

『新潮45』 1992年12月号掲載

古瀬幸広著

ワープロ一号機に賭けた男の戦争

ほんの十数年前まで、タイプライターのように簡単に日本語が打てる機械など夢のまた夢の話だった。その魔法のような製品を創り出したサラリーマン技術者の苦闘の歴史はあまりにも知られていない。

1

誰も信じてくれない製品

キーボードから仮名文字を入力し、変換キーを押すだけで漢字を含んだ日本語文書を簡単に作成できる装置。それが日本語ワードプロセッサ（ワープロ）である。この画期的な機械が初めて世の中に姿を現したのは昭和五三年九月のことであった。開発に成功したのは東芝である。

ワープロの革新性は、漢字が入力できる装置でありながら、両手を使ったタイピングが可能な点にある。それを実現したのはコンピュータを使った「仮名漢字変換」という技術であった。この技術の歴史は意外なほど古く、昭和三〇年代末まで遡れる。九州大学工学部電子工学科の栗原俊彦教授が、沖電気との共同研究として昭和三八年頃から手をつけていたのである。

しかし彼は、志半ばにして五〇歳の若さでこの世を去ってしまった。死因は心筋梗塞。昭和四八年一月四日のことである。そのあとを受け継ぐように、仮名漢字変換の研究開発に手をつけ、実用化に成功した企業が東芝なのである。

その道のりは決して平坦なものではなく、製品発表の直前まで問題を抱えていた。昭和五三年の春も終わりという頃のことだ。東京都青梅市にある東芝青梅工場では、もう呆れた、という表情で児玉皓次が愚痴を並べていた。

「どうもうまくない。誰もこの製品の価値をわかってくれないんだ。まったく、我が社は石頭ばかりで困る」

この製品とは、数カ月後にJW・10という名で発表された初の日本語ワードプロセッサのことである。児玉は話を続けた。

「どうだろう。いっそのこと完成してから見せては？　そうでもしないと、彼らには理解してもらえないよ」

2

「それはうまくない。「暴走だ」なんていわれてはつまらないし、逆につぶされる危険もある。なんとかわかってもらえる方法を考えよう」

返事をしたのは東芝総合研究所の森健一主任研究員である。「これまでの常識を覆す製品なんだから、口でいくら言ってもわかってもらえないのは当然。それほど凄い製品なんだ、というくらいに考えればいいじゃないか」ともつけ加えた。

東芝青梅工場に装置開発担当課長として勤務する児玉は、上司の溝口哲也ハードウェア開発部長とともに、研究所の森たちが完成させた新技術の商品化に携わっているところだった。既にプロトタイプも根幹部分は動いている。しかし「仮名漢字変換」と「日本語ワードプロセッサ」という聞き慣れない二つの言葉に、周囲は拒絶反応を起こしていた。

「キーボードから仮名を打ち込むだけで、自動的に漢字になるんですよ」

「誰でも手軽に素早く、日本語の美しい文書を作れるんです」

商品企画から営業にいたるまで、関係者を説得するために費やした言葉は数知れない。そのたびに、

「ワードプロセッサ？ いったいそれを誰が使うんだ。企業には文書課があり、邦文タイプストがいるじゃないか。いまさら違う方式のものを売ってもダメだよ」

「自分でキーボードを打って文書を作る人なんているのかい。だいたい手で書いたほうが早いじゃないか」

といった分厚い常識の壁に跳ね返されていたのである。

「こうなれば、現物を見せて納得してもらおうしかなさそうだな」森が口を開いた。彼は一見すると物静かな雰囲気だが、じつは勝気で闘志を内に秘めるタイプである。「速度競争をやろうじゃないか。都築事業部長を落とせばいいんだ。彼は頑固だけれど論理的な人だから、目の前で手書きより速く文書を作れるところを見せればわかってくれるだろう」

「なるほど」児玉も乗り気だ。「だったら、これまでまったく計算機とか邦文タイプライタなどに縁のなかった人間に実験台になってもらうほうがいいね」

「総務課の女性にキーボード操作を覚えてもらって、手書きと競争するというところでどうだろう。猶予はそうだな、一ヶ月だ。まったくの素人が一月の練習で手書きに勝てれば、この製品の価値をわかってもらえるはずだ」

昭和五三年七月、都築公男電算機事業部長が商品企画や営業の責任者とともに青梅工場にやってきた。総勢三〇名という大人数である。勝負をかけるにふさわしい舞台となった。

JW・10の製品化計画が持ち上がったのは、その前年の昭和五二年の秋のことである。たった半年間でここまでこぎつけたことになる。これほどの短期間で開発が進んだ背景には、連綿と続く仮名漢字変換研究の歴史と、何年にもわたる試行錯誤があった。壁を既に突破していたから、商品化がスムーズだったのである。

その「壁」とは、仮名漢字変換を使った日本語入力方式の実用化である。それを突破したチームは三人。中心人物は、川崎市幸区にある東芝中央研究所（後に東芝総合研究所、東芝研究開発センターと名称変更）で、文字認識の研究開発に携わっていた森健一である。

森は昭和一三年に東京都大田区で生まれた。少年時代に空襲を体験した世代である。遊び場所は大岡山の東京工業大学のキャンパス付近だったという。終戦後は疎開先の群馬県から東京に戻り、麻布高校から東京大学工学部応用物理学科へと進んだ。ちようど、東京大学理学部の高橋秀俊教授ら大学の学者が先頭に立ち、コンピュータで先行するアメリカをトランジスタとパラメトロンで日本が追い上げ始めた時期に学生時代を過ごしたことになる。

昭和三七年に東京芝浦電気に入社した森は、中央研究所の配属となり、コンピュータの磁気メモリーの研究開発に回された。

「でも、論文を調べると磁気メモリーはもうほとんどやり尽くされている分野だったので、面白くなかった。新しいことがやりたかった」

と森はいう。ほどなく、「文字認識を研究テーマにしたい」と上司に申し出た。彼は「新しいこと」にこだわる研究者向きの性格と、「ものを作る」ことに意欲を燃やす技術者向きのセンスの両方をあわせ持つ、いわば「研究エンジニア」であった。東芝入社後、一貫してコンピュータそのものではなく、コンピュータを使った応用製品の研究開発に携わって

いることからそれが窺える。

文字認識に取り組んでからの彼の最初の大きな成果は、上司の玄地宏（故人）の下で成功した昭和四一年の郵便番号自動読取装置の開発であった。文字認識の研究成果である。文字認識とはコンピュータに文字を読ませることをいう。これが意外なほど難しい。昭和三〇年代から四〇年代にかけては、文字認識が人工知能の最先端だったのである。

いまでこそ「人工知能」といえば、コンピュータに推論させるといった高度な処理を指しているが、そもそも「人間にしかできないこと」をコンピュータで実現することを人工知能という。人間は字が大きくても小さくても、多少崩れていてもそれを「字」として認識できるが、当時のコンピュータにとっては、対象を数字に限っても難問題だった。

森によれば、「最後まで悩んだのが、続き書きされた数字の認識だった」という。きちんと1文字ずつ書かれたものなら読み取れるところまでこぎつけたが、「100」などを続けて書かれるとうまくいかない。結局これは放棄し、代わりに独立した枠を封筒や葉書に印刷することにした。「枠を印刷しておけば、心理的に続き書きしにくくなる」と考えたのである。

数字が認識できれば、次はアルファベット、片仮名と認識対象を拡大していく。森は昭和四一年から六年間の予定でスタートした通産省工業技術院の大型プロジェクトで、東芝が分担した文字認識装置の開発に携わっていた。

このプロジェクトは「超高性能電子計算機計画」と名づけられたもので、予算は約百億円。目的は「我が国のこの分野の技術開発力の総力を結集した電子計算機技術の開発」である。具体的にいえば「IBMの最新鋭コンピュータに対抗できる国産コンピュータを作ろう」という官民一体の研究開発であった。

プロジェクトも終わりに近づいた昭和四十六年のことである。森は文字認識のさらに先を考えていた。コンピュータが文字を認識できるようになれば、次は言語を処理できることが望まれるだろう。たとえば機械翻訳が考えられる。そこまで高級ではなくても、日本語をコンピュータで処理できるようになれば、それだけで印刷会社や新聞社などに売り込むシステムが作れる。

彼は気になって調べてみた。意外なほど、日本語をコンピュータ処理するシステムは開発されていない。様々な問題が残されているようだ。同年三月三日付で森が書いた「日本語の機械処理システムについて」と題した書類に、それがまとめられている。

「日本語を機械処理することに対する潜在的な需要が大きいにもかかわらず、この発達をばんできた原因は次の点に要約される」という書き出しで始まり、「イ．入力装置の問題。ロ．出力装置の問題。ハ．システム上の問題」の三点について整理したものだ。

興味をひくのはイの入力装置の項である。その問題点として、次の五点が列挙されている。

- (1) 使用する漢字の標準化（コード、順序）が困難。
- (2) 字種が最低二千字は必要であるため鍵盤の数が多くなり、入力速度が極めて遅い。
- (3) オペレータのミスによる誤入力率が高い。
- (4) 漢字は正書法が一定せず、同音異義語が多いため、カナタイプを用いた入力装置の開発（かな漢字変換）が困難。
- (5) いずれの方法でもオペレータの訓練が長期間必要。

