

専門用語研究

Journal of the Japan Terminology Association

No. 5 1993-07

目次

第6回専門用語シンポジウム記録：専門用語の生態学

会長あいさつ	大塚 明郎	1
シンポジウム：専門用語の生態学		2
講演1 専門用語辞典の体系的構成法	長尾 真	3
資料：専門用語辞典作成法試論	長尾 真	10
講演2 用語の造語メカニズム	日高 達	20
講演3 専門用語の命名における諸問題	森岡 健二	26
パネルディスカッション	司会 横井 俊夫	33
パネリスト	長尾 真・日高 達・森岡 健二	

専門用語研究会
Japan Terminology Association

専門用語の生態学

1992年12月5日
東京・四谷 主婦会館

会長あいさつ

大塚 明郎

今回われわれの第6回目の「専門用語シンポジウム」を開きますに当たりまして皆さんご一同お集まり下さったことに、本当にありがたく感謝いたします。

私たち、おおげさにいえば「人間が人間である」というのは、言葉をしゃべるから人間である」という人間の定義に言葉が入ってくるというふうにいわれておりますが、実際、言葉のない生活というのは、われわれの生活にはあり得ない。どの方々も、ここにいる方ばかりでなくて、一人前の人間になれば職場にいる。職場にいと必ず職場の言葉というものを使うというのが当たり前になってしまっ、その時に特に使われる言葉を仮に「専門用語」といい、それを使うチャンスが多くなってくる。そうすると、その間にはいろいろな規則や約束なんかを作って使おうということになってくる。もっとも、私たち、人によって違いましよ

うが、少なくとも私にとって最も重要な約束事は、特に日本では学術用語集に集められた用語だろうと思います。ほかの言語体系を見ましても一つの言語集の中で日本語のようにきれいに、と私はいいたいのですが、一つの用語集をもっているという国は、極端にいわばないんじゃないか。われわれが一番いい用語集を持っているんじゃないかと私は考えるのです。それができたからこそ、現在の科学技術というか、実際の現場における科学技術化というものが今のように発達したと。つまりわれわれが今のように日本人が日本人らしく生きていけるというのは、この用語があったからできたんじゃないかと、そういうふう考える次第でございます。したがって、皆さんもご研究なさっていらっしゃるでありましょうし、またご意見もお持ちだろうと思います。今日は特に実際に使う場面においてどういうふうに使われる、どんなふうなものが使われているかということについて、日本を代表する方々のお話を聞くチャンスを持たたのではないかと思います。ご一緒に話し合いができればありがたいと存じます。

シンポジウム「専門用語の生態学」

●柴田 武（東京大学名誉教授）

今日は司会をおおせつかっているわけですが、私自身は専門用語の専門家でもございません。ましてや司会の専門家でもありません。なぜ司会をするかといいますと、いろいろないきさつがありまして、特に今日のシンポジウムの企画をいたしました。横井さんと一緒に企画をしたという、その関係でここに座っております。どうかよろしくお願いいたします。

今日は、すでに印刷物などでお知らせしてあるように、お三人の先生方、京都大学の長尾真先生、九州大学の日高達先生、それから上智大学名誉教授の森岡健二先生に、大きな題目として「専門用語の生態学」というようなタイトルのもとにお話をさせていただき、ディスカッションしようという計画でございます。実は、「専門用語の生態学」というのは私の責任でつけた題ではございません。生態学というのは、私は素人でどういうことかよく分かりません。ぼんやりとは知っていたつもりですけれども、この際百科事典で勉強しようと思って読みましたところが、生態学の定義がまだはつきりしてないそうであります。これはしめたと思ひまして、今日は新しい定義に従って、つまりこういうお話をさせていただく、それが生態学。物事のシステムそのものを生態的に見るのではない、あるいはできてきたいきさつを問題にするのではなく、われわれの専門用語というものがどういふふうに使われているか、それを見るのが生態学と定義いたします。その、どう使われているかということのうちでも、今日は特に専門用語をどう作るか、どう組み立てるか、というところに焦点を当てております。すでに講演の題をご覧下さってお分かりのように、長尾先生の場合は、構成法つまり作り方という題がついておりますし、日高先生の場合は、造語メカニズム、言葉をつくるというメカニズムのお話になるかと思ひますし、森岡

先生の場合も命名、名前をつけるという、そういうようなところで、お三人とも専門用語を作るという視点からのそれぞれの専門分野に関するお話が伺えるかと思ひます。

では、最初に長尾先生から「専門用語辞典の体系的構成法」と題してのご講演をお願いいたします。

以下の各講演は、シンポジウムでの講演原稿を加筆。

講演 1

専門用語辞典の体系的構成法

A Systematic Construction of a Terminology Dictionary

長尾 真* NAGAO, Makoto

ただ今ご紹介いただきました長尾でございます。私は専門用語の専門家ではありませんので、お話をさせていただくのは恐縮に存ずるわけですが、2年ちょっと前の1990年に、岩波書店から「岩波情報科学辞典」というのを出版いたしましたので、その時の経験について話をすることということで責任を逃れさせていただくと、こういうことで参上いたしました。

お手元にちょっとした資料「専門用語辞典作成法試論」というのがあるかと思います。これは1990年7月と書いてございますが、「岩波情報科学辞典」を5月に出しまして、まあ5、6年かかったわけですが、その時の経験といえますか、いろいろやったことをまとめておいたほうがいいかと思いついて作りました資料でございます、どこにも発表はしておらなかったんですが、今回こういうチャンスを与えていただきましたので、ちょっと古い内容で申し訳ないのですが、これをコピーしてさしあげるということにいたしました。

辞書のオーソリティーというのは、司会をいただいております柴田先生をはじめここにたくさんいらっしゃるのではないかと思いますので、辞書について云々するという事は誠に気がひけることなんですけれども、まあ辞書にはいろんな種類があるということをご承知のとおりです。ここで申し上げたいことは、説明の形態、あるいは冊子の形態ということでございまして、普通の一般用

語の辞書と専門用語の辞書はやはりかなり違う。百科事典的な要素もあるわけですが、そこには大きな二つの態度、すなわち大項目主義と小項目主義があります。最も典型的には「Encyclopedia Britannica」が大項目主義で有名な辞書でございます。「Encyclopedia Americana」はどちらかというと小項目主義で成功しているものである。中項目主義というものもあるかも知れませんが、まあ大きく分けるとそういうことになると思われます。そのほかに、特に技術分野におきましては、ハンドブック形式というのがありまして、これは学問体系に沿っていろんな説明がしてあってハンディに使えるというものでございます。辞書の形態としてはこのような違いがあると思われ

ます。それから、辞書のスタイル、Lexicographyといわれているものがあります。それにはやはり配列法というのが最も大事な問題として挙げられると思います。どういう配列法を採るかというのは、つまらぬ話に思えるかも知れませんが、ユーザーからみた場合に非常に大事な問題であります。最近「広辞苑」の逆引き辞典なんか出ておりますけれども、あれも一つの配列法であると考えられることができるわけでありまして。それから相互参照がどの程度なされているかという問題。これも辞書の性格を決める一つの要素です。作ろうとしている辞書によってどの程度相互参照をやるのが適切かということもある。それからどういう人たちが使おうとしているか、あるいはどういう人たちを目標にしてその辞書を作るかということが非常に

* 京大工学部教授

大きな問題になるというわけでございます。

もう7、8年前になりますが、「岩波情報科学辞典」を作るということになりました時、以上のようなことをまず考えまして、こういういろんな観点のどこに絞ってやるかということを決心したわけでございます。一般の言葉の辞書と専門用語の辞書は明らかに違う。これは、ここにいらっしゃる専門用語の専門家の皆さま方には、釈迦に説法だと思えますけれども、やはり何といひましても専門用語の場合は、一つの専門分野の中に概念というものが作られ、その概念に対してそれを表現する言葉が与えられる。それに対して一般用語の場合はまず言葉ありきということではないかと思えます。そういうふうになんか逆になっているということでもあります。概念の数だけ言葉があると考えてもいいかもしれません。そして概念は、学問の場合には定義がなされているとみれると思えます。一般用語の場合には、定義がなされているとはなかなか考え難いわけですが、専門用語、学問の用語の場合には定義がなされていると考えるわけです。その定義がどういう言葉によってなされるかという、やはり同じ学問分野の中の他の概念の組み合わせによってなされるということにならざるを得ませんから、学問は、概念相互間の非常に複雑なネットワークの形をしている。つまり、ある一つの学問フィールドはそのフィールドにおける概念の相互関係を表現する非常に複雑なネットワークであると捉えることができるんじゃないかと思うわけでございます。その時の関係というものをどういう相互関係——上位、下位概念とか、そのほかいろいろな関係があるわけですが、——として捉えたらいいか、ここが一番難しい問題でございませう。

そういうことをよく考えまして、辞書作りを始めたわけでございますが、まず何といひましても、辞書を作る場合には単語を集めないといけないということになります。特に情報科学のような新しい分野の場合には、既に確立された形の辞書というのはほとんどありませんので、どういう形で単語を集めるかというのは何もなしのところから考えなければいけないということになります。そこで、私がやりましたアプローチ、あるいは私が編集者

