

橋梁技術史資料を例としたパソコンの独日文書ファイルの活用

岩手大学工学部建設環境工学科 ○宮本 裕 岩崎正二 戸秀明
東北大学大型計算機センター 小玉邦子
函館工業高等専門学校土木工学科 三浦 登
一関工業高等学校土木工学科 安彦敏郎

German-Japanese File for Personal Computer as example in Study of Bridge Engineering History
by

Yutaka MIYAMOTO, Shoji IWASAKI, Hideaki DETO, Kuniko KODAMA,
Noboru MIURA and Toshiro ABIKO

1. まえがき

最近では大学の研究室においてもパソコンの普及がめざましく、学生の教育や研究指導にも広く使われている。参考文献1)によれば、コンピュータのできる仕事として、①文献などの情報収集②関連分野の研究者との研究連絡③数値計算、シミュレーション、データ解析④画像処理⑤結果のグラフ化⑥論文、報告書の作成などがあげられる。

この研究では、橋梁に関する文化史の資料を、パソコンを使って翻訳を行ったが、パソコン利用翻訳の特徴と問題点およびその意義について述べる。

日本はこれまで外国から科学技術の影響を受けてきた。したがって外国の科学技術史の出版物を日本語に翻訳することも、日本人が日本の技術史を考え、将来の発展のための基礎をつくるためには、重要なことと考えられる。技術の適切なフィードバックが明日の技術を生み出すこともあるから。

ここで述べるパソコンによる情報のファイル化とその応用は、今後産学官のいろいろな部署でさかんになると思われる。

この論文の全体の構成は以下のようになっている。

1. OCR入力による独文テキストファイルの作成
2. 独日文書ファイルによる翻訳作業
3. 独日翻訳ファイルの活用と改良

2. OCR文字入力システム

最近、日本のあちこちでOCR (Optical Character Recognition: 光学式文字認識、Optical Character Reader: 光学式文字読取装置) による「印刷文字読取システム」が使われ、ある範囲の活字などの印刷文字が読み込まれ文書ファイル化されている。文字データを計算機の対象として扱うことは、ワープロなどの普及で身近なものとなってきたが、このOCR文字入力システムは文字入力を強力に支援するシステムとして位置づけられるものである。

著者らは東北大学大型計算機センターに設置されたOCR文字入力システムを利用して、キー入力を絶対的に越える大量の文字入力に成功した。

この入力システムを活用すれば、データベースの構築、電子出版、機械翻訳、コンピュータネットワーク通信などの計算機利用分野を飛躍的に広げるばかりでなく、言語学や文学の研究をはじめとして技術関係の資料の整備やその活用など、印刷物化されている多くの文化の研究に効力を発揮することが期待される。

文字読み取り機能の流れは図-1のようになる。まず活字やタイプライタ文字、プリンターのドット文字などで印刷されている原稿をいったんイメージデータとして読み込む。次にパターン認識技術を駆使して機械可読な文字コードに変換し、目的のファイルに出力させる。

和文入力システムはPC-Star Jr. により、欧文入力システムはKURZWEIL K-5200によりそれぞれ文字認識される。和文については問題はなかったが、我々の研究に必要な欧文入力システムでは、多数の欧米の言語の中からドイツ語を選んだため、ドイツ文字特有のウムラウト文字の問題があった。すなわちKURZWEIL K-5200で画面に正しく表示され、レーザプリンターに正しく出力されても、フロッピーの標準の設定がIBM-PC用(2DD)にフォーマットされているため、この独文ファイルはNECのPC98では読んでも文字化けしてしまっただ。これはドイツ文字が日本語コードの下では漢字コードと重なっているためであり、PC98に半角カタカナ部分の文字をユーザが自由に定義できるような機能さえあれば簡単に回避できるのであるが、パソコンもワークステーションも欧文対応には消極的である。

結局、このシステムを米国から輸入販売しているサカタインクス(株)情報電子機材事業部の鶴巻幸平氏に頼んで、IBM-PCのウムラウト文字をNEC-PC上で処理できるウムラウト文字に変換するプログラムを作成していただいた。この変換プログラムによりNECのPC98で表現された特別の文字列をワープロ一太郎で読んで、一太郎のソフト上で置換の機能により、一太郎でのウムラウト文字に変換することができた。

