

Electric Kanji: an occasional series on the electronic management of Japanese text

Mark Ravina
Emory University

KEN LUNDE, a leading figure in the development of Japanese-capable software, is the author of a forthcoming book on the electronic handling of Japanese text. Ken is the Project Manager for Japanese Font Production at Adobe Systems, Inc., a leading manufacturer of computer-based font technologies. The following is excerpted from a press release:

Subject: Announcing *Understanding Japanese Information Processing* by Ken Lunde

Published by O'Reilly and Associates, Inc.
ISBN 1-56592-043-0

Publication date: September 1993
Pages: 435
Price: \$29.95

About this Book:

There are many complex issues surrounding the use of the Japanese language in computing. Unlike English, which has 26 letters in a single alphabet, Japanese has thousands of characters in threescripts. The issues around handling such an unwieldy collection of data are formidable and complex. Up to now, researching and understanding the relevant issues has been a difficult, if not unattainable task, especially to a person who doesn't read or speak Japanese.

Understanding Japanese Information Processing is a book that provides detailed information on all aspects of handling Japanese text on computer systems. It tries to bring all of the relevant information together in a single book. It covers everything from the origins of modern-day Japanese to the latest information on specific emerging computer encoding standards. Here is a sampling of the topics covered:

- The Japanese writing system
- Japanese character set standards
- Japanese encoding methods
- Japanese input
- Japanese output
- Japanese code conversion techniques
- Japanese code and text processing tools
- Japanese e-mail

In addition, there are over 15 appendices which provide additional reference material, such as a code conversion table, character set tables, mapping tables,

an extensive list of software sources, a glossary, and much more.

Both O'Reilly & Associates and the author are excited about the publication of this book because it fills a large void for information about the Japanese language for English-language readers. This book is able to both describe the Japanese language generally and cover very detailed specifics about implementing the language on computers. Even a person with only a casual interest in languages will find the book engaging. A programmer interested in writing a computer program which will handle the Japanese language will find the book indispensable.

For more information on this book, contact:

O'Reilly & Associates, Inc.
103 Morris Street, Suite A
Sebastopol, CA 95472 USA
(800) 998-9938
(707) 829-0515
Fax: (707) 829-0104
UUCP: uunet!ora!nuts
Internet: nuts@ora.com

Ken's book is an expansion and update of his online document JAPAN.INF, which is an invaluable resource for those working with Japanese text. JAPAN.INF discusses the handling of Japanese under UNIX, VM, DOS, Amiga and Mac. If you use anything more sophisticated than a Smith-Corona, Ken's guide will have a suggestion for getting Japanese on your machine. Many of the applications Ken cites are freeware or shareware, so cost need not keep you from breaking the romaji-barrier. Even if you're not ready to take the plunge into Japanese E-mail or wordprocessing, Ken offers a useful history of the various kanji systems (such as 常用漢字) and their various electronic counterparts. JAPAN.INF was last revised in early 1992 and therefore provides only cursory coverage of topics such as DOS/V, but it is still worthwhile reading. It is available via anonymous ftp. Look for the files JAPAN1.INF, JAPAN2.INF and JAPAN3.INF at the ftp sites ucDavis.edu (128.120.2.1) and msi.umn.edu (128.101.24.1). If parts of JAPAN.INF initially appears as gibberish, fret not. The gibberish is Japanese text, and the document describes how to get it back into Japanese.

E-mail in Japanese

Eudora, the splendid Internet mail manager for the Macintosh, now supports Japanese text. There are two versions: Eudora-J and Eudora-EJ. Eudora-J features Japanese menus, while EJ supports Japanese with an English-language interface. Japanese supports works under KanjiTalk, the Japanese Language Kit or SweetJam. Both Eudora-J and Eudora-EJ are free and available at several ftp sites. Check at ftp.cs.titech.ac.jp and miki.cs.titech.ac.jp. Although Eudora is all but self-explanatory, be sure to download the help files. For the very latest

information about Eudora, send mail to "eudora-info@qualcomm.com" You can subscribe to news about Eudora via the Internet. Send the following message to listserv@vmd.cso.uiuc.edu:

sub eudora your-first-name your-last-name

Be sure not to put anything else in the body of the message, even your signature.

If you can not use anonymous ftp, regular Eudora 1.3 is available by mail. (The availability of Eudora-EJ or Eudora-J by mail are unclear). Send a check for \$30 (payable to "QUALCOMM Incorporated") to:


QUALCOMM Incorporated
c/o Eudora Package
10555 Sorrento Valley Road
San Diego, CA 92121

You will receive the full Eudora distribution on diskettes, plus a printed copy of the manual. The \$30 is to cover the costs of materials, duplication, and handling. Eudora itself is free of charge.