の皆さんにお願いしたアプローチといひますのは、情報科学の分野を最初大きく五つか六つの分野に分けて、つまり理論とか情報のハードウェアの部分、ソフトウェアの部分、それから情報プロパーの部分とかいうふうに分けて、その中をまたいくつかに細分していくというものであります。初め二十幾つか三十かに分けて、そのあと六十くらいに分けて分けた。これは比較的容易に分けられましたけれども、ただ、どこまでの範囲を情報科学の範囲と考えて分野に入れるかということについては、編集者の人たちといろいろ議論しまして、まあ大体こういう範囲を入れようと決めたわけでございます。応用が中心になっている分野はあまり入れないでおこう。たとえば、社会学とか経営学とかいったところで情報技術が使われているが、そういうものは分野としては入れないといった具合です。その六十くらいの分野というのは、例えば計算機ハードウェアの中でもうちょっと細かくなっておりまして、計算機アーキテクチャーであるとかメモリであるとか、それから情報の分野でありますと、計算言語学であるとかパターン認識であるとか、情報学の分野であるとか、そういうふうな分野であります。それぞれの分野につきましてはかなりの本が出ております。つまり、辞典を作るということはある程度その分野が熟してきているということだと思われまますから、そこには必ず典型的、標準的な本があります。そういう本の中で最も教科書的でスタンダードといひますか reliable と思われる本を何冊か集めてきて、その本の各 chapter を調べる。そうすると、その分野がどういうふうなサブフィールドに分けられているかということが客観的な形で大体分かる。というようなことで、そこからまたサブフィールドを作ります。そうしていきますと何百かのサブフィールドができてくる。そして、そのあとは専門書をもっともってくる。その分野のスタンダードな教科書のほかに研究者のための本を何冊かまたもってくる。それからその分野の学術雑誌の中で典型的なものを見ます。大体そういう学術雑誌には survey 論文というのが何年かに1回載っておりますから、そういう survey に関する論文なんかを中心に集めて、

そういった論文や本の中に出てくる用語を網羅的に集めるというようなことをやりました。それから、そのほかにここに書いてありますいろんなハンドブックであるとか科学技術大辞典とかいろんなものがありますが、そういう用語集をもちろん見ます。

そういうふうにして単語を集めていきますと、自ずからその用語群が一種の階層性をなす。概念地図というような言い方をしましたけれども、簡単にいいますと Tree 構造をなす。これを「用語の木」と名づけましたけれども、基礎とハードウェアとソフトウェアと知識システム、それから情報と社会というふうに分けて、例えば知識システムですと人工知能とか計算言語学とかパターン情報処理とか、こういうふうに分ける。それからハードウェアの場合ですと、ハードウェアの案子、装置、方式、それからネットワーク的な通信の問題、ハードウェアの設計に関するテクノロジー、それぞれが一つの、まあ大学でいいますと1学期か2学期間の授業の内容になっております。こういうふうにして、大体2万語の単語を集めて tree の深さでいいますと、ちょっと忘れましたが8段階くらいまでの深さの用語の木を作りました。そしてそれを、みんなでディスカッションしまして、大体どのレベルまでの単語を採用するかを考える。tree の下へいけばいくほど概念が小さいとか、specific な概念になっている単語とみなすことができますので、辞書に入れるということを考えると、大きな概念の方から入れていった方がよいので、上の方からどの辺の枝までを採用するかといったことを考える。それを考えることによって、tree の作り方がある分野では浅いけれども、ある分野では深いというのを、修正することができまして、どの分野をとっても大体同じ程度の深さの tree が作れるように、tree を妥当な形に変更する。そして大体6段階くらいでしたか、ちょっと忘れましたが、そのくらいのところを採ります。これは出版社の考えかたによるというか、何ページくらいの分厚さの本を作りたいという考えがあるようですから、——まあこれは電子的な本の場合だともうそういう心配はほとんどないのですが、冊子体でやる場

合には値段とかいろいろな問題があるようですから——やむを得ずそういうことで4,500単語くらい採りました。もちろんその中には tree の下の方にある単語だけれども、ポピュラーであるとか、あるいは話題を賑わす単語であるとか、そういう単語も採る。それから、最近のいろいろな新しい journal の中とか「現代用語の基礎知識」とか「イミダス」とかいうものを調べて current な単語で大事なものを採ってきて tree の中に埋め込んで、これを採用するか採用しないかを決めるとか、いろんなことをやまして、まあ4,500くらいを見出し語にしました。しかし、4,500ではとてもわれわれは満足できない。見出し語は4,500にしたんですが、一つの見出し語の解説を書いている中にもっと specific な単語をどうしても使わざるを得ない事態がでてくる。そういう単語を使う時にはそこでその単語の説明、定義を与えることにして、見出しにはしないんだけど、索引からは必ず引けてその単語の定義が分かるというような書き方をする。そういう形で説明を与える単語を別に8,000語ほど指定しました。これは用語の木の下のほうに出てくるもので、ある見出し語の説明の中ではこれこれの用語にもふれ、そこで定義を与えるように書く人に提示をしたわけです。このようなものを入れると合計で13,000くらいの単語を見られるようにした。ですから索引から入ると13,000くらいの用語が何らかの形で定義が与えられているという辞書にしたわけです。用語の収集はこのような形で行いました。

そこで次にどういう定義の仕方をするか、つまり内容説明をどうするかということになるわけです。それにはどういう説明の仕方をするかが問題になり、定義とは何か、あるいは説明とは何かということを考えなければいけません。これについては平凡社の「哲学事典」の項目を調べるという非常に面白いことが書いてあります。あまり説明する時間はないんですが、「説明の理論」というのがあって、事象の説明はどのようにするものであるかということがちゃんと書いてある。これがベストのことかどうか分かりませんが、論理的に考えた時に説明とはどういうものであるかということは、確かにこれでよく分かる。つ

まり、ある事象の説明というのは、条件とそれに関わる法則とを明確にして、その結果どういことがいえるかということを推論することである。その説明にも、演繹的な説明であるとか、帰納的な説明であるとか、歴史的な説明であるとか、いろいろな説明の方法がある。このようなことが明確に書いてあります。それから、私も前からそう思っておりますが、説明を与える、あるいは定義を与えるというのは、necessary and sufficientな定義を与えないといけないわけですから、類似した概念、類似した単語とどういふうに似ていて、どういふうに違っているかという、その境界を明確にする。あまり強すぎる定義をすると、範囲が狭くなってしまって隣の概念との間にギャップができるというような、まあ、非常に抽象的な言い方で申しわけないのですが、きつすぎる定義とか弱すぎる定義でないように注意しなければいけない。つまり定義の限定範囲をどうするかということに常に注意しなければいけない。また、概念としても、個体概念であるか普遍概念であるか、どうい概念であるかということに明らかにしなければいけないとか、いろいろな定義の仕方に関する研究をしたというか、調査をいたしました。最も典型的には、内包的定義と外延的定義とその他というようなこととなります。そこで定義をする時に、循環を避けることも大切です。

こういうことをいろいろ考えました末、具体的には次のようにやりました。

ある単語、ある専門用語がありますと、その専門用語は情報科学という学問分野の中のどうい位置にあるかということを考える必要がある。つまり、用語の木の中のどの辺にあるんだということを知る必要がある。ということで、この単語はこれこれの下位概念であるとか、同義語であるとか、上位語、反義語であるとか、そういうことをまず与える。それから、この単語に類似した単語、あるいは類似した概念としてはどうい概念があるというようなことを考える。それから、内包的定義を書く。内包的定義だけですと非常に抽象的になってしまって具体的に何であるかということが分りにくいわけですので、具体例を書く。具体例を書くということは外延的定義を与えている

ことになる。外延的定義の場合は全部を網羅しないと本当はいけないんですけども、有限の概念というものはあまりないわけですし、有限でも膨大な数のものを全部羅列することはできませんので、仕方なく典型的な例を載せる。そういう形で外延的定義を与える。それから、関係的な定義を与える。これは他の概念とはどうい関係にあるか、これはどういものからできているかというようい部分-全体関係であるとか、どうい目的で使われる概念であるかとか、前後の順序関係とか、他との比較概念であるとか、他の概念との相互関係の中でどういふうにこの概念を捉えるかということにここで書く。それから、この単語は、どうい特徴によって他の類似概念から区別されているか、ということを考える。それから、歴史的な記述をしないといけない単語もありますので、歴史的な記述を書く。そういうものはすべてに書くわけではありませんが、あるものには書く。それから、この用語は、どうい context で使われるか、たとえば、「バグ」であるとか、「虫」であるという用語がソフトウェアで使われます。「バグをとる」という言い方は非常に特殊な言い方ですけども、ソフトウェアの世界では通常よくいわれることなので、そういう使い方をするとどういふうな定義をすることにしました。

4,500くらいい単語に説明を与えるのに、大体100人から120人くらい、一単語の説明を書いてもらった方まで入れますと150人くらいの方に手分けして書いてもらいましたので、それぞれの人が勝手に自分の思ったことをだらだらと書いてしまうとわけの分からないこととなりますので、説明を書く場合にはどういふうなスタイルで、どうい順序に、どういことに注意して書いて下さいという書き方についての詳しい指針を与えました。つまり、この単語は用語の木の中のこの位置にあり、その単語の上位概念はどういものであって兄弟概念はこんなものである、下位概念にはこんな単語がある、どうい世界の中で定義を与えて下さい、といった指針のもとに書いてもらいました。そういう意味では、どの定義を見ても大体同じような品質で書かれているということが出来ます。そんなふうにして書いていただいたも

のを全部チェックいたしまして、冗長な表現の場合はそれを削るとか、難しい文体についてはやさしく書き直し、表現の順序が狂っているものや不適切な表現はそれを直すことにしました。特に注意しましたことは、関連する概念の単語をひとりの人が書いていなくて違う人が書いている場合で、この場合は他の人が何を書いているか分からずに説明を書くことになるので、そういう関連する概念の単語を集めてその間の説明で矛盾が生じていないかどうかを相互チェックする。もちろんテキストは、コンピュータに入れたのですけれども、オンライン的にテキストハンドリングしながら見ることができなかつたので、紙に打ち出してもらって、ゲラを読んでチェックしたわけですが、その時に用語の木の配列の順番に打ち出して並べてもらう。そうすると関連した単語は近くに現れてきますから、ブルーフリーディングする時に非常に楽になる。そして、いろいろな単語の相互間に矛盾がないかどうかということをチェックしました。こういうふうなやり方をいたしまして、記述を確かなものとして、曖昧な表現はできるだけなくし、冗長なものは削ってすなりした読みやすい文体にする等々いろいろやりました本文を作りました。

