

GLOCOM Review

Volume 6, Number 5
May 2001

今号の内容

インターネット上の「書きことば」と言語計画

.....上村 圭介

2001年5月1日発行（第6巻第5号通巻63号）
発行人 公文俊平 編集人 土屋大洋
発行 国際大学グローバル・コミュニケーション・センター
Copyright (C) 2000 Center for Global Communications

GLOCOM Review は、国際大学グローバル・コミュニケーション・センター（GLOCOM）がその著作権を有するものであり、著作権法上の例外を除き許可なく全文またはその一部を複写・複製・転載することは法律で禁じられています。

インターネット上の 「書きことば」と言語計画

上村 圭介

目次

0. はじめに
1. 問題の諸相
2. 言語計画
3. 言語使用とインターネット
4. おわりに

要旨

インターネットは我々の「言語生活」のさまざまな側面において欠くことのできない重要な役割を果たすようになった。我々が日常行なう様々なコミュニケーションが、インターネットの助けを借りて行なわれるようになったのである。しかし、地球上のすべての言語話者がインターネット上の言語生活の中にいるわけではない。もちろん、その理由の一つは情報技術が十分に普及していないということもあるが、もうひとつに、インターネット上で、より根本的にはコンピュータ上でその言語を使うための前提である文字が、すべての言語について利用できないことが挙げられる。

この問題は、長年これまでコンピュータと文字の問題として考えられてきた。しかし、本稿では、この問題をより大きなインターネット上の「書きことば」の問題としてとらえる。そうすることで、この問題が単に言葉を書きあらわす技術だけにとどまらない「言語計画」上の課題であることが明らかにされ、本稿が、単に「文字が表示できる」「できない」という枠を越えて、情報技術と新しい書きことばについての議論をする土台を提供できることを期待する。

0. はじめに

地球上で使われている言語は、少なく見積もって数千語、多く見積もれば1万語を超えと言われる。しかし当然ながら、これら数千から1万強もの言語の置かれた状況は一様ではない。一部のヨーロッパ起源の言語のように、母語話者間の日常的コミュニケーションだけでなく、非母語話者も巻き込んだ政治、教育、通商、外交といった場で使われる言語がある一方、家庭内での話しことばとしてだけ使われる言語もある。

しかも、その中で文字をもつ言語は限られている。言語のほとんどは文字をもたない言語である。ところが、コンピュータ上で、そしてインターネット上で読んだり書いたりできる文字をもつ言語となると、その限られた言語の中から、さらに少数の言語に限られてしまう。Ranka Bjeljac-Babic (2000) がいうように、コンピュータで文字が扱えないということは、今やその言語が文字をもたないことに等しい。しかし、その理由は、コンピュータが英語圏で開発されたということにだけに単純に結びつけられるべきものではない。コンピュータで、ある言語の文字を表示できるということは、コンピュータと言語の問題のうちの最初の一段階の問題が解決されたというにすぎない。

一方で、コンピュータと言語、あるいはインターネットと言語の問題は、新たな注目を浴びつつある。例えば、2000年に開催された九州・沖縄サミットで「デジタルデバイドの解消」を目的に設立された Digital Opportunity Task Force (dot force) である。この日本のNPO 諮問委員会の中で、Kumon, Aizu and Peake (2001) は、デジタルデバイド解消のためのアジェンダの中に、言語と文化についての施策を盛り込むことを提案している。コンピュータと文字の関わりは、インターネットと言語という、より広い視野の中で議論され始めている。

本稿では、まずコンピュータと文字のどこが問題となっているのかをまとめた上で、応用言語学の一分野である「言語計画 (language planning)」について概観し、コンピュータと文字の問題が、単に技術的な問題ではなく言語計画の課題であることを示す。そして最後に、言語計画や社会言語学がこれまでに明らかにしてきたところに照らしあわせ、コンピュータが言語にどのような影響を与えるかを考察したい。

1. 問題の諸相

この問題を言語計画上の課題ととらえる前に、まずコンピュータと文字において何が問題となっているのか、ここで改めて概観する。コンピュータで文字が扱えない、ということにはいくつかの段階がある。ここでは、コンピュータで文字が扱えないということがどういうことなのかをいくつかの段階に分けて示

そう。ただし、ここで言うコンピュータの多言語処理とはあくまでコンピュータで不自由なく多言語の文字が利用できることとしておく。組版、スペルチェック、ハイフネーションといった言語的機能は、それより高次の機能として、本稿での考察からは省く。

