

情報時代と日本語

榊 島 忠 夫 (大阪府立大学総合科学部教授)

近畿病院図書室協議会10周年記念おめでとうございます。その記念講演をさせていただくことは、大変光栄です。私は日本語を、特に文章とか文字とかを専門にしておりますので、そのことについて少しお話ししようと思います。

今から20年ほど前に、大阪の会社を対象に「事務能率を上げるために、文書作成について研修を行う予定があるか」という調査をしたことがあります。その時にほとんどの会社は、そういうことは考えていないと答え、ある文書課長は、そんな金があれば他にまわすとはっきり言われました。その後、「創造性の開発」がブームになりまして、私もひっぱり出されたことがあるのですが、それが下火になってきますと、今度は文章を書くことが大切だという論があちこちで起こってきまして、文章をどう書くかの研修にひっぱり出されることになりました。ちょうどこの会が創立されたころ、つまり10年ほど前から、また世の中の変化がひどくなってきました。ある時、私の家内が（関西医大の図書館に勤めておりますが）、カタログを持ってきて、こういう物を買うつもりはないかと、私の前におきました。見るとワードプロセッサでした。前から文書を能率よく作るための機械に関心がありましたので、それではと買ってみました。私がワープロを使い出したのは関西では一番か二番目かのよう

で、使っていることがたちまち日本中に伝わりまして、以前に機械翻訳などを工学部の人達と少しやったことがあるのですが、カナ漢字変換、つまりカナを入れて、それを漢字に換える方式の辞書を作られました。今も、お祝いの手紙やお悔みの手紙を自動的に書くシステム、またそれをもう少し進めて、論文やレポートなどを書く時に、こういうふうな順序でこう書きなさいというふうに、コンピュータを助けにして書くシステムの開発を考えております。さらに、それに従って書けば、後でその論文の要約が自動的にできるというようなシステムが、売り出されることになると思います。私は現在そのようなことをやっております。

日本語の多文字体系表記

さて、日本語について考える時に、我々日本人は当たり前だと思っておりますけれど、日本語の文字・表記は世界に類を見ない珍しいものです。例えば英語ですとローマ字（ラテン文字）だけを使いますね。ところが日本の文字は、漢字を使いますし、仮名も使う、それも平仮名と片仮名とを使います。さらにローマ字も使います。例えば「T シャツを着る」という短い一文の中に四種類の文字が出てきます。「T」というローマ字、「シャツ」という片仮名、「を」「る」という平仮名、そ

れに「着」という漢字です。

漢字なら漢字という文字の集合を文字体系といいますが、日本語の表記は、いろんな種類の文字を使って表記する、多文字体系表記であるわけなのです。しかも、その文字の読み方にもいろいろ問題があります。最近パーソナルコンピュータに「竹取物語」を入れまして、検索処理が自由にできるプログラムを作りました。「竹取物語」の中には「御」という漢字が出てきます。「オン」という字といましたが、実際にこれを何と読むのかわからないのです。「オン」とも読めるし「ミ」とも読める。それで京都大学を停年でお辞めになられました岩波文庫の「竹取物語」の校訂者の阪倉篤義さんにどう読むのかを聞きましたところ「オン」か「ミ」かははっきり決めがたいということでした。しかし我々が読む時には、意味が通るからオンだろうかミだろうかとあまり考えずに適当に読んでいます。日本語の表記のあり方だけでなく、我々の文字に対する反応自体も特異です。これを全部仮名に変えようと思うと、「オ」とか「ミ」とか「オン」とかを決めなければなりません。恐らく英語をしゃべる人ですと、それはきっちり決めないと気がすまないと思います。