もう一つ、この辞書を作るに当たりまして一番私がやりたかったことは何だったかといいますが、かねがねいろいろな百科事典であるとか、英和辞典であるとか、英語辞典であるとかを活用しております時に、以前から感じていたことですが、通常の辞典というのは、ほとんど見出し語からしか引けない。百科事典なんかは索引から引けますけれども、索引から引く場合でも、一つの単語に対して索引から引けるというのはほんの2、3カ所しかない。つまり、まあ悪くいうとお座なりの索引しか付いていない。標準的な教科書の場合でも後ろに載っている索引というのはかなりお座なりの索引でしかない。ですから、たとえばある本の中のどこかにこういうことが書いてあったという記憶があり、それをもう一度確かめたいと思うんだけど、索引から見ても目次から見ても、どこから見ても、それがどこに載っていたか、どうしてもそれを引っ張ってこれないということは、

非常に多いわけでございます。そういうことをできるだけなくして、せっかくたくさんの方が心血を注いで書いて下さった内容なんですから、どういう方向からでも、ユーザーの欲しいと思う内容に関しては、書いてあればその書いてある所に必ずアクセスできる、そういう方法を確立する必要があります。これは、情報科学の問題であり、情報科学を扱うこの情報科学辞典で、それをやらなければ情報科学としては全然面白くないと思って頑張っているような工夫をしました。その一つが用語の木という先ほどお見せしたものです。学問体系全体を tree 構造で見せて——もちろん一つの単語は必要に応じて tree の何カ所かに出てきますが——これによって学問の全体構造がトップダウン的に分かる。それからもう一つは、一つの用語の説明の中にふんだんなポインターを付けて関連する項目にいくらかでも飛んで行って相互参照ができるようにする。それからもう一つは索引を充実する。そのために語基によるKWIC索引を作りました。13,000単語の語基ごとのKWIC索引を作りますと、大体40,000行くらい出てきます。一つの専門用語は大体2つくらいの語基が多い場合は4つくらいの語基になりますので、全部やりますとそれくらい出てきます。これをそのまま出しますと、ぼうだいなページ数になり、大変ですので、不必要な語基であるとか、いろんな減らす工夫をやりまして、28,000行くらいに減らして載せていただくことにいたしました。英語の用語表現についても同様になりました。その結果、ある内容を知りたいという時に、非常に気楽に引けます。たとえば「並行処理」と「並列処理」というのが一応区別されています。そういうことがこのKWIC索引を見るとすぐ分かります。また、「超並列コンピュータ」ということが最近よくいわれますけれども、「超並列コンピュータ」というのは何だろうと思って引こうとしても、それが見出し語からは出てきません。「岩波情報科学辞典」の見出し語では「高度並列計算機」という単語で出ておまして、「超並列コンピュータ」でいくら引いても出てこないで載ってないと思われるで大間違いで「並列コンピュータ」ということを引きたいんだから、KWIC索引の「並列」というところを見

てもらうと、ちゃんとそこには「高度並列計算機」とか「何とか並列・・・」とか、「並列」に関する用語がすべて出てきますから、「高度並列計算機」というのがどうも自分が引きたいと思っている「超並列コンピュータ」じゃないかと思って引くことができる。

こういうことにしたんですが、本文が800ページの本に対して用語の木とKWIC索引が確か400ページくらいになったかと思います。全体としては、膨大なページ数の本になってしまいました。これは無駄だと思ふ方もおられるかもしれませんが、私にいわせると、本を買う人は、索引とかそういう部分にもおおいにお金を払う価値があると思っていただきたいのです。まあどこまでこのようなことが成功したかどうか分かりませんが――。

このようにして得た結果が全部計算機の中に入っておりますので、その後、これを私どものワークステーション（計算機）に乗せました。テキストについては先ほど申しましたように、文のスタイルがある程度統一されて書かれています。したがって、このような典型的な文体というのは計算機ですぐ分かるわけです。たとえば、「・・・は・・・ともいう」という表現は、同義の別表現をいっているとか、「・・・」という文体が出てきたら、これは内包的定義を与えているところだなとか、これは外延的定義を与えているところだなとか分かります。そこで、自動的にその部分を取り出してこれとこれはシソーラスの同義関係の単語であるとか、これは外延的定義の単語だから下位概念としてもってきたらいいんだとかいうようなことが自動的に分かることになりまして、自動的に用語の木とかシソーラスを作ることがある程度できました。これは人間が手で作ったものと大体良く合います。

それから、この800ページの内容のテキストをハイパーテキストシステムに自動変換しました。たとえば「集積回路」という単語で引きますと「集積回路」の本文テキストが表示画面にぱっと出ます。それを読んでいる最中にその中のある単語が分からなかったら、その単語はどういう意味かなとポインターでクリックしますと、辞典の中

のその単語の説明されている部分が別のウィンドウに表示される。たとえば、「集積回路」の説明文中に現れる「ウェーブプロセス」というのはどういうものかぱっと出せるというように、テキストの中の任意の単語、専門用語を指しますとその説明が出てくる。それから、この単語は用語の木の中ではどの位置にあるかということも同時に別のウィンドウに出すことができる。こうしますと、情報科学全体の中で今自分が読んでいる用語はこういう位置づけの概念のものである、他の概念とはこのような関連性がある、といったことが理解できる。こういうことが自由にできます。それから、tree 状に用語の木が作っておりますので、この tree のある sub-tree だけについての本文を全部集めてこれを用語の木の順に打ち出しますと、いわばその分野の教科書ができたことになります。このようにいろいろな使い方ができます。

そのほかに、この「岩波情報科学辞典」の内容だけではどうしても難しすぎて理解ができない、もうちょっと何か例題とか何かないかという人にとっては、「岩波情報科学講座」という24巻のシリーズがもう10年くらい前に出ておまして、その索引を全部計算機の中に入れておりますので、それを参照することができます。「岩波情報科学辞典」のある単語を読んだんだけど、もうちょっと何とかならないかという場合には、「岩波情報科学講座」の第何巻の何ページにそういうことが詳しく説明してあるということがわかるわけです。たとえば、「スタック」という単語の場合には第何巻の何ページに出てくるかというのがぱっと分かりますのでそれを参照する。残念ながらこの「岩波情報科学講座」は当時はまだ計算機で作っておりませんでしたので、そのページを示すだけであって、その内容がぱっと計算機の表示画面に出せるところまではいっておりませんが、それはテキストを入れればできるようになる。

また、全文検索というのができます。ハイパーテキストシステムでいくらでもポインターをたどっていけるということのほかに、先ほどいいましたように、こういう単語がどこかに出てないかなというのを調べたい時に、索引を見ても、用語の木を見ても、見出し語を見ても、どこにも載って

いない。だけど辞書のどこかに載っている可能性があるという場合には、全テキストを総なめに走査する。そうすると出てくる可能性があるわけです。これは時間がかかると思われるかもしれませんが、岩波情報科学辞典は、本文テキストが220万文字くらいの本ですけれど、それを全文走査するのに5、6秒でどこどこに出てきたかがすべて分かります。そのテキストの部分すべて見ることができます。また、年号の入っているテキストを全部出せとって全部出させますと、それをもって情報科学の歴史をたどることができるのか、いろいろな使い方ができるようになっています。これは計算機でテキストを操る時の非常なメリットだと思っています。

「岩波情報科学辞典」は計算機に入っておりますので、これにこのような強力な検索機能をつけてCD-ROMにして出すといいのじゃないかと、岩波書店も考えているようですけれども、ちょっと分かりません。研究の上ではそういうCD-ROMを出せるぎりぎり手前のところまでやったわけです。このように私は「岩波情報科学辞典」に関して何年間もの時間を費やしまして、自分としての理想的な形のもを実現したつもりだったんですが、1990年の5月に発売されました、岩波書店は5万冊も10万冊も爆発的に売れると思ったらいいんですが、まあ今のところまだ2〜3万冊で、期待したほど売れていないらしいです。しつこくお願いして岩波書店にページ数をたくさん費やさせ、コストをかぶらせてしまったことに対して自責の念をおぼえています。なぜ売れなかったかということをつらつら考えますと、先ほどいいましたように定義の部分をかなりかっちり書いた。少し高踏的というか少し難しく書きすぎた。どのページをめくっても同じようなニュアンスで読めるというのは、ある意味ではいいんだけど、どこを読んでも非常に密度が高いので苦しいというか、ちょっと遊びの雰囲気を入れないといけないのかなあと……。それから用語は13,000選んだのですけれども、情報科学の分野というのは日進月歩でして、どんどん単語も増えていきますし、非常に軽い単語でもやはり引きたい人はたくさんいるわけです。学問的にはほんの1行で済むような

話のことでも見たい人はいるわけで、そういうことに興味を感じる一般的なユーザーもたくさんいらっしゃるようですので、そういう単語を排除したのもちょっとまずかったかなと...私個人はそういう単語も網羅したかったのですけれども、まあ本のキャラクターだとか、いろいろな理由でそれができなかった。いろいろと反省しております。将来、電子版を作るとすれば、もうちょっと違う考え方もできるのではないかと思います。

以上のような考え方を一般用語の辞書にもってこれるかどうかということについては、なかなか難しい問題ですけれども、一般用語についても膨大なオリジナルテキストを用意して、それを計算機で分析したりすることによって、似たようなやり方でひょっとして面白いことができるんじゃないかなという気持ちはもっております。