1.1 文字の対応表

コンピュータで文字を扱うためには、ひとつひとつの文字と、その文字のための電子的表現（いわゆる文字コード）の対応表が用意されていなければならない。日本語の文字コード規格として長く利用されている JIS X 0208 では 6000 字を超える漢字のほかに、かな、数字、記号、その他の文字が記載されている。しかし、逆に言えば文字コードの対応表に記載されていない文字は、コンピュータ上で扱うことができない。日本語に限らず、文字コードを割り当てて対応表に記載する文字として何を選択するかというのは、極めて重要な問題であり、例えば日本語の場合は JIS X 0208 をめぐって、既に記載されている文字の妥当性と、記載されていない文字を追加するための方法という二つの面から大きな議論が起きた。

しかし、一方で、より重要なことは、その対応表が個々の利用者間で共有されていなければならないことである。対応表が共有されていなければ、例えば通信の送り手が "Internet" を送ったつもりでも、受け手にはそのことが伝わらず、例えば、"ln23s(s)" のような単なる記号の羅列としてしか受け取られないということが起きてしまう。世に流通しているコンピュータシステムの中には、BTRON をベースに開発された『超漢字』 <<http://www.chokanji.com/>> のように世界中の文字を扱うことができるものも存在する。しかし、このようなコンピュータシステムの利用者は、まだ限られている上、今後そのシステムの利用者が拡大するかは疑問の余地がある。また、インターネットを使うことがメッセージをやり取りすることである以上、メッセージの送り手と受け手の両方でその言語が使えなければ意味がない。

このように、文字と電子的表現の対応表である文字コードは、コンピュータで言語を処理するための基本的な要素である。日本語の場合には、文字コードが前述の問題を抱えているとは言え、すでに標準化された文字コードの体系があるが、クメール語やミャンマー語のように標準化された文字コードが存在しない言語では、この作業が急務となる。原田（2000）では、クメール語処理の最大点の問題点は、統一された文字コードがないことだとしているが、これは他の言語でも同様だろう。

1.2 文字のデザイン

文字と記号の対応表ができたとしても、それではまだ十分ではない。その文字がコンピュータの画面上に表示される場合に、またはプリンタから印刷される場合に出力される文字の形を電子的なデータとして与えなければならない。このデータは一般的には「フォント」と呼ばれている。

フォントを作成するには、その言語と文字についての専門的な知識や、デザイン上の技巧が必要であり、誰にでも容易にフォントを作成できるというわけではない。また、アジアの諸言語で使用される文字は一般的に数が多いため、この作業に必要な時間と労力がその分増加する。また、単に数が多いだけでなく、西アジアや南アジア起源の文字体系では、文字は、表記方向に対して直線的に並べられるだけでなく、文字を構成する筆画が他の筆画の周囲に付加されることで一つの文字として機能することがある。このような場合は、単独で使われる場合と、他の文字を構成する一つの筆画として使われる場合とでデザインや位置関係を変更しなければならず、フォントの作成はより困難となる。ヨーロッパ諸語では多数ある無償フォントが、アジア諸語ではそれほど多くない理由の一つはこのような文字体系としての複雑さにあると思われる。

また、文字の形を記述したデータがあっても、それだけでコンピュータにその文字を表示することはできず、コンピュータ側にその文字を表示するための機能が要求される。この点については後述する。

1.3 文字の入力方法

日本語であれば通常使われる日本語キーボードの一つ一つのキーには、英語のアルファベット、日本語のかな文字、その他の記号などが割り当てられている。この割り当ては、部分的に若干の相違があるものの基本的には共通である。キーボードは、利用者が入力したいと思う文字に対応するコードをコンピュータに指示するという基礎的な役割をもっているが、クメールやミャンマーのような言語の場合には、共通したキーボードの割り当てがないと言われている。

入力方法についてはもう一つ問題がある。それは、「文字変換」と呼ばれる部分である。日本語入力でなじみの深いフロントエンドプロセッサ (FEP) インプットメソッド (IM) などのソフトウェアは、この文字変換のための機能を提供する。これはキーボードで入力しきれない文字を入力するための方法である。日本語や中国語など、表語文字 (表意文字) である漢字を使用した言語では、入力すべき文字の数が非常に多いが、せいぜい 100 程度のキーしかもたないキーボードから直接すべての漢字を入力することができない。そこで、入力すべき漢字の読みを入力し、その読みに対応する漢字の候補を表示し、その中から目的の字を選択し入力する「漢字変換」とよばれる手法が一般的には用いられて