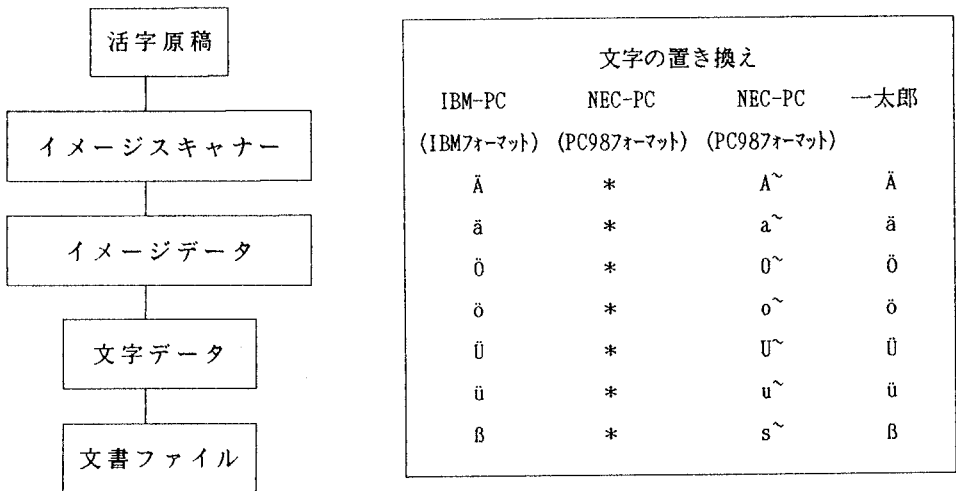


図-1 文字読み取り機能の流れ

図-2 ウムラウト文字の置き換え

(* 通常使わない文字に変換される)

3. ドイツ語文献「橋の文化史」についての意義

ここでとりあげる「橋の文化史」⁴⁾はミュンヘンにある世界的な科学技術史博物館であるドイツ博物館の、橋コーナーの展示の解説と橋の技術史をまとめた本である。この本の特徴としては、橋の技術の歴史だけを述べたものではなく、社会の発展とあいまって、その時代に要求された技術がどのように開発され発展していったかということが記述されている。また技術の発展が社会変革におよぼした影響も述べられている。

たとえばローマ帝国の維持と発展のためにはローマ帝国の領土内の道路の完備が必要条件であり、ローマ

帝国の道路建設にける情熱はなみなみならぬものがあった。ローマ帝国が道路の帝国とよばれるゆえんである。そして当時のローマ帝国の道路の特徴であるまっすぐ野を越え谷を越えて続く道路について考えるなら、橋を架ける技術がその道路機能を決定する主要因であったのである。

また別の例として中世の居城都市であり、遠隔地貿易都市であったレーゲンスブルクの発展の基礎となったレーゲンスブルクの石橋があげられる。レーゲンスブルクはドナウ川のほとりに位置し今でこそ静かな宗教都市にすぎないが、当時は中央ヨーロッパを南北に結ぶ道（ヴェネツィアからブラハ、キエフにいたる）とヨーロッパを東西に結ぶ道（フランスやベルギーからヴェルツブルクを通り、ウィーンからソフィアを経由してコンスタンチノープルにいたる）の交わる町がレーゲンスブルクであったのである。この町の発展のために多くの技術上の困難を克服して石橋を建設する経過は技術史の観点からも興味が深い。