漢字と日本語

一体なぜこのような珍しい表記になったかを考えてみますと、過去千年の間の日本人の態度が積み重なって、こうなってきたのではないかという気がします。日本には神代の文字があったという説を、平田篤胤あたりが出ておりますが、調べてみますとほとんど偽物です。我々の日本語には今、アイウエオという五つの母音がありますが、千年以上も昔には、どうもイとエとオにもう一つ別の発音があったらしいということが、万葉集を調べるとわかります。ところが神代の文字といわれているのは「いろは」と同じ発音しか表していないのです。そうしますと、平安時代以

後の発音を表しているということになり、偽物だとわかるわけです。

日本人は、四世紀頃に中国から漢字を借りて、それで文章を書くということを行いました。なにしろ中国語と日本語とは全然言語が違います。漢字を輸入しても日本語に合わないのです。なぜかといいますと、日本語には助詞・助動詞がありますし、それに活用する部分があります。「書く」「行く」なら「カキケコ」と活用します。中国語は単語を並べて文法的関係を示し、助詞や助動詞は使いませんし、活用もありません。そういうことですから、日本語を漢字で書くということは甚だ難しいわけです。

どのようにして日本人が漢字を日本語に合わせたかといいますと、それには二つの方法があります。その一つは、漢字が持っている意味を文字に当てて読み、逆に日本語の意味を漢字に当てて書く、訓読み・訓書きです。たとえば万葉集に(例文1)のような歌があります。

例文1 「^{この} ^か ^は ^ゆ ^{ふね} ^は ^{ゆく} ^{べく} ^{あり} ^と ^い ^へ ^ど
従此川 舟可行 雖在
^{わた} ^り ^せ ^{ごと} ^に ^ま ^も ^る ^ひ ^と ^あ ^り
渡瀬別 守人有

これは純粹な大和言葉に漢文的な書き表し方を与えたものです。昔から学者が苦労して読んでいるのですが、「従」というのは「ヨリ」で、昔は「ユ」ですから、「従此川」を「コノカハユ」と漢文風にひっくり返して読みます。「フネハユクベク」は「舟」に「ハ」を補って、「舟可行」と書き、「アリティヘド(雖在)」、「ワタリセゴトニ(渡瀬別)」、「マモルヒトアリ(守人有)」。五七七が歌の規則ですから、そのように当てたのだろうと考えて読むことができます。これは漢文風な書き方を日本語に当てた書き方です。

しかし東国の方言などで作られた歌を書こうと思うとこの書き方は不可能になります。訓書きは意味を表したもので、発音を表してはいません。そこでもう一つの方法として、

漢字の意味を捨てて、音を借りる書き方が使われています。

例文 2 「^ま久^は波^な波^は ^う津^つ呂^ろ布^ふ等^と伎^き安^あ里^り
^あ之^し比^ひ奇^き乃^の ^や麻^ま須^す我^が乃^の禰^ね之^の ^な我^が久^く波^は安^あ
^り家里^り」

このように漢字の音を借りて書き表す方法です。これは中国人が既に昔からやっております。仏教が渡ってきた時のインドの仏の名前や日本人の名前などもこのような方法で中国では書いております。朝鮮語と日本語はかなり構造が似ているので、訓読み・訓書きは恐らく朝鮮で行ってきたのを、日本人が学んだのだらうと思います。

このように日本語を漢字で書き表そうとすると、漢文風に書くか、それとも一字一音式で書くかになりますが、漢文風に書くと何と発音すればいいのかわかりにくい。一字一音式で書くと書くのが面倒くさいということになります。そこで日本人はその両方を取って、次のような書き方をしました。