どうも長々と失礼いたしました。

●柴田 武

どうもありがとうございました。ここで今のお話を私なりにまとめますと、「岩波情報科学辞典」ですか、その辞典の考え方から、そしてまた、その組み立てから、最後にその販売政策に至るまで、自己評価までなされる非常に全体がシステムティックなお話を伺えたかと思えます。いろいろご質問があるかと思えます。実は私自身二つくらい聞きたくてしようがないんですけれども、司会者自らそういうことをしちやいけませんので我慢いたします。後のディスカッションの時に質問をしていただくことにいたしまして、次の日高先生の方のお話に移りたいと思います。どうかよろしく。

専門用語辞典作成法試論

A Methodology of Constructing a Terminology Dictionary

長尾 真

1. はじめに

辞典は知識の宝庫である。百科事典には世の中のあらゆることが書かれている。言葉の辞書には単語のあらゆる意味用法、用例が載っている。しかし1つの事典に盛り込まれたこの膨大で貴重な情報が自由に活用されるようになっていくかどうかということを考えてみる必要があるだろう。現在の辞書は見出し語からしか引くことが出来ない。例えば、look ahead という句は look という見出し語で引いてみて、なければ ahead で引いてみるしか方法がない。何か他の語の説明文中や、引用例中に出ているかもしれないが、それを捜し出すことは不可能である。百科事典の場合には見出し語で引く他に、索引があってそこから引くことも出来るが、その場合の索引語の数は非常に限られている。自分の関心をもつある事柄が膨大な事典の中で見出し語や索引からたどり着けない所に埋もれて書かれているということもかなりあり、それらは偶然に見発見されるという幸運に巡りあわない限り、永遠に見捨てられた情報であり、そのような情報であれば元から書かれていない方が良かったのではないかという疑問さえ生じるのである。

電子化時代の辞書を考えるとき、まず考えるのはこのことである。せっかく蓄積された情報は、いかなる角度、観点からでも要求に適合するものであればかならず取り出せるようにすることである。今日の情報処理装置の能力からすれば、それは決して不可能ではないはずである。次に期待されることは、本という頁に制限されることがなく記憶容量を十分に取れるため、出来るだけ詳細で

豊富な情報を入れて利用者の高度な要求に答えるということである。しかし、十分な注意を払わなければ冗長で理解しにくい情報を沢山入れてしまい、全体を煩雑で質の悪いものにしてしまう恐れがある。第3番目に考えられることは、辞書作成作業を出来るだけ計算機で行なうことによって、辞書全体として一貫した、矛盾を含まない良質のものとする事が出来ること、また、辞書作成作業を効率よく行なえることである。第4番目に、与えられた辞書を、単に分からないことが生じるたびに引くという形で使うのみでなく、計算機力を借りることによって、個々の利用者の目的に合わせて利用者が辞書を再編成することが出来る可能性を持っているということである。例えば、動詞ばかりの辞書、cooking に関する用語ばかりを集めた辞書、慣用語辞書といったものが、元の辞書の内容が十分に豊富であれば、うまく作れる可能性があるだろうということである。またそれを可能とするように元の辞書の内容に配慮することである。

これらのほかにも電子化辞書を考える場合に可能なことは色々ある。それらは OED 2nd edition の作成過程を見るまでもなく明らかなことであろう。本論文では筆者が岩波情報科学辞典(1990年5月刊)の編集に携わった経験から、専門用語の辞書を計算機技術を用いて作る場合の考え方を詳しく述べよう。

2. ことばの辞書と専門用語の辞書

ことばは自然発生的に作られて来ている。基本

的な単語がどうしていつ作られたかが明らかになるという場合は殆どないといってよい。従って、言葉の意味は通常厳密に定義することが出来ない。ことばの意味は、そのことばの使われ方、そのことば、あるいはそのことばを含む文が現実世界でどう解釈され、何を指すかといったことの集積によって規定され、それらに共通に存在する性質によって説明される。従って、自然言語のことばの意味を明らかにするための辞書における用例は重要な意味をもつ。

専門用語もことばであるから、上に述べたことがあてはまることは言うまでもない。しかし、専門用語の場合にはより重要な側面がある。それは1つの専門分野の内容は、そこに存在する概念の集合と、それらの概念相互間に成立する関係性によって表現されるということである。専門用語はこれらの概念それぞれに与えられたラベルであるとみなす事が出来る。自然言語の場合と全く逆であって、まず概念があって、それに用語が与えられると考えられる。専門分野の概念は作られたものであり、外界世界に具体的対照が存在する場合が多いため、それらには通常厳密な定義が与えられており、その定義は専門用語の組み合わせによって行なわれる。これが概念相互間の関係性を表現しているものであることは言うまでもない。

このように、1つの専門分野の内容を明らかにすることは、そこに存在する全ての概念を明らかにし、それらの概念の相互関係性を明確にすることである。あるいは、逆に、この相互関係性の明示が概念の定義であると見ることも出来る。即ち、1つの専門分野は多くの概念の複雑なネットワークという構造体を成しており、そのネットワークのノード間を結合する枝はそれぞれの概念間に成立する関係を表現していると見なすことが出来る。従って、専門用語の辞書は、この複雑なネットワーク構造体を部分的にも全体的にも出来るだけ理解しやすく、引きやすい形に表現するものでなければならない。

3. どのような辞書を作るか

辞書を作る場合には、どのような辞書を作るかについてはっきりとしたイメージを最初から持って

いることが大切である。それは次のような要素を含む。

分野：
 一般の言葉の辞書：現代語、古語、・・・
 一般の用語の辞書：百科的用語、・・・
 専門分野の用語の辞書：専門分野の範囲の設定、・・・

説明の形態：
 大項目主義
 小項目主義
 ハンドブック形式
 あらゆる用語の羅列と簡単な説明

説明の形式：
 定義の仕方
 用例
 派生語
 図、音

説明のレベル：
 学問的立場
 一般人向け解説
 学習者（中学・高校）向け

辞書のスタイル：
 配列法 アルファベット、
 主題別、・・・
 相互参照の程度
 索引の密度、索引の作り方

使用目的：
 深い内容の学問的理解
 一般的な浅い理解
 用語の使用法を学ぶ
 用語のスベルや発音のチェック

見出し語：
 単語
 句
 複合語
 学問的概念

時代性：
 古語
 現代
 最新用語

辞書にはこれまで種々のものが作られて来ている。しかし現代は特に辞書に対する関心の高まってきた時代であるといえる。ことばの辞書だ

けでなく、あらゆる種類の分野において種々のスタイルの辞書が作れるようになってきている。これを社会現象論的に考察してみることは非常に面白いことである。世界が1つとなり、文化が爛熟してくれば来るほど1つの分野内でも非常に多くのことが取り上げられるようになるし、他分野との交流が盛んとなり、種々の物事を出来るだけ大きな全体の中で捕らえ、位置付けたいという要求が強くなって来ている。それに伴って物事の全体像を出来るだけ詳しく把握し表現し、いかなる疑問に対しても対応出来るようにしたいということになってくる。例えば、毎日の新聞に報道されるニュースは単に政治・経済や、社会的、日常的なことだけでなく、医学、宇宙、科学、工学などあらゆる分野に渡っているし、また、政治・経済の分野においても、それぞれの分野のかかなり深い知識、学問的内容が取り扱われており、それらを適切に理解するためにはかなりの専門的知識が必要となって来ているのである。

このように今日の辞書に要求されるものは、個々の事柄の深い内容であるとともに、それが他の種々の事柄とどのような関係を持っているのか、それぞれの分野や学問全体の中でどのような位置付けにある事柄であるのかといった、全体と部分の全てが見え、理解出来るような機能である。例えば、ことばの辞書と百科の辞書の合体されたようなものはその一例であり、古くは *Petit Larousse*、近くは *Grand Dictionnaire Encyclopedic Larousse* がその典型的なものである。

次に問題となるのは辞書の利用の仕方である。上に述べたように今日の辞書は、いわば辞書に対する素人の人が思い付いた事柄でぱっと引いて欲しい情報が得られるということを実現しなければならない。例えば、新聞に現われたある単語をとにかく辞書の引き方もろくに知らずに引くといったことにどう対処するかということが大きな課題となる。ある専門分野の学問を習得している人であればその分野の辞書の見出し語にとられている語を想像して、求める項目に正しくたどり着くことは容易であろうが、そうでない人の場合には大変である。こういった人達が辞書を引くのは2つの場合に分けられる。その1つは本や新聞などで

出会った用語の意味が分からないという場合である。その用語が辞書の見出し語になっていれば問題はない。そうでない場合にすぐに諦めず、索引を調べて、そこにあるかどうかをチェックすることが出来れば、その人は辞書についての能力を持っている人ということになる。しかし、一般のことばの辞書で2つ以上の単語からなるある言回しがそれらのいずれの単語の説明からも発見できないという時は絶望である。ひょっとすると全然関係のない単語の説明文中に引用例として載っていたりするかも知れないが、それを発見するのは偶然に寄らねばならない。用語が辞書の見出しにないときに、シソーラス的機能があってそれを頼りに別の表現を発見して辞書を引けるようになっていれば良いが、それにはしっかりしたシソーラスを辞書に付けておく必要がある。

そのようなシソーラスが付いていない場合にも同様なことが出来ないかを考えることも必要であろう。日本語の場合、用語は複合語であることが多く、その場合には複合語を構成する基本語（語基という）によってKWIC索引を作るのは1つの解決法である。そうすれば、例えば、超並列計算機という語はどこを探してもないが、並列という語基の部分調べてみると、高度並列計算機という語に出会うので、これが探していた語ということが分かる。このような工夫をすることによってシソーラスがなくてもある程度のことは可能なのである。

4. 辞書作りの準備

辞書を作るためには、その辞書で取り扱う概念の範囲や分野を確定し、見出し語としてとる語の決定をしなければならない。これは注意を要する難しい仕事である。

用語の収集のためには用語が使われている適切なテキストを集めることから始めねばならない。そのテキストとしてどのようなものが適切であるかは一般の言葉の場合も、専門用語の場合にも同様に慎重に検討する必要がある。ここでは専門用語の辞書作りのためのテキストについて述べる。