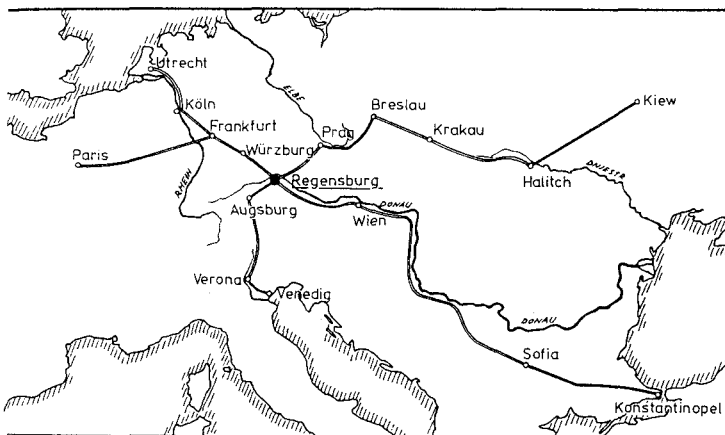


図-3 12世紀当時のレーゲンスブルク（原図：橋の文化史）

さらにもう一つの例をあげるなら、ドイツ鉄道のための橋梁建設における巨大アーチ橋の技術と経済上の問題がある。ザクセン政府がライプツィヒを鉄道交通の中心とすべく、ザクセンとバイエルンを結ぶ鉄道の建設の条約に調印したのは1841年のことであった。鉄道会社の株は投資家によりすぐ売り切れたのであるが、その後深い渓谷に架ける橋梁工事の困難さがわかってくと会社の資金が不足になり、ザクセン政府が資金援助をしなくてはならない状況になった。また橋建設のためにそれまで経験と勘にたよっていた技術を、本格的に数学を使って定量的に設計を行うことができるようになったのが、技術史上の一つのエポックである。

図-3のレーゲンスブルクが東西南北の交易の中心に位置していたように、ドイツは中世からその位置もヨーロッパの中心であって、ヨーロッパ文明をたどるときは決して無視できない地域であった。中世においては神聖ローマ帝国の範囲が現在のドイツの領土をはるかに上回る広さであったので、ドイツ語も現在よりも広い範囲で使われていた。したがってヨーロッパの技術史を研究する者にとってドイツ語の文献は必読のものであろう。現代においても、ドイツ統合により名実ともヨーロッパの指導的な位置につくであろうと思われるドイツと、ドイツ語文化はますます重要になると考えられる。

4. 文献「橋の文化史」に関する独日文書ファイルの作成

土木技術史の資料としての「橋の文化史」の活用をはかるために、原文のドイツ語とその訳である日本語の両方を併記した文書ファイルを作った。今回は主に翻訳作業の効率と正確さを高めるために、次の章で述べるような活用のしかたをはかった。

この論文で述べる翻訳のための文書ファイルの作成とその活用については、ドイツ語の場合はもちろんそ

他の外国語の翻訳の際にも役に立つものである。ただしコンピュータであつかうために、入出力のためのメディアの種類とかデータ構造などの条件および対応するソフトプログラムなどに制約を受ける。この研究はハードウェアとして日本電気製のPC98、ソフトウェアとしてはジャストシステム社のワープロソフト一太郎を使った。

5. 独日文書ファイルによる翻訳作業

このようなドイツ語の見直しの情勢の中で、土木技術史の資料として貴重と思われる「橋の文化史」を翻訳した。文書ファイルの作り方としては、原文のドイツ語の文章を段落ごとに入力してから、すぐその後に日本語訳を入力していった。

翻訳作業を能率化するために、この独日文書ファイルを次のように活用して翻訳作業をした。

①ある用語がどこで何度使われていたか数えることができる。

この検索機能は索引を作るときに強力な武器になる。この検索機能を使って、用語の場所と訳語のチェック作業を効率よく正確に行うことができた。これまでは人間が用語を拾いながら索引を作ってきたようであるが、機械によって作られた索引は人間より徹底しているため、より完成度が高いと思われる。

たとえば日本語訳では同じでも、ドイツ語が違う場合がある。石工（いしく）はSteinhauer とSteinmetzの場合があり、出てくる頁が違う。厳密さを追求するため、全部原書で確認し、索引も別項目で整理した。

②人間の記憶・整理能力の不完全を補う。

たとえばある用語を訳すとき、日時を改めた場合に訳語が変わっていることがある。実際に訳をすすめていくうちに理解が深まり、最初に考えていた訳語がより適切なものになっていることがある。この訳語の統一をはかるため、全体を訳してからもう一度全体の文章について最善の訳語に直すことができる。ちなみに翻訳専門家も、いちおう正しいであろう rough translation をまず作ることをすすめている。この後全体を眺めて順次翻訳の修正を行いながら、完全な訳に近づけることが望ましい手順であろう。

例として、くさび形石(Keilstein)とくさび石(Schlußstein)を訳していたとき、どちらも「くさび」という言葉を使っているため読者に誤解を与えるおそれが出てきた。くさび形石とはアーチを構成する1つ1つの石のことであり、アーチの円形を形作るため1つ1つの石はくさびの形をしているからそう呼ばれる。一方くさび石とはつまりキーストンのことであり、支保工の上にアーチを積み上げて最後にはめ込んでアーチ作用をとらせる石のことである。したがってここでは、くさび形石(Keilstein)と要石(くさび石、キーストン) Schlußstein というように訳しわけた。

