

ジャイナ教聖典のデータベース
— 特殊フォントで表現されたデータのPDF —
Database to the Canonical Texts in Jain
- PDF of the Text in Special Font -

逢坂 雄美
Yumi OUSAKA

仙台電波工業高等専門学校, 仙台市
Sendai National College of Technology, Sendai 989-3124

あらまし:

中期インド・アリアン諸語(Middle Indo-Aryan)で書かれたジャイナ教(アルダ・マガダ語)古文献等の解析のために計算機リソースを開発してきた。これまで、ジャイナ教聖典の5冊の聖典に関するマッキントッシュ版電子ブックを数年前に作成済みである。この電子ブックには、当該研究に必須の基礎資料、電子化テキスト・韻律図式と解析結果・詩脚の正順索引と逆順索引・語彙の正順索引と逆順索引が収録されている。最近のウインドウズマシンの使用状況からみて、電子ブックをこのマシン上でも読むことができるならば、その有用性は大幅に上昇するであろう。特殊フォントを使用しているために、従来のツールではウインドウズ用電子ブックを作成することができなかったが、最近不十分ながらもそのツールが提供されるようになった。ウインドウズ用電子ブック作成とその問題点について議論する。

Summary :

We investigated computer analysis of the canonical texts in Prākṛit (called Ardha-Māgadhī), that is, the early Jainist ones in Middle Indo-Aryan (MIA). By using Macintosh we have already published the electronic books for the five important canonical texts on WWW. This book contains the essential data on the study of MIA, that is, the electronic text, metre analysis, pāda index and word index. Recently Win-

dows PC is used for the study of MIA. It would be very useful if we could produce the electronic book on Windows. New computer tool made us prepare this electronic book on Windows. We will discuss how to produce this book in detail.

キーワード: 中期インド・アリアン, アルダ・マガダ語, ジャイナ教, パーリ語, 初期仏教, 電子ブック

Key words: Middle Indo-Aryan, Ardha-Māgadhī, Jain, Pāli, early Buddhism, Electronic Book

1. はじめに

中期インド・アリアン諸語(Middle Indo-Aryan)で書かれた古文献は、仏教(パーリ語)とジャイナ教(アルダ・マガダ語)の聖典がその例に挙げられるように、世界の文化の貴重な宝庫であり、現代文明の重要な源流の一つである。初期仏教・ジャイナ教は厳格なカースト制度のバラモン教に対抗する形で同時期に発生したため、相互に関係を持ちながらその教理等を深化させてきた。それゆえ、今日では初期仏教を研究をする上でジャイナ教の研究は不可欠であり、その逆も真なりであると認識されている。しかしながら、これらの諸聖典は難解な言語で書かれていること、さらに質的・量的にも膨大

であることもあって、今後、厳密な研究成果が待たれる分野である。

このような状況下において、我々は中期インド・アリアン諸語に関する正確な知識を有する研究者と情報科学の研究者との密接な共同研究により、図1のような5つの計算機リソースを開発した。

- I. アルダ・マガダ語又はパーリ語のローマ字化システム構築（これをPali96フォントと呼ぶ）。
- II. ローマ字化テキスト作成：ジャイナ教の重要な5つのテキスト（Uttarajjhāyā, Āyāraṅga, Sūyagada, Dasaveyāliya, Isibhāsiyāim）及び仏教の4聖典（Dhammapada, Theragāthā, Therīgāthā, Suttanipāta）等をPali96フォントを使ってマッキントッシュに収録した。
- III. 聖典の韻律解析プログラム：電子化テキストファイルが韻律解析の入力として使用される。
- IV. 語彙の正順索引・逆順索引とパーダの正順索引・逆順索引の作成プログラム。
- V. 聖典の文法解析プログラム。

これらのツール等は、当該分野の研究において極めて重要な役割を果たす基礎資料を作成するに当たって、威力を発揮する。中期インド・アリアン諸語の文献の研究を飛躍的に進展させるためには、韻律解析・文法解析・語彙と構文論等の系統的な研究が不可欠である、と認識されている。テキストの韻律解析は批判的校訂本作成に不可欠であり、語彙の索引は翻訳に当たっての手助けとなる。また、語彙の逆順索引は文法構造解析に必要で、パーダ（詩脚）の正順または逆順索引はテキストの正しい読みの確立と並行詩脚を見いだすのに重要である。