いる。このような「二重分節」を行なうことで、現在のコンピュータはキーボード上のキーの数よりも多くの文字を入力することを可能にしている場合が
おおい。

また、文字変換は、書きことばとして漢字を使用している言語だけに必要となるのではない。前述したように、前後の文字の環境によって、あるいは語彙的
弁別のために、ひとつの字を構成する筆画の配置が変化する文字体系の場合に
も、文字変換は有効に機能する。例えば、クメール文字がその一例である。音
節文字であるクメール文字は、子音文字を中心として周囲に母音文字、その他
の子音文字が付加されることで一つの音節を表わす。しかし、クメール文字の
場合には、一つの音節文字を構成するそれぞれの筆画が視覚的に配置される順
序と、その文字が表わす音節内の音素の発音上の順序が対応しない場合がある
と言われる。そのような場合には、発音の通りに入力すると、筆画が正しく配
置されない。一方で、筆画の通りに入力すると、入力の手順が発音を反映しな
い。このような複雑さを吸収するための仕組みとして文字変換が期待されてい
る。

なお、国際大学グローバル・コミュニケーション・センターが1998年から主宰
している「アジア太平洋フォント研究会（APフォント研究会）」では、カンボ
ジアにおけるクメール語の文字処理についての問題について検討してきた。活
動の詳細は次のウェブページに掲載されている。<[http://www.glocomnet.or.jp/
appf/](http://www.glocomnet.or.jp/appf/)>

1.4 文字とソフトウェア

文字の問題は、実はそれ以上の問題も内包する。現在の多くのソフトウェア
は、大抵の場合バイリンガルである。しかし、その意味は「英語ともう一つの
言語」を使えるという意味でのバイリンガルであるため、英語を含まない二つ
の言語を併用することは容易でない。

このことは、コンピュータのオペレーティングシステムやソフトウェア自体が
マルチリンガルでないということも理由の一部であるが、それ以上にソフト
ウェア開発者のマルチリンガル意識（能力である必要はない）が不十分である
ことが大きな要因である。英語話者の開発者が作成した文書作成ソフトウェア
は、大抵の場合、日本語の入力を受け付けず、そのままでは日本語の文書作成
に使うことができない。逆に、日本語話者の開発者が作ったソフトウェアでは、
中国語が処理できないということも起こりうる。

文字とソフトウェアの問題は、主として「表示ができない」と「入力ができ
ない」とこの二つの側面をもつ。しかし、コンピュータソフトウェアのレベ
ルでの多言語処理の機能は、近年大きな進歩を見せている。例えば、マイクロ
ソフトの Internet Explorer では、バージョン3の時点から、多言語サポートのた

めのソフトウェア部品が提供されており、これを導入することで、複数の言語によるウェブページを表示し分けることが可能となっていた。アップルコンピュータのオペレーティングシステムは、バージョン 8.5 以降、導入の時点で多言語表示のためのソフトウェア部品を組み込むことが可能になっている。また、無償で利用できることや、その性能から近年注目されている Linux 系のオペレーティングシステムでもソフトウェアが多言語処理のための仕組みを備えている。

このように、オペレーティングシステムや、ソフトウェア開発環境によって多言語処理を可能にする仕組みが内蔵されるようになるなど、多言語対応ソフトウェアの開発をめぐる状況は現在ではかなり改善されている。しかし、実際にはこれだけで問題が解決できるとは思われない。一つは、ソフトウェアが可能にする多言語処理が不完全であるかもしれないということが挙げられるが、それ以上にソフトウェアを作る「プログラマー」の意識や既存システムとの整合性への配慮によるところが大きいだろう。

この点を考える上ではコンピュータの「2000年問題」を参考にできるだろう。「2000年問題」とは、西暦の処理を4桁でなく2桁で行っていた一部のコンピュータソフトウェアが、西暦2000年以降の日付けや年次に基づいた計算を正常に処理できなくなり、連鎖的にその影響が社会生活に及ぶとされた問題である。コンピュータ業界は、2000年を迎える前に、この「2000年問題」への対応を迫られた。危惧されたような極端に大きな被害をもたらす事態には幸いなことに至らなかったが、それでも社会生活の一部には「2000年問題」に起因する小規模の障害が発生した（この点についての報告は、首相官邸ホームページの2000年問題に関するページ <<http://www.kantei.go.jp/jp/pc2000/index.html>> や日経 BizIT のホームページなどで今でも見ることができる <http://bizit.nikkeibp.co.jp/it/y2k/anchor/index.html>）。

