there was a total of 120 kanji which could be used together with the 1850 kanji in the TK List for choosing how to write the given names of children. Since these 120 name kanji had been selected in relation to the TK List, however, it became necessary to reconsider the question of kanji of this type when it was clear that the familiar TK List would before long be replaced by a new list. Accordingly, early in 1979

the CAAC again began investigations into this topic; its report was submitted to Justice Minister Okuno Masaaki in May 1981. The main recommendations of the CAAC report may be summarised as follows: (1) restrictions on the use of kanji for given names should continue to operate; (2) all 1945 kanji in the JK List should be permitted for use in given names; (3) the number of kanji for use exclusively in writing given names should be increased from 120 to 166;<sup>94</sup> (4) variant forms (of both JK List kanji and name kanji) to be permitted in writing given names should be clearly specified.<sup>95</sup>

The CAAC's recommendations in relation to name kanji were approved by the Justice Ministry, and put into effect from October 1, 1981—the date from which the JK List itself came into effect—by amending the operational regulations for the Family Registration Law. The final topic in Section six of the Preamble relates to the use of the JK List by such bodies as the Cabinet's Bureau of Legislation, the Monbushō, the Justice Ministry, newspaper companies, and the Japan Broadcasting Corporation. Entitled *Kakushu no kijun nado* (Standards, etc.), this final passage expresses the hope that the JK List will be followed by the Bureau of Legislation, the Monbushō, and so on in place of the TK List, making modifications to the kanji in the new List as appropriate to each area.<sup>96</sup>

Following on from the Preamble, there is an explanation of the JK List (*Jōyō kanjihyō no mikata*) which deals with the format of the List, but this need not concern us here, except to note two points. Firstly, the form of certain individual kanji as they appear in the famous *K'ang hsi tzu tien* 康熙字典 dictionary (completed 1710) is noted in parentheses where this differs substantially from the main form in which a particular kanji appears in the JK List, e.g., 国(國) 桜(櫻). Secondly, certain irregular readings for kanji are given (and hence approved) in the "Remarks" (*Bikō*) column beside kanji in the List; e.g., 衣 *i/koromo* "garment, clothes", which combines with 浴 *yoku* "bathe" to make up 浴衣 *yukata* "unlined cotton kimono".<sup>97</sup>

The final part of the JK List consists of an appendix (*Fuhyō*) which sets out the way of writing in kanji 110 miscellaneous words which involve irregular readings, e.g. *kotoshi* "this year" (alternatively, depending on the point of view, this appendix may be regarded as a list of irregular readings for such kanji combinations as 今年 ).

The JK List was the subject of much debate after the JK List report was made public in March 1981. Although there was relatively widespread support for the List, there was opposition expressed in some quarters. The Japan Teachers' Union, for example, was against the new List replacing the TK List, largely because of what was perceived as a potential increased

burden on school-pupils, but also because it would mean the abolition of the Separate TK List, which provided a firm framework specifying the kanji to be taught at different stages of the period of compulsory education.<sup>98</sup> The Japan Teachers' Union was also opposed to the concept of the JK List as a guide (*meyasu*), a point concerning which the major newspapers also expressed reservations.<sup>99</sup> On the other hand, the status of the JK List was welcome by many as being less restrictive than the TK List. For the Japan Literati Association, however, the List was still too restrictive, and the increase in number of kanji compared with the TK List was too small.<sup>100</sup>

## 5. In Conclusion

In the early decades of the 20th century various attempts were made at script reform, but these tended to be frustrated either by traditionalism or by events of the time. After the Second World War the time was ripe for change, and orthographic reforms in the direction of simplification of the script were seen as providing the basis for all other reforms. The major orthographic reforms of 1946, though partly resulting from external pressure, would hardly have been possible without the insight and experience gained from the protracted exchanges of views and deliberations over proposals for orthographic change during the period 1900-1945. The TK List itself was, though, compiled and put into effect in a very short space of time after the War. In the case of the JK List, circumstances permitted a less hurried approach, and deliberations over the new list extended over a period of some eight years, taking into account the experience gained through operation and use of the TK List.