少し長くなるが、ワープロ開発の背景を知るためにも順に説明していこう。まず、コンピュータで文字を扱うには、文字にコードを割り当てなくてはならない。アメリカではA～Zまでの大文字と小文字のそれぞれに、7桁の二進数（7ビット）で構成される文字コードが割り当てられ、規格化されている。

同じように漢字にもコードを用意すればいいのだが、それが一筋縄ではいかないというのが(1)である。ちよつとした漢和辞典でも一万字、『諸橋大漢和辞典』には五万字超の漢字が収録されている。このうち、いったいどれだけの字を選びだせば、日本人の言語生活に必要な十分なのかはつきりしていない（これはいまだにそうである）。

当用漢字だけに絞るといふ思い切った手もある。しかしそれでは、人名、地名などの固有名詞をカバーできず、とても新聞社や自治体に売り込めるものとはならないだろう。さらに、漢字を選びだしたとしても、複数の読みがあるから、順序もすんなりとは決まらな

い(ちなみに現在の J I S 漢字コードは、7ビットの倍の14ビットを使い、使用頻度の高い二九六五字を第一水準としてその音読み順に並べ、旧字体や頻度の低い漢字三三九〇字を第二水準とし、漢和辞典の並びにしてコード化している)。

そもそも漢字は数が多い。基本的なものに絞りこんでも二千字は必要だ。だから装置がどうしても大きくなり、操作も遅くなる。邦文タイプライタがいい例である。タイプライタとは名ばかりで、実際にはタイピングのできないものだった。これが(2)である。

邦文タイプライタより能率のいいものとして、漢テレ鍵盤(漢字テレックス鍵盤の略)もあつた。これは大がかりな邦文タイプライタに通信機能をつけたものであり、専任オペレータが両手両足を使って操作する。

キーボードには六〇〇強のキーが並んでおり、その一つ一つのキーに漢字や平仮名などが四字割り当てられている。四字のうちどれを選択するかは、両足のペダル操作による。

①両足とも踏まない、②左足のペダルを踏む、③右足のペダルを踏む、④両足のペダルを踏む、で四種類の選択ができるというわけだ。

両手両足を使うので、熟練が必要だが邦文タイプライタよりは早い。しかしそのぶん、ミスも増える。これが(3)の指摘である。英文タイピストの誤入力率は0・1パーセントに過ぎないのに、漢テレ鍵盤ではオペレータの誤入力率が五パーセントにも達していた。両手両足を使う弊害だといえるだろう。

それ以外の方法としては、九大の栗原が手をつけた仮名漢字変換が考えられるが、日本語はそもそも正書法が定まっていないから難しい。まったくいい加減な言語だ。「いい加減」とも「いいかげん」とも書くし、同音異義語が厄介で、「いい下弦」と変換してしまう可能性もある。

そして、漢テレ鍵盤式にせよ、仮名漢字変換式にせよ、オペレータを長期間訓練して育てる必要があるのが大きな問題だった。その後、ワープロを開発することになる人間が、仮名漢字変換までもその中に含めているのは矛盾しているようだが、当時は大型コンピュータを使うことが常識だったことを考えれば納得できるだろう。

人目を忍んで国内留学する

「とてもできない」という常識の裏にこそ、開発のチャンスはある。森はこの分野の研究開発に目をつけた。しかし、「当時の東芝には自然言語処理に詳しい研究者が一人もいなかった」という。それも当然の話で、当時は日本全国を見渡しても、コンピュータを使った自然言語処理に手をつけている企業は稀だったし、大学でも本格的に研究しているのは、九州大学の栗原俊彦教授と京都大学の長尾真助教授の二人くらいのものであったのである。まだそういう時代だった。

森はまず、人材を育成するところから始めることを考えた。白羽の矢が立ったのは九州

大学院修士課程を修了し、昭和四六年に入社したばかりの部下・河田勉である。彼を選んだのは、栗原教授がいた九州大学工学部電子工学科の出身であり、栗原の研究内容についても知っていたかららしい。

「そもそも入社して最初に森さんに会ったとき、「君は九州大学の出身だったね。だったら栗原先生の研究は知ってるでしょう」と声をかけられたんです。驚きました」

と河田はいう。栗原は日本語処理の先駆者であり、天才と賞される豊かな発想と、奇人といわれる行動力を持った学者であった。

「もちろん栗原先生の研究は知っていました。「考え事をしながら身支度をしたため、ネクタイをうっかり二本締めて大学に来た」とか、嘘のような伝説の尽きない方でしたしね」

という河田は、昭和二年に北九州市で生まれ、小倉高校から九州大学工学部電子工学科に進んだ経歴を持つ。が、当時の栗原はほとんど授業をやっておらず、また大学院では回路工学の古賀利郎教授についていたので、栗原の直系の弟子というわけではない。

「少しは知ってるんだから、なんとかなるだろう」というのも乱暴な話だが、彼は突然「これから一年間、京大の長尾先生のところで勉強してこい」と言われることになった。昭和四七年初頭のことである。大学への国内留学だ。

「東芝では前例がなく、少しモメた」そうだが、森は上司を説得し許可をとった。なにしろ日本語処理の研究である。海外に留学先があるわけではない。国内留学が人材育成の唯

一の、そして最短の方法なのであった。

留学先の候補としては、九大の栗原研究室もある。しかし、栗原の仮名漢字変換研究は、もともと沖電気との共同研究としてスタートしたものだ。そこに東芝が後から割り込むわけにはいかない。残る選択肢は京大なのであった。

「長尾先生と京都大学がよく受け入れてくれたと、いまでも感謝しています」

と河田。彼が留学した昭和四七年は、東大紛争からまだ三年。学生運動の余波が強く残っている時代である。国内留学の試みも「産学協同反対！」というヒステリックな反応の渦に巻き込まれる可能性があった。だから河田は「念のため名簿上に名前を掲載しなかった」し、そのため図書館も使えずに「毎日、教官用の入口から入る」という生活を続けていたのである。

京都大学での河田の研究テーマは日本語の構文解析だった。

「栗原先生たちがソフト的な理想論を追求していたのに対して、長尾先生はかなり工学寄りの方でした。研究テーマにしたのは、国語の教科書をコンピュータに理解させる試みです。将来的には論文などの詳録を自動的に作るシステムを作りたいかった。でもコンピュータに肝腎の漢字が入力できないので随分困った。留学の最後に、形態素解析の論文を書きました」

彼はまさにその後、自分自身が解決することになる問題で苦勞しながら、一年間を過ご

していた。形態素 (morpheme) とは「意味を持つ最小の言語単位」を指す。文を細かく切っていったときに、それ以上切ってしまうと意味がなくなる、という単位のことである (さらに細分すると音素になる)。

仮名漢字変換は形態素解析の応用である。が、この段階ではまだ、自分が学んだ形態素処理技術がその後、ワープロ開発という形で活かされることになるとは河田は思っていない。まったくの基礎研究の段階なのだ。

「正直いって、この頃は「仮名漢字変換」というフレーズを耳にしても、現在のような形でもものになるとはまったく思っていなかった。直観的にも、「できる」という感じはまったくしなかったんですね。まして自分たちがやることになるとは夢にも……」

という河田。彼の国内留学は、自身にとっても森にとっても、いまだはつきりとした目的のない旅なのであった。

しかしその最中に、河田は長尾が師事した坂井利之教授の研究室で、新しく旅のパートナーとなる人物に出会った。修士二年で文字認識の人工知能化などの研究を行っていた天野真家である。当時の天野は河田と二つ違いの二四歳。名古屋出身の彼は、名古屋工業大学工学部電気工学科を昭和四六年に卒業し、京都大学大学院工学研究科に進学。そこで坂井教授についていた。

「坂井先生と長尾先生の『文字認識』(共立出版)を読んだのが、坂井研に行きたいと思っ

たきっかけだったのだけれど、文字認識は既にメーカー各社が本格的に開発を進めていた。こうなると人と金のリソース面で大学は勝てないので興味が薄れてしまっただけ。自然言語処理をやるうかなと思っただけです。ちょうどそこに、河田さんがやってきた」

と天野。河田と出会った彼は、その夏、東芝に就職が内定した。「就職試験に通ったら、希望の部署に配属させてもらえる「配属予約」という慣習があり、確実に自然言語処理研究に携われるのが魅力だった」のだという。

まだほとんどの企業が自然言語処理には手をつけておらず、迂闊に就職するとまったく関係のない分野に回される危険がある。天野はあくまで自分の興味にこだわった結果、就職先に東芝を選ぶことになり、翌昭和四八年四月から河田の同僚となった。

森と河田、それに天野という主人公の三人がこれで揃ったことになる。三五歳の研究開発リーダーに、二七歳と二五歳の部下という若いチームだった。

じつにいい加減な日本語の文法

いよいよ東芝が本格的に研究開発を開始する前に、どうしても触れておかななくてはならない業績がある。NHK総合技術研究所(現NHK放送技術研究所)の取り組みだ。話は四年前の昭和四四年に遡る。

当時のNHKは、海外特派員からのニュースの受け取り場面で問題を抱えていた。特派

員はテレックスを使い、ローマ字書きでニュース原稿を送信してくる。これが読みづらく、理解するまで時間がかかる。臨時ニュースでは一分一秒を争うだけにもどかしい。そこで、「コンピュータと仮名漢字変換を使って、読みやすくできないか」