専門用語の場合にはその専門分野に関する本が多く出版されているのが普通である。専門用語の

辞書を作るということは、その分野がかなり発展し、多くの専門家、研究者、学生等がその分野にいるというのが普通であり、そのような分野は一応学問として体系化されているということを前提とすることができるだろう。そうすれば、その分野には多くの標準的な本や、権威のある本が既に存在しているはずである。ただし、作ろうとする辞書の覆う範囲を1冊の本が覆っているということはずまない。そうであれば、辞書を作らずにその本を参照すれば良いからである。通常1つの専門用語の辞書が覆う範囲はいくつかの細分野の集まりになっている。従って、それらの細分野を明確化する努力をすることがまず必要である。そして1つの細分野が1冊の本、特に標準的な、大学あるいは大学院で使用する教科書になる程度に全体を細分化することが必要である。

次にそれぞれの細分野において出版されている本を調べ、そこから標準的な教科書的な本で出来るだけ詳しいものを2、3冊と、その細分野における専門家向けの学術書、研究書を数冊選ぶ。特に後者の場合、専門家向けの本は1つの細分野のさらに細分野となっていることが多いので、いくつかの専門書によって細分野全体をうまく覆うよう配慮する必要がある。

このようにして集めたテキスト(本)から専門用語を選ぶのであるが、少なくとも自然科学系の専門用語辞書の場合には、次のようなステップを踏むことが望ましい。先ず、標準的な教科書の目次を調べ、その章構成、節構成などから、その分野の主要な概念は何かを理解する。多くの場合、そのような主要な概念に対応する用語は章や節の名前の中に現われている。このようなことを top down 的に章の名前から節の名前といった順に詳細化して行き、大きな概念に対する用語(これを章概念の用語という)から小さな概念に対する用語(節概念の用語)にまで至り、これらをこの順に木構造に組み上げる。大きな概念の用語は木の根の方にあり、小さな概念の用語は木の枝の方に置かれる。これを用語の木と名付ける。この用語の木は1つの専門分野の概念の体系を示している。

さて教科書の節やさらにその下の句切りの後は文章となっていて、その文章の中にいくつもの専

門分野の概念が説明され、それらの概念に対する用語が示される。そこでこれらの用語は節概念の用語の下に順次並べて行く。すなわち用語の木の枝をさらに先に伸ばして葉を付けるわけである。これらの説明文の用語相互間にも階層性が発見できる場合もあり、そのような場合にはその階層性を枝の形に表現し、枝をさらに伸ばして行く。しかし多くの場合説明文の用語群は階層的関係にあるのではなく、節概念の用語に対して異なった関係性により結合された下位概念の語となっている。例えば、半導体には非晶質半導体や有機半導体、化合物半導体といった下位概念語があるが、半導体のもつ性質としては帯構造、フェルミ単位、さらに pn 接合といった概念が説明されており、これらも用語の木の半導体という語の先につながる語とするのである。従って、一般に用語の木の2つの用語を結ぶ線には、どういう関係がこれら2つの語の間に成立しているかが記入されることが望まれる。しかし関係には上位下位関係、部分全体関係、作用関係、因果関係など多くの関係があり、これらを正確に決定することは非常に難しい。

このようにして教科書に現われる用語を木の形にまとめ上げ、その後教科書の索引を調べてそこに現われる全ての重要語が用語の木の中に適切に位置付けられているかどうかをチェックする。もし用語の木の中に現われない語があったら、それをテキスト本文によってチェックし、用語の木のどこに位置付けるのが良いかを決定する。次に専門家のための専門書に現われる用語について、それらが既に作った用語の木のどこに入れるのが良いかを考え、用語の木を充実させて行く。このようにして用語の木を成長させて行くと、かなり深いレベルを持つ木となる。この用語の木を作ることとは簡単そうだが実はそれほど簡単ではない。著者によっては教科書の構成を他の書とは全く違ったものとしている場合があるし、また教科書中の章、節、パラグラフといった階層的構造がかならずしもそれら章、節、パラグラフを代表する概念の階層構造と一致するわけではない。このような関係は深刻に考えれば考えるほど決定に迷いが生じてしまう。この用語の木はおおむね上位下位関係、同義語関係、類義語関係などを与えるシー-

ラスと類似しているが、用語の木はどちらかといえば教育という観点からの教科書的配列となっている。従って、用語の木にはシソーラスにあるような全体を支配する原理というものがないが、後に述べるように、ハンドブック作成などに対して大きな効果を発揮する。

このようにして作った用語の木に、必要な用語が漏れていないかどうかをチェックするフェーズが次に必要となる。作ろうとしている専門用語辞書の分野に同種の辞書があればそれを参照することは当然であるが、JIS等の標準用語集や最近2、3年のこの分野の専門雑誌を調べ、基本的な用語に抜けないかどうかをチェックするとともに最近の新しい用語を拾い上げる努力をする。

専門分野の用語を取り上げる時に常に問題となるのは複合語の問題である。一般用語の場合と違い、専門分野には複合語が非常に多い。それは学問の発展に従って、その学問分野に新しい概念が増加して来るが、それに1つずつ全く新しい語を作っていたのではたまらないので、その概念に関連する概念の語を組み合わせで新しい概念の語とすることが多いからである。専門分野の複合語を取り上げればほとんどきりが無いが、専門分野の用語は概念に対応しているから、1つの専門学問分野の中で独立した個別概念として存在するものについては、たとえその概念に対応させられた用語が複合語であっても、それは1つの用語として取り扱うべきものである。

さて、以上のようにして作られた用語の木に現われる用語を全て辞書の見出し語とするかどうかは別の問題である。辞書は商品として適当な大きさである必要があるし、利用者を何処に設定するかによって用語選定の考え方も違って来る。学問的な立場で大学の学生や研究者、企業における専門職業に就いている人達を対象とする場合と、より広く社会の一般人で新聞などに出てくる未知の語について知りたいという人達を対象とするかによって用語選定の基準は当然異なる。ここでは大学の学部学生以上の人達を対象とし、あまり膨大な辞書にならないようにするという立場に立つと、既に作った用語の木のどのレベル以上の用語を見出し語として取るかという立場から用語の選定を

行なうことが妥当であろう。あるレベル以上の語はある大きさ、ある重要さのレベルの概念にほぼ対応しているので、このような方法で用語選定を行なうと、比較的レベルの揃った用語選定を達成できる。つまり、あまり細かい専門的すぎる語は取らず、どの語についてもほぼ同じような重要性をもつ語の集まりとすることが出来るのである。さらに、この用語の木のもう少し深いレベルの語も集め、見出し語とはしないが、説明文中に出て来てそれを読めば意味が簡単に分かるという形のものとし、これらは索引から引けるようにするといった工夫も可能である。このようにすれば極端に多くの見出し語とならず、なおかつ多くの語が説明されている読み易い辞書となる。このようにせず、全ての索引語を見出し語として出すと、説明が簡単すぎて、どのような分野でどのような意味をもって使われる語であるかがはっきりしない、いたってばらばらの感じを与える辞書となってしまふ可能性がある。

従来用語の選定というのはどのようにして行なわれてきたのであろうか。ことばの辞書には長い歴史があり、過去の辞書の土台の上に新しく出て来た語を集め、取捨選択して見出し語を決定しているのが多い。ソーンダイク以来、多くのテキストデータ中に同一単語が現われる頻度を計算し、それによって語の重要性の rank づけをして用語選択の1つの基準とするようになって来ている。これはことばの辞書作りの場合における1つの客観的な尺度として優れたものといえよう。

しかしこの方法を専門用語辞書作成のときに採用できるかどうかは分からない。今日まで専門分野の大量のテキストについて用語の頻度分布を取って調べた例は聞かないし、我々が最近完成した情報科学辞典約220万文字（約50万単語）のデータを調べても、かならずしも頻度によって用語選択の基準になるということは言いにくい。従って、専門分野辞書の用語選択には上記のようなアプローチをとる方がはるかに安定であり、結果に信頼がおけるということが出来る。特に、用語を用語の木という体系として捕らえることが出来、その学問分野の構造と全体が top-down 的に把握しやすいという利点がある。そこでこのような捕らえ

方を計算機を用いて自動的、あるいは半自動的に行なうことが可能なのかどうかということが問題となるが、これについては現段階では解答を与えることは出来ない。用語の体系化は非常に知的な作業であることから、その自動化の困難性を想像することが出来るだろう。

5. 辞書内容の記述

専門分野の用語はその分野を構成する概念に対して与えられた名前である。そしてその学問分野は個々の概念の相互関係付けられた構造の全体として捕らえるのが妥当である。従って、概念相互間の関連性を明確にしながら用語の説明をすることが必要となる。辞書項目は膨大であるからこれをごく少数の限られた人数で記述することは出来ない。いきおい数十名から100名といった人数が必要となる。そういった多くの人達に出来るだけ明確な内容説明で個人差の生じないものを書いてもらうためには幾つかの工夫をすることが必要である。著者が情報科学辞典で取った方法を以下に紹介しよう。

まず第一に、説明文の書き方の1つのひな型を与え、出来るだけそれにそって見出し語の内容を表現してもらうようにした。そのひな型を表1に示す。ここで()内にはその項目に応じた()内に示す情報を必要かつ十分な形に表現することを示している。即ち、先ず、見出し語が学問全体の中の何処に位置するかを示すために、その語の用語の木の中での位置コードを与える。次にその同義語でよく使われるもの、略語、略記などがあればそれを書く。次に見出し語が学問体系の元で何処に位置するものであるかを示すことを目的として、その語の上位概念を書く。次にその語の内包的定義を与え、続いて外延的定義として典型的な例を幾つか示す。そして、その見出し語の類義語、関連語、反義語等とどこがどのように違っているか、どのように使い分けられるかといった区別概念を書くことによって、その語の意味をより明確に理解・把握させるようにしている。続いて、その語の概念あるいは対象が何から構成されているかを列挙させ、そのあと、それが持つ性質、属性、機能、目的などを書かせる。そして、ある種