上記のツールを使用して作成されたデータ

は、各出版社から単行本等して出版されたもの【1】と、ウェブ上で公開しているものがある（<http://www.sendai-ct.ac.jp/~ousaka>）【2】。ウェブ上で公開している計算機リソース（フォントシステム・電子ブック・プレインテキスト・逆順索引・特殊フォント用エディタ）を表1にまとめてある。

表に示しているように、マッキントッシュ版電子ブック（ジャイナ教聖典の5冊の電子ブック（Eブック））は既に作成済みである。この電子ブックには我々の作成した特殊フォントPali96を使用して作成された電子化テキスト・韻律図式と解析結果・詩脚の正順索引と逆順索引・語彙の正順索引と逆順索引が収録されている。最近のウィンドウズマシンの使用状況からみて、電子ブックをこのマシン上でも読むことができるならば、その有用性は大幅に上昇するであろう。

表1. 我々のウェブサイトアップロードされたデータ。表中の記号○、×は該当する計算機システムで使用可能なデータと使用不可のものを示している。

データ	マッキントッシュ	ウィンドウズ
特殊フォントPali96とキーボード入力法	○	○
ジャイナ教の5冊の聖典のコンピュータブック	○	●
ジャイナ教の5冊の聖典のプレインテキスト	○	○
初期仏教聖典の語彙の逆順索引	○	○
特殊フォントPali96用簡易エディターEDISANとその使用書	×	○

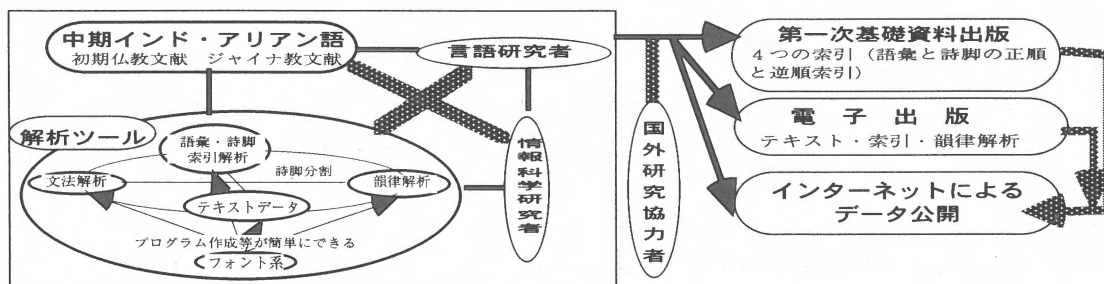


図1. 計算機リソースと研究組織

特殊フォントを使用しているために、従来のツールではウインドウズ用電子ブックを作成することができなかったが、最近不十分ながらもそのツール (Acrobat4.0) が提供されるようになった。表1の中で、●の部分、ウインドウズ上のジャイナ教聖典の5冊の電子ブック (Eブック)、つまり Portable Document File (PDF) は、このツールを用いて最近作成されたものである。本論文では、なぜこのウインドウズ用電子ブックを作成するのが遅れたのか、この作成に関連した問題点について議論する。

当該問題の理解には、既に作成しているマッキントッシュ版電子ブックとの比較が不可欠である。第2節では、この比較を主眼点にして、問題の原因となっている特殊フォント及び電子ブックに収納された内容について議論する。第3節ではウインドウズ版電子ブック作成及びその問題点について述べる。第4節では今後の課題等についてふれる。