という話がNHK総合技術研究所で持ち上がったのである。仮名漢字変換に目をつけたのは、昭和三十八年に京都大学工学部を卒業して入所した若手研究員・相沢輝昭である。

相沢にはこの研究に取り組むだけの素地があった。彼の卒業論文は『計算機による仮名漢字変換』というもので、九大の栗原と同時期にまったく同じテーマに興味を持っていた。

指導教官は、河田も師事した長尾であった。相沢ははにかみながら、

「タイトルは立派ですが、至極単純な論理の仮名漢字変換システムで、実験用に百語ほどの辞書を作っただけ。しかもデバッグしきれなくて動かなかった」という。

デバッグとはコンピュータのプログラム上のミス（バグ）を修正することである。卒業論文という限られた枠内での取り組みに過ぎず、なおかつ未完成なままであったが、この経験が数年後に生かされることになる。

なにより仮名漢字変換システムの試作に取り組む意欲を持っていたこと自体がそうである。昭和四〇年代に入ると、専門家の間に「仮名漢字変換」という言葉も概念も知られるようになった。しかし、大勢の反応は鈍く、ほとんど無視されていたのである。

コンピュータで私たちが用いている言語を処理する「自然言語処理」の分野では、今も

昔も機械翻訳が花形だ。険しい山ほど登る価値があると考えるのは当然のことだろう。機械翻訳こそが真冬のエベレストなのである。

ここまではとてもよく理解できる。問題はその一方で、形態素解析の応用である仮名漢字変換はただひたすら力技を必要とするだけの砂漠で、学問的には面白味がまったくなく、やってもとても実用にならないと考えられていたことだ。

いったん入り込むと、なかなか出口が見えないのが形態素解析である。とくに仮名漢字変換となると、変換の頼りにするべき日本語の文法がじつにいい加減であるうえ、基礎的な資料がほとんどないという貧しい状況なのだ。

仮名漢字変換の基本中の基本となる変換用の辞書作りでも苦勞が多い。私たち素人は「国語辞典をそのままコンピュータ化すればいいではないか」と簡単に考えてしまいがちだ。いまでも、「まったくこのワープロは馬鹿で全然変換しない。『広辞苑』でも一冊まるごとワープロ辞書にすればいいのに」

という文句を聞くこともある。ところが、あれほど分厚い『広辞苑』でも、ワープロ辞書として見ると大きな不備がある。国語辞典は「難しく意味のわからない言葉を引く」ためのものであり、「小学生」のような説明不要の語彙が収録されていない。日本人の暗黙の了解を得られる言葉は無視されている。

ワープロ辞書は日本人にとって、まさに「書くための辞書」である。だから「小学生」

も「原付」も放りこんでおかないと、「こんな簡単な言葉も変換しないのか」と文句を言われてしまう。人名・地名など固有名詞も補わなくてはならない。同じ「五万語」でも、ワープロ辞書と国語辞典とは中身が大きく異なっているのだ。

国語辞典を穴のあくほど見つめても、正しい候補を選ぶ際の鍵となる接続情報がほとんど掲載されていないことも悩みの種だった。これも国語における暗黙の了解事項とされているものの一つである。品詞を頼りに「はしらない」を「走ら+ない」(動詞「走る」の未然形+助動詞「ない」と分析できても、「はしらないことになった」となると、「ことに+なった」が接続可能かどうかという資料がないのである。「ない」がどんなものに接続可能なのかの一覧表などが欲しいところだ。これがあれば、「柱ない」が誤りであることもすぐにわかる。

「国文法の書物を見ると、部分的には接続情報が出ていた。でもほんの一部だけ。結局、栗原先生以来、仮名漢字変換の研究者は何と何が接続できるのか、という接続表に取り組んできたようなものです」

と天野はいう。さらに用例情報が充実していれば、「寒心、感心、関心、歎心」という四つの候補があったとしても相当絞り込んで変換できる。「関心を持つ」とは言うが、「関心を買う」とは言わない。こういった情報も、国語辞典に網羅されているわけではないのだ。

相沢は卒論時の経験からこの問題をよくわかっており、九大の栗原が集めたデータに目

をつけていた。栗原は閃きを次々に形にしていく猪突猛進型の学者である。「工学部の教授が、国語学者の真似事をするとは何たることか」といった陰口などまったく意に介さず、文学部生をアルバイトに使い、自分たちで接続表のもとになる基礎資料を集め、変換用単語辞書として蓄積していた。

それを現在もお受け継いでいる稲永紘之は、

「ある日栗原研究室に顔を出すと、アルバイトの女子学生たちが「だらけ」のつく言葉を必死にリストアップしていたんですよ」

と話す。部屋の全員が、必死に「椅子だらけ」だの「男だらけ」だのと口にしながら、〇×をつけている。『広辞苑』でも見出し語として拾われているのは「灰だらけ」「血だらけ」「泥だらけ」のわずか三例である。では、「灰、血、泥」にしか「だらけ」が接続しないかというところではない。ほかの語にも接続するのだ。こういった用例データも栗原は集めていた。

試作の了解をとった相沢は、沖電気の黒崎悦明に連絡をとった。沖電気と九大が共同で仮名漢字変換の研究を行うにあたって九大に派遣され、日本初の仮名漢字変換についての論文を栗原と共同で書き上げた人物である。

「形式上は沖電気さんとうちとの共同研究という形にして、データを磁気テープでいただけないでしょうか」と相沢はもちかけた。

「わかりました。お受けしましょう」

と黒崎。話はいまから思うと意外なほどスムーズだ。仮名漢字変換の研究者同士の、組織の壁を超えた連帯感があった。さらにつけ加えるならば、誰もが「仮名漢字変換はすぐに実用化できるものではない」と考えていたから、鷹揚だったということもある。

相沢は栗原たちが蓄積した変換用の単語辞書など、仮名漢字変換に必要な基本的なデータ類を正式に継承し、その上で研究開発を進めた。本格的なスタートはNHK技術研究所に待望の高性能大型コンピュータ・IBM360/モデル40Gが導入された昭和四四年からである。相沢は情報処理研究班の同僚・江原暉将とともに、ニュース文を対象にした仮名漢字変換システムの研究開発に取り組んだ。

「栗原教授は、当時私の上司であった副島末好さんの高校の同級生だったこともあり、親切にいろいろ教えてもらいました。研究室にお邪魔し、変換に使う単語辞書や文法ルールをどういう考えで作ったのか尋ねたことを覚えています。「ここにすべて書いてあります」とドサツと資料を机の上に置いてにやりと笑われた。軽く1メートルくらいにもなる資料の山です」

と相沢は振り返る。その結晶である単語辞書は「いまから見ても非常に整ったものだった」という。この単語辞書、変換処理方式、文法分析のための情報などの基本部分を受け継ぎ、分ち書きされた仮名文によるニュース原稿を、自動的に漢字かな交じり文に変換

するシステムを試作したのである。

橋本文法（橋本進吉博士が整理した国文法体系）を参考に日本語を文節で区切り、その文節を自立語と付属語に分け、単語辞書との最長一致法と品詞別の接続表で文節単位の文法処理（局所文法処理）を行って変換するというのが、栗原方式の仮名漢字変換の基本的な考え方である。

これをIBM360上で実現するにあたって、相沢たちが苦労したのは、「なんといいってもコンピュータのメモリーが少ないこと」だった。IBM360は使用素子をトランジスタからICに変えた第三世代のコンピュータとされるもので、世界的にヒットした最新鋭の大型コンピュータである。

それでも現在では考えられないほどメモリーの容量が小さい。いきおい、変換に用いる文法情報をいかにして整理し、圧縮していくかに知恵を絞ることになる。

「とくに厄介だったのは、自立語と付属語の接続をチェックするための接続表でした」

と相沢はいう。この接続表は仮名漢字変換の基本的な文法処理を担う重要なものだ。たとえば、「うぐきます」という入力があった場合、コンピュータは様々な文法的可能性を考え、最後に接続表を参照しながら、

「雨後・きます」は接続不能だが、「動き・ます」なら可能」

という判断をして「動きます」と変換する。さきほど「走らない」の例でも述べたよう

に、文法的な接続情報は、確からしい変換結果を選ぶための鍵なのである。しかし相沢は、「活用形も含めた単純な表を作ると、接続表が402×402の大きなマトリックスになってしまう」ことに頭を悩ませた。約20Kバイトもの容量を費やす量だ。

現在のコンピュータにとつては、なんてことのない容量だが、当時は頭の痛い問題であった。このままではとてもコンピュータに入りきらない。また入ったとしても、実用的な処理速度を確保するのが難しい。

研究開発の目標は、ニュース原稿を漢字かな交じり文に変換するシステムを作ることなのだ。それが人間よりはるかに遅いようでは、開発の意味がなくなってしまう。相沢は何度もマトリックスを眺めた。そしてふと、「大きなムダをやっていることに気がついた」のである。