2000年を迎える前には、多くのコンピュータシステムでは西暦を2桁で表さなければならぬ制約は取り外されていた。そのため、実際には「2000年問題」の原因となったプログラミング慣行は必須ではなかったと言えるが、実際には西暦を2桁で処理する慣行は2000年に至るまで継続していた。もっともこれは、単にプログラマーの教育不足や怠慢に帰せられる問題ではなく、既存システムとの整合性などの要因を考慮した結果、そのようなプログラミングにならざるをえなかったことが原因だという見方もある。

ともあれ、「2000年問題」は、この種のトラブルを回避するためにプログラマーの意識などの見えない「制約」を理解することが重要であることを示した。2桁で西暦を表わすのではなく十分な桁数によって西暦を表わすことで、「桁あふれ」に伴う障害を回避できることをプログラマーが理解していなければならないということである。

このことはまさにソフトウェアの多言語処理についてもあてはまる。現在のオペレーティングシステムは、改良され多言語処理を容易に行なうことができる機能をもっている。しかし、その機能をプログラマーがソフトウェアに組み込むには、その機能の存在や意味をプログラマーが理解していなければならない。例え

ば、風間（2000）は、多言語対応が充実している Java 言語の環境であっても、プログラマが十分その意味について理解しなければ、その機能がソフトウェア上に実現されないことを指摘している。

2. 言語計画

コンピュータと文字の問題は、1. で指摘したような重層的ないくつもの問題群から構成され、今でもそのすべてが解決されたとは言えないが、ソフトウェアの発展により技術的解決の可能性については拓けてきている。次に問題となるのは純粋に技術的ではない側面となるだろう。0. で紹介したように、言語がインターネットで「書きことば」をもつことの重要性に言及する意見や、Digital Opportunity Task Force (dot force) のように、インターネットと言語の問題をデジタルデバイド解消のための一ステップとして取り上げようとする場もある。家辺（1998）や上村（1998）のように、「デジタルテキスト」という視点、あるいは電子時代の「異本作り」という視点から、インターネットとことばの関係を理解するべきだという主張も見られる。コンピュータと言語の問題は、既に単に文字を表示できるかどうかを超えた場での議論をする段階に達したと言えるだろう。

本稿では、そのための枠組みとして言語計画を取り上げる。言語計画という視点をこの議論に導入することにより、技術的な側面だけからこの問題に迫るのではなく、広く言語が社会に与える影響、また社会が言語に与える影響という面から、コンピュータと文字、そしてインターネットと言語の問いを検討する機会が与えられることになるだろう。

2.1 言語計画の諸領域

公用語の選択、言語教育、文字政策など、社会におけることばの問題を解決するために言語やその言語がおかれた社会的環境に働きかける過程を社会言語学では「言語計画 (language planning)」と呼んでいる。言語計画は、ある言語を話す話者集団 (speech community) が抱える言語的な課題を解決するために、その言語がおかれている様々な条件に働きかけ、課題を一定の方向に導こうとする試みである。例えば Eastman (1983) は「言語計画とは、あるスピーチコミュニティが、その言語に対してもつ意図、目的、目標を実現するために行なわれる意志決定である」と述べる。典型的には、共通語や国語の選定という広範囲に影響をもたらす「マクロ」の働きかけが言語計画の例として挙げられることが多いが、スピーチコミュニティとは、国や地域の境界に隔たれたマクロな話者

集団である必要はない。例えば家庭内での言語使用についての選択というような「ミクロ」のレベルでも言語計画は行なわれている。

ここで言語計画についてもう少し細かく検討しよう。言語計画のマクロの枠組みについての全体像をまとめたものとして Kaplan and Baldauf (1998) がある。これにのら、言語計画の役割についてまとめよう。

この中では、言語計画は、席次計画と実体計画という二つの働きかけに大別され、さらに次に示すように、いくつかの小さな働きかけに分けられる。

- status planning (席次計画)
 - language selection (言語の選択)
 - language implementation (言語の普及)
- corpus planning (実体計画)
 - codification (規範の制定)
 - graphisation
 - grammatication
 - lexication
 - elaboration (規範の更新)
 - terminology modernisation
 - stylistic development
 - internationalisation

席次計画 (status planning) とは、社会におけるある役割を遂行するための言語の選択に関わる過程を言う。もっとも典型的な席次計画は公用語 (official language) や教育のための言語 (language for education) の選択の過程で行なわれる。席次計画は、異なる言語の間で行なわれることもあるが、ある言語がもつ複数の変種 (方言) の間で行なわれることもある。Kaplan and Baldauf (1998) では、選択された言語 (変種) を社会に普及させるための過程もこの席次計画の中に含めているが、後述の実体計画と並んで言語計画の三つの領域を構成するとする見方もある。

