Le japonais avec Linux

Thierry Bézecourt

Le japonais avec Linux
par Thierry Bézecourt

Ce document est destiné aux utilisateurs de Linux qui désirent lire ou écrire des textes japonais. L'objectif n'est pas de dénombrer toutes les manières possibles de configurer son PC, mais simplement de donner quelques solutions qui marchent.

Vous n'aurez peut-être pas besoin de toutes les informations contenues dans ce document. En effet, le support de la langue japonaise est de mieux en mieux assuré dans les distributions de Linux, donc vous pourrez peut-être sauter certaines étapes. Voyez par exemple un exemple de mise en place avec Mandrake.

La version que vous lisez actuellement est celle du 25 janvier 2002. La dernière version est disponible sur http ://www.thbz.org/japonais/chapitres/book1.html (version HTML en ligne), sur http ://www.thbz.org/japonais/japonais-linux.html (version HTML en une page, pour téléchargement ou impression) ou encore sur http ://www.thbz.org/japonais/japonais.ps (version PostScript).

thbz.org (http ://www.thbz.org)

Table des matières

Contributions et remerciements

Toute remarque et contribution est la bienvenue (mailto :japonais@thbz.org), en particulier les corrections et les ajouts d'informations. Si vous avez des recettes plus simples que les miennes pour configurer votre machine, dites-le moi aussi.

Merci aux personnes suivantes, qui ont apporté une collaboration substantielle à ce document : Matthieu Dazy, Roland Mas, Julien "Mendoza" Messager, Christophe Raverdy, Thierry Grandjean.

Chapitre 1. Lire des pages Web japonaises

Avec Netscape, commencez par afficher une page Web japonaise, par exemple http ://www.yahoo.co.jp. Elle présente probablement une majorité de caractères incompréhensibles. Sélectionnez le menu "Affichage" de Netscape, puis "Character Set", puis "Japanese (auto-detect)". La page devrait s'afficher à nouveau avec des caractères japonais.

Si les caractères affichés sont toujours bizarres (points d'interrogation, caractères accentués européens anormaux, etc), voyez l'annexe sur les polices de caractères pour vérifier que vous avez les polices de caractères nécessaires et pour les installer si nécessaire.

Chapitre 2. Un exemple de mise en place avec Mandrake 8.1

Les informations comprises dans cette section ont été fournies par Thierry Grandjean. Merci Thierry !

2.1. Généralités

Configuration :

Mandrake 8.1 (kernel 2.4.8-26mdk)

KDE 2.2.1

J'ai essayé d'installer les packages qui me semblait utiles pour l'utilisation du Japonais (kon, kterm, pages de man et autres doc en version japonaise).

J'ai aussi créé un utilisateur (environnement KDE) dont la langue est le japonais pour valider ces installations.

2.2. Emacs

Puis, pour utiliser emacs, installation de :

emacs 20.7-17mdk

emacs-leim 20.7-17mdk

emacs-X11 20.7-17mdk (pas strictement indispensable, mais j'apprécie le concours de la souris)

emacs 20.7-17mdk

Lorsque l'on relance emacs après l'installation d'emacs-leim le menu Mule est actif. Dedans on choisit Set language environnement puis japanese.

J'ai ouvert le fichier whatis (version japonaise) sans problème. L'input method est positionné a japanese par défaut, mais il faut la valider avant de commencer la saisie.

2.3. XEmacs

Fichiers installés

xemacs 21.4.4-4mdk

xemacs-mule 21.4.4-4mdk

En fait, c'est xemacs-mule qui affiche le menu Mule supplémentaire sous Edit. Il s'agit des deux "produits" différents. Sous le menu Mule le menu Set Coding system of buffer file permet de choisir japanese.

Par contre, contrairement à emacs, il ne m'affiche pas directement les idéogrammes de mon fichier whatis (mais il détecte le bon système de codage). Je n'ai pas trouvé comment changer cet état de fait. Par contre, je n'ai pas installé le CJK-emacs-4.2.0-4mdk. Mais il est plutôt dédié au formatage de texte avant impression.