橋本文法の基本となる「文節」は、

(接頭辞) 自立語 (接尾辞) + 付属語

という構成をとる。この自立語(先行単位)に、付属語(後続単位)が接続可能かどうかをまとめたのがこの接続表だ。表を見つめているうちに、相沢は、

「ある先行単位Aに、活用する後続単位Bが接続する場合、Bのすべての活用形がAに接続するか、あるいはBのいかなる活用形もAに接続しないかのいずれかである」

ことに気がついた。つまり、Bの活用形の一つ一つに接続表を作り、先行単位Aとの接

続を確認する必要はなかったのである。これで大幅に接続表を圧縮することができ、IBM360上に仮名漢字変換システムを実現するメドがついた。

それだけでなく、相沢たちは、サ変動詞の接続表を付属語のサ変語尾に統合して圧縮をかけている。これが面白い。「くする」というサ変動詞を登録しなくても、「サ変語尾のつく名詞」を分類しておけばいいのだ。国文法の常識にはない「サ変名詞」の登場である。

たとえば、「こうあん」には「考案、公安、弘安、公案」などの候補があるが、「こうあんする」は「考案」だけだ。サ変名詞という情報は同音異義語の選択にも役に立つ。加えて、断定助動詞「だ」を形容動詞語尾に統合したりといった地道な整理統合を行い、最終的には五分の一ほどに圧縮することに成功したのであった。

それでも開発はなかなか進まない。彼らの苦勞を理解するには、当時のコンピュータ事情を頭にいれておかななくてはならない。時は昭和四〇年代である。コンピュータ処理はすなわち「一括処理」(batch processing)であった。「それ以外の選択肢は考えられなかった」という時代の出来事なのである。

一括処理に対する言葉は「会話処理」(interactive processing)である。部長の机の未決箱に書類を入れ、処理されるのを待つのが一括処理であり、部長と二人でああでもない、こうでもないと相談しながら処理するのが会話処理である。

コンピュータにとつては、前者の一括処理のほうが都合がいい。データの入力からいち

いち人間の相手をしていては処理が遅くなる。だから、辞書データの打ち込みでも、変換のテストでも、まずデータとプログラムを鑽孔テープやパンチカードなどに打ち込む。この段階ではまだコンピュータを使わない。

テープあるいはカードが出来上がったら、それをコンピュータにセットする。ここから処理が始まるのだ。が、入力結果も変換結果もすぐには見られない。漢字プリンタのない時代の話だ。コンピュータから出力されるのは漢テレ鍵盤に用いる紙テープなのである。次にこれを漢テレ鍵盤にセットして、やっと印刷結果を手に行けるのであった。

漢テレ鍵盤の印字速度は一秒間にたったの二文字である（いまなら安いパーソナルワープロでも二百字／秒のものがある）。この遅さも頭痛の種だ。単語辞書データをチェックするにも、印刷だけかなりの時間がかかる。そのうえ機械が不安定で、

「ちよつと機械から目を離すと、すぐに紙テープが切れて印刷が止まっている」という状態だった。切れたテープをつぎはぎする紙細工さえ、彼らの重要な仕事の一つだったのである。

それでも開発から数年で、試作システムは動き始めた。このシステムはニュース文の仮名漢字変換という限られた用途向けのものとはいえ、研究所の見学者に文章を言ってもらい、その場で変換してみせることもできた。

「とはいえ、「実験のご成功をお祈りします」が「ご性交を」になったりして、冷や汗をか

いた恥ずかしい思い出もある」

と相沢。最後まで苦労したのは単語辞書の整理だった。栗原たちが作った単語辞書はよくできていたが、逆にあまりにも語彙が豊富すぎて困ったのである。語彙が多いということとは、同音異義語が多く、誤変換の可能性が増えるということでもある。慌てて辞書のデータベースを印刷し（といっても秒速二字の漢テレ鍵盤だから、数日ばかりだ）、ニュース原稿にあまり使われそうにない単語を削除して、「辞書を絞りこんだ」のであった。

本格的な研究開発のスタートから五年が経過した昭和四八年一〇月一六日、相沢と江原の二人は、試作結果とその評価を「計算機によるカナ漢字変換」と題する論文にまとめ、雑誌『NHK技術研究』に提出した。この論文は、あくまでも大型コンピュータを使った仮名漢字変換システムについてのものだが、その後のワープロ開発に大きな影響を与えた重要なものである。

このシステムの仮名漢字変換は、栗原俊彦が考案した最長一致法をそのまま踏襲していたが、(1)変換に用いる文法情報を当時のコンピュータに実装できるように圧縮し、小型コンピュータでも仮名漢字変換を実現できる可能性を示した点、(2)変換処理速度や変換率など様々な能力測定を行い、今後の技術開発の進むべき道を示した点に、彼らの功績がある。

そしてなにより、論文が素晴らしい。相沢たちが精魂込めて作り上げた品詞の分類とその接続表や、単語辞書の内容の一部など、仮名漢字変換を実現する技術の中身がすべて公

開されている。ノウハウも技術上の問題点もよく整理されており、これを読むだけで、次の日から誰もが単語辞書を含めた仮名漢字変換システムの構築に乗り出せるほどの内容となっている。

昭和五八年にパソコン用ワープロソフトを開発したある開発者など、ソフトの開発後にこの論文に出会い、深く溜め息をつき、落ち込んだという。「十年も前に書かれたものなのに、自分がやった苦労がすべて書かれていた」からである。

「あの時期に出た論文としては、白眉のものですよ」

とワープロ開発に携わった人間たちの誰もが口を揃える。彼らが残した論文は、その後、ワープロ開発に乗り出した多くのメーカーのエンジニアのバイブルとして機能したのである。

しかし、これほどの業績を上げながら、NHK技研自身がこの研究を進展させることはなかった。試作システムが出来あがってまもなく、NHKはテレックスを使った原稿送信から電話送稿へと切り換えたため、彼らの研究は「続ける意味をなくしてしまった」のである。

最初から最後まで変わらない物静かな口調で、相沢はこう語る。

「当時、もしもソフトウェア特許がとれる時代であったなら、私たちがやったことはワープロの基本特許として申請できたらうと思います……でも、ああいう時代だったんですか

ら、権利化できなかったことは仕方がない。自分たちは生かせなかったが、ワープロ誕生に多少なりとも貢献できた。それで十分です。私たちは本来、放送に関係する技術開発を行うべき人間であって、製品開発を目的としているわけではないのですから」

彼の最大の失敗は、マイクロコンピュータがこれほど発展するとは「夢にも思わなかった」ことであった。仮名漢字変換に手をつけるのも、諦めるのもあまりにも早すぎた。

相沢たちが論文を書いていたちょうどその頃、海の向こうのアメリカではインテル社の嶋正利とフェデリコ・ファジンが、ミニコン（ミニコンピュータの略）の機能を一つのLSIにまとめた8ビット処理のマイクロプロセッサを開発中だったのである。翌年「8080」という型番で発売されたこのLSIが、数年後にパーソナルコンピュータを生みだし、十数年後にはNASAがアポロ計画で使った大型コンピュータより性能のいいものを、家庭の机の上に載せたのである。

予想価格はなんと四億円の代物

NHK技術研究所が仮名漢字変換の試作システムをほぼ完成させていた昭和四八年四月、国内留学を終えた河田勉が天野真家を伴って、東芝総合研究所（東芝中央研究所から名称変更）に戻ってきた。

この頃、森は文字読取装置（OCR）と画像処理という二つの研究グループを統括して

おり、京都大学から戻ってきた河田と入社したばかりの天野は、いったんOCRグループに配属された。天野のほうは、それからしばらく新人研修である。

当面の仕事は、漢字OCRの研究開発だった。その裏で、二人は京大でやっていた自然言語処理についてあれこれ考えながら、研究所のコンピュータ環境の整備などをやっていた。

「TOSBACのOCというミニコンが部屋にあつて自由に触れましたが、開発環境がまったく整っていなかったんですよ。大学の研究室よりも貧弱だった」

というのは天野である。電源を入れたあと、紙テープからプログラムを読み込む前近代的なコンピュータだった（とはいえ、当時の小型コンピュータはこれが普通である。フロップディスクが発売され始めたのは昭和四七年のことだ）。

これではとても使いにくいので、天野はカセット式の磁気テープ装置からプログラムを読み込めるようにしたり、言語処理に向けた基本ソフトを作ったりしていた。天野は器用で、何もかも自分でやってしまうタイプのエンジニアである。

翌昭和四九年夏、河田は森と、森の上司である情報システム研究所長・玄地宏に呼び出された。

「京都大学での研究を活かした成果のうち、企業に貢献できるものは何か？」

と問われたのだという。いつまで遊んでいるんだ、と言わんばかりの口ぶりである。河

田は、

「仮名漢字変換を使ったマルチ端末日本語入力システムを作りたい」

と申し出た。マルチ端末日本語入力システムとは、一つのコンピュータに複数の端末をぶらさげ、仮名漢字変換で日本語情報を入力しようというものだ。仮名漢字変換には大型コンピュータが必要だが、端末を複数にすれば、一台あたりの単価を下げられるという発想である。

「それはいい。社内の技術評価展示会に、インキでもなんでもなく「動くもの」を出すことが大切だ。研究途中であるうとも、成果を見せれば皆がサポートしてくれるんだからな」

玄地は森と開発に成功した郵便番号自動読取装置の経験を披瀝し、河田を励ました。この日から、森はOCR開発の裏のサブチームとして日本語処理を選び、河田と天野の二人を仮名漢字変換システムの研究開発に取り組ませたのである。