2. マッキントッシュ版電子ブック

ここでは、最初に特殊フォント Pali96 の特性について述べ、次いでマッキントッシュ版電子ブックについて議論する。

2.1 フォントシステム, Pali96

計算機解析に当たって、各聖典の刊行本をそのままコンピュータ処理することが出来れば、それが最良の方法である。しかし、デーヴァナーガリー文字 (梵字、卒塔婆等に使用されている文字) で書かれたテキストを、自動的に読み込むコンピュータツールが存在しないために、現時点ではこれは不可能である。デーヴァナーガリー文字テキストからローマ字化テキストへの変換には膨大なタイプ入力が必要され、多大な時間と経費が要求される。この作業中に、予期できないタイプミス等がローマ字化テキストに忍び込み易い。また、種々の索引を作成した場合、その校正には膨大な時間を要する。以上のことから、認識性の高い高品質の中期インド・アリアン諸語の為のローマ字化フォントを作成できるかどうか、その研究を遂行する上で重要な役割を果たすことになる。さらに、適切に構築されたフォント体系はテキスト解析プ

ログラムの作成を容易にする。これらの条件を全て満たすようなローマ字化フォント (表2参照) を作成した。この表には、通常のローマ字フォントに無い特殊文字 (例えば 'm' 等) のみを掲載している。特殊フォントは全て1バイトコードに割り当てており、拡張アスキーコード上に定義されている。

この拡張アスキーコード上に定義されたキャラクタの属性保持が、このフォントの取り扱いでは最も肝要である。マッキントッシュ上では、全てのワープロソフト等でこの属性を正確に保持しているが、ウインドウズシステムでは基本的にこの属性が保持されていない。このことが、原因でウインドウズ用電子ブックが作成されなかった。

ウインドウズシステムではワープロソフト等でも当然同様なことが発生している。このため、我々はマッキントッシュで作成したテキストを、ウインドウズでも使用できるようにするために、このフォント属性を保持できる簡易エディタ EDISAN を作成している。

表2には、ウインドウズマシンにおける特殊フォント Pali96 のアスキーコードへの割り当てと共に、キー操作法を示している。このフォント系は、基本的にはマッキントッシュの場合と殆ど同一の入力法を採用している。つまり、マッキントッシュの Option key を、ウイ

表2. ウインドウズPCにおけるパーリフォント Pali96 のキー操作法とアスキーコード割り当て。記号 <A>, <S> はそれぞれ Alt key, Shift key を示している。正規の文字は通常通りにタイプできる。特殊文字は Alt key, Shift key と通常のキーを組み合わせることにより入力できる。第3列と6列の数字はアスキーコードを表している。

ā	<A>+n	136	ṭ	<A>+t	160
ā	<A>+a	140	ṭh	<S>+<A>+y	231
ī	<A>+<S>+`	147	d	<A>+d	182
ī	<A>+i, i	148	ḍh	<A>+i, U	243
ū	<A>+e, u	156	n	<A>+n, n	150
ū	<A>+u, u	159	ṭh	<A>+y	180
m̄	<A>+m	181	dh	<A>+z	189
kh	<A>+k	251	ph	<A>+p	185
gh	<A>+g	187	bh	<A>+b	186
ñ	<A>+c	169	l̄	<A>+l	194
ch	<A>+j	198	l̄	<A>+<S>+l	241
jh	<A>+k	198	ś	<A>+s	167
ñ	<A>+<S>+i	246	ṣ	<A>+x	197