例文 3 「^はる^のに^すみ^れつ^みに^とこ^しわ^れそ
^のを^なつ^かし^な ^ひて^よね^にける
野平奈都可之美 一夜宿二来」

アンダーラインを引いた字は音を借りて書いたもので、アンダーラインを引いていない字は意味を表したものです。ここで既に二種類の文字を使ったということが出来ます。例文 2 の音を借りた書き方を仮名と呼びます。仮名というのは仮りの「名」で、仮りに使った文字ということ。これに対して、意味を表す文字を、本当の文字、本当の漢字という意味で、真名といいます。両方とも形は漢字ですが働きがちがいますので、例文 1 の方を真名、例文 2 を仮名と我々はいっております。万葉集自体にもう既に、意味を表す文字（真名）と音だけを表す文字（仮名）の二種類を併用することを、例文 3 のように行っているのです。先程申しました日本語表記の特徴、漢字・平仮名・片仮名・ローマ字を使う多文字体系に至る、最初の出発点がここに出ているわけです。

平仮名と片仮名

しかしこのような書き表し方では、どれが「意味」を表し、どれが「音」を表す文字かは見ただけではわかりません。この万葉集が読めるのは五七五七七という区切があるからで、それに当てはめて読めるわけです。東大の電子工学の教授の話ですが、旅行をしていて旅館に泊まった時に、床の間に字が書いてある。それをどのように読むかといいますと、全体の文字数が五で割れるか七で割れるかという計算をします。五で割ってみて割れたら、五つごとに文字をまとめて、それをなんとか読んでみる。七で割れたら七つごとに区切って読む。こう言っておりました。中国文学をやっている人に聞きましたら、俺達もそんなことをやっているよと言っており、専門家も似たようなものだなと思いました。

しかし、歌ではなく文章を書いた場合には、どこが仮名でどこが真名か、わからなくなってきました。

ところで、当時一般の人々は用事をたすためには仮名を使っていました。法隆寺に仮名の落書があります。漢字は画数が多いものですから、これを仮名として書くときには、なるべくこれをくずして草書で書くようになり、草書から平仮名ができてきました。例えば、「保」という字を草書で書くと「ほ」になります。特に平安時代には、女性は漢文を読まないのが普通でした。紫式部などは、他の女房達から彼女は漢文読みだと言われていましたが、これはほめたのではなくて、悪口を言われたのです。女性は本当の漢字を知らないものですから、自由自在にくずして行って、今の平仮名ができたわけです。

片仮名の方は学者が創り出しました。お経、仏典、漢籍などを勉強するとき、今と同じように書き込みをしますが、その時に、やはり漢字で書き込まなければなりません。そこで漢字の形を省略して書きます。例えば「保」

を省略して下の部分を使って、「ホ」と書くわけです。このように学問の世界で片仮名を漢字の中に入れて、明治時代まで使われています。一般の人や女性は平仮名の中に漢字を少し入れて書くということをしてきました。

日本人は漢字を中国から輸入して、何とか日本語を漢字で書き表そうと努力してきましたが、平仮名・片仮名ができるまでに約500年かかっています。日本のインテリは漢文を勉強して、それを読み、書くのが本当の文章だと思っておりました。まして、今の国語審議会や文部省のようなものがなかったので、日本語を書くために固有の文字を創り出そうというようなことは、インテリは全然考えなかったのです。むしろインテリではない人達や、漢文の勉強を始めようとする人達の間で、平仮名・片仮名が自然発生的に出来上がったのです。このために、片仮名、平仮名が完成するまでにほぼ500年程かかっているわけです。

漢字は表意文字、意味を表す文字であり、表意文字だけでは日本語を書き表わせないために、表音文字、音を表す文字を創ったのです。それも自然発生的に500年もかけてできたものです。しかも、表音文字は1種類でいいのに、平仮名と片仮名という2種類もできてしまいました。全く無用なことです。ところが日本人というのはおもしろいもので、国語政策を昔から全くやらないのです。漢字があってそこに平仮名と片仮名ができました。世界的にみますと、表意文字から表音文字ができると、たいてい表意文字、意味を表す文字を捨てて、表音文字だけを使うようになっております。ところが日本人は、一度手に入れたものは絶対に捨てません。仮名ができたけれども、漢字は捨てない。平仮名と片仮名ができて、それを一つに整理しようという相談は誰もしないのです。もっとひどいことには、平仮名の内部、片仮名の内部にもいわゆる変体仮名が沢山あります。平仮名の変体