ついで実体計画 (corpus planning) が挙げられる。実体計画はさらに「規範の制定 (codification) 」という側面と「規範の更新 (elaboration) 」という二つの側面がある。規範の制定においては、言語のその時点での姿に基づいて規範が記述され、公的に承認される。規範の制定は、さらにその言語の文字の規範を定める graphication、文法を記述する grammatication、その言語の語彙を確定する lexication という三つの過程に分けられている。

規範の更新 (elaboration) は、記述され制定された規範を、社会的要請にあわせ充実させていく過程である。その中には、科学技術や社会制度の変化によって新たに必要となる新しい用語を選び出すための terminology modernisation (用語の現代化)、その言語が使用される場面に適当な文体を与える stylistic development (文体の創造) がある。また、ここで言われる internationalisation (国際化)

は、ソフトウェア開発で 'i18n' と略称される「国際化」とはことなり、現代の英語のように、スピーチコミュニティを離れて使用されるようになった言語に対して、そのような役割にふさわしい新たな規範を与える作業である。

2.2 インターネット時代の graphisation

graphisation は、ある言語を表記するための書きことばを確定する作業のことである。伝統的な言語計画において graphisation は、ある言語を書き表すために使用する文字体系 (script) の選択、または場合によっては文字体系の変更が最初の課題となる。文字体系とは、例えば、ラテンアルファベット、キリルアルファベット、デーバナーガリーといった広義の「文字」である。graphisation の過程では、他にも、正書法の規範化といった問題を扱う。

現代日本語でも、graphisation は行なわれてきた。代表的なものは、旧かな遣いから表音的な新かな遣いへの変更や漢字字体の変更といった、文字体系と正書法の変更だが、ローマ字運動やかな文字運動なども graphisation の領域に属するものである。また、当用漢字表、常用漢字表の制定、教育漢字の制定および変更も、graphisation に関わる働きかけと言える。

さて、コンピュータである言語の文字を使用可能にするということは、コンピュータやインターネットにおいて、その言語に書きことばを与える作業、あるいは、その言語の書きことばの活動領域を拡大する作業であると言ってよいだろう。このような視点からは、1. で検討したコンピュータと文字がもたらす具体的な問題点 (文字の対応表、文字のデザイン、文字の入力方法、ソフトウェア) を解決しようとする試みは、ある言語の電子的な「書きことば」を確定するための、graphisation に関わる作業だと言えることができるだろう。もちろん、音声による入力を可能にする音声認識の技術や、テキスト読み上げの技術は、コンピュータにおける話しことばの処理も大きな課題である。しかし、本稿の文脈ではコンピュータやインターネットにおける多言語処理とは書きことばの処理を意図している。

2.3 「国語国字問題」から言語計画へ

コンピュータと言語の問題は、これまでとかく扱える文字の数という視点に解消されがちであった。これは、現代日本語における言語計画の主たる対象が漢字制限の是非をめぐる「国語国字問題」であったことと無関係ではないと思われる。しかし、インターネットと言語の関係が単に文字の数の問題以上の意味をもつようになると、国語国字問題という枠組みでは不十分だろう。ここで、

言語計画という視点から、インターネットと言語の問題をとらえなおす必要が出てくるものと思われる。

確かに数年前に比べれば、現在主流の商用コンピュータシステムは多言語処理が容易に行なえるように改良されている。しかし、これらは、すべてコンピュータシステムとその関連ソフトウェアを販売するというビジネスが成立するために十分な数の話者をもつ言語の場合である。話者の数が十分でない言語では、このような商用コンピュータシステムのサポートを期待しにくい。既に述べた通り、インターネットで読んだり書いたりできる文字を持たない言語は、今や文字をもたないに等しい。この不均衡な状態を解消するために、われわれは何らかの働きかけを行なう必要がある、という認識が強まっている。そもそも、このような問題意識に基づいてコンピュータと言語の問題を理解することは、この問題を言語計画における課題としてとらえていると言ってよいだろう。また、文字コードの標準化を初めとした諸領域は、実際的な作業としても、ある言語にインターネットでの「書きことば」を与える試みであり、これはまさに電子時代での言語計画なのである。