表 1

(用語) (見出し語)

- 〈同義語〉 (同義語)ともいう。(略記)と略記することがある/多い。
 (反義語)に対立する概念である。
 〈内包的定義〉 (上位概念)に属すもの/の1つ/であり、(内包的定義)である。(上位概念)に属す類似の概念としては(並列概念語)があるが、(用語)は(区別的特徴)において、これらと区別される。類似の語として(類義語)があるが、(区別的特徴)によって区別される。
 〈外延的定義(例示)〉 (用語)の典型的な例としては(例示)がある。
 / (用語)は(列挙)からなる。/ (用語)には(列挙)が属している。
 〈構成的定義〉 (用語)は(列挙)からなり、(列挙i)と(列挙j)とは(関係ij)という関係をもっている。
 〈性質・属性、機能・目的的特徴による定義〉 (用語)は(性質・・・目的)という性質/属性/機能/目的/をもち、(使用環境)の中で(目的)のために(用法)のように用いられる。(用語)は(他の用語)との間に原因・結果/部分・全体/前後・順序/比較/同一特徴/材料・生成物/の関係にある。
 〈発生的定義〉 (用語)は(歴史的/必要性/原因)から(人)によって(年)に作られた/出て来た/ものである。その後(経過)を経て、(用語)の現状は(現状)である。
 〈文脈的定義〉 (用語)は(文脈)という形で使われる。

の単語については歴史的記述、その語の使われ方等についても記述するよう要請している。

このような説明文中に現われる他の見出し語には'印を付し、また説明文の途中や最後に→Aという形で当該部分、あるいはこの見出し語についてはさらにAという見出し語を参照することによって関連情報を得ることが出来、より深い理解が得られるという事を示している。さらに見出し語にはしていないが簡単な説明を与えた方がよい語については、説明文中に“—である○○”という形の説明を行ない、説明された語○○は索引語として採ると言うようにする。

以上のいくつかの段階は人手による作業となり、

計算機が助けることは出来ない。100名もの人にオンライン接続されたワークステーションを配ることは出来ないし、また、著者はどこで著作作業をするか分からないから端末の使用効率も良くないと想像されるからである。書かれた原稿はただちに計算機に入れられ、そこから後は計算機の助けを借りられる。その1つは不等に難しい未定義語が説明文中に使われていないかどうかというチェックで、他は相互参照されているいくつかの用語の説明の間に矛盾が含まれていないかのチェックなどである。この後者のチェックを自動的に行なうことは出来ない。計算機が出来ることは相互参照されている説明文を同じ頁に表示して、人が相互比較して読めるようにすることだけである。しかし、この仕事は大切である。類似の2つの専門用語がどこで違っているかということは常に明確に読者に知らさねばならないものであるが、多数の著者がそれぞれ独立に説明を書いている時に、それらすべての間に無矛盾性を保つということは非常に難しい。従って、一度説明文が書かれた後、このような形で関連用語をクラスターして統一的に見直す作業が必要となる。Longman Dictionary of Contemporary English が実現したように説明文をあらかじめ予定された2000語程度の単語によって記述することが出来れば理想的であり、そのためのチェックに計算機を使うことも出来るが、専門用語辞典の場合に、そこまでの注意とチェックを行なうことが可能かどうかは辞書作りの経済的、日教的負担の問題である。専門用語の定義と説明に他の専門用語を使うということは避けられないことではあるが、その use と used for との上下関係が一貫していて、循環的定義になっていないということには注意を払う必要がある。これはある程度自動的にチェックすることが出来るが、正確には関連説明文を人が比較対照してチェックする方法しか現在のところない。

用語の説明文を種々の用語説明について出来るだけ均質であるように要請することが大切であるが、それを多くの著者に執筆してもらう場合に確実に実現する方法を考えることは大切である。情報科学辞典の作成において採った方法は次のよう

なものであった。その第1は説明内容として書くべきことを既に述べた表1のように示すことである。その第2は各見出し語に対して、用語の木のその語を含む部分を著者に示して、学問体系のこの位置付けにおいて説明を書いていただきたいということを明示することである。特に、その用語の上位概念語、下位概念語、同列に並ぶ関連語を著者にはっきりと示し、これをはっきりと意識して執筆してもらうことが大切である。次にその用語の説明に当てる概略の語数を示すことが必要である。こうすることによって辞書全体の中でその用語の重要性をどのように位置付けているかを執筆者に知らせることが出来るわけである。情報科学辞典の場合には、説明文の長さをA, B, C, Dの4段階に分け、Aは400文字、Bは200文字、Cは100文字、Dは50文字程度といった指示を与えた。これは辞書の編集者の考え方により決めることになる。

説明文の書くべき内容とその順序は表1のように与えたが、説明文の文体が人によってまちまちであっても読者は読みづらい。また、限られた紙面の中に出来るだけ多くの情報を入れるという目的からは冗長な表現を採らないよう配慮することが必要である。また、あまりにも簡単化した文章の場合には、省略が多すぎたりして読者が文の内容を理解しにくいということも生じうる。そこで一度書いた説明文章を見直し、種々の点で文章を推敲し直すという過程の重要性を強調し、そのための方法を具体例を通じて著者に示すことを行なった。このようにして出来るだけ正確で読み易く、しかも無駄のない均質な説明文を作ることが可能となった。勿論、第1校か第2校かの段階で編集者が文章を読み、種々のコメントを著者に与えて文章を書き改めてもらったり、編集者が直接文章を書き改めたりする仕事を行なうことも必要である。

以上のような作業をスムーズにするためには、著者や編者は個々の見出し語、用語とその説明文だけを孤立的に書いたり読んだりしてはならない。全ての説明文は用語の木の順に従って、配列されて編者に提供されることが必要である。用語の木のある部分に関する説明文が用語の木の配

列に従って並べられると、それはあたかもその部分に関する1冊の本を与えられたかのように読むことができ、常に、全体像の中での個々の内容の理解という立場でチェックが出来るのである。工学分野によくあるハンドブックはこのような形式の物である。従って、小項目主義の辞書でアルファベット順に配列をする場合には個々の項目を読むだけで自己完結的にその内容がわかるようにすることが大切であるが、これを用語の木の配列順に読めばハンドブックあるいは大項目主義の辞書を読んでいるのと同じ効果を与えられるわけである。これは読者にとってはもちろんのこと、辞書編集の途中における種々の内容のチェックのためにも望ましい形式である。

6. 電子化辞書の活用の方法

現代のように忙しくて多様化した時代になると、いろんな人が知りたいということだけを簡単に手軽に知ることのできる小項目主義の辞書が喜ばれる傾向がある。従って、辞書の配列は小項目主義のあいえお順(アルファベット)というのが最も便利な形態であるといえる。しかし、学生や時間のある人はある部分について体系的に知識を整理したいという要求をもつ場合があり、そういった場合には大項目主義の辞書の方が良い。そして工学の専門分野では大項目主義の説明に加えて、現実社会における状況や具体的なデータを付け加えたハンドブック形式が広く使われている。

このような種々の立場からの要求に対しては、小項目主義の項目説明と用語の木とを組み合わせることによって対処することが可能である。用語の木のある部分木は大項目主義のある項目の内容説明の配列にほぼ対応しているので、この用語の木の構造に従って各項目の説明をテキスト的に章・節...の形に配列し直せばよい。こうすることによってその分野の大項目主義による説明文章が作れたことになるし、また、適当な項目にたいしてこれを行えば、その項目を表題とする一冊の教科書的な本が出来上がった事になる。こうした辞書の再編成は電子化辞書の時代には比較的容易であり、個人向けの本としての再編成が自由出来るという面白い辞書の利用法が考えられるので

ある。そういったことを容易に実現するためにも用語の木は大切である。そして用語の木そのものも個人の考え方によって自由に変更することが出来るような電子化辞書活用ソフトウェアを提供することが必要であろう。

辞書を活用する場合のもう1つの重要な問題は曖昧検索といわれていることの実現である。辞書を引く機会には幾つかの場合がある。その1つは何か文章を読んでいる最中に意味の分からない単語・用語が出て来た時に辞書を引くという場合である。もう1つは知りたいということがぼんやりと頭の中にあって、その概念が明確でなく、ましてその概念に対応する用語を正確に定められないという時である。他には単語の綴りを確かめたり、対応する英語表現を知りたいといった場合などである。これらの内、第1の場合には引きたい単語が明確に目の前に存在するので引き易い。辞書を直接引いてその単語に出会う確率は非常に高い。しかしたまには見出し語としてその語が載っていないという場合もある。その場合に簡単に諦めてしまわずに何とかして該当する項目に至り就く方法を辞書が提供できるかどうかが最も大切なことなのである。専門分野の専門用語であっても、かならずしも用語の統制が厳格に行なわれているとは限らない。電子計算機は計算機とも言うし、また、コンピュータ、コンピューター、電算機とも言う。従って、これらいずれからも電子計算機にたどり着くことが出来るようになっている必要がある。即ち、広い意味でのシソーラスが必要なのである。

第2の場合には状況が非常に曖昧であり、質問したい本人が何をどのように質問したらよいかさえはっきりしていない場合である。そのような時には種々のヒントとなる情報を提供し、その人の考えを明確化する手助けを出来るようになっていなければならない。例えば、ふと思いついた部分概念や関連概念からブラウジング等の方法で、それに関係する他の可能な概念を提示し、自分の欲しがっていたものをそこに発見させるといった手法が必要となるのである。関連概念を示すという立場からは用語の木は有効である。自分の知りたいことが何であるかははっきりしなくても、それ