この段階でもなお、正式の研究開発プロジェクトではなかった。まだ何が出来るかよく見えてない状態である。河田と天野の仕事は、いわば「闇」の仕事であった。東芝の研究所ではこれを、「アンダー・ザ・テーブルの研究」といい、「人と物と金のうち二〇パーセントを投入してよいことになっていた」という。この状態が昭和五一年まで続いた。

闇の仕事でも、旅にとりあえずの目的地ができた。二人が本格的に仮名漢字変換に取り

組んだのはこの日からである。最初に手をつけたのは変換用の単語辞書作りであった。「九州大学の辞書は喉から手が出るほど欲しかったけれど、もともと沖電気が研究費を提供したのだから遠慮せざるを得ない。自分たちで作るしかなかった」

と河田はいう。最初は国語辞典を参考にして単語辞書作りを始めた。ほどなく、栗原たちが愕然としたように、彼らも結果に呆然とした。日常的に使う単語がまるで収録されていないのだ。状況報告を求める森に、

「日本語」という見出し語すら国語辞典にはないんです。ちょっと困っています」

と報告する河田。森はすぐ、

「よしわかった。データを毎日新聞社の人に頼んでやる」

と言い、記事送信に使った漢テレ鍵盤用の紙テープを調達してきた。この送信済の紙テープには、漢字かな交じり文で書かれた新聞記事の文字データが記録されている。ここから漢字表記の単語を切り出せば、国語辞典に収録されていない言葉も拾えるというわけだ。「それでも新聞用語というのには限られていますからね。急激に進むというわけではないんです」

と天野はいう。和英辞典やアクセント辞典など様々な辞書類にもあたった。和英辞典は書くための辞書だから、「小学生」も載っている。国語辞典より和英辞典のほうが、ワープロ開発には貴重な資料となった。

「でも、いちばん役にたったのは、なんといつてもNHK技研の相沢さんたちの論文です」と二人は口をそろえる。「自分たちは研究開発を中止したが、論文を公表したことでワープロ開発に貢献できた」と相沢が言うのも決して大袈裟なことではない。とくに「論文の末尾につけられた品詞の分類表が役に立った」そうだ。

まだ二人の頭に「ワープロ」はない。彼らが目指しているのは仮名漢字変換を日本語入力に応用する、大型コンピュータを使ったシステムなのである。仮名漢字変換の実現には、文法解析を瞬時にこなす高速のコンピュータに、大量の単語辞書データを記憶する大容量の記憶装置が必要だ。昭和四〇年代には、それを実現できるのは大型コンピュータだけだった。

河田が入社したときに森から渡された書類を見ると、当時の技術者が持っていた「常識」が窺える。この書類の最後には、当時の森が想定した日本語の機械処理の製品像が書かれているのだ。

最も安い製品が、裁判所や議会の記録用システムで、予想価格は五百万円となっている。入力方式は当時の東芝が研究していた「ラインプット」(速記畑の川上晃氏が開発。カナタイプを使った二ストローク入力で漢字を入力する方式)、あるいは「分解発音入力」(カナタイプを用いて漢字・熟語の音訓読み、部分読みの情報で入力する方式)であって、仮名漢字変換ではない。

仮名漢字変換を想定しているのは大型コンピュータを使った電算写植や日本語情報処理システムで、予想価格は四億円である。仮名漢字変換を採用すると、製品価格は数億円になってしまう——そう考えられていたのがこの時代なのであった。

初めて出たワープロという名前

それからの河田と天野は、コンピュータのハードウェア屋と基本ソフト屋、それに入力業者とをわけあわせたような仕事に追われた。

「できるかどうかよくわからない。環境も整っていない。でも、だからこそやり遂げたという気持ちにもなった」と天野。河田は「漢字を表示するディスプレイ装置も手作りでした」という。二人は、互いの得意分野をあわせると、ハードウェアからソフトウェアまで自在に操れる最強のコンビだった。

周囲にあるのは英数字（とせいぜいカタカナ）しか表示できないものばかり。これでは話にならないので、河田はソニーテクトロニクス社製の蓄積管タイプのブラウン管を調達してきた。それでもコンピュータとディスプレイを結ぶインターフェイスが弱く、「朝のうちにデータを入力しておく、午後にやっと表示される」という気の長い環境だった。一面面の表示に、六時間ほどかかるのだ。

そこで、中を改造してベクトル描画で漢字を表示できるように手を加えた。「これは長尾

先生に教えてもらった方法なんです」と河田。これでやっと、一面面が十秒程度で表示できるようになり、随分と仕事が捗るようになった。

単語辞書のデータを入力するのも大変な仕事だ。入力できる機械は一台しかなく、最初のうちは二人で「四時間交替で打ち込んでいた」のである。しかし、とても二人の手に負えないので、入力会社に依頼してパンチカードを作ってもらった。

しかし、人間の操作にはミスがつきものである。単語辞書のデータがミスタイプでけっこう間違っている。これを修正するのがまた大変な労力を要する仕事だった。辞書データの入力ミスを修正するだけでも、二人がかりで三ヶ月かかった」のだという。

二人の課題は、仮名漢字変換の変換率の向上だった。目標通り、端末を使って仮名漢字変換で日本語のデータ入力をするシステムを開発しても、変換率が悪いと使い物にならない。

「当時の変換システムでは、せいぜい変換率は五〇パーセントというところでしょう。これをどうやって八〇パーセントに向上させるか、というのが最初の目標だった」

と天野はいう。彼らは再び国語の基本である橋本文法に立ち戻ってシステムを再検討した。九大の栗原が考案した最長一致法の見直しも行った。

最長一致法は、昭和四〇年頃の非常に貧弱なコンピュータでも、仮名漢字変換を可能にした素晴らしいアルゴリズムである。たとえば、「サカナイ」という入力に対して、まずサ

カナイで一致する単語を探し、なければサカナと切る。「魚」がヒットしたところで接続情報を調べると、「魚」に「い」がつくのはおかしいことがわかる。次はサカと切り「坂ない」だ。これも接続できない。こうして「映かない」まで辿りつくのである。

本当なら、文法的な様々な可能性をすべてチェックし、そのなかで最も確からしい変換結果を選ぶようにしたいところだ。ところが、当時のコンピュータの能力はそれを許さなかった。そんな面倒なことをやっていると、朝から晩まで延々と変換を待つことになる。最長一致法は当時の性能の低いコンピュータでも仮名漢字変換を実現するための、工学的な解決手段なのだ。

それから約十年、パンチカードや紙テープを使うのは相変わらずだったが（つまり入力機能が貧弱だったが）、コンピュータの計算能力は飛躍的に進歩している。もう最長一致法ではなく、正攻法でやってもいい時代だ。また、そうしないと変換率の向上も望めない。

二人は相沢論文を参考にしながら、その不備を徹底的に補いつつ、自分たちでコンピュータ向きの文法を作りなおした。そして、最長一致法に頼らず、すべての可能性をチェックして（全数チェック）、そのうち最も正解らしいものを第一に表示するアルゴリズムとした。

本格的に取り組んでから約一年半後の昭和五一年三月、いよいよ彼らのシステムの実力が試される日が来た。研究所内で大型コンピュータを使った変換の実験をやったのである。

プログラムはフォートランで書かれたパンチカードの山で、一括処理で仮名漢字変換を行うものだった。

コンピュータにプログラムとデータを読み込ませ、あとはひたすら待つ。打ち出されてきた変換結果に満足した森は、「これならやれそうだと確信を持った」という。そして河田と天野の労をねぎらう一方、即座に新しい課題を与えた。

「大型コンピュータを使っている話にならない。ミニコンに載せなさい」

森には小型化への強い意志があった。しかし、河田と天野の二人にとっては、一瞬、顔が青くなるような指示である。

「え、ミニコンに載せるんですか？ 64Kバイトしかないんですよ」

反射的に天野が抵抗した。ここでいう64Kバイトとは、主記憶装置の容量のことである。コンピュータはプログラムと処理データを主記憶装置に読み込んで処理をする。デスクワークにおける机のようなものだ。

机が十分に広ければ、様々な資料類を机上に置けるから処理が速くなるが、64Kバイト（英数字にして約六万五千字分の容量）といえば、新幹線の食事用テーブルのようなものである。文法書とノートの両方を置ける広さではない（まして辞書を置くスペースはどこにもない）。マニアのおもちゃと思われるいた誕生したばかりのパソコンが同じく64Kバイトだったと書けば、だいたいの想像はつくだろう。

しかし、森の指示は変わらない。「不可能だと思えることに挑戦してこそ、研究の価値があるのだ」とたしなめた。二人はすぐに気をとりなおし、今度はミニコンに搭載するために、河田はプログラムに工夫を凝らし、天野は文法をさらに整理する毎日となった。

森はこう指示する一方で、翌月には「日本語処理の研究」と題する研究企画書を提出し、これまで延々とアンダー・ザ・テーブルの研究としてやってきた二人の研究開発を、正式のプロジェクトとして発足させたのである。

ここからの歩みは急である。周囲の環境が整ってきたということもある。様々な面で同時進行的に開発が進んだ。たとえば、まったく別のチームが漢字プリンタの開発に手を付けていたこともあり、合同で研究開発を行うことになった。さらに河田と天野の二人に、非常に腕のいいプログラマの武田公人が加わり、中心メンバーは三人となった。

ミニコン（ミニコンピュータの略）で仮名漢字変換を実現した例は過去にない。いずれも大型コンピュータを用いたものばかりだ。「完成すれば売り物になる」と河田は当時の書類に書いている。