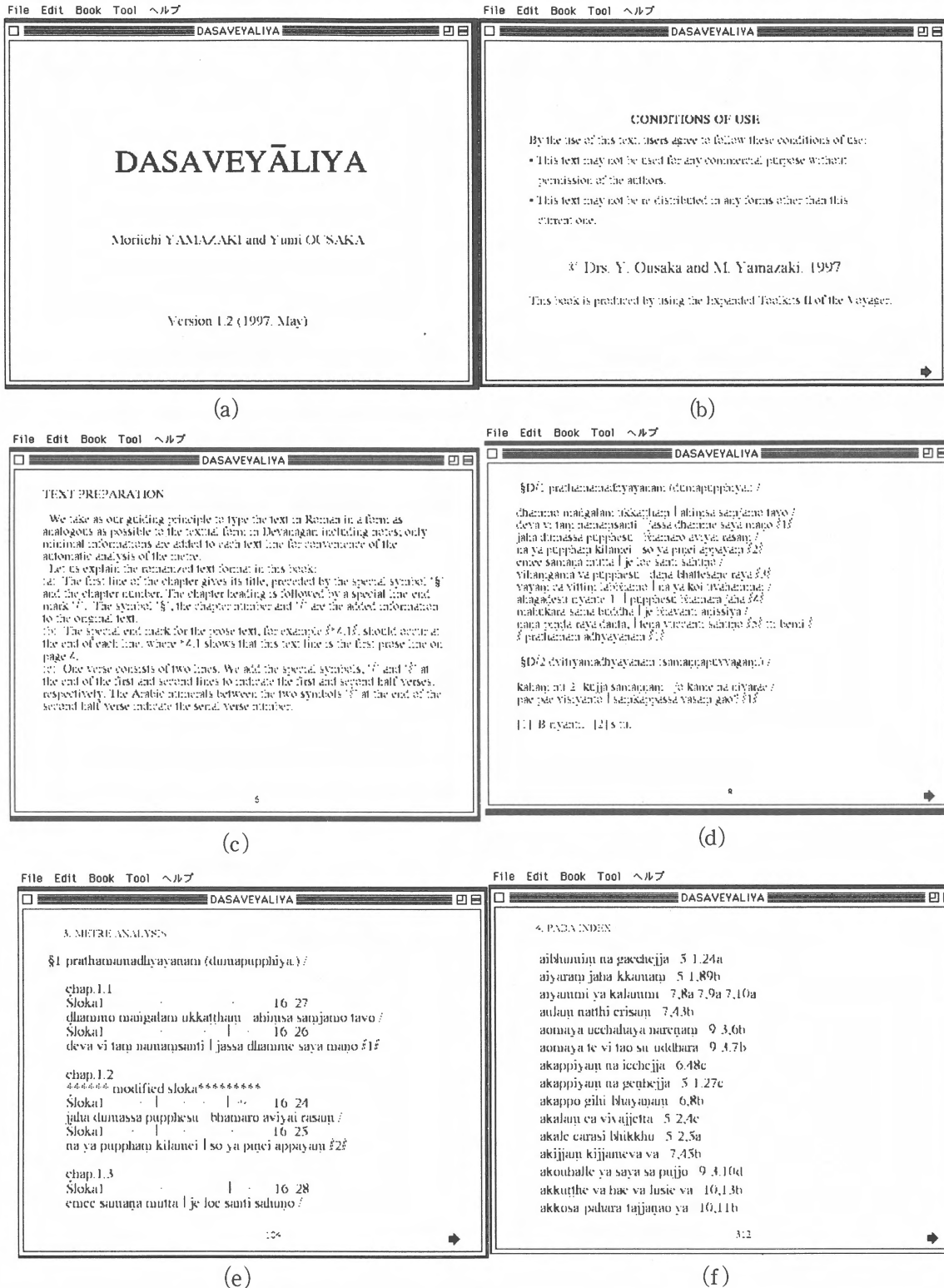


図2. マッキントッシュ版電子ブック

ンドーズマシンではAlt keyに置き換えているだけであり、Shift keyの使い方は全く同じである。また、アスキーコード割り当ては両方の計算機で同じにしている。このことにより、基

本的には、マッキントッシュで作成したテキストは、ウインドーズマシンでも同一コードで表現できることになる。実際、ウインドウズマシン用のジャイナ教ブレインテキストは、マッキ

ントッシュ用テキストをウインドウズマシン用に変換して、作成している。(マッキントッシュとウインドウズマシン用のフォントは共にTrue Typeフォントの形式で作成され、我々のウェブサイトにアップロードしてある。)

2.2 電子ブック

ジャイナ教の重要な5つのテキスト(Uttarajjhāyā, Āyāraṅga, Sūyagaḍa, Dasaveyāliya, Isibhāsiyāim) 及び仏教の4聖典(Dhammapada, Theragāthā, Therīgāthā, Suttanipāta)等をPali96フォントを使ってマッキントッシュに収録した。仏教の4聖典の電子化テキストはパーリ文献協会の前会長ノーマン氏より提供されたものであり、ノーマンフォントで作成されている。このテキストを我々のフォントで書き換え、索引作成に活用できるようにした。著作権に抵触するため、電子化テキストは残念ながら公表することができない。