Chapitre 3. Logiciels à télécharger : PJE

Un grand nombre de logiciels nécessaires ou utiles pour lire et écrire du Japonais sont regroupés dans la distribution PJE (Project Japanese Extensions). Toutes les informations nécessaires sont dans le HOWTO de PJE (http ://www.teaser.fr/~craverdy/PJEHOWTO_FR.TXT), traduit en français par Christophe Raverdy. Voici un extrait de ce HOWTO qui vous donne une idée de ce qu'on trouve dans PJE :

> PJE contient des logiciels qui permettent d'afficher, saisir, éditer, traiter et imprimer des caractères et phrases en japonais. KON est inclus pour permettre une utilisation depuis la console. kterm et rxvt sont inclus pour servir depuis le système X-Window. Les systèmes de saisie comprennent Wnn 4.2, Canna et, SKK pour une utilisation avec mule. Il y a également jvim, un clone de vi qui fait fonction d'éditeur, Mule (MUlti Lingual supported Emacs) peut aussi être utilisé. Une version japonaise de Tex est jointe pour le traitement de texte. Vous pouvez créer/éditer des images avec Gimp ou tgif.
>
> Vous trouverez également des logiciels pour lire les forums de discussion sur Usenet, lire et écrire du courrier électronique. Il y a également des programmes de conversion entre les différents systèmes de codage japonais (SJIS, JIS, EUC).

Chapitre 4. Le Japonais et Emacs

Si vous n'êtes pas déjà familier avec des logiciels comme kterm, kinput ou xjdict, vous aurez peut-être intérêt à installer Emacs juste pour le confort d'utilisation du japonais qu'il offre.

4.1. Emacs : où, quoi, comment

Emacs est un éditeur de texte muni de multiples extensions, qui permet d'envoyer des mails, de lire les news sur Usenet, de parcourir des pages Web, de jouer à Tetris, et même d'éditer des fichiers.

De nos jours, il existe deux versions principales d'Emacs : XEmacs et GNU Emacs (ou Emacs tout court). J'utilise XEmacs au bureau et GNU Emacs à la maison, et je pense qu'ils se valent à peu près, le choix entre les deux se faisant surtout pour des raisons idéologiques. XEmacs, doté de fonctions graphiques évoluées, est peut-être plus facile à utiliser pour un débutant.

Pour obtenir et installer GNU Emacs, voyez http ://www.gnu.org/software/emacs/. Il est recommandé d'utiliser au moins la version 20.6.

Pour obtenir et installer XEmacs, voyez http ://www.xemacs.org.

Ce document traite des deux logiciels, mais avec un traitement de faveur pour GNU Emacs, parce que c'est celui que j'utilise chez moi, et que je peux donc faire des expérimentations dessus comme je veux.

4.2. Lire du japonais sous GNU Emacs

Ce qui suit permet de lire du japonais sous GNU Emacs, mais aussi d'autres langages utilisant des caractères non latins, comme le chinois ou le russe. J'utilise SuSE 8.1.3, mais ce document devrait être valable pour toutes les distributions (pas trop anciennes) de Linux. Certaines notes sont spécifiques à

Debian, qui semble avoir un bon support pour l'internationalisation : une grande partie des tâches décrites ci-dessous disparaissent avec cette distribution, et il n'y a pas besoin de trafiquer le fichier .emacs (sauf pour les problèmes de fontset, évoqués plus bas).

4.2.1. Polices de caractères

Voyez l'annexe sur les polices de caractères pour vérifier que vous avez les polices de caractères nécessaires et pour les installer si nécessaire.

Si, malgré cela, Emacs n'affiche pas les caractères japonais, c'est qu'il utilise peut-être le jeu de polices par défaut, qui ne gère que des polices basiques, et non le jeu de polices fontset-standard, qui peut traiter les polices japonaises. Il faut alors lui dire d'utiliser le bon jeu de polices soit sur la ligne de commande (emacs -fn fontset-standard) soit par une resource X (rajouter la ligne Emacs.font : fontset-standard dans votre fichier ~/.Xresources, et relancer X).