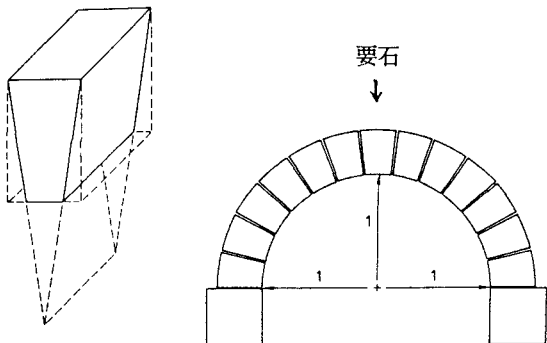


図-4 くさび形石と要石(原図:橋の文化史)

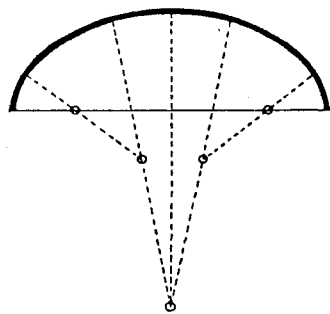


図-5 籠手型アーチ(原図:橋の文化史)

また別の例として、籠手型アーチ(Korbboogen)がある。半径の異なるいくつかの円弧アーチを組み合わせることができるアーチのことである。「多心アーチ」、「多中心アーチ」、「籠アーチ」、「籠手アーチ」とそのつ

ど訳していったが、アーチの形が籠 (Korb) の手の形をしているからこの名前をつけたことに気づき、「籠手型アーチ」に統一した。

もう一つの例としてガール橋を建設したローマ人たちが使ったクレーンについてであるが、簡単なウィンチ (Haspel) と重い石を持ち上げるための巻上機 (Göpel) の区別が不明で、訳も不統一でそれらに相当する言葉、クレーン、ウィンチ、巻き上げ機などを適宜用いていた。しかし他の疑問もあって著者との質疑応答の機会を得て著者から説明を受けた後で、簡単な機械はウィンチに、比較的高度な伝動装置を使った機械は巻上機にそれぞれ区別して訳することにした。