DOS users should consult Ken Lunde's JAPAN.INF for suggestions on Japanese E-mail. For DOS/V, John Morris reports good results with a program called 秀Term. It is a public-domain application, but we have not yet found an ftp site.

Japanese Language Kit for the Macintosh

This summer Apple released the Japanese Language Kit, its long awaited Japanese module for WorldScript. The module allows Japanese text entry under System 7.1, the latest release of the Macintosh operating system. Unlike KanjiTalk, the established Japanese interface for the Macintosh, the Japanese Language Kit allows for Japanese text entry with a standard English language interface. System level menus remain in English, although the menus for bilingual applications can be switched from English to Japanese. Files and folders can be given Japanese names. Because the Japanese Language Kit is a set of system extensions, not a separate operating system, users can run virtually any English application without compatibility problems. Most applications designed for KanjiTalk, such as EG Word, will also run under System 7.1 with the Japanese Language Kit.

Turning the Japanese entry mode on and off is a snap. The  + SPACE key combination toggles between the US keyboard and the Japanese entry keyboard. Although the latest upgrade of EG Bridge (version 5.1) is compatible with System 7.1, the language kit includes its own conversion system, Kotoeri ことえり. Kotoeri offers the standard input modes: ひらがな, カタカナ, R o m a n, かな (half-width, one-

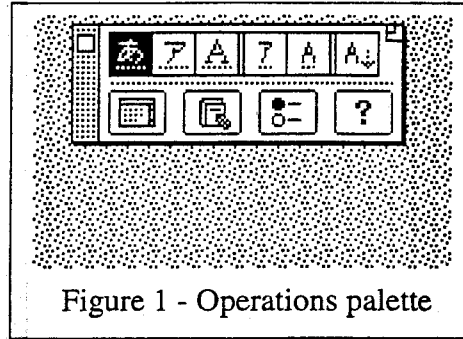


Figure 1 - Operations palette

byte *katakana*) and Roman (half-width, one-byte Roman). Hiragana and katakana can be converted to kanji by tapping the space bar. The user can select the input mode through keystroke commands or through

a handy operation palette (see figure 1). The palette, like all graphic interfaces, is unnecessary if you remember the keyboard commands, but invaluable when you forget the command codes. Kotoeri offers some welcome advances over EG Bridge, 2.1 変換 and other, older conversion utilities. It has a character palette option (figure 2), which allows the user to choose characters by radical, a handy option when a memory lapse leads you to the wrong *on* and *kun-yomi*. The character palette can be brought-up with a keystroke or from the operation palette. Kotoeri allows users to create multiple user dictionaries and the dictionary creation utility offers minor improvements over other system. Unlike EG Bridge, Kotoeri allows you to enter and view multiple dictionary entries at a time. But you may have little need for this option. Kotoeri's main dictionary is surprisingly robust. Under EG Bridge, for example, I had to manually enter 候 as a reading for そろろう. Under Kotoeri, this reading was included in the main dictionary. Finally, Kotoeri comes with bilingual documentation: parallel text in English and Japanese. No one who has struggle with a Japanese-language manual will overlook the value of a manual in idiomatic English. The parallel Japanese text also serves as a good primer for computer jargon.

The Japanese Language Kit comes with a complete set of fonts: two Postscript fonts (SaiMincho and ChuGothic), two TrueType fonts (HonMincho and KakuGothic) and one bit-mapped font (Osaka). The new TrueType fonts are a major step forward, since, unlike Japanese Postscript fonts, they do not require a special printer or special software for smooth scaling. (Japanese Postscript fonts require a special rasterizer — the standard Postscript interpreter will not do). With TrueType kanji, the computer, not the printer, scales the fonts to any size. The results are splendid, both on screen and on the printed page (see figure 3). The fonts consume an enormous amount of disk space, nearly 20 MB for both fonts, but this is not surprising given that a full set of kanji fonts contains nearly seven thousand kanji. The TrueType fonts also print surprisingly fast.

One liability of the new system is that it does not have a kanji-creation utility, such as the 外字登録 control panel in EG Bridge. Although my only occasion to use the 外字登録 utility was to create the special glyph for *yor*, some specialists find the standard JIS character set insufficient, particularly for proper nouns. Unfortunately the 外字登録 utility will not run under system 7.1. You can register non-standard