The JK List represents a move towards kanji usage which is slightly more complicated than that based on the TK List, and in this sense may be said to go against the uneven but nevertheless overall trend during the period 1900-80 towards simplification of the writing system. In terms of its status and the total number of kanji, the JK List may be said to be the product of compromise—compromise both between the views of the Japanese people in general in relation to kanji list, and between different groups within the Deliberative Council which compiled the List. The guideline status of the JK List has both its good and bad points, but it could be that this will be the answer to the need to provide some guidance on kanji use in writing ordinary modern Japanese texts on the one hand, yet without imposing excessive restraints on the other. Whether this device proves to in fact be a workable solution, however, is something which only time will tell.

1. For an account of reform (both proposed and actual) of Japanese script and written styles in the latter part of the 19th century, see Twine 1983.
2. Yano Fumio, author of *Nihon buntai moji shinron* (New Discourse on Japanese script and style, 1886), for instance, was of this view.
3. The Kokugo chōsa iinkai (Japanese Language Investigative Committee; formed 1902), for example, was in favour of a phonogram script (referred to in the Committee's investigative policy as *on'in moji* [phonemic script], with the term *fonoguramu* added in parentheses; see Maruya 1983:54), but saw reduction in the number of kanji in use as a necessary interim measure.
4. That is to say, the *hiragana* and *katakana* syllabaries each consisting of 48 signs, as to be found in Nelson 1966, Appendix 7. The principal effect of this particular measure was to standardize *hiragana* signs, of which many variant forms (*hentaigana*) were still widely used. For the actual text of this regulation, see Yoshida & Inokuchi 1962:951-952.
5. For the list of kanji, see Yoshida and Inokuchi 1962:110-112, or Inokuchi 1982:393-395.
6. The traditional system of SJ kana usage was determined largely by Motoori Norinaga (1730-1801) in his *Jion kanazukai* (Sino-Japanese kana usage; 1776), which was based on the Chinese work entitled *Yün ching*—a set of sound tables printed in 1161. Concerning the SJ kana usage reform of 1900, see Shinkōsha 1973:150-153; for the actual text of the Monbushō regulation, see Yoshida and Inokuchi 1962:479-483.
7. It was, however, still permissible to follow the traditional SJ kana usage in relation to these three groups of kana; see Shinkōsha 1973: 153.
8. Viz. あう (e.g., in the case of 嬰鳥, etc.); あふ (凹, etc.); あう (應, etc.); あふ (押, etc.); わう (王, etc.); むう (翁, etc.).
9. This body was set up as a national body to investigate and report on problems and aspects of the Japanese language. For a table of the various Japanese language committees and their reports as mentioned in the main text of this article see Appendix 1.
10. Maruya 1983:56-57.
11. The major task of bringing a degree of standardization to the form of kanji had been begun about ten years earlier by the Investigative Committee. Part of the work of the Committee can be seen in its *Kanji yōran* (Outline of kanji, 1908).
12. Kana used in this way are called *furigana* in Japanese. As the system of compulsory education began to take effect, the consequent rise in the general level of reading led to the blanket use of *furigana* changing to a selective use (Shinkōsha 1974:9). Selective use of this kind continued widely until the orthographic reforms of 1946 (to be described below).
13. Maruya 1983:44.
14. Upon the formation of the Hara Takashi Cabinet, a colloquial written style was adopted in Monbushō directives issued by the then Education Minister, and also in public notices issued by the railways and post offices; see Ōno & Shibata 1977:279.
15. For the text of the statement, see Shinkōsha 1974:11-12.
16. There had been some criticism of, and dissatisfaction towards, the Investigative Committee on account of its having sometimes conducted research into arcane topics rather than into more immediate problems; see Ōno & Shibata 1977:280.

17 Ōno & Shibata 1977:280.

18. That this was the field of application of the 1923 *Jōyō kanjihyō* is clear from the explanation of this list by Hoshina Kōichi which appeared in the Official Gazette (*Kanpō*) of May 12, 1923 (see Inokuchi 1982:39-40).

19. The joint announcement which appeared in newspapers on August 5, 1923 was entitled simply *Sengen* (Declaration); for the text, see Shinkōsha 1974:13.