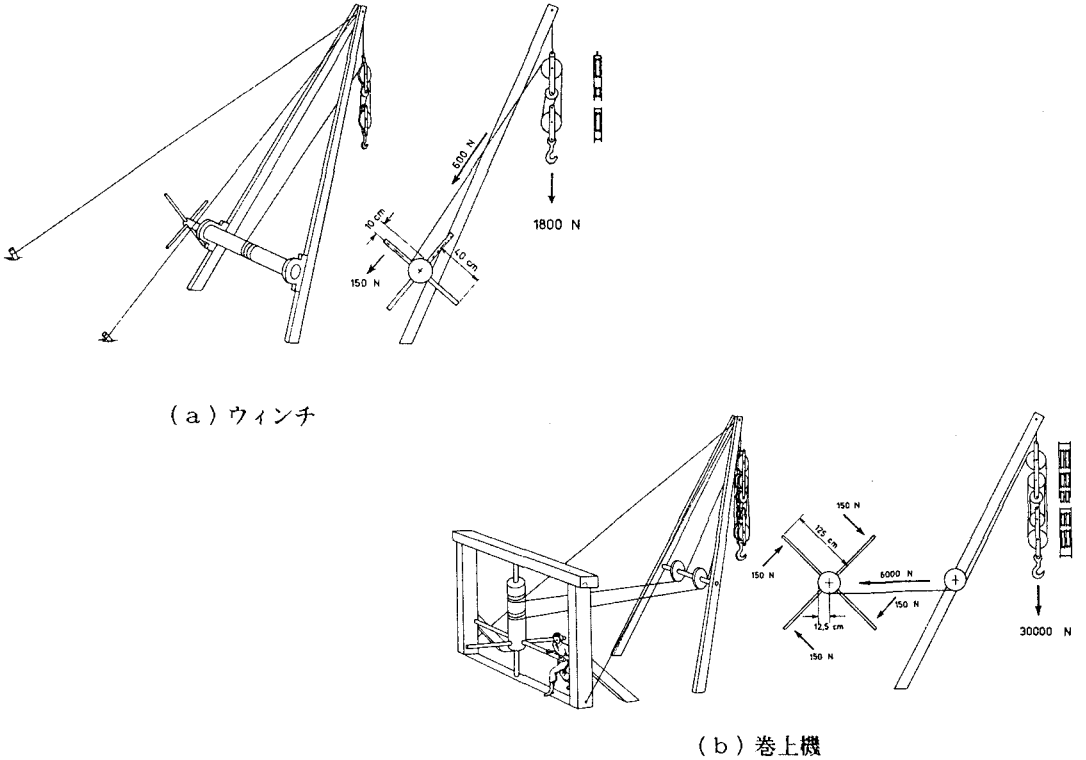


図-6 ウィンチと巻上機 (原図: 橋の文化史)

③自分用の辞書 (用語集) の補強

国内にあるドイツ語の辞書は一般に専門用語の記載が少ない。特に比較的新しい用語はなおさら載っていないことが多い。これは英語と比べてドイツ語の辞書は需要も少ない上に、辞書を作る立場の人間の数が絶対的に不足していることも原因であろう。この研究で述べるような専門用語集の努力の積み重ねが、より良い辞書の作成につながるものである。最近土木学会でも土木用語を解説した「土木用語大辞典」の出版作業を進めているが、英語の他にドイツ語もあつかう用語もあるので、各個人の研究メモの集大成といったことをしなくてはならないが、その作業を合理的に行うためにも各自の情報を活用するのが望ましいだろう。

パソコンで行う翻訳システムも最近ぼつぼつ見られるが、その際に自分用の辞書を作っておくことも作業をする上で重要である。このような辞書は同方面の研究者間で情報交換をすることも有益と考えられる。

なお実際に苦労して翻訳した例を上げると、ドイツ語の文中にときどき英語系 (おそらくラテン語から来たものであろう) の用語が外来語的に使われている場合がある。Polariskop (偏光スコープ)、Arena (円形競技場) などがそれである。

またドイツでは何でもない概念でも日本にいと全然理解できない概念がある。西洋文化を、西洋文化の内部で生活していない日本人が理解するのが困難であるのと同じことである。その例として、フランクフルトの町にある、馬車のための屋根のついた乗り場に、多くの人々が集まっている図があって、その図の説明に「Station der Fürstlich Thurn und Taxis'schen Fahrpost」というのがある。Fürstlich は「王侯の」という意味であり、Thurn und Taxis'schen は固有名詞のようでもあるが、我々の常識では歴史的にフランクフルトにそのような地名があるとは思われない。後で直接著者について聞いたところ、それはドイツ人なら子供でも知っていることで、1505年に Thurn und Taxis という名前の領主 (Fürst) が馬で手紙を運ぶ郵便制度を作り、時の皇帝から郵便業務をする権利を認められたことからきていることで、1505年から1866年まで彼の一族が郵便業務を独占していたが、その後ドイツの国有郵便制度に代わったというわけである。したがって上記の「Fürstlich Thurn und Taxis」という一族の名前は、郵便業務を表す言葉として使われていたので、この図の説明として「郵便馬車の乗り場」ということにした。⁵⁾

6. 独日翻訳ファイルの活用

このようにして作った独日文書ファイルの活用の方法として、具体的に以下のようなものが考えられる。

①ドイツ語と日本語の対応する構文集

和独つまりドイツ語論文を書くときに、それに対応する日本語の似たような表現を、この文書ファイルからさがせばよい。当然、独和翻訳の作業中にも、この文書ファイルをたぐって的確な日本語訳をつけることができる。

「Deshalb hat man es ein Imperium der Straßen genannt (Abb.42), und die Feststellung ist sicher nicht übertrieben, daß das großartige Straßennetz das riesige römische Reich entscheidend zusammengehalten hat.¹⁵⁾

「それゆえ人々はローマ帝国を道路の帝国とよんだ〔図42〕。大規模な道路網が巨大なローマ帝国をまとめるのに決定的な役割を果たしたといっても決して過言ではないのである。¹⁵⁾

この場合「ist sicher nicht übertrieben 決して過言ではない」という構文を利用する場合に便利である。

「Das Erstaunen des Reisenden und Fuhrmanns vor einer solchen Ingenieurleistung kann man sich nicht groß genug vorstellen. Für die Stadt und das Bistum bedeutete die Brücke neben den wirtschaftlichen Vorteilen den nicht hoch genug zu bewertenden Ausdruck eines Machtanspruchs.」