4.2.2. Affichage du japonais dans Emacs

Vous devriez à présent être en mesure de visualiser un fichier contenant du japonais encodé au format EUC-JP ou ISO-2022-JP. Le dictionnaire de Jim Breen, est au format EUC-JP (voyez la section Liens pour le récupérer). Utilisez la séquence suivante :

C-x RET c euc-japan C-x C-f "nom_du_fichier"

C-x RET c permet de spécifier un jeu de caractères (dans cet exemple "euc-japan") qui sera utilisé lors de la commande suivante, c'est-à-dire ici lors de l'ouverture et de l'affichage du fichier.

C-x RET c doit aussi être utilisé avant C-x C-s lorsqu'on souhaite sauvegarder un fichier dans un jeu de caractères donné. Pour plus d'informations, voyez la section "Specifying a Coding System" dans la documentation en ligne de GNU Emacs.

Si, au lieu de caractères japonais, vous voyez s'afficher des caractères européens (par exemple des guillements à la place du hiragana "ka"), c'est que le buffer courant est configuré en affichage mono-octet. C'est sans doute dû à un conflit avec une instruction de votre fichier d'initialisation d'Emacs (.emacs) : si vous avez mis (standard-display-european 1) dans ce fichier, essayez de le supprimer. Vous pouvez aussi activer l'affichage de caractères multi-octets dans un buffer d'Emacs avec la commande M-x toggle-enable-multibyte-characters

4.3. Ecrire en japonais sous GNU Emacs

Ceci s'applique à GNU Emacs 20.6. Cela demande de recompiler Emacs, ce qui n'est pas difficile, et c'est vraiment très agréable de parvenir à entrer des kanjis sous Emacs. D'ailleurs, si ça se trouve, votre version de Emacs a été compilée avec Leim : voyez d'abord la section Entrer des caractères japonais, et revenez à la section Installation de Leim si cela ne marche pas directement.

4.3.1. Installation de Leim

Il faut recompiler Emacs avec Leim. N'ayez pas peur, ce n'est pas difficile de compiler Emacs. Leim est un système qui permet de rentrer des caractères japonais (le chinois est aussi supporté) d'une manière conviale, à peu près comme l'IME sous Windows. Leim est disponible aux mêmes endroits que GNU Emacs, par exemple sur http ://ftp.gnu.org/pub/gnu/emacs (ftp ://ftp.gnu.org/pub/gnu/emacs). Il faut prendre la même version de leim que celle d'Emacs avec laquelle vous voulez l'utiliser, par exemple leim-20.6.tar.gz si votre version d'Emacs est 20.6. Compiler Emacs avec Leim ne pose pas de problème. Tout est indiqué au début du fichier INSTALL de la distribution d'Emacs : en gros, il suffit de décompacter l'archive de Leim au même endroit que celle d'Emacs et de relancer la compilation d'Emacs : Emacs reconnaîtra automatiquement que Leim est présent.

4.3.2. Entrer des caractères japonais

Une fois que vous avez compilé et installé GNU Emacs avec Leim, vous pouvez entrer du texte japonais dans n'importe quel buffer. Pour cela, créez un nouveau buffer (ou allez dans le buffer *scratch* qui est créé par Emacs au démarrage), puis tapez M-x set-input-method. Emacs vous demande alors une "méthode de saisie" : tapez japanese. A présent, lorsque vous tapez ka, le caractère hiragana correspondant sera affiché automatiquement. Si vous voulez à la place le caractère katakana, tapez K. Si vous voulez l'un des kanjis se prononçant "ka", tapez sur la barre d'espace une ou plusieurs fois, jusqu'à ce que le caractère désiré s'affiche. D'autres commandes sont disponibles : tapez C-h m pour plus d'informations.

> Tuyau: Vous pouvez basculer facilement entre le mode de saisie normal (français) et un autre mode de saisie avec C-\ (toggle-input-method). Cette commande, par défaut, utilise le mode de saisie choisi lors du dernier appel à set-input-method ; pour qu'elle fonctionne dès le lancement d'Emacs, rajoutez la ligne suivante dans votre .emacs (ou utilisez M-x customize-variable) :
>
> (setq default-input-method 'japanese)

> Note: Un autre mode de saisie, nommé SKK

4.4. Lire et écrire du japonais sous XEmacs

Merci à Julien "Mendoza" Messager qui a fourni le contenu de cette section.