20. This statement—*Kanji seigen ni kansuru sengen* (Declaration regarding restrictions on kanji)—appeared in newspapers of June 1, 1925; see Shinkōsha 1974:14-15 for the text.

21. In the event, restrictions on kanji use in advertisements were never put into effect (Shinkōsha 1974:15).

22. The revisions of 1931 involved the deletion of 147 kanji originally in the list, and the addition of 45 others which were necessary for writing Imperial rescripts. For the text of these revisions (including lists of the kanji which were deleted and added at that time), see Inokuchi 1982:41-42.

23. For the main points of the 1924 proposals, see Shinkōsha 1973:173-174. For the proposed changes of 1931, see *ibid.*, 177-178.

24. For a list of articles dealing with this issue in newspapers and journals of the time, see Shinkōsha 1973:179-180.

25. Concerning the various connotations of this term, see Miller 1982:92-94.

26. Ōno & Shibata 1977:286-287.

27. 關辭變繼 are forms of 關辭變繼 to be found in the *K'ang hsi tzu tien*, a dictionary which became a standard reference work on kanji in Japan as well as China from soon after it was compiled in 1710.

28. Inokuchi 1982:46.

29. *Ibid.*, 54.

30. *Ibid.*

31. For the text of Minami's report, see Inokuchi 1982:63-65.

32. This list did not in any case come to be widely used (Inokuchi 1982:120).

33. For the main points of the New Kana Usage List of 1942, see Shinkōsha 1973:188-189.

34. For a list of the main articles which appeared concerning the New Kana Usage List, see Shinkōsha 1973:190-191.

35. Maruya 1983:61.

36. In November 1945 the *Yomiuri shinbun* went so far as to carry an editorial headed *Kanji o haishi seyo* (Abolish kanji).

37. Ōno & Shibata 1977:288.

38. Maruya 1983:115.

39. The proposed list consisted of 1046 of the 1134 *jōyō kanji* in the List of Standard Kanji, together with 249 of the lower frequency *jun jōyō kanji* in that List; kanji in the latter category were selected on the basis that they would be widely used in the future (Inokuchi 1982:73).

40. The kana usage committee was headed by the language scholar Andō Masatsugu (1875-1952).

41. Inokuchi 1982:75. Nine of the original 1295 kanji were deleted (也 —used for the literary copula *nari*, 俸 *hō* "salary", 慾 *yoku* "desire", 棉 *men* "cotton", 硯 *suzuri* "inkstone", 聯 *ren* "join", 覺 *kō* "expire", 輯 *shū* "gather", collect, 輿 *kago* "palanquin"), and a further 564 kanji added.

42. The modern kana usage rules had also been approved by the Deliberative Council earlier in the same year (September 21.). For the text of these rules, see Yoshida & Inokuchi 1962:667-684.

43. The scope of the TK List and the modern kana usage rules were somewhat different from that of the material on which they were based. The 1134 *jōyō kanji* had been selected as high-frequency kanji in the context of a list containing a total of approximately 2500 signs, while the New Kana Usage List for Sino-Japanese was intended, as its title suggests, to have application to SJ only.

44. The kanji in the Separate TK List are (or were) popularly known as the *kyōiku kanji* (kanji for [use in compulsory] education). To many, *Tōyō kanji beppyō* (Separate TK List, might seem to be an odd title, given the role of the List. Some light is shed on this point by Miyake Takerō, who, in writing on this topic, reports that at the time that the TK List itself was being compiled, it was envisaged that the number of kanji in general use would gradually be further restricted to about 1000 signs. The Separate TK List was, in other words, apparently compiled partly with the aim of providing for the future a list of "core" kanji for general use. The Separate TK List in its final form consisted of 881 signs (a further 115 kanji being added in 1968, also to be taught in the six years of primary education), but an early draft of this list contained 1060 kanji. See Kokugo gakkai 1955, *Kyōiku kanji* entry.

45. Inokuchi 1982:95.

46. *Ibid.*, 95-96.

47. On December 29, 1947, it was specified in the Rules for Operation of the Family Registration Law that only kanji in the TK List, or *katakana* or *hiragana*, be permitted for writing the given names of children born after January 1, 1948. See Inokuchi 1982:223.