「旅行者や運送業者たちがこのような技術の成果を目の前にしてどのように驚いたか、その驚きを今日の人々はいくら想像しても想像しきれないということはないであろう。町と司教区にとって、その橋は経済的な利益という意味の他に、どんなに高く評価しても評価しすぎることはない彼等の権利要求の表現という意味もあわせて持っていたのである。」

この場合「nicht groß(hoch) genug いくら・・・しても・・・しすぎることはない」という構文を利用する場合に便利である。

「Er beobachtete gleichzeitig, wie das Wasser sich in Strömungen teilte, um das Hindernis zu umgehen.⁶⁰⁾

「同時にまた、水は障害物を回避するために二つの流れに分かれるということもわかった。⁶⁰⁾

この場合「um ... zu するために」という構文が使われている。

なお上記文章の最後の上付き文字の60は参考文献の番号であるが、出版社の都合で参考文献は割愛した。もし関心のある研究者がおられたなら、筆者らに問い合わせただけで資料をお教えしたい。

②原著者への質問や読者へのアフターサービス

著者らの翻訳した「橋の文化史」の中の2、3の語句について、名古屋大学名誉教授成岡昌夫博士から質

問を受けた。これに対して著者らは原文と訳文からなる独日ファイルを検索し、問題の語句が使用されている箇所をすべて検出し、翻訳作業の確認を完全に近い形ですることができた。そして成岡博士に疎漏なく対応することができた。

また翻訳出版後に読者からの問い合わせなどで、原著者との連絡の必要が出てきた。その際原著者の最新情報にもとづき、これらの独日文書ファイルの修正作業も必要におうじて処理しているが、それは合理的かつ迅速に行うことができた。

③引用文献付きの電子出版（FD とか CD-ROM）も出版社との間で了解がつけば可能である。

このようにしてできた独日文書ファイルは、他の研究者のための資料となる。すでに辞書や文学全集などのファイルをおさめたフロッピーディスクやCDが出版されている。これらの市販の電子出版物は膨大な時間と人手をかけて作成しているところに他の追従を許さない価値があるわけであるが、著者らが作成・活用している独日文書ファイルは、修正などができる手作りのよさと、自分の研究テーマに沿うという柔軟性があり、しかも利用しだいで市販の電子出版物と遜色ない活用がはかれるものである。

ただし現実には電子出版をパソコンレベルで考えた場合、容易にコピーできることから著作権という問題があり、まず研究者間での情報交換といった意味での活用をしたほうがよいようである。

7. ワープソフト一太郎について

ここで使ったワープソフトはジャストシステム社の「一太郎」である。そもそも著者らが独日併記の文書ファイルを作成しはじめたのは、6年前であった。その頃に独文の特殊文字（ウムラウトやエスツェット）を外字として使えるワープソフト「太郎」があった。ただしその頃の「太郎」では、ウムラウト文字やエスツェットとして、全角の外字文字しか使えなかった。それから「太郎」はいくつかの機能が追加されて名前も「一太郎」に変わった。その後著者らの希望に応え、ジャストシステム社は半角文字の特殊文字が使えるようにバージョンアップを行い、さらにそれらの半角の特殊文字を含む単語の検索置換という機能を開発した。これによってウムラウト文字を含む単語や構文を、容易に検索したり置換したりできるようになった。一太郎は1 M Byteのフロッピーディスクいっぱい文書ファイルも編集でき、多数のウィンドウによる複数文書の同時編集などの機能をはじめ、種々の文書作成のための有効な機能を持つソフトである。現在は光電社の「KOA-Techno Mate」も独日文書ファイル機能があるが、一太郎の持つ多くの機能は持たないようである（たとえばウィンドウを開いて3つ以上のファイルを同時にあつかえない）。

もちろん一太郎はパソコン雑誌に指摘されているように多くの欠点がある。あまりにも多くのユーザの希望に応えるため、いわゆる重いソフトになり、使ってみて応答が遅い。機能が多すぎてとても全部の機能は使いきれないし、また使う必要もない。

したがって、この研究のような独日併記文書ファイルを作成できる、そうメモリーを必要としない動作の速いワープソフトがあれば、そのほうが望ましい。いずれにせよ独日併記の文書ファイルが作成できるソフトなら何でも、この研究でのべたことは可能である。