がどのような分野のものであるかは、殆どの場合言うことが出来る。従って、その上位概念の語を根とする用語の木を示すことによって、その部分木の中を探して自分の必要とするものを捜し出すわけである。

漠然とした概念そのものをはっきりさせることは難しくても、その部分概念、要素概念に対応する部分的な用語を思い浮べることが出来るということが多い。そのような部分的な語のことを語基と呼んでいる。従って、この語基によって索引が出来れば欲しい情報に行きつける可能性が高いということになる。そこで我々の情報科学辞典の場合には約13000語の索引語を語基によるKWIC索引とし、同一語基の部分を読めることによって自分の知りたい用語を発見することが出来るようにしてある。これは非常に有効な方法である。例えば、超並列コンピュータを引きたいという時、このままでは項目にない。その場合には、2つの方法がありうる。まずコンピュータあるいは並列コンピュータを引いて、その説明文中に超並列コンピュータのことが書いてあるかどうか、また書いてなくてもそれに関係する事項へ参照指示があるかどうかを調べることである。第2は例えば、並列という語基によって索引を引いてみることである。その場合には、並列という語基を含んだ索引語が表2に示すように全て1か所に集められているので、この表を見ることによって超並列コンピュータはないが、高度並列計算機という用語はあるということが分かり、目的を達成することが出来るのである。

各項目の説明文章の中で年号の含まれる文をすべて取り出して年代順に並べることによって、その専門分野の年表を作ることも出来る。また、説明文中の人名を全て取り出すことによって専門分野の研究者の業績の表とすることも可能であろう。説明文の最初の部分に置かれた用語の上位概念語、下位概念語を取り出して、その関係性を保ったままで配列することによって、用語の木とは少し違った形のシソーラスを作ることも出来るだろう。こういったことは全てが完全に自動的にには行なえないだろうが、人間が少し介入する程度で作れるだろうから、計算機でこういった情報の再整

理と種々の観点からの表現をすることは非常に有効なことである。辞書を計算機に入れることによって、冊子体の辞典では不可能だった種々の事が可能となり、辞書に含まれている情報を十分に活用する道が開かれたことになる。後はどのような処理の仕方をすればどのような新しい情報が発見できるかというユーザの活用能力の問題となるのである。特に見出し語が用語の木のどこにあるかがポインターで指示されているので、その語の学問全体の中での位置付けや、関連する概念の内容を知ることが出来るし、見出し語の説明文中に現われる主要な専門用語については関連して参照すれば参考になる用語が示されているなど、用語間の相互参照性が十分に考慮されているので、後は読者の興味と能力によって、どのような活用も可能となる。このようにして、岩波情報科学辞典は辞典に盛り込まれた情報は100%使われるような構造になっており、これを冊子体で作るには相当な苦勞があった。これを計算機に入れてより自由な使用が出来るようにソフトウェア開発を行なっているが、冊子体の場合に工夫したあらゆる言語処理的な内容が電子辞書の場合にも必要なことは明白である。電子化すればすぐ何でも出来るという考え方は誤りであって、高度の言語情報処理技術とその専門分野の学問体系が十分に分かっているならば、質の良い辞書活用システムを実現出来ないことは当然のことである。

表 2

復帰	平面	96a46
切除	平面	393b10
	平面	グラフ 686a34
極大	平面	グラフ 80a21
	平面	グラフの判定 686b7
	平面	磁化 292a23
	平面	走査法 196b37
	平面	ディスプレイ装置 650a11
	平面	波解 409a1
	平面	プロッター 262b10
切除	平面	法 393b3
	並行	計算 285b17
	並行	実行 285b41
	並行	処理 614b21, 684a14
同時	並行	処理 198b40
分散	並行	処理 677b6
	並行	処理制御, 684b1
	並行	処理制御方式 684b20
事後検証	並行	処理制御方式 283b1
事前解析	並行	処理制御方式 292b1
楽観的	並行	処理制御方式 283b3
	並行	性 685a42
	並行	プログラム 685a35
	並行	プロセス 380b37, 685b1
	並行	文 688a19
and	並列	389b4
or	並列	815a38
積	並列	389b3
和	並列	815a37
高度	並列	アーキテクチャー 39a19
	並列	アナログ処理 602b13
	並列	オブジェクト指向言語 686b25
	並列	加算器 112b25
	並列	画像処理装置 686b43
多重プロセッサ	型	並列 画像処理装置 687a11
	パイプライン型	並列 画像処理装置 687a6
完全	並列	型画像処理装置 686b51
局所	並列	型画像処理装置 686b47
	並列	句の意味性 687a25
	並列	計算 604b18
	並列	計算機 396b4, 687b2
高度	並列	計算機 39b2, 229a28
	並列	こみ集め 687a39
	並列	最外戦略 144a31
	並列	最内戦略 144a14
	並列	算法 687b1
	並列	実行 389b5, 815a39
	並列	写像 396b15
	並列	処理 456a16, 687b39
	並列	冗長 342a35
神経回路網の	並列	情報処理 348b39
神経場における	並列	情報処理 349b5
	並列	制御 456a37, 806a50
	並列	性 685a43
	並列	整列 688a21
	並列	接続 463a2
	並列	チャンネルモデル 463b6
	並列	デジタル処理 602b13
	並列	伝送 688a35
	並列	ネットワーク 387b30
	並列	ハッシング 684b3
	並列	比較型AD変換器 63b4
	並列	プログラミング言語 806a51
	並列	プログラム 685a39
	並列	プロセス 688a43
	並列	分解 688b7
	並列	分散処理 657b39
	並列	文 688a48

講演 2

用語の造語メカニズム

Probabilistic Mechanism of Terminology

日高 達* HIDAKA, Toru

長尾先生のお話は、専門用語の内容の定義、意味の定義をどのようにがっちりするか、また、関係ある専門用語間のお互いの参照関係をどういふふうに取り入れていくかというお話でした。私の話は「用語の造語メカニズム」となっていますけれども、実は専門用語については全然何の経験もございません。ただ、一般の言語処理をしている時に辞書にはない単語が出てくるのがしばしばあります。特に派生語はその数が膨大なために、辞書の見出し語に採り入れられているものは派生語全体のごく一部にすぎない。とにかく相当に未知語、未登録語があるというわけです。それで私の場合は、その文字列の格好からこれはどうも一般の用語の形に似ているなあということをやりに、未知語の推定をやるという実験を試みたわけです。

これまでの未知語の処理というのは、主に字種情報をたよりにしています。「大方の文節は漢字で始まる自立語に始まり、ひらがな書きされる語尾変化部や付属語で終わる」という字種情報をたよりに未知語の推定をやっている。もちろん前後の分脈の利用もされますけれども、単語それ自体としては字種情報をたよりにしている。それをもう少し進めて、単語らしさというものを字種情報だけでなく、辞書における単語の見出し語の傾向から推し計ることができないかということを考えてみたわけです。日本語の場合、漢字というのは音をもっていますけれども、同時に意味ももって

います。例えば、「広葉樹林帯」といいますと、「広」と「葉」という漢字が合わさって「広葉」というのができる。しかもその主要な意味は一般にそのどちらかをもってまして、この場合は「葉」の方、つまり「広葉」の head に当たる意味は「葉」がもっている。それから「広葉」と「樹」がつながりまして「広葉樹」となり、その主要な意味は「樹」がもっている。そういうことを繰り返していつているわけですね。ですから大体後の文字が主要な意味をもっているように思われます。しかし、全部そうとは限りません。例えば「読心術」という語の場合、心を読むという意味で「読心」の主要な意味は前の文字の「読」であります。それから「読心術」というのは「読心」の主要な意味としての「読」と「術」が結ばれて、その全体としての主要な意味は「術」になっている。「乱気流」といいますと、これは「気流」というものがまずできまして、これの主要な意味は「流」になります。これに「乱」という接頭語が付いて「乱気流」という単語になるわけですが、その主要な意味は「流」がもっていると考えられます。このように語構成は、一般には context-free 的な形態をもっている。このような語の結びつきは、図1に示すような木構造で表現することができ、さらに、木構造は文脈自由文法の枠組みで統一的に取り扱うことができます。ここで、丸で囲まれた文字は文脈自由文法の nonterminal symbol と考えますと、木構造から文脈自由文法の書き換え規則が得られるわけです。こうして得られた文脈自由文法から辞書の見出し語を生成することが

* 九州大学工学部教授

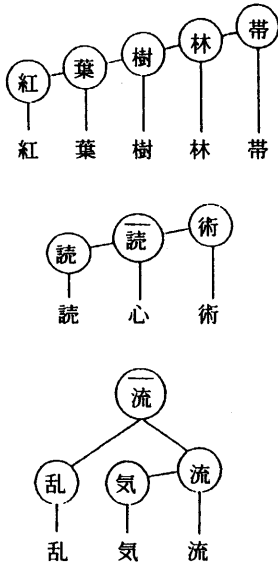


図1. 語の構造

できますが、これではまだ、どのような単語または木構造が生じやすいかどうかを反映していません。

全体でどういった傾向をもつかということは、ある種の統計的な性質を考えることに相当しまして、例えば今作りました語に対応する木構造を辞書の見出し語について全部解析して作ってやる。次に、木構造を文脈自由文法の書き換え規則に変換してやります。その中にそれぞれの書き換え規則がどういった割合で出てくるかを調べ、確率文法化してやるわけです。つまりこれは、例えば、 x が何かに書き換わる時に特に α に書き換わる確率は見出し語の中でどのくらいあるかといった傾向を確率化してやるわけです。そうしまして、逆にこの語の木構造を構成する時にいろいろな書き換え規則が使用されるわけですが、それに確率が付いていますので、それを全部掛け合わせた値をこの木構造が発生する確率だというふうに読み直すわけです。つまりこの確率の計算をしまして、見出し語に対応するすべての木構造が発生する確率が極大になるように確率を設定してやるわけです。すでに確率文法で開発された方法が