4.4.1. Installation d'un XEmacs dans lequel on pourra entrer du japonais

> Note: Remarque : pour trouver les fichiers mentionnés ci-dessous, vous pouvez par exemple aller les rechercher avec le moteur de recherche filewatcher.org (http://filewatcher.org)

Les fichiers suivants contiennent tout ce qu'il faut :

xemacs-21.1.8-elc.tar.gz

xemacs-21.1.8-info.tar.gz

xemacs-21.1.8.tar.gz

xemacs-mule-sumo-2000-01-24.tar.gz

xemacs-sumo-2000-01-24.tar.gz

Ensuite il faut compiler XEmacs en faisant :

```
./configure –with-mule
make
make install
```

Il reste alors les fichiers "sumo" à installer.

4.4.2. Configuration

Pour automatiser tout ca, vous pouvez établir des raccourcis claviers avec les commandes suivantes dans votre fichier .emacs :

```
;; Configuration SKK
(global-set-key "\C-x\C-j" 'skk-mode)
(global-set-key "\C-xj" 'skk-auto-fill-mode)
(global-set-key "\C-xt" 'skk-tutorial)
(autoload 'skk-mode "skk" nil t)
(autoload 'skk-tutorial "skk-tut" nil t)
(autoload 'skk-check-jisyo "skk-tools" nil t)
(autoload 'skk-merge "skk-tools" nil t)
(autoload 'skk-diff "skk-tools" nil t)


(autoload 'skk-isearch-mode-setup "skk-isearch" nil t)
(autoload 'skk-isearch-mode-cleanup "skk-isearch" nil t)
(add-hook 'isearch-mode-hook
        (function (lambda ()
                (and (boundp 'skk-mode) skk-mode
                (skk-isearch-mode-setup)))))
(add-hook 'isearch-mode-end-hook
```

```
        (function (lambda ()
                (and (boundp 'skk-mode) skk-mode
                (skk-isearch-mode-cleanup)
                (skk-set-cursor-color-properly)))))

;; Tutorial SKK en Anglais (c'est mieux que tout en Jap! :))
    (setq skk-tut-file
"/usr/local/lib/xemacs-21.1.9/mule-packages/etc/skk/SKK.tut.E")
    (setq minibuffer-max-depth nil)
```

Pour activer SKK il suffit maintenant de faire C-x C-j (idem pour le désactiver). On voit "Kana" en hiraganas s'afficher en bas. Il suffit alors de taper son texte en romaji pour qu'il s'affiche en Hiraganas.

Pour taper des Katakanas, il faut appuyer sur q ("Kana" en bas passe en katakana). Idem pour repasser en mode Hiragana.

Pour taper en romaji, il faut appuyer sur l. C-j pour repasser en mode Kana.

Pour taper en caractère romains grande taille (je ne sais pas comment ca s'appelle) il faut appuyer sur L (majuscule). C-j pour repasser en Kanas.

Pour les Kanjis c'est plus subtil. Mettons sur je veuille taper "Umi" (Mer). Je tape alors en mode Hiragana "Umi" avec le "U" majuscule, puis un espace. SKK me propose alors un Kanji. Si c'est le bon, il suffit de faire C-j. Sinon encore espace jusqu'à obtention du bon kanji.

Pour lancer le tutorial SKK (très bien foutu) : C-x t.

4.5. UTF-8 sous GNU Emacs

Cette section décrit comment faire de l'encodage et du décodage de fichiers et de messages en UTF-8 sous GNU Emacs. Emacs, en effet, ne supporte par défaut que ISO-2022-JP et EUC-JP.

Cette section doit beaucoup à l'excellent document Unicode Encoding for GNU Emacs (http ://www.cs.ust.hk/faculty/otfried/Mule/) d'Otfried Cheong.

4.5.1. Installation

UTF-8 n'est pas supporté par défaut par GNU Emacs 20.6, mais il est facile de le rajouter. Pour cela, il faut télécharger Mule-UCS. J'ai récupéré la version 0.74 sur ftp ://ftp.m17n.org/pub/mule/Mule-UCS/. Si vous utilisez Debian, vous pouvez le trouver dans le paquet mule-ucs.