3. 言語使用とインターネット

2. では、新たな「書きことば」を与える試みについて述べた。書きことばを与えられるということは、その言語が使用される場面が広がりうるということの意味する。しかし、実際上は、単に書く手段が与えられるということと、その言語が、話者の言語生活の中で実際に書きことばとして受け入れられることは直接結びつかない。例えば、社会言語学の分野では以前から二言語併用 (diglossia) という状況があることが指摘されている。Fasold (1984: 34) がまとめるところによれば、二言語併用とは、あるスピーチコミュニティにおいて、家庭、公の場、報道、文芸、学校といったコミュニケーションの場に応じて、上位変種 (high variety) と下位変種 (low variety) が使い分けられることをいう。これは個人の能力の有無を越えた社会的規範によって、言語の使われ方が制約されていることの例である。つまり、技術によって書きことばの技術的手段が与えられた言語であっても、そもそもその言語が書きことばとして使用される領域が自動的に作り出されるわけではないということの意味する。また、ある言語に書きことばを与えるということは、その言語がそれまでになかった使われ方をするということであり、その結果、それまでにはなかった語彙や用法が必要となる場合がある。一般的にこのような場合には、書くための手段を開発することと並行して、その言語の書きことばのための場面やジャンルを、ある程度人為的に作り出すことが必要となる。2. で紹介した、'terminology modernisation' (用語の現代化) や 'stylistic development' (文体の創造) は、このような目的でも行なわれる。

インターネットの上で、ある言語が使えるということ、あるいは使えないということが、その言語のあり方にどのような影響を与えるのだろうか。二言語併用は、使い分けの規範が比較的安定した状態であるが、それでも二つの言語の力関係は、ある一点で均衡しつつけるわけではなく、それぞれの言語がもつ威信 (prestige) やコミュニケーション上の機能などによって変化していく。

もっとも穏やかな均衡の変化は、二言語併用の場合のように、ある特定のコミュニケーション場面においてだけその言語が採用されず、それ以外の場面では、以前と同様に使われつつけるという場合である。安定した二言語併用の場合には、このような言語の住み分けが長期間にわたって持続するということもあるが、このような安定状態はむしろ稀である。また、このような言語の住み分けは、書きことばや公の場での使用といった社会的権威の高い場面で使用される上位変種と、そうでない日常的な場面において使用される下位変種という住み分けである。

ところが、インターネットは、公の場でのメッセージの伝達から、親しみ深い日常的な会話まで幅広いコミュニケーション領域で使われる。最近とくに盛り上がりを見せている、ICQ、Yahoo!メッセンジャー、AIMなどのp2pインスタントメッセージアプリケーションや、ウェブ上や専用アプリケーションによって自由にユーザ同士のコミュニティを作り出すことができるグループ形成アプリケーションの急成長をみると、インターネットは、より社会生活に浸透した、日常的なコミュニケーションの手段となりつつある。そういう視点からも、二言語併用的状況がうまく機能しないことも予測され、話者にとってはその言語ができるだけ多くのコミュニケーション場面で利用できることが望ましいだろう。

また一方で、すでに述べたように、ある言語に書きことばが与えられたということだけで、その言語が、書きことばを要する場面すべてにおいて使用されるということの意味するのではない。日本語でも、東京方言をもとに様々な方言変種を取り込んで成立した共通語は、書きことばが必要とされる様々な場面において使用されるが、それ以外の方言変種は、文字によって書き表わすことが可能なはずであっても、そのことを越えた社会的な規範によって、通常は文字に書かれず、書きことばとして流通することはない。かつてワールドワイドウェブが登場したころ、英語でなければインターネットに載せる意味がないと思われていた時期があった。これは、インターネットという使用場面において、スピーチコミュニティの言語とは別の言語が使われるべきだという規範意識が存在したことを意味する。現在では、このような規範意識はかなり薄くなっていると思われるが、ある言語に対して、1.で述べたような、インターネットで書きことばとして使うための技術的解を与えたとしても、それだけでその言語がインターネットの言語として使用されるとは限らない。社会的な規範、文体、場面といったその言語が書きことばとして使われる文脈も同時に与えなければならないのである。このような問題は単純に電子的 graphisation を実施しただけで解決される問題ではなく、例えば、文体の創造や用語法の現代化を通じて、そ