48. *Ibid.*, 224-226.

49. For the list of 92 name kanji, see *ibid.*, 227. At the same time that these name kanji were introduced (May 25, 1951), the Rules for Operation of the Family Registration Law were amended so as to permit the use of these additional kanji in writing the given names of newborn children.

50. The recommendations of the Deliberative Council's Kanji Section were known as the *Tōyō kanjihyō shingi shiryō* (Material for deliberation concerning the list of kanji for current use) or *Tōyō kanji hosei shiryō* (Supplementary material on the kanji for current use). The recommendations made by the Kanji Section also proposed that the reading *ka* be recognized for 個 in addition to *ko*, and the reading *hi* for 火丁 besides *tō*; it was also proposed that the simplified form 火丁 be adopted in place of 火登. See Inokuchi 1982:112-113.

51. Nihongo kyōiku gakkai 1982:502.

52. For the text of the *okurigana* rules, see Yoshida & Inokuchi 1962:903-922.

53. For the text of this report, see Inokuchi 1982:120-127.

54. *Ibid.*, 129-130.

55. As Yasunaga Minoru (1981:24-25) notes, it would in a sense have been preferable to have begun by establishing the different kanji to be included in the list, and then proceeding to the question of *on* and *kun* readings and *okurigana*, but apparently at that time there was insufficient material on which to base such a list.

56. Many of the additional readings in the new *Onkun* List result from a different format which is more explicit than the 1948 List. According to Kokugo gakkai 1981 (*Tōyō kanji* entry) if the 1973 List were to be rewritten according to the

format of the 1948 *Onkun* List, then the total number of the two types of readings in that case would be 3485 instead of 3938.

57. Yasunaga 1981:26.

58. *Ibid.*

59. *Ibid.*, 27.

60. Inokuchi 1982:145.

61. Each Session of the Council was of two years' duration, at the end of which there was a change of membership.

62. Yasunaga 1981:28.

63. Inokuchi 1982:148.

64. *Ibid.*, 151-152.

65. Yasunaga 1981:29.

66. For example, in the case of 煉 *ren* "refine (metal)" (a kanji outside of the list), should this be "standardized" in form to 煉 through analogy with kanji in the list such as 練, or left in the form 煉?

67. The type most commonly used in Japan.

68. Yasunaga 1981:29.

69. The general rationale underlying inclusion of these kanji is explained below with reference to the JK List (1981).

70. It was decided by the Deliberative Council that the Justice Ministry should report on this topic, in view of the ramifications any changes to name kanji had on the Rules for Operation of the Family Registration Law (Inokuchi 1982:162).

71. Yasunaga 1981:30.

72. Opposition to the Proposed List by the Literati Association was expressed in its *Jōyō kanjihyōan ni tsuite no ikensho* (An opinion concerning the proposed list of kanji for general use).

73. An organisation dedicated to popularising the use of *katakana* for writing Japanese (*Kokuritsu kokugo kenkyūjo* 1983:264).

74. There were 19 such kanji, viz. 且但虞脹勺畝匆侯爵帥朕錘銃薪奴翁隸婆嚇

75. 朕 *chin* "(imperial) we", 虞 *osore* "fear", and 又 *mata*—included in

the TK List because of their use in the text of the postwar Constitution - are examples of this type of kanji.

76. The TK List Preamble begins with the following statement: "This list shows the limits of the kanji to be used in the laws and ordinances, official documents, newspapers, magazines, and society in general." For the full text, see Inokuchi 1982:78.

77. The TK List Preamble states that *furigana* "are as a rule not to be used" (*furigana wa gensoku to shite tsukawanaï*).

78. Much of the two years' work by the Deliberative Council between March 1979 and March 1981 centred upon further discussion of the term *meyasu*. See *Asahi shinbun*, March 24th 1981: *Jōyō kanjihyō - umi no kurushimi hachinen* (The JK List — eight years of struggle).

79. Inokuchi 1982:173.

80. The selectional criteria for kanji in the TK List were as follows:

1 As far as possible, pronouns, adverbs, conjunctions, interjections, auxiliary verbs, and particles are to be written in kana.