Pour installer Mule-UCS, suivez les instructions qui sont dans les fichiers README et README.Unicode du répertoire Lisp. En bref, il faut lancer la commande suivante pour compiler les fichiers :

emacs -q –no-site-file -batch -l mucs-comp.el

Puis il faut copier le répertoire lisp de la distribution en-dessous du répertoire site-lisp associé à votre installation GNU Emacs. Puis il faut rajouter ce répertoire au chemin reconnu par Emacs. Pour cela, éditez votre fichier d'initialisation d'Emacs (~/.emacs), et rajoutez ceci (en adaptant le chemin indiqué si nécessaire) :

```
(setq load-path
      (append (list
       "/usr/local/share/emacs/20.6/site-lisp/mule-ucs"
      load-path))
```

Enfin, il faut rajouter la ligne suivante dans le même fichier pour intialiser Unicode :

```
(require 'un-define)
```

4.5.2. Test

Relancez Emacs, et vérifiez que utf-8 est reconnu comme système d'encodage. Pour cela, choisissez "Describe Coding Systems" dans le menu "Mule", et tapez utf-8. Si Emacs refuse cette valeur, c'est que Mule-UCS n'a pas été installé correctement. Dans ce cas, vérifiez que vous n'avez pas lancé Emacs avec l'option –unibyte et que la variable EMACS_UNIBYTE n'est pas définie dans votre environnement.

Si UTF-8 est reconnu par Emacs, ouvrez un nouveau fichier, entrez dedans du texte (par example en mélangeant des caractères français accentués et du japonais), puis sauvegardez-le avec la séquence suivante :

C-x RET c utf-8 C-x C-s

Cette commande sauvegarde le fichier en l'encodant en UTF-8 : C-x RET c utf-8 spécifie que la commande qui suit (c'est-à-dire C-x C-s) doit utiliser l'encodage UTF-8.

Vous pouvez vérifier que le fichier a bien été encodé en UTF-8 en l'ouvrant sous Netscape (si c'est de l'HTML ou du texte simple), et en choisissant le jeu de caractères UTF-8 dans le menu "View".

Vous pouvez désormais utiliser le jeu de caractères UTF-8 de la manière que les autres jeux de caractères.

4.6. UTF-8 sous XEmacs

Pour cette section, je vous renvoie au Unicode-HOWTO (ftp ://ftp.ilog.fr/pub/Users/haible/utf8/Unicode-HOWTO-4.html) de Bruno Haible (traduit en français (http ://www.freenix.org/unix/linux/HOWTO/Unicode-HOWTO-4.html) par Samuel Tribehou) qui décrit tout cela.

Il semble toutefois que le patch mentionné dans le Unicode-HOWTO mérite quelques "ajustements" avec certaines versions de XEmacs. Matthieu Dazy a dû faire ce qui suit avec XEmacs 21.1.8 (sous Solaris 2.6) :

> une fois le patch appliqué, éditer src/file-coding.c et y remplacer les deux 'xxDEFUN' par 'DEFUN',
>
> ajouter –cflags="-DENABLE_COMPOSITE_CHARS" à la configuration.

D'autre part, il faut veiller à charger xemacs-ucs avant Gnus lors du démarrage de XEmacs, sans quoi Gnus ne reconnaît pas le charset utf-8.

4.7. Lire et écrire du japonais sous Gnus avec GNU Emacs

4.7.1. Lire

Si Emacs a été correctement configuré, et que Mule-UCS a été installé, cela devrait fonctionner sans problème. Pour configurer le jeu de caractères selon le forum, voyez la section "Summary Buffer->Charset" dans la documentation de Gnus. D'une manière générale, Gnus se débrouille bien si l'article a les en-têtes standards "Content-Type" et "Content-Transfer-Encoding". Pour voir les en-têtes d'un article, mettez le curseur sur la ligne de l'article dans le buffer "Summary", et tapez t.

Il semble que, sous XEmacs 21.1, il faille passer à Gnus 5.8.4. Il y aurait des problèmes d'incompatibilité avec la version 5.8.3.