仮名は江戸時代に、版本、印刷をするために木を彫って作りますが、その関係で少しずつ整理されてきております。

漢字・平仮名・片仮名の併用

学者は漢字の中に片仮名を混ぜ、庶民は平仮名の中に漢字を混ぜて書いていましたが、片仮名と平仮名とを混ぜて書く文章が現れました。江戸時代の一般町人が読む町民の文学です。例文4は「東海道中膝栗毛」の文章です。

例文4 北八「ソリヤアおめへないらのおこつたのだ。豆をくやあなおる 弥二「エエわるくしやれずと、はやく出してくれろへ 北八「そんならまじ目に、ソレ田まちの反魂丹、手を出しな

この文章を見ますと、「ソリヤア」とか「エエ」とか「ソレ」とか、発音そのものに近いといひますか、言葉の中で言葉らしくないとこゝろに片仮名を使っています。日本人の庶民の中から、漢字・平仮名・片仮名を混ぜて書くことが起こってきたのです。学者は文字をいろいろ知っていますが、実はなるべく一種類の文字、漢字だけで書くのを正しい書き方と考へていました。何種類も文字を入れて書くのは、庶民的な書き方なのです。

このように、日本人というものは、文字を手に入れるとそれを捨てないで、また統一しようとしないうで、それぞれを上手に使ってきました。これはどういふことかと考へてみますと、例文4を見てもわかるように、「ソリヤア」とか「エエ」とか、感動詞や接続詞の類や、話し言葉でよく出てくるものは片仮名で書き、意味・概念を表す部分は漢字で書き、それ以外は平仮名で書いています。実はこれは、言葉の意味とか、言葉の種類、ニュアンスなどをできるだけ視覚的に書き表そうという傾向の表れなのです。日本人は言葉を単に記号によって書き表せばすむというのではなく、言葉の意味やニュアンスを視覚的に書き

表そうとする態度を持ってきたのです。

外来語（外国語）の取り入れ

それからもうひとつ、日本人の持っている態度を考えてみますと、外国のものを取り入れるのが非常に好きなのです。街に出ますとイタリア料理とか、スペイン料理、中華料理など、ありとあらゆる料理店があります。何でも食欲に取り入れようとし、手に入れたらそれを上手に使おうとするところがあります。中国から漢語を取り入れたのが、まず外来語を取り入れた最初です。それからオランダ語が入ってきて、そして英語などを取り入れています。同じ言葉なのに、「ガラス」（オランダ語）というときとおった物質で、「グラス」（英語）というときとコップであるとか、「ガム」（英語）はチューインガムで、「ゴム」（オランダ語）というときと伸びたり縮んだりするもの。そのように輸入した時期や国によって同じ語を使いわけています。

例文 5 「この分野の研究の中心的存在であるエール大学のコンピュータ・サイエンス学部長で人工知能研究所長の Roger Schank 教授が1980年に作ったのが Cognitive Systems Inc. (CSI) だ。CSI は、英語を理解できるデータベースのフロント・エンドを開発し販売している。」

この例文 5 は漢字・平仮名・片仮名・ローマ字を併用していますが、これは文字を混ぜて使ったというだけでなく、外国語をそのまま取り入れているのです。大正時代から昭和の初めにかけて、モダン語辞典・新語辞典という辞典が沢山出ています。そのモダン語・新語の半分以上が外来語なのです。日本人は常に外来語をよく取り入れて使ってきました。中国の留学生に日本語を教えていましたら、「さじ」というのは老人語ですかと言われて、びっくりしたことがあります。いや老人の使う言葉ではないかというと「スプーン」というではないか。若い人が「さじ」と言ってるの