2 Names of flora and fauna are also to be written in kana. By analogy with this, the names of utensils are also to be written in kana.

- 3 Complicated kanji are to be omitted.
- 4 Kanji with a small range of use are to be omitted.
- 5 Kanji having *kun* readings only, or mainly *kun* readings, are to be omitted.
- 6 In the case of different kanji having the same *on* reading and a similar meaning, one of the two is omitted.
- 7 Low-frequency kanji used for writing SJ words are to be omitted where there is an alternative word which involves TK List kanji.
- 8 Kanji used mainly in government offices are to be omitted.

For the original text of the above, which includes examples in each case, see Inokuchi 1982:79-80.

81. For the *on* and *kun* readings of these kanji, and examples of usage as set out in the JK List, see Kokuritsu kokugo kenjyūjo 1981, *Shiryō: Jōyō kanjihyō*. For the English equivalents, etc., for the JK List kanji, see W. Hadamitzky and M. Spahn, *Kanji and Kana* (Charles E. Tuttle Co., Rutland & Tokyo, 1981).

82. See *Yomiuri shinbun*, March 24, 1981: *Shinkanji jijō 1* (Circumstances relating to the new kanji: 1).

83. Inokuchi 1981:155.

84. *Ibid.*

85. See *Yomiuri shinbun*, March 31, 1981: *Shinkanji jijō 8* (Circumstances relating to the new kanji: 8).

86. Inokuchi 1982:147.

87. *Ibid.*, 157.

88. *Ibid.*, 168.

89. *Ibid.*, 173.

90. The following are examples of minor variations in different varieties of Ming type for the one kanji: 印 印, 奔 奔, 空 空. The following are instances of minor variation between (a) kanji in Ming type and (b) their handwritten equivalent: (a) 北 (b) 北, (a) 史 (b) 史.

91 *Nihongō kyōiku gakkai* 1982:512-513 for the full text of this additional section.

92. Minor differences between kanji in Ming type (the form in which they appear in the TK Forms List) and their handwritten equivalents were the cause of some uncertainty and confusion in schools. See *Asahi shinbun*, March 24, 1981: *Tegaki moji sawagi heru?* (Less commotion over handwritten kanji?).

93. See *Asahi shinbun*, March 24, 1981: *Shinkanjihyō—kyōiku genba no taio wa* (The new kanji list and how to deal with it in the teaching situation).

94. See *Asahi shinbun*, March 24, 1981: *Shinkanjihyō—kyōiku genba no taio wa* (The new kanji list and how to deal with it in the JK List.)

95. At the time that the CAAC made its report, this aspect of name kanji was unclear, some variant forms being permitted for use in entering given names of newborn children in the official family registers, while others were not; see Inokuchi 1982:236-237. The rules for Operation of the Family Registration Law as amended on October 1, 1981, as a result of the CAAC report clarified this situation, since they set out, in the form of an appended list, those variant forms for JK List kanji and name kanji which are permitted for writing given names in the official registers. 205 variant forms are recognised for this purpose—195 being for JK List kanji, and the remaining 10 for name kanji. See Inokuchi 1982:245 for the complete list.

96. Inokuchi 1982:179-180.

97. Some of these same irregular readings (such as 浴衣 *yukata* itself) are also given in the appendix (*Fuhyō*) which follows the actual list of 1945 kanji.

98. See the long statement that was made by the Japan Teacher's Association on the day that the Deliberative Council's JK List report was made public (Inokuchi 1982:217-220).

99. The *Mainichi shinbun*, for example, commented that "it was impossible not to feel considerable concern regarding the fact that the JK List had slipped out of a framework of control through the vague expression *meyasu*"; for this and other editorials on the JK List by the major newspapers, see Inokuchi 1982:217-220.

100. For the Japan Literati Association's *Jōyō kanjihyō ni tsuite no seimei* (Statement on the JK List), see Inokuchi 1982:199-202.