4.7.2. Ecrire

4.7.2.1. Jeu de caractères et encodage du message

Il faut d'abord comprendre la différence entre les deux concepts suivants :

Le jeu de caractères d'un message correspond à la manière dont les caractères (lettres européennes, kana ou kanjis, etc) sont transformés en octets. Plusieurs conventions existent, comme ISO-8859-1 (qui ne traite que des caractères européens), ISO-2022-JP ou EUC-JP (pour les caractères japonais et américains), ou enfin UTF-8 (qui convient pour les caractères utilisés dans toutes les langues majeures, dont le français et le japonais). La section sur les jeux de caractères vous indique comment choisir (ou plutôt comment laisser Gnus choisir) le jeu de caractères à utiliser.

Le format d'encodage pour le transport correspond à la manière dont le message, quel que soit le jeu de caractère qu'il utilise, doit être transformé avant d'être envoyé sur Usenet. Les deux principaux formats d'encodage utilisés sur Usenet sont "Quoted-Printable" et "8bit". Il est fortement recommandé de n'utiliser que "8bit", car "Quoted-Printable" est un format ancien qui n'est plus vraiment utile de nos jours. La section sur l'encodage vous indiquera comment dire à Gnus d'utiliser le bon format d'encodage.

4.7.2.2. Spécifier le bon jeu de caractères

Je ne connais pas de manière de spécifier explicitement le jeu de caractères (champ "Content-Type" dans l'en-tête du message) que l'on veut utiliser quand on envoit un message, mais ce n'est pas grave, car Gnus sait se débrouiller pour choisir celui qui convient. Pour cela, il regarde le contenu du message et cherche un jeu de caractères qui puisse encoder tout le buffer, ou au moins une partie. Les cas qui nous intéressent sont les suivants :

Si le message contient uniquement des caractères ASCII et des caractères européens : tout sera en ISO-8859-1.

Si le message contient uniquement des caractères ASCII et des caractères japonais : tout sera en ISO-2022-JP.

Si le message contient des caractères ASCII, des caractères européens et des caractères japonais : tout sera en UTF-8 si Mule-UCS a été installé, sinon le message sera divisé en plusieurs éléments Mime.

Autrement dit, il n'y a rien à configurer. Mais faites d'abord un ou deux tests dans fr.test (news :fr.test) avant de vous lancer...

4.7.2.3. Spécifier le format d'encodage

Comme l'encodage Quoted-Printable est plus ou moins interdit (ou du moins réprouvé) sur Usenet, on peut spécifier à Gnus de n'utiliser que le 8bit (c'est-à-dire, pas d'encodage) dans le corps de texte des messages envoyés (mais pas dans les en-têtes, car c'est mal). Ça se fait par la variable gnus-group-posting-charset-alist. La configuration recommandée est : autoriser tous les jeux de caractères sans encodage dans le corps du message, et aucun dans les en-têtes (ceci est valable aussi bien pour l'e-mail que pour les news). Pour ce faire, on peut ajouter les lignes suivantes dans son .emacs ou dans son .gnus :

```
(setq gnus-group-posting-charset-alist
      (quote ((message-this-is-mail nil t)
              (message-this-is-news nil t))))
```

On peut aussi utiliser le buffer Customize (M-x customize-variable gnus-group-posting-charset-alist), si on parle anglais.

4.8. Lire et écrire du japonais sous Gnus avec XEmacs

Matthieu Dazy rapporte que, après avoir configuré XEmacs pour UTF-8 comme indiqué précédemment, la lecture de messages UTF-8 sous Gnus ne pose pas de problème. Par contre, lors de l'envoi de message, Gnus persiste à diviser le message en plusieurs parties Mime lorsque caractères européens et japonais sont présents en même temps, au lieu d'utiliser UTF-8. Si vous y parvenez, merci de me dire comment vous avez fait (mailto :japonais@thbz.org) !