一方、ジャイナ教の聖典には著作権に関する制限がないので、電子化テキスト(図2(d)参照)のみならず、韻律解析結果(図2(e)参照)・詩脚の正順索引(図2(f)参照)と逆順索引・語彙の正順索引と逆順索引を全て電子ブックに収録した。この電子ブックは、Voyagerの日本語版Expanded Bookツールを使用して作成している。このツールでは、文字指定を細かくできる。この指定は煩雑ではあるが、図2(a)より分かるように、特殊フォントも正常に表示されていることが分かる。細かいことを言うと、このツールを使った検索機能のダイアログボックスに特殊フォントを表示できず文字化けが起こる欠点がある。これは、この機能に対して通常フォントのみが使用できるようになっている、ことによる。しかし、基本的には殆ど問題ないと言える。

言語研究上必要な基礎資料がテキスト資料と共に、1冊の電子ブックにまとめられているので、非常に有効な電子化ツールである。この電子ブックもフォントと同様、我々のウェブサイトから自由にダウンロードして使用できる。

3. ウインドウズ版電子ブック

ウインドウズ版電子ブックを作成するために

は、これまで議論してきたように特殊フォント属性を保持する処理がまず第一に必要となる。これは、我々の作成した簡易エディタEDISANにより実行される。最初にこのエディタについて述べ、その後Acrobat4.0を使用して作成されるPDFについて議論する。

3.1 特殊フォントPali96用簡易エディタ, EDISAN

これまで、我々はプラークリット語・パーリ語・サンスクリット語用のフォント”Pali96”をマッキントッシュ上で作成してきた。このフォントは、拡張ASCIIコード上で定義されたフォントである。これを元に、中期インド・アリアン聖典の解析ツールの構築等をマッキントッシュを使用して行ってきた。

これらの研究ツールをウインドウズマシンでも使用できるように拡張中した。最初に、特殊フォントをウインドウズ用True Type Fontに拡張し、その後このフォントを使って作成されたテキストを移植した。

日本語版ウインドウズOSを搭載しているパーソナルコンピュータ上で、適当なワードプロセッサを使用してこのテキストを読みると文字化けを起し、正確に読むことができないと言う厄介な事態が発生した。これは、2バイトコードフォントを標準にする日本語ワードプロセッサには、特殊フォントを表示するのに不都合な点があることに起因する。

このバグを回避できたら、印刷と画面上の表示が一致する、所謂WYSIWYG(what you see is what you get)表示のテキストを利用できることになり、言語学的な研究の能率も向上し、1バイトの特殊フォントを基に作成された大事な基礎資料を有効に活用することができ、さらに、語彙の索引作成等のプログラム開発も容易にできるという、多くの利点がある。なお、我々のフォントと同種の特殊フォントとしてはNorman Fontがあるが、これに対しても同様な状態になっている。

ウインドウズマシン上では特殊フォント(Pali96 or Norman)の簡便な入力法がなく(入力法がない文字もある)不便であるということから、3年ほど前に、特殊フォント用の簡便なテキストエディタ(EDISAN1.0)を初めて開発し

た。その後改良を重ね、現在はEDISAN3.0である。

EDISAN 開発用のツールは特殊フォントを全て正しく表示できる VisualC++ を使用している。C++ 上での文字コードは普通のアルファベットには ASCII コード、日本語には SJIS コードを使っており、2 バイト文字である日本語を表示させたいときは、文字コードの前後に SJIS コードであることを示すコードを書くことになっている。つまり、VisualC++ では 2 バイトコードが 1 バイトコードよりきちんと分離されており、127～255 番目の拡張 ASCII コードを含めて、0～255 番目のすべてのコード上の 1 バイト文字を正常に表示できるようになっている。この結果、拡張 ASCII コード上に定義されている特殊フォントを VisualC++ では正常に表示できる。

マイクロソフト社より、ある程度のワープロ機能を備えた Word Pad のプログラムが公開されている。我々はこの Word Pad に、特殊フォント入力機能等の必要機能を付加した。この方法により、Word Pad の通常の機能を活かした上で、簡易エディタを迅速に作成できた。