## References

- Inokuchi Yūichi. 1982. *Meiji igo no kanji seisaku* (Kanji policies since the Meiji Period). Nihon gakujutsu shinkōkai.
- Kokugo gakkai. 1955. *kokugogaku jiten* (Dictionary of Japanese language studies). Tōkyōdō.
- Kokugo gakkai. 1981. *Kokugogaku daijiten* (Comprehensive dictionary of Japanese language studies). Tōkyōdō.
- Kokuritsu kokugo kenkyūjo. 1981. *kokugo nenkan* (Japanese language yearbook).
- Kokuritsu kokugo kenkyūjo. 1983. *Kokugo nenkan* (Japanese language yearbook). Shūei shuppan.
- Maruya Saiichi. 1983. *Nihongo no sekai* (The world of the Japanese language), v.16: *Kokugo kaikaku o hihan suru* (Criticism of national language reforms). Chūōkōronsha.
- Miller R.A. 1982. *Japan's modern myth: the language and beyond*. New York & Tokyo: Weatherhill.
- Nelson A.N. 1966. *The modern reader's Japanese-English character dictionary*, 2nd ed. (revised). Rutland, Vermont, & Tokyo: Weatherhill.
- Nihongo kyōiku gakkai. 1982. *Nihongo kyōiku jiten* (Dictionary of Japanese language education). Takishūkan shoten.
- Ono Susumu & Shibata Takeshi. 1977. *Iwanami kōza Nihongo* (Iwanami lecture series on the Japanese language) v. 3: *Kokugo kokuji mondai* (Problems of the national language and script). Iwanami shoten.
- Shinkōsha (ed.) 1983. *Fukkoku bunkachō kokugo shirizu* (Cultural Agency Japanese language series, reprint), vl: *Kokugo mondai* (Problems of the national language). Kyōiku shuppan.
- Twine, Nanette. 1983. "Toward simplicity: script reform movements in the Meiji Period". *Monumenta Nipponica*, 38:2. 115-132.
- Yasunaga Minoru. 1981. "Jōyō kanjihyō ga umareru made" (Background to the JK list). *Gengo seikatsu*, 355:24-31.
- Yoshida Sumio & Inokuchi Yūichi (ed.) 1962. *Meiji ikō kokuji mondai shōan shūsei* (Collected proposals concerning problems of the Japanese script since the meiji Period). Kazama shōbō.

**Appendix I. Chronological Table of Committees on the Japanese Language, Reports, Regulations, etc. (1900-1981)**

Reports, etc.	Regulations, Official Lists of Kanji, etc.
1900 <i>Japanese Language Investigative Committee set up (1902)</i>	Monbusho regulations on standard kana signs, SJ kana usage, and restriction of kanji in primary texts (1900) Above regulations rescinded (1908)
<i>Investigative Committee disbanded (1913)</i>	
"Proposed Modifications to Kanji" (1919)	
1920 <i>Interim Committee on the Japanese Language formed (1921)</i>	
"List of Kanji for General Use" (1923); pronunciation-based kana proposed for both SJ and NJ (1924) Modified version of above proposal (1931)	
<i>Interim Committee disbanded; Deliberative Council on the National Language formed (1934)</i>	
"Proposed Modifications to the Form of Kanji (1938)	
1940 "List of Standard Kanji" (1942; 2669 kanji); "New Kana Usage List for Sino-Japanese" (1942)	For weapon names and related texts, Army Ministry restricts number of kanji to 1235 (1940), and adopts pronunciation-based kana usage (1941)
"TK List" (1946; 1850 kanji); "Modern Kana Usage" (1946)	TK List, Modern Kana Usage (1946)
"Separate TK List" (1947); "TK <i>Onkun</i> List" (1947)	Separate TK List, TK <i>Onkun</i> List (1948)
"TK Forms List" (1948)	TK Forms List (1949)
1950 "Proposal & Declaration Concerning <i>Kanji</i> for Given Names" (1951: 92 kanji)	Separate TK List for Name Kanji (1951)
"Rules for <i>Okurigana</i> " (1958)	Rules for <i>Okurigana</i> (1959)
1960 "Reform of the National Language" (1963)	