Chapitre 5. Traduire du Japonais en Anglais et inversement sous Emacs

Voici l'exemple de scénario qui fait rêver l'étudiant en japonais : lire des pages Web japonaises ou des newsgroups nippons, et obtenir instantanément la traduction des kanjis que l'on ne comprend pas, juste en tapant sur une touche (ou deux). C'est possible sous Emacs : il suffit de lire la page Web avec W3, ou les news sous Gnus (deux logiciels intégrés dans Emacs), et d'utiliser edict.el pour obtenir la traduction des mots obscurs. Cette section décrit l'installation et l'utilisation de edict.el.

Il faut d'abord récupérer le dictionnaire edict de Jim Breen : edict.gz (ftp ://ftp.cc.monash.edu.au/pub/nihongo/edict.gz). Décompressez ce fichier et copiez-le quelque part sur votre disque dur. Dans ce qui suit, je suppose qu'il est installé dans /usr/local/japonais/edict.

Récupérez ensuite edict.el, un programme en Elisp maintenu par Stephen J. Turnbull. La dernière version, au moment où j'écris ces lignes, est semble-t-il http ://turnbull.sk.tsukuba.ac.jp/Tools/XEmacs/packages/edict/edict-1.06-src.tar.gz.

Les instructions du fichier README sont tout sauf claires (par exemple, il se réfère à un fichier auto-loads.el qui n'est pas fourni avec la distribution et qui, de toute manière, n'est pas nécessaire). Ce qui suit devrait être suffisant. Extrayez d'abord les fichiers dans le répertoire site-lisp/edict de votre installation Emacs. Puis rajoutez dans votre fichier .emacs les lignes suivantes :

```
;;; Japonais : edict

;; Permettre à Emacs de trouver les macros Elisp. Remplacez le
;; répertoire par celui où vous avez installé edict.el.
(setq load-path
    (append (list "/usr/local/share/emacs/20.6/site-lisp/edict")
    load-path))

;; Japonais : edict.el
(setq edict-dictionaries
    '(("/usr/local/japonais/edict" . euc-jp)))
```

```
;; Charger la librairie edict.el (c'est rapide, pas la peine
;; d'utiliser le mécanisme d'autoloads)
(load-library "edict")

;; Raccourci clavier : 'M-_' pour traduire du japonais vers
;; l'anglais.
(global-set-key "\e_" 'edict-search-kanji)

;; Raccourci clavier : 'M-*' pour traduire de l'anglais vers le
;; japonais
(global-set-key "\e*" 'edict-search-english)
```

C'est fini à présent. Relancez Emacs, ouvrez un fichier contenant du japonais, ou tapez du japonais dans un buffer. Puis sélectionnez un mot japonais à la souris et tapez M- _ (c'est-à-dire la touche ALT, suivie de la touche "underscore") : la traduction du mot devrait apparaître dans un petit buffer, juste au-dessus du buffer courant. Ensuite, sélectionnez un mot anglais et tapez M- * : la traduction en japonais apparaît.

> Note: Une fois que vous avez fait fonctionner edict, vous pouvez récupérer d'autres dictionnaires sur le site de Monash (ftp://ftp.cc.monash.edu.au/pub/nihongo/00INDEX.html), comme par exemple le dictionnaire des noms propres (enamdict) ou celui des noms de lieux (j_places). Installez-les dans le même répertoire que edict, et rajoutez leurs noms dans la variable edict-dictionaries :

```
(setq edict-dictionaries
      '(("/usr/local/japonais/edict" . euc-jp)
("/usr/local/japonais/enamdict" . euc-jp)
("/usr/local/japonais/j_places" . euc-jp)))
```

Annexe A. Polices de caractères

La première chose à faire est de vérifier que vous avez des polices de caractères japonaises avec la commande suivante (dans une fenêtre xterm) :

xlsfonts | grep jis

Si cette commande affiche des noms de polices de caractères, c'est que vous avez probablement ce qu'il faut. Sinon, vous pouvez sans doute les installer à partir du CD-ROM de votre distribution Linux. Par exemple, sous Debian, vous pouvez installer les paquets xfonts-intl-*. Vous pouvez aussi récupérer le package Intlfonts (http ://metalab.unc.edu/pub/gnu/intlfonts) d'Emacs, mais faites attention à sa taille : l'ensemble fait plus de 20 méga-octets, mais vous pouvez récupérer uniquement ce qui est relatif au japonais dans le sous-répertoire split/.