を聞いたことがない。年寄りが「さじ」と言っているというのです。そう言えばそうだと思いました。私は「幻燈」という言葉を教室で使って学生に笑われたことがありますし、私が子供の頃に乗合自動車と言っていたのはバスになってしまいました。古い言葉が次々に外来語に置き換えられています。しかし、変わらない語もあります。「万年筆」はなぜ残っているのでしょうか。万年の筆ですよ。考えてみると実に古くさいですね。「鉛筆」。鉛の筆も変わらない。どうして鉛筆と万年筆は変わらないのか不思議です。

このように日本人は新しいものをどんどん取り入れていく、そして取り入れたものはあまり捨てないで、十分に利用していくという甚だエネルギーに満ち溢れた態度を持っています。

コンピュータと漢字

そういうところにコンピュータが現れてきました。このコンピュータが日本語にどのような影響を与えるであろうかということが、これからの問題ではないかと思えます。

漢字が日本語に入ってきた時から明治の初めまで、日本語の正式な文章は漢文でした。ところが幕末・明治維新の頃に外国語に接して、彼等が20数文字で文章を全部書き表しているのを見てびっくりしてしまうわけです。そこで漢字をやめてはどうかという意見が起って来ました。文章も言文一致にしようということで、明治の初め頃から日本語の文字をなるべく簡単にし、文章もニュアンスや感情が表せるような書き言葉にしたいという動きがありました。それは簡単に言えば、漢字をやめようという方向をたどっております。漢字を制限する気持ちは無くても、文章を話し言葉に近づけて、古くさい漢語をやめていきますと、だんだん漢字が減っていくことになります。そして戦後になって漢字が制限され当用漢字ができました。その後、コンピュー

タの出現が、漢字をやめようという理由のひとつになりました。コンピュータでは仮名しか使えなかったからです。これに対してコンピュータで漢字が使えないから漢字をやめるというのは本末転倒で、コンピュータが漢字を覚えるべきだと言う人がいました。乱暴なことを言うなと思っておりましたら、それがだんだんその通りになってきました。ワードプロセッサで仮名を入れてキーを押せば漢字に変わる仮名漢字変換が可能となりました。ワードプロセッサを使っているうちに、私は次第に漢字というものはコンピュータで書くものだという気がしてきました。漢字は、なかなか書きにくいもので、例えば「シノギョケズル」というのをどう書くかと聞かれると、ほとんどの人が書けない。ところが「鎬を削る」と字を見せると読める人が多いのです。画面に仮名を出してキーでそれを漢字に変換する。その時に変換された漢字が正しいかどうか判別できれば漢字は正しく書けるわけです。漢字はワードプロセッサで書くとうまく使えます。しかし、ワードプロセッサを使っておりますと、次第に機械を使わなければ手書きでは漢字が書けない、という状態になってきます。

日本人は、先程から申し上げましたように、表意的に、つまり意味とかニュアンスを表す書き方を非常に好みます。ですから恐らく、漢字をやめてローマ字だけとか、仮名だけの表記には絶対にしないと思います。そうするとちょっと矛盾が起こってきまして、手書きでは漢字は書けないが、読むのは読める、しかも漢字で書いていないと読みにくいということになります。そうしますと、漢字というのは手書きの文字ではなくなってきました。手書きで書く時には、漢字の部分は片仮名で書いて他は平仮名で書く、そして機械に持ってきてそれを漢字に換える、ということが起こってくるのではないのでしょうか。そうすると、漢字は減少する方向に向かうのではないかと