が、果たしてそれをどう使うのか。彼らはただ単に大型コンピュータ上のプログラムをミニコンに置き換えるだけでなく、どのような「つくり」にするのか、その概念をはっきりさせる必要に迫られた。

「ワードプロセッサ」という目標にたどり着いたのはこの時である。英文ワードプロセ

サが既に登場しており、これも参考になった。また、ミニコンを使えば、一括処理ではなく会話処理の仮名漢字変換システムを作れるから、同音異義語もユーザーが画面上で選択すればよい。

一挙に現在のワープロの形に近づいた。T O S B A C 4 0 C というミニコンを使って翌昭和五二年三月に完成した試作機は、現在のワープロの原型と評価するにふさわしい特徴を備えている。登録した辞書数は約二万三〇〇〇語であった。

彼らはこの結果を、同年仙台で行われた電子通信学会情報部門全国大会で発表している。論文のタイトルは「カナ漢字変換機能を備えたワードプロセッサ」であり、天野真家、河田勉、武田公人の連名だ。これが「仮名漢字変換」と「ワードプロセッサ」が合体した本邦初の論文であった。

次なる目標は商品化である。いまから思えば意外なことだが、彼らにとってはこれが最も大変な作業だった。森たちは研究所の人間であって、モノを作るのは事業部である。事業部に依頼されて行う研究もあるが、今回のようにシーズを研究所が育てた場合は、事業部の人間の説得から始めなくてはならない。

ところが、これまで東芝が商品化し、ヒットを飛ばしてきた電気釜や、電気ゴタツのようにはわかりやすい商品ではないのだ。「仮名漢字変換」という言葉すら、ほとんどの人にとっては初耳のものである。

森が社内売り込む一方で、河田らは試作機を使って研究所内の報告書を作成しながら、改良を加えていった。「歌いながら歌詞を入力したりして、変換しないものがあると文法をいじる」という毎日である。変換しないものを見つけては、接続表を補強し、とんでもない誤変換を見つけたら、なぜそうなるのかをプログラム上で追って、修正するのである。「なんといってもびっくりしたのは、「って」という助詞ですよ」

と天野。文法書には載っていないが、「私って、ばかね」の「って」は紛れもなく助詞ではないか。その処理には非常に困った。「って」を助詞として登録し、「わたしって」を「私って」と変換できるようにしてもいいが、そうすると「川ってしまった」(変わってしまった)「橋っていく」(走っていく)のようなとんでもない変換もしてしまう。副作用が大きいのだ。結局、「私って」の変換は諦めた」という。

実際に動く機械があると、いろいろなことが見えてくるものだ。報告書を打ち込むうちに、ふと同音異義語の選択には法則性があることも発見した。

同音異義語は非常に厄介なものだが、いくつかある候補のうち、ある一つを選択すると、次にはその候補を選択する率が非常に高い。いったん「校正」を選べば、次もたいてい「校正」であって、「公正」や「構成」は滅多に選ばないことに気がついたのである。同音異義語の選択は文脈依存性が非常に強いということだ。

そこで「最新学習機能」をつけた。構成、公正、校正、攻勢、厚生、後世……とある候

補のうち、たとえば「厚生」を選ぶと、次にはこれを最初の候補として表示する機能である。これによって、「変換率が劇的に向上した」のである。もしもこの機能がなければ、いまだにワープロは使い物になっていないだろう。

既に試作機を作る段階で、候補の表示(つまり辞書の並び)を使用頻度によって並び換える頻度順学習機能を思いついたので、短期と長期、二つの学習機能を考案したことになる。

「もちろん栗原さんを初めとして、仮名漢字変換そのものは先行する研究と技術開発がいくつあったが、問題はすべて変換率が低かったこと。東芝はそれを九〇パーセントまで向上させたから、ワープロが実用になったのだ」

と森は胸を張るが、その変換率向上の鍵を重要な順に並べると、(1)文法を整備しなおし、全数チェックのアルゴリズムを採用したこと、(2)最新学習機能、頻度順学習機能を考案したこと、(3)自立語が連なる複合語も変換できるようにしたこと、(4)単語辞書を整備したこと、であったと評価できる。

もっとも、先行した研究開発との差は一括処理での仮名漢字変換と、会話処理でのそれとの差だという見方ができないわけではない。純粹に工学的に見れば、彼らの功績は大型コンピュータでしか実現できないと思われるいた仮名漢字変換をミニコンで実現し、会話処理化した点にある。

そしてそれ以上に重要な功績は、仮名漢字変換を使った日本語入力と、ワードプロセッシングとを結びつけたことだ。そもそもはテレックスで送られてくる記事の反訳システムとして考案されたものを、日本人の書く道具として発展させたのである。

遂に商品化へのゴーサイン

この新技術に対する社内の大勢の反応は鈍かった。しかし、試作機の噂を聞きつけ、慌てて駆けつけた人物がいる。当時東芝青梅工場ハードウェア開発部長だった溝口哲也だ。昭和五二年一〇月二日のことであった。

こういう素早い反応は工学的なセンスによるものだが、それだけでなく当時の青梅工場が転機を迎えていたことも関係している。青梅は大型コンピュータやミニコンを開発、生産する東芝のコンピュータ部隊の本拠地であったが、当時は瀬戸際に立たされていた。結局、翌昭和五三年には大型コンピュータ分野から撤退している。溝口たちは必死に新規事業を模索していたのである。

「アッ、さすがだ、と思いました」

と森はその日の思い出を語る。一通りのデモを終えたところで、溝口は研究所に来る途中に川崎駅で買った新聞を差し出し、
「この社説を打ってみてくれ」

と言ったのだ。一瞬、場に緊張が走った。玄地のいう「インキでも動くもの」は通用しないぞ、ということだ。もはやこれはデモンストレーションではない。本番である。一同、固唾を呑んで画面を見つめた。キーボードを操作するのは天野である。

多少の錯誤はあるものの、確実に社説が再現されていく。その瞬間、森たちは自分たちの成果に改めて自信を持ち、溝口は商品化の決心をしたのであった。

翌日、溝口は青梅工場で田中範夫副技師長に会って詳細を報告し、商品化プロジェクトチーム発足の了解をとりつけた。この段階でも、渋々承知させた、というような状態だったようである。

それでも一一月には青梅工場を中心に日本語ワープロ開発体制が出来あがった。それに伴い、河田、天野と武田の三人は頻繁に青梅に通うことになった。とくに天野と武田は、もう「通勤先を青梅に変更したような年間だった」という。青梅工場からは植田恭、寒川善明、月村治雄、羽田野修、武藤治夫、吉井清、吉井二郎の七名が参加した。総勢一〇名のチームである。

東芝が「同音語選択装置」と題する特許を出願したのはこの頃だ。天野、河田、武田、森の連名で一二月二四日に出願した。特許請求の範囲は、

《日本語文をその読みに従ってカナで入力し、対応する漢字混り文に変換するカナ漢字変換システムにおいて、

カナ漢字変換結果及び同音異義語を表示するディスプレイ装置と、同音異義語の中の1単語に優先権を付与するための入力手段と、この手段により優先権が付与された単語を記憶する記憶手段と、同音異義語が生じた場合には前記記憶手段に優先権の付与された単語が存在するかどうか調べる手段とを有し、

前記記憶手段に優先権の付与された単語が存在する場合には、その1単語を自動的に選択することを特徴とする同音語選択装置》

である。最新学習機能を使って、画面上で変換していく仮名漢字変換システムはすべてひつかかる。これがワープロの基本中の基本の特許なのである。

権利は確保した。しかし、商品開発となると、決めなくてはならないことが山のようにある。まず価格だ。試作機の価格はそのまま合計すると二千万円程度のものであった。この価格では、とても売れそうにないことが試作機段階のヒヤリング調査でわかっていて、目標を五百万円に定めた。

次はハードウェア設計である。まず心臓部のCPUには、ちょうど試作機でも使ったミニコン・TOSBAC40CのCPUがLSI化されていたので、これを使うことにした。昭和五二年のことだから、そろそろマイクロプロセッサをCPUに使ったパーソナルコンピュータがアメリカでも、日本でも登場していたが、当時はまだマイクロプロセッサの能

力が不足していたことと、完成した機械を核に医療情報システムなどを開発する案があったため、ミニコンベースのCPUを使ったのである。

ここまではすんなり決まった。問題があったのは単語辞書の記憶装置とディスプレイ、それにプリンタである。

仮名漢字変換では単語辞書を頻繁に参照するから、読み書きの速い記憶装置に単語辞書を置かなくてはならない。速さだけなら半導体メモリーがベストだが、辞書はデータ量が膨大にあるので高価になりすぎる。

価格と性能のバランスではフロッピーディスクが優れている。それでも読み書きが遅いし、容量も十分ではない。結局、これも当時はとても高価なものだったが、10Mバイトのハードディスクを使うことになった。パソコンにはまだカセットテープを記憶装置として使っていた時代の話である。ユーザーが作成した文書の保存用として、8インチのフロッピーディスクドライブを一基用意した。