図3にエディタの構成画面を示している。通常のワープロソフトの画面とほぼおなじであり、使用方法も殆ど同じである。このエディタを使用して、リッチテキスト形式のマッキントッシュファイルを読み込むことにより、ウインドウズ用の特殊フォントの属性を固定したファイルを作成できる。このエディタも我々のウェブサイトに掲載してある。

3.2 ウインドウズ用 PDF 作成

ウインドウズ用にも Voyager より Expanded Book ツールが発売されている。マッキントッシュ版で作成された電子ブックが、このツールを使って見ることができるならば、ウインドウズ用にわざわざ作成し直す必要がない。しかしながら、このウインドウズ用ツールでは特殊フォントを使ってマッキントッシュで作成された電子ブックを開くことができなかった。これはウインドウズ用ツールでは、拡張アスキーコード上に展開された特殊フォントの取り扱いが正確に定義されていないことによる。

これに代わる PDF 作成ツールも又これまで提供されていない。今年になり、Adobe 社よりウインドウズ版及びマッキントッシュ版 Acrobat 4.0 が提供された。両版のツールを購入し、チェックしてみた

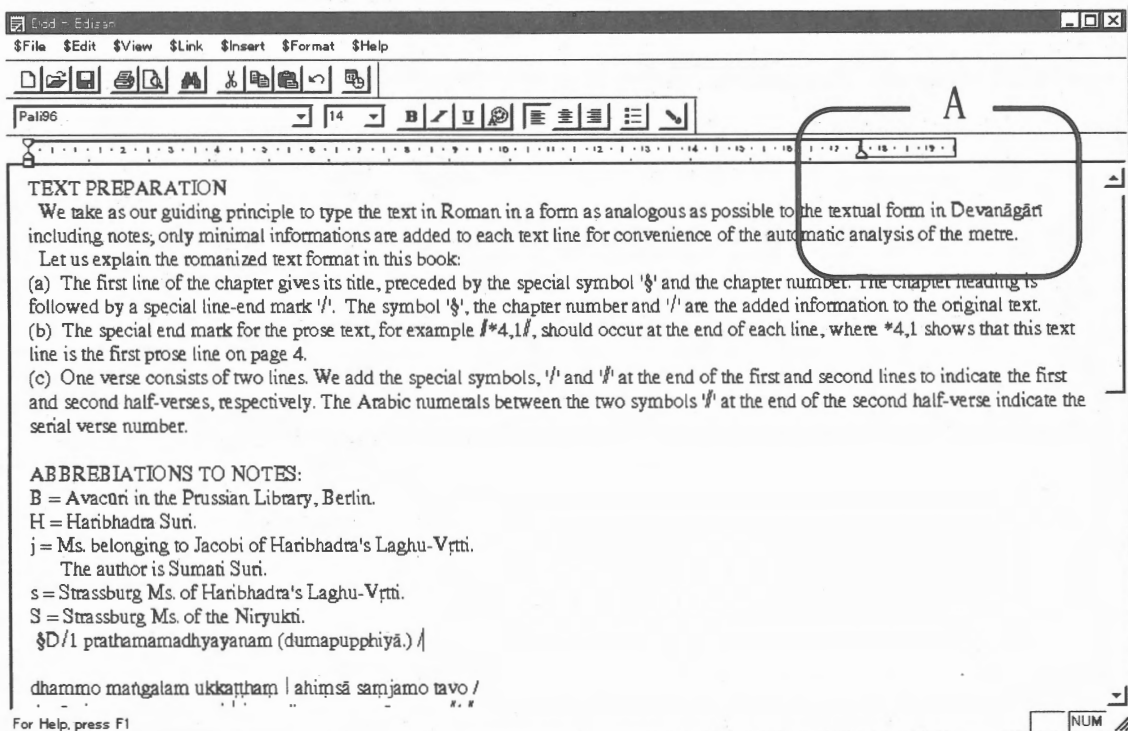


図3. 簡易エディタEDISAN3.0

ところ、大きな違いがあった。

マッキントッシュ用では、特殊フォントで作成されたファイルのPDFを作成したところ、画像として取り扱われてしまった。その結果、PDFのメモリーは数十倍も大きくなってしまった（メモリーが数十倍になったことから、この現象に気が付いたのであるが）。同じ操作を、ウインドウズ版でも確かめてみた。この場合は、正常にPDFを作成でき、ファイルサイズは数十%程度縮小できた。その結果は、不思議にもバージョン4.0のマッキントッシュ版Acrobatでも正常に読むことができた。