8. 文書データの特徴について

本研究の独日併記文書データファイルの特徴と今後の問題点を箇条書にすると、以下のようになる。

①日本電気製PC98上のワープソフトを利用して作った。

②「橋の文化史」の独日ファイルは作業の便宜性を考え、全体を5つのファイル(128KB, 166KB, 103KB, 132KB, 139KB)に分けて保存されてある。また図の説明はこれとは別のファイル(128KB)として存在する。それぞれのファイルにおいて、各段落ごとにまずドイツ文その次に訳文としての日本文が交互にならべられている。収録されている用語の主なものには約300語ひろいあげられているが、用語全体ではこの数倍はあろう。構文の件数も相当あるわけであるが、これらのデータを効率よく利用する方法については今後の研究

課題となるであろう。

③ワープロソフトとして、一太郎でなくても半角のウムラウト文字(Ä Ö Ü ä ö ü)やエスツェット文字(ß)のあつかえるワープロソフトさえあれば、最小限本研究で述べたことはできる。

④ハードウェアの進歩が早いいため、長期の保存・活用に対してはメディアのライフサイクルを考慮して媒体変換をする作業が必要であろう(8インチから5インチそして3.5インチとフロッピーの主流が変わってきた。将来も媒体技術は進歩するであろう)。

⑤OSも変更される可能性があるので、ファイルの互換性を保つことが重要である。つまり重要なことは過去の遺産資料が将来も使用可能となるよう、ハードとソフト上で連続性をとれることである。

⑥OCR入力については、現在のところ学習機能を使わないと90数パーセントの認識率ではあるが、ときどき予期せぬ箇所で誤認識されるので注意が必要である。人間が機械の不備をカバーしないといけない。

9. あとがき

ここに述べたことは、独日文書ファイルの活用を考えたものであるが、データ構成を整えるなら、いわゆるデータベースとして利用することも可能である。

ここではパソコンのファイルに入ったデータをいかに使いこなすかという観点で、ワープロソフトさえあれば誰にでもできる使い方を紹介したものである。

またこの方式は日英、日仏、日露、日中、日朝など多くの組み合わせの文書ファイルにも適用できる。ただしそれぞれ必要に応じて外字が必要になる。著者らが述べたかったことは、独日文書ファイルの活用だけではなく、翻訳研究資料におけるパソコンの活用であり、情報処理技術を活用した文献研究一般のテクニックとそれによる研究発展システムの将来性である。

なお、共著者の仕事の分担は以下のとおりである。

宮本はこの研究の全体を企画し、原文のOCR入力と翻訳の作業を分担した。岩崎と出戸は独文ファイルの修正と和訳ファイルの編集を分担した。小玉は独文の文字データをOCR入力する際に、専門的立場でアドバイスをするとともにシステムの改良作業をした。²⁾³⁾ 三浦と安彦は橋梁工学および工学技術史の専門用語を調べ、必要な情報を提供した。

「橋の文化史」については、独文ファイルは宮本らにより大部分はキーボード入力をされたが、OCR入力装置が東北大学大型計算機センターに設置されてから、残りの独文データはOCRから読まれ、いくつかのドイツ文字をあつかう際の問題点が解決されていった。したがって次のドイツ語文献の翻訳作業からはより効率の良いこの一連のシステムが適用されることが期待される。

参考文献

1. 佐宗哲郎、理学研究者のためのコンピュータ環境、東北大学大型計算機センター広報 SENAC、Vol.24 No.1、1991.1
2. 小玉邦子・高橋洋一・加藤昇・高橋秀知、印刷文字読取システムについて、東北大学大型計算機センター広報 SENAC、Vol.24 No.2、1991.4
3. 小玉邦子・孫寧・根元義章、印刷文字読取システムの構築と認識結果の検証、第13回全国共同利用大型計算機センター研究開発論文集、1991年10月、pp.16-21
4. Bert Heinrich、宮本裕・小林英信共訳、『橋の文化史』、鹿島出版会(1991.6.5)
5. 宮本裕、ドイツ漫遊記、KTS・COMMUNICATIONS Vol.14、川田テクノシステム、平成3年5月
6. ブックレビュー、土木学会誌、1992年3月号 pp58
7. 宮本裕・岩崎正二・出戸秀明、パソコン文書ファイルを使った翻訳研究の能率化、土木学会第47回年次学術講演会、平成4年9月