それでも当時のハードディスクは能力が低く、素直に使うと変換が遅くなる。ハードディスク上の辞書を検索するのに時間がかかるからだ。そこで河田がプログラムに工夫をした。彼はこういうハードウェアと密着した設計が非常に得意だ。人間のキーボード操作のほうが遅いことに着目し、たとえば「へんかんは」の「へ」を打ち込んだ段階で、ハードディスクから該当辞書部分をメモリー上に読みだしてしまおうようにした。これを辞書の先

読みという。

ディスプレイとプリンタは新規に調達あるいは開発しなくてはならなかった。ワープロは日本語を入力し、日本語を表示し、日本語を印刷する機械である。しかし、この当時までコンピュータは英語あるいはカタカナで使うのが常識であったため、漢字を表示・印刷できるコンピュータ用装置がほとんど開発されていなかったのである。

画面に文字を表示する原理は、電光掲示板のそれと同じである。文字を点の集合（ドットマトリックス）で表現する。英字だと8×8ドットもあれば字を表現できるが、漢字は画数が多いため、最低限16×16ドット、欲をいえば24×24ドットが必要だ。

これが画面表示文字数に関係してくる。英字を横40字×縦20行表示可能な画面でも、それに16ドット構成の漢字を表示させると20字×10行表示になってしまう。画面解像度の高いディスプレイが欲しい。彼らはディスプレイメーカーを駆け回って、なんとか24ドット構成の漢字を41字×14行表示できるブラウン管を調達することができた。大田区の中央無線からである。

プリンタはもっと厄介だった。ハードウェア設計上の最大の難関であったといってもいい。ワープロは文書を作成する機械であるから、仕上がりが大切な要素となる。

「16ドット構成だと「鷹」のように画数の多い漢字は、字画を省略しないと表現できない。嘘字になってしまうのはまずい」

という議論が噴出した。最低限必要なスペックは24ドット構成だ。これなら「鷹」も画数を省略せずに表現できる。

さらに彼らは用途を考えた。誰が買って、どう使うかという議論である。価格からいって、主な納入先は企業や官庁だろう。作成される文書は契約書や公文書である。字画が略されているのは論外だし、複写式の用紙に印刷することができ、改竄に強いものでなくてはならない。さまざまな種類のプリンタがあるが、それに対応できるのはドットインパクト方式のものだけだ。

ドットインパクト方式は、細いワイヤを縦に並べ、それを磁気コイルで動かしてインクリボンを押くことで印字する。衝撃がかかるので、複写式の用紙に印字し、同時にコピーを作ることが可能だ。用途にはびつたりの方式だが、24ドット構成となるとこれまでの約三倍の密度である。そのぶん、ワイヤを細くしなくてはならないし、磁気コイルの構成も問題となる。

開発チームは必死に目的にあうプリンタを捜し求め、東芝社内のあちこちに相談した。24個の磁気コイルを縦に並べるのは難しくても、ちょうど百合の花のように円形に開いて配置し、そこから絞ってやることにした。これならなんとかなる。

問題はワイヤの素材だ。円形にコイルを並べた場合、印字ポイントまでワイヤは円弧を描くことになる。この「曲がり」に対応できるほどしなやかで、なおかつワイヤの先を細

くでき、あまり磨耗しないものでなくてはならない。

「一時は諦めかけたけれど、東芝の関連会社でタングステン合金のワイヤを開発していたところを見つけたので、なんとかなった」

と森はいう。これは劇的な進歩である。なにしろ、それ以前の「漢字プリンタ」はそれだけで約五百万円の価格をつけていた。JW・10にふさわしい安価で使いやすく、表現力の高い漢字プリンタが開発されたのだ。

ハードウェアはなんとかなりそうだった。次の課題はソフトウェア開発である。

基本的なソフトウェアは試作機の段階で出来上がっている。しかし商品化となると、機能をもう一度見直し、仕様を再検討する必要がある。キーボードも特別設計しなくてはならない。最初に全員で合宿を行って方針を煮詰めた。

「まず漢字キーを押したあととキーボードから仮名で「読み」を入力し、再び漢字キーを押すと、画面に候補がざっと表示されます」と天野はホワイトボードに書きながら説明をすすす。ワープロの核心部分の説明である。後ろのほうから声があがった。

「でもそれは、候補がクルクル回って、キーを押すたびに次が表示されるほうがいいね」「なるほど」と天野は感心したという。現在のような「次候補(変換)キーを押すたびに、次の候補が表示される」というアイディアはこの合宿で得られたものだった。

幸いにして、搭載する漢字については悩む必要がなかった。彼らが開発にあたっている

最中に、通産省工業技術院がコンピュータ用の漢字についての文字コード規格を発表したからだ(「情報交換用漢字符合系」。制定は昭和五三年一月一日付)。いわゆるJIS漢字コードである。

このコードが制定される前までは、漢字を各社バラバラにコード化していた。河田たちが開発段階で使ったのは漢テレ鍵盤コードであるし、紙面制作のコンピュータ化に乗り出していた新聞社も独自のコードを持っていた。

しかし、文字コードのような基本的なものが各社バラバラでは、情報交換に支障をきたすことは明らかだ。彼らはJIS漢字コードの採用を決め、24ドット構成の印刷用文字をデザイナーに発注した。

「書類がコピーされることを考えて、「コピーされたときに見やすいデザインにして欲しい」と依頼した」

と河田はいう。他社がワープロに参入したあとも、長く東芝は独特のデザインの文字フォントを持っていたが、これはコピーを考えてのことなのであった。

漢字変換部分は研究所の河田たちの担当である。変換方式は漢字指定方式と文節変換方式の二つを用意した。

キーボードの下に「文節／漢字」というキーと「ひらがな」というキーを用意し、漢字の始まりで「文節／漢字」キーを押し、ひらがなの始まりで「ひらがな」キーを押すのが

漢字指定方式である。これは原稿を見ながら打つ清書用に用意したものだ。

いまでは複文節変換が当然となっており、操作性の悪い漢字指定は、いまではほとんど使われていない。このモードを用意したのは、

「文節」ことに押ししてください」と言っても、そもそも一般の人間に「文節」という概念がわかるのか」

という議論の結果だったようだ。「漢字にしたいときはこう、ひらがなのときはこうです」と説明しやすく、理解しやすいモードを用意したのである。

文節変換は現在一般的な方式である。文節ごとに「文節／漢字」キーを押して変換していく。画面に表示されたものが目的の漢字かな交じり表記ではない場合は、「次候補」キーを押してその場で選択することができた。これが会話処理をとったワープロの革新である。「本当はいままでいうAI辞書なども考えていたのだけれど、プログラムが大きくなりすぎる上に、たいして変換率が向上しないので諦めたんですよ」

と森。AI辞書とは用例辞書の俗称である。「歛心」ならば「買う」であって、「飼う」ではない。「夏」なら「暑い」だし、「お湯」は「熱い」だ。この関係を「共起関係」という。実際に用例辞書がワープロに搭載されたのは一九八〇年代後半だったが、その歴史は古い。工学系の仮名漢字変換技術者たちは、自然言語処理に踏み出したときの教科書の一つ、言語学者N・A・チョムスキーの“*Aspects of the Theory of Syntax*”から、これを着

想していたのである。

この発想がありながら、本格的に搭載しなかった理由には、変換方式を複雑にすると、反応速度が落ちるということもあったようだ。とはいえ、「だい2かい」は必ず「第2回」に変換するなど、一部には共起関係（第と回）を利用した変換も取り入れている。

反応速度というのは、たとえば「文節／漢字」キーを押したあと、候補が表示されるまでの時間のことである。この東芝のチームが優れているのは、質と速度に徹底的にこだわった点だ。さきほどの24ドット印字へのこだわりが「質」の部分である。

速度面でもこだわりを見せた。作成している文書の一頁目で削除を実行した際の反応速度や、変換の候補表示などである。文字フォント（字の形を示すデータ）も例外ではない。

JIS漢字コードで定められた六三九字分の文字フォントを頻度順にランク分けし、よく使う漢字のフォントは電源投入と同時にハードディスクから半導体メモリーに読み込むようにして、反応速度の向上を図っていたのである。

第一号機は六三〇万円に決定

この間、森は溝口や児玉とともに、社内の説得を続けていた。なんとかして、この商品を世に送りだしたいという気持ちばかり先走るが、頑として首を縦に振らない人物がいる。電算機事業部長の都築公男だ。

製品のプロトタイプ作りもかなり進んだ昭和五三年六月、森たちは都築に大見得を切った。

「機械のことをまったく知らない女性をこれから訓練するから、その成果を見てくれ」

というのである。実証主義に訴える作戦だ。それから一ヶ月後の七月、都築が青梅工場を訪れた。女性が森たちのいう「ワードプロセッサ」を操作するところをその目で確かめるためである。

「それでは、見せてもらおうか」

ドスのきいた低い声で都築が口を開く。都築の到着前から、マシンには既に電源が入り、準備が整っていた。周囲はすべて息をひそめながら野次馬と化している。うなるように回る冷却用のファンの音がやけに耳につく。

「はい」と小声で返事をし、急造のわかタイプリストが試作機に向かった。キーボードに置いた指が緊張でかすかに震えている。「漢字／文節」と書かれたキーを押し、「はいけい」と打つ。再び漢字／文節キーを押し瞬間、「拝啓」が画面に表示された。仮名漢字変換を使った入力である。会議室にいる大半の人間にとって、これまで一度も目にしたことのない不思議な光景だった。