の言語が電子的コミュニケーションを担えるように方向付けていく必要があるだろう。

言語の使い分けについての社会的な規範という観点からすると、インターネット通販業者である Espanol.com が実施した調査の結果は興味深い (“ Online Hispanics Favour Sites in English ” <http://www.nua.ie/surveys/?f=VS&art_id=905355509&rel=true>)。この調査では、アメリカに在住する 500 万人の英語とスペイン語のバイリンガルのインターネットユーザの間では、スペイン語のサイトよりも英語のサイトを好む傾向があるという結果が出たという。一方で、そのようなユーザであっても、スペイン語の書籍や音楽への関心は高いという。つまり、スペイン語のモノへの関心はあるものの、インターネット上ではスペイン語よりも英語のほうが便宜性や信頼性が高いというよう分裂した判断が行なわれているのだと見ることができるだろう。

また、2000 年 12 月に発表された調査結果 (“ 12.3 million web users in Canada, just 2.8 million are Francophones. Language divides Canadian web users ” <http://www.nua.ie/surveys/?f=VS&art_id=905356228&rel=true>) によれば、二言語併用の国であるカナダ (人口約 2700 万人) に 1230 万人いるとされるウェブ利用者の中で、フランス語話者は 280 万人だけだという。カナダは、フランス語話者の数は英語話者と比べて多くないが、2000 年 5 月から 10 月にかけての利用者数の変化を見ると、英語話者の利用者が 22 パーセント増加しているのに対し、フランス語話者の増加はわずかに 2 パーセントだという。ここでも、英語話者とフランス語話者の間で、インターネットへの期待が異なっていることが見て取れる。

これらの調査結果だけで何かを言うことは難しいが、これはすでにインターネット上で、言語に対する期待や威信が変化してしまっていることを意味するのではないだろうか。

今やインターネットは英語話者だけのものではなく、近年インターネットユーザにおける非英語話者の数の増加は著しい。Bangkok Post (2001) によれば、ガートナーグループが行なった調査で、2000 年に 4180 万人とされたアジア太平洋地域におけるインターネットユーザは、2001 年末までには 72 パーセント増の 7200 万人に達すると予想されている。

さらに、インターネットで使うことのできない言語は文字をもたないのと同じだ、というのは、単にその言語がインターネットで使われないということにとどまらない。現代の社会生活においてインターネットによるコミュニケーションのもつ意味は非常に大きくなっており、実際には、その言語がインターネットで書きことばとして使われないだけでなく、インターネット以外の社会生活にも、その影響は及んでいく。その影響としてもっとも大きいのは、言語の死 (language death) である。インターネットにおける書きことばの不在が、その言語への評価や期待を低下させ、その結果、言語の死にまでは至らないとしても、使用場面の限定、話者数の減少という状態を導き出す。

国連環境計画 (UNEP) が、2001 年 2 月にナイロビで行なった報告では、地球上に 5000 ~ 7000 あるとされる言語のうち、2500 もの言語が 21 世紀中に消滅す

ると警告している（『朝日新聞』2001年2月12日付、ほか）。また、The International Clearing House for Endangered Languages <<http://www.tooyoo.l.u-tokyo.ac.jp/ichel.html>> には、話者の減少により、言語の世代間継承が困難になった危機言語（endangered languages）のリストがある。これによれば、次に示す数のが言語の死に瀕しているか、その可能性が高いとしている。

- ・ 南アメリカ: 45 言語
- ・ アフリカ: 186 言語
- ・ アジア・太平洋: 79 言語
- ・ 北東アジア: 42 言語
- ・ ヨーロッパ: 83 言語

言語の死の直接的な原因は、話者の減少であるが、間接的には、その言語の使用場面や機能が限定されることや、その言語がもつ威信や期待の低下がある。今後、インターネットによるコミュニケーションが社会生活の大きな割り合いを担っていくとするならば、インターネットにおける「書きことばの不在」は、これらの少数派言語の消長について少なくない影響をもつだろう。

あらたなメディアの出現で言語の使用領域が拡大したという点では、放送や出版も似たような役割を果たしている。しかし、出版、放送というメディアとインターネットとは、個人に与える増力（empowerment）の度合いが大きく異なる。インターネットは、広く社会に訴えかける力をもつ一方で、一人一人に与えられた表現の手段でもある。それに比べて、出版や放送は、ある程度の規模をもつコミュニティにおける表現の手段である。

出版や放送においても、言語使用は決して自由ではなかった。例えば、Clyne（1991:12）では、第一次世界大戦から1970年代までのオーストラリアにおいて、LOTE（language other than English）の公的使用が制限されてきたことが紹介されているが、この例に見られるように、出版や放送というメディアにおける言語使用も、決してフリーハンドの自由を得ていたわけではなく様々な理由によって制限されてきた。とはいえ、インターネットは、よりパーソナルなコミュニケーションの担い手としても使用され、ある個人にとっての表現手段としての意味をより強くもつため、出版や放送などと比べても個人の日常レベルの言語生活に与える影響がより大きいと言える。その意味でも、インターネットにおける書きことばの創出はすべての言語にとって重要な課題であることが分かるだろう。