Annexe B. Liens

Cette section contient un petit nombre de liens, parmi ceux qui m'ont le plus aidé à configurer ma machine ou qui me paraissent les plus utiles.

B.1. Documentations

Linux-nihongo (http://hikari.tlug.gr.jp/~craigoda/writings/linux-nihongo/linux-nihongo.html) : l'excellente documentation de Craig Oda, au format PDF, sur l'utilisation du japonais sous Linux. A lire si vous voulez utiliser le Japonais dans une xterm, par exemple.

FAQ Windows (http://www.rhialto.co.uk/jwinfaq/jwinfaq.txt) : une FAQ sur l'utilisation du Japonais sous Windows.

Japonais et informatique (http://www.teaser.fr/~craverdy) : Christophe Raverdy propose une procédure pour installer tout ce qui est nécessaire, sous Windows et surtout sous Linux avec SuSE.

FAQ de fr.lettres.langue.japonaise (http://www.chez.com/tierney/chinagirl/faq.htm) : vous trouverez là aussi un grand nombre d'informations utiles. Vous pouvez poser des questions sur le forum Usenet fr.lettres.langue.japonaise (news:fr.lettres.langue.japonaise), mais, si vous comptez envoyer des messages en japonais, consultez d'abord la page de Jean-Marc Desperrier (http://perso.club-internet.fr/jmdesp/news/) pour bien configurer votre logiciel de news.

B.2. Logiciels

The Monash Nihongo ftp Archive

(ftp://ftp.cc.monash.edu.au/pub/nihongo/00INDEX.html) : le grand site FTP consacré au Japonais. On y trouve divers logiciels et dictionnaires, dont edict de Jim Breen. Indispensable.

Jim Breen's Japanese Page (http://www.dgs.monash.edu.au/~jwb/japanese.html) : un site complémentaire du précédent, avec des liens et des commentaires.

Home Page of SKK (http://skk.kuis.kyoto-u.ac.jp/skk/index-e.html) : pour récupérer le mode de saisie SKK pour Emacs.

B.3. Unicode et polices de caractères

Unicode-HOWTO (ftp://ftp.ilog.fr/pub/Users/haible/utf8/Unicode-HOWTO-4.html) : un HOWTO intéressant sur la configuration de Linux en Unicode, écrit par Bruno Haible. Il existe une traduction en français (http://www.freenix.org/unix/linux/HOWTO/Unicode-HOWTO-4.html) par Samuel Tribehou.

UTF-8 and Unicode FAQ (http://www.cl.cam.ac.uk/~mgk25/unicode.html) : une description technique de UTF-8.

Unicode Encoding for GNU Emacs (http://www.cs.ust.hk/faculty/otfried/Mule/) : un guide complet pour utiliser Unicode sous Emacs, par Otfried Cheong. Il indique comment récupérer des polices de caractères (européennes seulement), comment installer Mule-UCS, et comment utiliser toute la puissance d'Unicode sous Emacs.

Projet GNU Unifont (http://czyborra.com/unifont/) : un projet GNU visant à créer une police complète pour tous les caractères définis par Unicode.

Installing New Fonts to X Window System (http://sizif.mf.uni-lj.si/linux/cee/x11/font.install.html) : un guide à consulter pour installer une police de caractères si vous n'avez pas le mode d'emploi.

Annexe C. A propos de ce document

Ce document vous rappelle quelque chose, vous avez déjà vu d'autres docs plus ou moins techniques avec une apparence proche de celle-ci. La raison est simple : j'ai écrit le document non pas directement en HTML ou en PostScript (selon la version que vous consultez en ce moment), mais en SGML-DocBook. Les versions HTML et PostScript sont ensuite générées par le système de feuilles de style de Norman Walsh. Si ça vous intéresse, allez voir sur http ://www.docbook.org et http ://www.nwalsh.com/docbook/dsssl/index.html. C'est bien pratique, surtout si on utilise Emacs (http ://www.gnu.org/software/emacs/emacs.html) et PSGML (http ://www.lysator.liu.se/projects/about_psgml.html) pour écrire le code SGML lui-même.