都築は黙りこくってじっと実演を見た。キーボードから入力された仮名文字列が、順にキー操作で「時下ますますご清祥のことと……」と漢字かな交じり表記になっていく。書

きなぐるよりは遅いかも知れないが、丁寧に書くよりは明らかに速い。

みるみるビジネス文書が仕上がっていく画面を見すえたまま、「よし」と都築が口を開いた。視線を森たちに移し、ぎろりと睨む。

「今度のデータシヨウに出品しよう。その前に報道発表だ」

誰もが一瞬、自分の耳を疑い、少し間を置いて会議室が蜂の巣をつついたような騒ぎとなった。そうと決まれば、準備しなくてはならないことが山のようにある。そもそも「今度のデータシヨウ」まで、あと三ヶ月しか残されていないのだ。

この「勝利」ほど、森にとつて感慨深いものはなかっただろう。七年前のある人物との出会いを思い出す出来事であった。

森が彼と出会ったのは、円が固定相場制から変動相場制に切り替わり、「貧乏」はもう売り物にならない」と永野重雄（当時日商会頭）が文藝春秋に書いた昭和四六年のことである。一九六〇年代の急速な経済成長にともなう「余波」が、世界から日本に押し寄せている時期であった。

国際社会における日本の立場の変化は、ニュースの増加となって現れる。この頃、各新聞社の編集・制作部隊は真剣に悩んでいた。国内のニュースも、海外からのニュースも激増しているが、新聞制作の現場がそれに追いつかない。

もともと速報性では、テレビやラジオにかなわないのだ。いまのうちに何か手を打って

おかないと、新聞は時代に取り残されてしまうのではないか。たとえば日本経済新聞社と朝日新聞社は、この危機感から新聞紙面制作のコンピュータ化に取り組んでいた。

そして毎日新聞社は、東京本社電信課を中心に、情報伝達の革新を目指して、漢テレ鍵盤を使って記事を全国の支社に配信するシステムを構築できないかと考えていた。ちょうど森が日本語処理に興味を持ち、河田が入社した頃のことである。リーダーは部長級の技師・弓場晋太郎と課長の柴田隆也の二人（いずれも故人）であった。

弓場たちはこの企画を東芝に持ち込み、共同開発プロジェクトとしたのである。不思議な符合だった。開発が始まって半年もたたないうちに、弓場は森の部屋に案内されることになった。昭和四六年の秋のことである。弓場が森に持ち込んだのは、

「新聞記者が直接、日本語をタイプできる機械を作れないか」

という相談であった。漢テレ鍵盤を使った高速記事送信システムで、記事伝送は効率化できる。それならいっそのこと、大きくて操作の複雑な漢テレ鍵盤の代わりに、容易に日本語の文章を作成可能な機械が欲しい——毎日新聞社のグループは、そういう結論に達したのである。

弓場は自分たちの思いつきを東芝の担当者に話した。「それは面白そうだ」と紹介されたのが、森研究主務なのであった。ちょうど日本語の機械処理について調査し、手を出そうとしていた森にとっては願ってもない話である。毎日新聞社まで出向いて熱心に「取材」

をした。

弓場は森に、「外国人記者は、いつでもどこでもタイプライターで記事を書き、テレックスですぐ送信できる。でも日本人はまだ手書きだ」と熱っぽく現状を語る。そして森は、自分たちが目指すべき装置の条件を最終的に「次の三つに絞った」のである。その三条件とは、

(1)手で書くより速く文章作成ができること

(2)持ち歩けること

(3)電話回線を使って、作成した文章を伝送できること

であった。都築の説得に成功したこと自体が、弓場が望んだ第一の条件を満たした証拠である。

「じつはそれからしばらくの間、いろいろと試作していたんだよ」

と森はいう。最初に目をつけたのはタブレット式の機械だった。これは邦文タイプライターでなじみのある方式だし、原理も簡単なので、それほど訓練が必要ではないという良さがある。

「漢字をタブレットに並べ、その上を金属製のペンでなぞることで漢字を入力する試作機を作って実験を繰り返した」

と森はいう。この試作機は金属ペンが近づくとコンデンサの容量が変化する原理を応用

して、ペンでさつと触れるだけで、どの字を指定したかを認識する仕組みになっていた。機械式にガチャンとタイプする邦文タイプライタよりは入力が速い。

「邦文タイプライタは熟練者でも一分間に三〇字くらい。この試作機だと、一分間に六〇字くらいの入力はできた」

のだそうだ。じつに二倍である。しかし、ここで壁が来た。いくら訓練してもこれ以上は速くならない。目的の字を探して指定することに時間がかかり、「手で書くより速く、という条件を満たせなかった」のだ。

手書きだと一分間に七〇字は書ける。これを超えられないようでは、弓場たちが熱望している「欧米人にとつてのタイプライタに匹敵する機械」は作れそうにない。

次に目をつけたのは漢字の構造であった。漢字は偏と旁に分解できる。旁のほうは形が複雑だが、偏はわりと単純なものが多い。偏を指定し、次にその字の「読み」で絞り込む。残った候補のなかから一つを選ぶ方式を考えただのである。いわば偏指定音訓入力だ。ところがこれも、「何度やっても目標速度を達成できなかった」のであった。

この失敗の経験からワープロ開発は始まったとわかっていい。河田を国内留学にやり、仮名漢字変換に成功すると、次は小型化を進めた。ワープロ開発に到るまでの森の強力な開発マネジメントの背景には、常に弓場の一言があったのである。

JW・10の開発にあたって、最後まで難航したのは編集機能部分のプログラミングだった。画面いっぱい文字を表示し、それに削除だの挿入だのといった編集を加える。これが意外なほど難しい。たとえば、画面の五行目で三文字削除したとしよう。当時は、五行目の最後に三文字分の空白ができるというのが一般的だった。六行目から繰り上がっていかないのである。

いまではどんな安いワープロでもこれくらいのはやつのけるが、当時のマシンの性能には負担の大きい処理だった。

「メモリーの少ないなかで、実用速度を確保しながら、これを実現するのが大変だった」と天野はいう。少しでも遅いと、たまに開発の様子を見にくる森に叱られるのがおちだ。

徹夜続きの毎日である。都築の強引な一言がここで効いてきた。その様子を見かねて、武田も天野のプログラミングを手伝ったのだという。

「最後までもめたのは、プリンタの位置やキーボードの細かなレイアウトでしたね」

と関係者は口を揃える。日本で初めての商品だ。モデルはどこにもない。プリンタを右に置くか左に置くか、文節／変換キーはどこに置くのがよいかといった議論が最後の最後まで続いた。最終的には、担当者間で「俺はここを折れるから、そっちもそこを折れる」といった妥協もあったようだ。

価格もギリギリで六三〇万円に決定した。報道発表は昭和五三年九月二六日のことであ

る（出荷は翌年二月）。型番はJW・10。開発から現在に至るまで東芝が登録した特許は国内八九件、海外二九件で、他に公開した国内特許が一〇六六件ある。まさに画期的な機械であった。

しかしそれでもなお、この新しい道具が、世の中をどう変えていくことになるのかを正確に予測した人はごく稀だったのである。発表翌日の一般紙の冷やかな報道ぶりがそれを物語っていた。ほとんどがベタ記事である。むしろ彼らの本当の戦いは、ここから始まるのだった。

（文中敬称略）

現職（本文に登場順）

栗原 俊彦／昭和四八年心筋梗塞のため逝去（五〇歳）
児玉 皓次／(株)東芝 情報処理・制御システム事業本部長附
森 健一／(株)東芝 情報機器事業本部技師長
溝口 哲也／(株)東芝 パソコン・ワークステーション事業部長
都築 公男／元ユニソル社長（元富士ゼロックス専務）
高橋 秀俊／昭和六〇年心不全のため逝去（七〇歳）
玄地 宏／平成三年呼吸不全のため逝去（五六歳）
長尾 真／京都大学工学部教授
河田 勉／(株)東芝 研究開発センター情報通信システム研究所第二研究所長
古賀 利郎／九州大学工学部教授
坂井 利之／京大名誉教授・龍谷大学理工学部長
天野 真家／(株)東芝 研究開発センター情報通信システム研究所主任研究員
相沢 輝昭／NHK放送技術研究所主幹研究員
稲永 紘之／九州芸工大学講師
黒崎 悦明／沖ソフトウェア株式会社専務取締役
江原 暉将／NHK放送技術研究所主幹研究員

副島 末好／西日本工業大学教授（元佐賀大学教授）
嶋 正利／ブイ・エム・テクノロジー（株）副社長
川上 晃／
武田 公人／（株）東芝 研究開発センター情報通信システム研究所主任研究員
田中 範夫／（株）東芝 記録媒体事業推進部長
植田 恭／（株）東芝 青梅工場ワープロ開発部課長
寒川 義明／（株）東芝 青梅工場パソコン製造部主任
月村 治雄／（株）東芝 青梅工場ワープロ設計部長
羽田野 修／（株）東芝 青梅工場品質保証部
武藤 治夫／（株）東芝 青梅工場ワープロ開発部
吉井 清／（株）東芝 青梅工場ワープロ開発部長
吉井 二郎／（株）東芝 青梅工場ワープロ開発部課長
弓場晋太郎／平成二年狭心症のため逝去（六四歳）
柴田 隆也／昭和四九年脳腫瘍のため逝去（四五歳）