いう気がします。

現在までの日本語の表記は使用が多い文字から言いますと【平仮名・漢字・片仮名・ローマ字交じり】になっています。ところが情報化が進んできますと世界中の情報が多量に伝達されますから、明治時代には漢字を使って外国の言葉を翻訳しましたが、今は翻訳している余裕がない。特にコンピュータ関係の文章などがそうなのですが、政治でも経済でも、世界全体に同じ概念を使うようになりますと、いちいち日本語に翻訳するよりも、外来語のまま取り入れた方が間違いがないということになります。そうしますと、日本人の外来語好きがそこに加わって、日本語の中にどんどん外来語が増えていきます。外来語・外国語を書く場合に、片仮名で書くかローマ字で書くかということになりますと、外国語のできる人は片仮名で書いたのではおかしな発音になるので、ローマ字で書こうということになるかもしれません。そうしますと、漢語が外来語に置き換えられていきますので、これから使われる文字は【平仮名・ローマ字・漢字・片仮名交じり】の順に変わっていくのではないのでしょうか。ただし、文字の間で勢力の移り変りはあっても、漢字がなくなるということはないのではないかと思います。

コンピュータと表意的表記

日本人が書く文章は、今後ますますにぎやかになっていきます。例えば、ワードプロセッサを使うと、文字の大きさをいろいろ変えることができます。今日お配りした資料は、私が我が家のワードプロセッサで書いたものですが、ここで、標題の「情報時代と日本語」と名前の「樺島忠夫」に使った大きな文字、その下の「大阪府立大学」の小文字、本文の文字、それに万葉集の例文にふったルビの4種類の大きさの文字を使っています。ついながら申しますとルビというのは世界的に珍しいもので、文字に文字で注をつけると

いう非常に不思議なものです。このようにワードプロセッサを使う人には、どうしても文字の大きさを場合によって変えたいという気持ちになるものです。それから、図や写真を入れたいという気持ちもあります。もう少し進むと、意味の違いを視覚的に表すために、文字に色をつけることになるでしょう。カラープリントができるようになると、若い人達は、ポスターや雑誌に見られるように、文字を色を変えながらプリントし、赤い文字、緑色の文字、黒い文字でニュアンスを表すようになると思われます。例えば、怒った時には赤、驚いた時にはまっ青というように書きたがるかもしれません。ますます日本語の文章は、華かで、多彩になってゆくのではないかと思います。

そんなふうにかかれた文章をコンピュータで検索しようと思うと大変なことになります。的確に検索するには、一つの言葉に漢字と平仮名と片仮名とローマ字を全部当てて検索しなければなりません。その上に、色の区別やアンダーラインの有無などを入れますと、大変なことになります。これが面倒だから文章をすべて仮名に換えて入力しようとする、先程も言いましたように、竹取物語を仮名に換えようとしても、どう読めばよいかわからない、という漢字の問題が起ってきます。それからもうひとつは、日本人の外来語好みの問題です。論文を検索するために必要なディスクリプタを決めましても、それに属するものとして、外来語も漢語も大和言葉も当てて

おかなければなりません。こういう複雑なことが必要だから日本では技術が進歩するのだらうと思いますが、日本語の検索では非常に時間と金がかかるのではないのでしょうか。なんとか日本語を標準化する必要があります。標準化というと日本人はだいたい嫌うのですね。昔から文字を整理しなかったように、なりゆきを自然にまかせて、表現の自由を尊ぶ、楽しむというところがあります。日本人は大変な馬力で新しいものを取り込み、そうして出来た種々雑多なものを上手に使い分けているのですが、それが場合によっては、情報の処理にとってやっかいな問題となってくるわけです。そのへんをどうしたらよいかということは、文学者に任せておいたのでは仕方がないので、図書館のように、文字・言葉を扱う人に考えてもらうのがいいのではないのでしょうか。図書館の人は社会に対してそういう点からはあまり物をおっしゃらないようですが、日本語の知的な生産・処理のために、日本語の在り方について御意見を述べられてもいいのではないかと思います。

10周年、この間には情報とか、日本語をめぐる問題にずいぶんいろいろなことがありました。情報を生産したり、受け取ったり、またそれを処理したり、検索したりする。そのために日本語はどうあるのがよいだろうか。そういうこともこれから大いに研究して、我々のほうにもお聞かせいただければ、ありがたいと思います。