- |      |  |  |
|------|--|--|
| 1970 | <p>“Rules for <i>Okurigana</i>”<br/>(revised) (1972)</p> <p>“Concerning the Addition of<br/>Name Kanji” (CAAC, 1975)</p> <p>“New Kanji List (Draft)”<br/>(1977; 1900 kanji)</p> <p>“Proposed JK List” (1979;<br/>1926 kanji)</p> | <p>Rules for <i>Okurigana</i> (revised) (1973)</p> <p>List of Additional Name Kanji<br/>(1976: 28 kanji)</p> |
| 1980 | <p>“JK List” (1981; 1945 kanji);<br/>CAAC Report on Name Kanji<br/>(1981)</p>  | <p>JK List (1981)</p> <p>Amendment to Family Registration Law<br/>(approving 166 name kanji)</p>             |



## The Authors

A.E. Backhouse is senior lecturer in the Department of Japanese at Monash University (Clayton, Victoria 3168, Australia) in Melbourne. His main field of research is Japanese lexicology, and he has also published articles on Japanese grammar and other areas of Japanese linguistics. His current work is in lexical semantics and includes a semantic study of personality terms in Japanese.

Nanette Twine is a senior teaching fellow in the School of Modern Asian Studies, Griffith University (Nathan, Queensland 4111, Australia) where she teaches Japanese language and literature. Her major research interests are the Genbunitchi movement (which resulted in the development of a modern colloquial style in written Japanese); Meiji period script reform movements; and Meiji period literature. She has translated Tsubouchi Shoyo's *Shosetsu Shinzui* into English as *The Essence of the Novel*.

J. Marshall Unger is associate professor of Japanese at the University of Hawaii (1890 East-West Road, Honolulu, HA 96822). Most of his publications deal with the early history of the Japanese language, but over the past four years he has been working on computer-based education in Japanese, Chinese, and Korean. His most recent production, however, is a piano performance: *Sonata for Viola and Piano* by the modern American composer Nancy Van de Vate on Orion Records.

C. Seeley is senior lecturer in Japanese in the Department of Asian Languages, University of Canterbury (Christchurch 1, New Zealand). His research field is the history of the Japanese writing system from the earliest times down to the present, and his publications include articles on forms of writing in early Japan, and the historical kana usage of Keichu (1640-1701).

## Special issues of **VISIBLE LANGUAGE** still in print

### **XVII 3 Lettrisme: Into the Present**

Guest editor: Stephen C. Foster

### **XVII 2 The Renaissance of Die Hermeneute**

Guest editor: Charles Robert Kline, Jr.

### **XVII 1 Calligraphers: What Gives Me the Most Trouble**

Guest editor: Gunnlaugur SE Briem

### **XVI 1 The Symbolism of Mathematics**

Guest editor: Richard R. Skemp

### **XV 4 Aspects of Cuneiform Writing**

Guest editor: Marvin Powell.

### **XV 2 & 3 Visual Cues in Word Recognition and Reading**

Guest editor: Keith Rayner.

### **XV 1 Spatial Arrangement of Text**

Guest editor: James Hartley.

### **XIV 4 Dynamics of Writing**

Guest editor: Peter Wason.

### **XIV 3 Freud and Visible Language**

Guest editor: Mary Lydon.

### **XIII 4 Graphic Design Education**

Guest editor: Sharon Poggenpohl.

### **XIII 3 Behavioural Studies of the Handwriting Skill**

Guest editor: Alan W. Wing.

### **XII 1 The Interface of Reading and Listening**

Guest editor: Dominic Massaro.

### **XI 4 Beyond the Empire of Signs**

Guest editor: Steve Ungar.

### **XI 2 At the Edge of Meaning**

Guest editor: Aaron Marcus.

### **X 4 Spelling Research**

Guest editor: Richard Venezky.

**Information on Ordering copies** is given with the general information on *Visible Language* elsewhere in this issue. A list of the complete contents of these special issues (and all other issues) is available on request.

**If you are interested in serving as guest editor** for a special issue devoted to your specific research interest, write the editor outlining the general idea you have in mind and listing a half dozen or so topics and possible authors. If you would rather discuss the idea first, call the editor: Dr. Merald E. Wrolstad at 216/421-7340.