以上のように同一バージョンでも、マッキントッシュとウインドウズ版では極端な違いが現れた。Adobeに問い合わせたところ、どちらのAcrobatとも、特殊フォントを正確に取り扱えるようには設計されておらず、ウインドウズ版ではたまたまよい結果になったのであり、動作保証はできない、とのことであった。なお、マッキントッシュ版Acrobatの以前のバージョンでは、少なくとも画像化されることはなかったのに、新バージョンでは画像化されるようになった。

その後、ウインドウズ版Acrobatで作成されたPDFの動作を何種類かのそれぞれのマシンで調べたところ、何ら異常が現れなかった。それゆえ、このファイルを作成し、ウェブ上にアップロードして公開することにした。

特殊フォントで書かれたウインドウズ版ファイルのPDF作成は以下のように行う。①電子ブック作成用データ(図2に示された全てのファイル)はマッキントッシュ上で作成されて

いるが、全ファイルをMS-Wordのリッチテキスト形式で保存する。この操作は、マッキントッシュ上で行う。②次に、ウインドウズマシンでの操作を行う。①で作成されたファイルを、簡易エディタEdisan3.0で読む。そのとき、フォントをPal i96に指定しておく。その後、ファイルをそのままセーブする。この結果、拡張アスキーコード上のフォントの書式が正確に保持される。これは、前小節で議論したように、リッチテキスト形式のファイルを作成したことによる効果である。この後、適当なワードプロセッサを使ってPDF表示形式のファイル内容に整形して、PDFを作成すればよい。なお、この整形されたファイルでは各行末でのハイフネーションに注意する必要がある。このことについては次節で議論する。

図4はジャイナ教聖典 *Dasaveyāliya* のPDFを示している。図の左側のしおりより分かるように、電子化テキスト、韻律解析・詩脚索引・語彙索引等の必要な基礎資料を全て収録している。このしおりは目次の機能を果たしていると同時に、各項目へのジャンプ(リンク)機能をも兼ねている。この結果、このPDFは比較的使いやすいものとなっている。なお、詩脚の逆順索引表示においては、その性質から、詩脚部分の表示には後揃えタブが適し、ライン情報には前揃えタブが適している。図5から分かるように、このPDFではこのようなタブ揃えを実現している。マッキントッシュのExpanded Bookツールではその機能がないので、このようなタブ揃えは不可能である。又、図6は韻律解析結

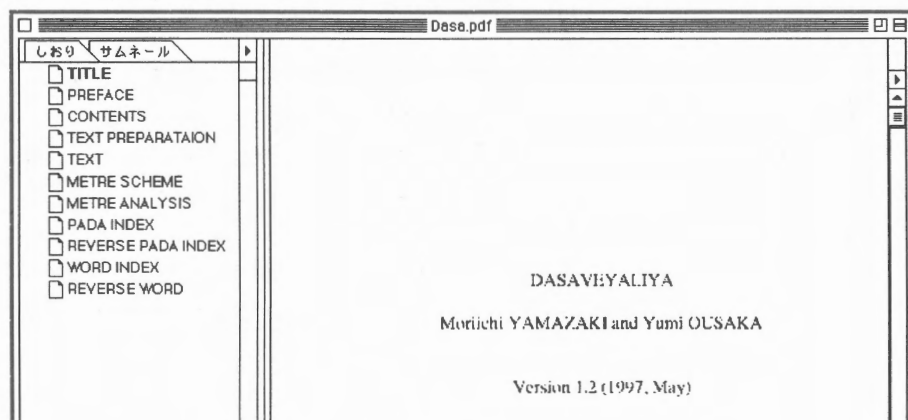


図4. ジャイナ教聖典のPDF例

果である。この例からも、特殊フォントが正常に表示されていることが分かるであろう。現在、ジャイナ教の重要な5つのテキスト(Uttarajjhāyā, Āyāraṅga, Sūyagaḍa, Dasaveyāliya, Isibhāsiyāim)のPDFを作成し、ウェブ上にアップロードしている。