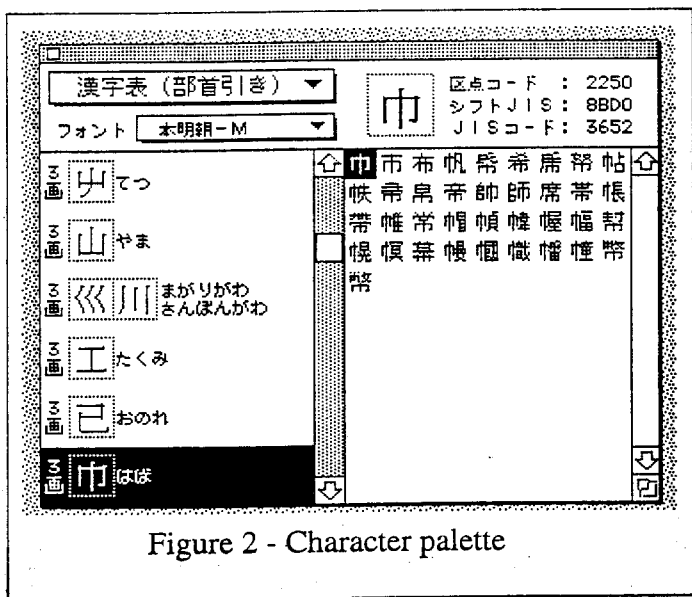


Figure 2 - Character palette

characters under 7.1 with a utility like ATM-J from Adobe Systems. Users interested in creating large numbers of characters might do well to look into a more advanced font creation application like Fontographer.

The Japanese Language Kit allows users to enjoy the much ballyhooed advances of System 7, such as publish and subscribe, TrueType fonts, better font and DA management, aliases, etc. The hardware requirements of System 7.1 with the Japanese language extensions are large but not exceptional. Like most Apple software, the system is backwards compatible. You can run the Japanese Language Kit with 7.1 on a 68000 with 4MB RAM, but things will happen excruciatingly slowly. A 68030 processor with at least 5 MB RAM is far preferable,

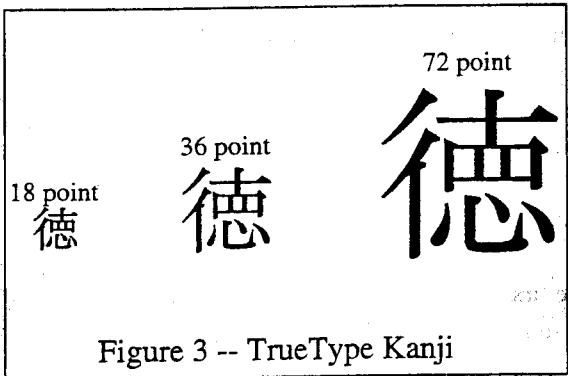


Figure 3 -- TrueType Kanji

and 8 MB is best for memory-hungry applications. The Japanese Language Kit bodes well for the future of multilingual computing: it is the first Asian language module for WorldScript, the multilingual environment for the Macintosh. At present there are modules for Eastern European languages, Hebrew, Arabic, Farsi and Japanese: all these languages can be handled simultaneously with a "WorldScript aware" application like Nisus or the latest release of WordPerfect. When the Chinese and Korean extensions arrive, WorldScript should fulfill the

needs of the most demanding polyglot.

The Japanese Language Kit lists for \$249.00, but it retails for roughly 20% less. It is available from both retailers specializing in Japanese software (such as Qualitas 510 848-8080) and large software houses (such as MacConnection 800-800-2222). Apple recently announced academic pricing for the Japanese Language Kit. It is now available through educational channels for \$149. For the educational discount, contact the dealer from which your school regularly buys. If you bought KanjiTalk 6.0.7, you can upgrade it to JLK for \$129. Contact Apple directly.

Japanese Capability for DOS

The following notice was posted to the Japanese Teachers and Instructional Technology mailing list. We reprint it here without recommendation. John R. McRae of Cornell, however, reports good results using the demo with *Nota Bene*.

Sponsored by Oxford University, a Japanese front-end processor for IBM-compatible PC has been developed and its version 1.0 is being released.

JitsuYou JKDOS is a so-called add-on software to standard MS-DOS/PC-BIOS of IBM compatible PCs. It provides the possibility of input and output of Japanese (Kana/Kanji) on the level of MS-DOS so that standard western software can be used for Japanese without any modifications.

Before starting an application, first start JitsuYou JKDOS by entering JY in the JY directory. JY is a TSR program which loads itself, the screen-driver and printer-driver into memory.

Following this you can use the character-oriented DOS applications like MS-WORD, WordPerfect, dBASE, Borland's language products etc., and input and output Japanese directly. JitsuYou JKDOS 1.0 supports EGA/VGA/Hercules graphics cards, 24 pin, HP II compatible and PostScript printers.

Together with ShiYong CCDOS 2.0, it is possible to mix Japanese, simplified Chinese, classical Chinese and most European languages in a single file and on a single printed page.

Available for a test/demo installation:

```
ftp.lrz-muenchen.de[129.187.10.35], in
/pub/culture/east-
asia/sw/IBMPC/J/JitsuYou
/pub/culture/east-
asia/sw/IBMPC/C/ShiYong
```

Hu Bo -- ucc02au@sunmail.lrz-muenchen.de (the author)