4. おわりに

これまで、インターネットにおける言語使用の問題は、単に文字の表示や入力ができることにとどまらない影響をもつことを見てきた。不幸にして日本における言語計画の大半は、漢字問題に費やされてきた。漢字廃止論、漢字制限論、漢字復興論というグラデーションはあるにしろ、日本における言語計画（特に、マクロなレベルの言語計画）は、一貫して文字を対象としていたと言える。その影響もあってか、コンピュータと言語、特にコンピュータと日本語との関わりは、常に文字コードを中心に論じられることが多かった。もちろん、筆者は文字コードが言語のコンピュータ処理にもつ根源的な重要性を過小評価しているわけではないが、コンピュータと言語の関わりを、文字の多寡だけに還元してしまうことは望ましくない。既に主張してきたように、コンピュータと言語の関わりは、文字の多寡だけでなく、インターネットという文脈において、新たな「書きことば」の問題として問い直す必要がある。

家辺（1998）は、「デジタルテキスト」という新たな問題設定を行なうことで、「文字が出るか、出ないか」という矮小化された議論からコンピュータと言語の問題を問い直している。

「……文字コード議論が白熱する中で、デジタルテキストの問題をことごとく文字の問題に還元するような論調が見られるのは奇妙なことである。文字はデジタルテキストという問題領域の一部分でしかなく、文字のあり方はテキストが作る文脈に大きく依存している。また文字は独立した存在物ではなく、人間の読み書きというプロセスに組み込まれたものである。……」（150ページ）

このようにコンピュータと言語の問題は、より大きな言語的営みの問題としてとらえることで、より適切な議論の場を生み出すことができるだろう。

コンピュータと文字の問題は、これまで主として技術的視点から解決されようとしてきた。これは、たとえば言うなら、印刷術の発展において、活版技師だけがその問題を論じてきたようなものである。しかし、コンピュータと文字の問題の解決には技術的視点だけでは十分でない。コンピュータの画面上で文字が表示できるようになるというだけでは、コンピュータと文字というより大きな問題のほんの一部を解決することに過ぎない。この問題は、まさに、言語計画の問題として、単なる技術者だけでなく、スピーチコミュニティ、話者、政治的意志決定者といった様々な参加者の価値観を考慮して解決されるべき問題なのである。

上村圭介（かみむらけいすけ）
国際大学 GLOCOM 主任研究員

参考文献

- ・ 『朝日新聞』 2001年2月12日「5000～7000のうち2500の言語『絶滅』の危機、国連環境計画で報告、グローバル化が背景」
- ・ Bangkok Post. 2001. 'Asia Pacific Net Subscriptions To Top 72m In 2001 - Study' (2001-01-03). <http://www.newsbytes.com/news/01/160014.html>
- ・ Clyne, Michael. 1991. *Community Languages: The Australian Experience*. Cambridge, England ; Melbourne : Cambridge University Press.
- ・ Fasold, Ralph. 1984. *The Sociolinguistics of Society*. Oxford: Blackwell
- ・ 原田 至朗 (2000)「コンピュータ上におけるクメール語文字利用の現状と改善試案」『東洋文化研究所紀要』第139冊
- ・ Kaplan, Robert B. and Richard B. Baldauf, Jr. 1998. *Language Planning from Practice to Theory*. Multilingual Matters.
- ・ 上村 圭介 (1998)「電子化テキストの諸問題 『中国の環境研究文献集をもとに』」『GLOCOM Review』1998年10月号(33号)
- ・ 風間 一洋 (2000)『Cafe Babe - Java プログラミングノート：国際化と日本語処理』アスキー
- ・ Kumon, Shumpei, Izumi Aizu and Adam Peake. 2001. "Comments on the Draft Report of the DOT Force 2.0". <http://www.glocom.ac.jp/dotforce/j/20010417glocomcommentto20.html>
- ・ 'Okinawa Charter on Global Information Society'. 2000. <http://www.dotforce.org/reports/it1.html>
- ・ Ranka Bjeljic-Babic. 2000. '6,000 languages: an embattled heritage'. http://www.unesco.org/courier/2000_04/uk/doss01.htm
- ・ 家辺 勝文 (1998)『デジタルテキストの技巧』ひつじ書房