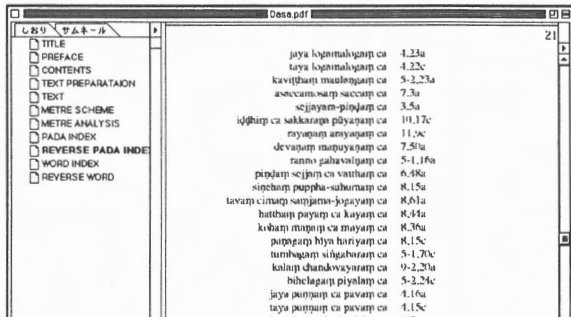


図5. PDFの詩脚索引例

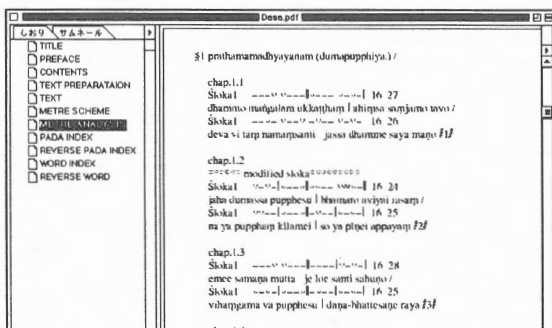


図6. PDFの韻律解析例

4. 討論

前節の最後から2番目の段落で、PDF作成直前のファイル表示において、行末でのハイフネーションに注意を要すると指摘しているが、これは本質的に、簡易エディタEdisan3.0の機能的欠陥に起因すると推測される。この現象は、既に図3に明示されている。図3の、Aで示される部分から分かるように、水平ルーラーの右端よりも更に右側にテキストが表示されている。これはテキストの1行分を、表示可能画面の1行分におさめる機能が正確に動作していないことを示す。この結果、恐らく、単語としての認識も正確に処理されないであろう。今

後、この欠陥の解消について考察する。

謝辞

この研究遂行に当たって適切な助言をしていただいた、山崎守一博士と宮尾正大教授(室蘭工大)に深く感謝いたします。この研究は2000年度文部省・科学研究費補助金、基盤研究(B)の助成の下に遂行したことを付記し、深謝致します。

参考文献

- 【1】 M. Yamazaki, Y. Ousaka and M. Miyao, Indexes to the Dhammapada, Pali Text Society, Oxford 1995.
- Y. Ousaka, M. Yamazaki and K.R. Norman, Index to the Vinaya-piṭaka, Pali Text Society, Oxford, 1996.
- M. Yamazaki, Y. Ousaka, K.R. Norman and M. Cone, Index to the Digha-nikāya, Pali Text Society, Oxford, 1997.
- M. Yamazaki and Y. Ousaka, Theragātha: pāda index and reverse pāda index, Monograph Series 12 of Philologica Asiatica, Chūō Academic Research Institute, Tokyo 1997.
- M. Yamazaki and Y. Ousaka, Therīgāthā: pāda index and reverse pāda index, Monograph Series 13 of Philologica Asiatica, Chūō Academic Research Institute, Tokyo 1998.
- M. Yamazaki and Y. Ousaka, Suttanipāta: pāda index and reverse pāda index, Monograph Series 14 of Philologica Asiatica, Chūō Academic Research Institute, Tokyo 1998.
- M. Yamazaki and Y. Ousaka, A Pāda Index and Reverse Pāda Index to Early Pāli Canonical Texts, Kosei Publishing Co., Tokyo 2000.
- 【2】 下記の索引はウェブ上に掲載してある (<http://www.sendai-ct.ac.jp/~ousaka>).
- Reverse word index to the Vinaya-pitaka, M. Yamazaki and Y. Ousaka (1997)
- Reverse word index to the Digha-nikaya, M. Yamazaki and Y. Ousaka (1997)
- Reverse word index to the Mahanidessa, Y. Ousaka, L.S. Cousins and M. Yamazaki (1999)
- Reverse word index to early Pali canonical texts, Y. Ousaka and M. Yamazaki (2000)