りますので、これを利用して確率文法化してやるということです。ところがこの方法は、例えば名詞で作りますと大変な労力が必要なんです。実は、語の木構造は人間が作ってやらなければいけない。どれが head であるかというようなことをいちいち見ていくのです。head がない場合もあります。例えば「春夏秋冬」といった並列語の場合、どの言葉も head ではないわけですね。そういうものはその文字列全体を一つの文字みたいに見まして、それだけでも単語である、head であると考えれば良いわけです。7万とか8万とかある名詞からそれを作るとなると大変な労力です。しかし多くの場合、日本語では、大体後の語が head になっている傾向が強い。そう仮定しまして、自動的に単語の木構造を作り上げる。確かに例外もありますけれども、こう仮定して語の木構造を人手で作る労力を全部省いたわけです。そうしますと文法が文脈自由文法の中でも特に簡単な正規文法になります。

今、単語は漢字からなっていると申しました。漢字は意味も持っているし、音も持っている。ところがやはり、仮名も混在しているわけですね。日本語では、意味を担う最小単語は原則として漢字ですけれども、仮名も混じっているわけです。そこで、造語単位をどう決めるかということですが、「意味がある最小単位で、しかも音も持っている」最小単位を造語単位にしたい。一般には漢字1文字とそれに続く平仮名の字です。平仮名1文字1文字は単に音しかなく、意味は持ちません。そこで、漢字1文字とそれに続く平仮名列を造語単位にすれば、多くの場合、音と意味を持つことになる。例えば、非常に簡単な例でお話します。図2に示すような10単語の世界を考えてみましょう。名詞の見出し語にこういう10単語だけが登録されているとします。「文」「国語」「文法」「学問」……そういうものが登録されていた。そうしますと、この中で16個の造語単位ができます。もちろんこれは全部漢字でできていますので、1文字ずつが造語単位になります。そうしてさっきの文脈自由文法の書き換え規則を図的に表しますと、図3のようになります。これで initial から final state までにたどっていくものが造語され

単語集合W (10単語)

= {文/ぶん, 国語/こくご, 文法/ぶんぽう, 学問/がくもん, 言語/げんご, 問題/もんだい, 調子/ちょうし, 英語圏/えいごけん, 文語調/ぶんごちょう, 調教師/ちょうきょうし}

造語単位集合Z (16造語単位)

= {文/ぶん, 国/こく, 語/ご, 法/ほう, 学/がく, 問/もん, 言/げん, 題/だい, 英/えい, 圏/けん, 調/ちょう, 子/し, 教/きょう, 師/し}

造語される未知語 (19語)

問, 学問題, 文語, 文語圏, 文語調子, 文語調教師, 英語, 英語調, 英語調子, 英語調教師, 国語圏, 国語調, 国語調子, 国語調教師, 言語圏, 言語調, 言語調子, 言語調教師, 調

図2. 小単語世界

ることになるわけです。そうしますと、実は元々の見出し語から19個もの未知語が造語されるわけです。この中には「問/もん」とか「調/ちょう」とかあまり単語らしくないものも含まれているわけです。それで何とか造語機能を少し弱くしなければいけない。特に短い単語が新しく造語されないように抑えなければいけない、ということを考えてわけです。つまり、登録語といいますか、元々辞書にあった10個の登録語はすべて造語されなければいけないが、2造語単位以下の未知語はなるべく造語されないようにしよう。2造語単位以下の単語はほとんど辞書に登録されているからです。使用頻度も非常に高いものでして、2造語単位以下のものを造語してはちょっとまずい。こういうことから、造語機能を少し抑えることができないかということで、単語の意味に加えて単語の位置—つまり、単語の最初の語になり得るか、完結する最後の語になり得るかというポジションの情報を造語単位に加味しました。これは語の最初になり得る文字ですよ、これは語の最後になり得る文字ですよ、1語だけで最初にもなり得るし完結性も持っている語ですよ、語の中間の文字となり得

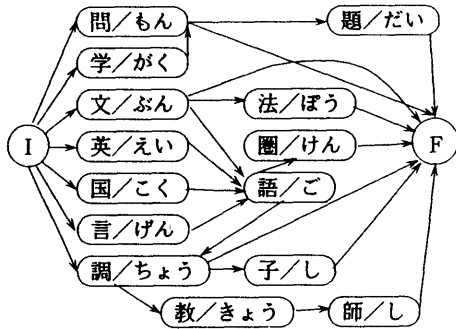


図3. 書換え規則

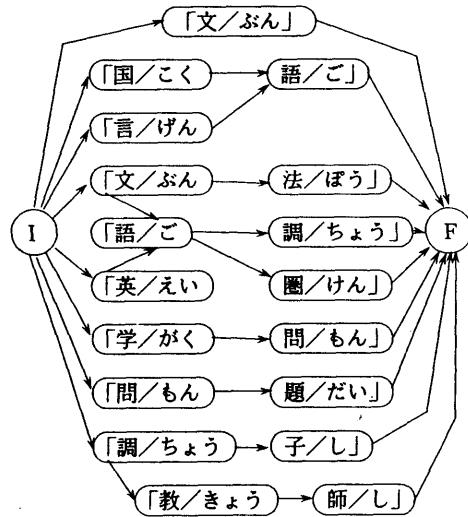


図4. 制限された書換え規則

る語ですよと、そういう四つに分類いたしました。そうしまして、さっきと同じように見出し語の集合から文法を作りますと、図4のような文法ができあがります。これでやりますと新しい語、未知語の発生は非常に抑えられまして、例えばここではたった二つですね。「文語圏」と「英語調」というのができるだけとなるのです。そういうふうなことをやまして、とにかく2造語単位以下の短いものは新しく生成されないように、生成機構

を抑えたわけです。私どもで作りました自立語辞書、これにはいろいろな品詞が入っていますが、その中に名詞が74,454語あります。それからさっきの context-free の確率文法を作り上げたわけです。ある程度機械的に作り上げたということです。ちなみにその見出し語の中で1造語単位だけからなっている語は9,000、2造語単位からなっているのは43,000、全体合計して74,454個の見出し語からさっきの確率文法を作り上げたわけです。

造語モデルを作りましたので、次にその造語モデルがいいか悪いかということの評価しなければいけません。どういうふうな立場で評価するかというのは大変難しいことで、ここでは非常に単純に、少なくとも次のような条件を満足することを造語モデルに要請します。それは、第一に、辞書の見出し語としてあったものは全て造語されなければいけない。第二に、2造語単位以下の未知語は造語してもらっては困る。これは、2造語単位以下単位はほとんど登録されていると考えてよろしい、したがって新しく造語してはいけないということです。第三に、造語モデルから計算されます生起確率、これが見出し語については大きいけれども、見出し語にない、造語モデルから新しく生成される単語については、小さくなる。それから、単語の長さでいいますと、¹日本語の見出し語を造語単位に分解していきますと、大体3造語単位以下の単語が多い。4造語単位とか5造語単位の単語は極端に数が減ってきます。したがって、第四に、長い単語の生起確率が非常に小さくなっていくように設計されているかどうかということで評価をしてやればよい。つまり、これだけの性質をもたないものは造語モデルとしては役に

立たない。少なくともこういう四つの要請は満たしておかなければいけないということを造語モデルの条件と考えました。この造語モデルから、発生頻度の高い順に200万語発生させていきました。

図5は、語長別に累積出現率を表示したものです。横軸は、造語された単語が生起確立の高い順に左から右方向に並べられていると見てください。これを見ますと1造語単位のものがまず全部現れる。それから2造語単位の長さのものが確率の高いものから低いものまでだんだん発生していき、2造語単位の単語がほとんど発生した段階から次に確率の高い3造語単位のものが発生してきます。それから、4造語単位の単語が発生し、次に、5造語単位の単語が発生していくという傾向を示しています。つまり、短い単語の方が発生確率が高い。しかも2造語単位のものは登録語しかございません。3造語単位のものには登録語もあるし、未知語もあるわけです。ここいら辺から新しい未知語が発生していているわけですが、とにかく2造語単位のものは発生確率が高く、長い単語になるほど発生確率が低いのだということが窺えます。これをさらに詳細に分析しますと、大体生起確立の高い最初の5万個くらいの単語はほとんど登録語で占められる。それからだんだん新しい3造語単位以上の単語が発生してきている。大体7万くらいに達しますと、ほとんど

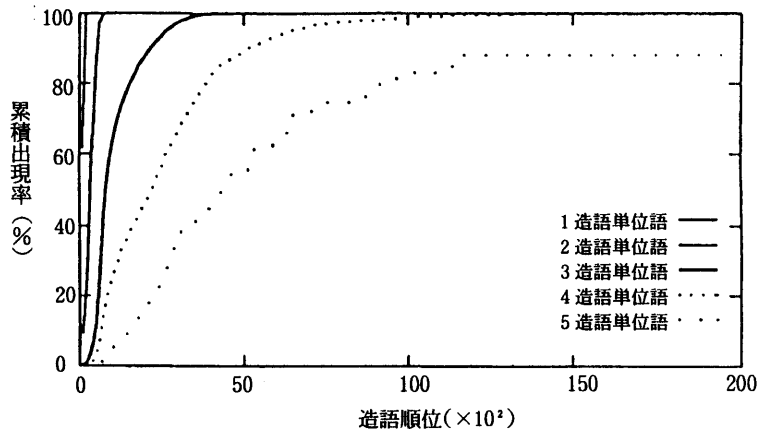


図5. 構成造語単位数別登録語累積出現率

ど見出し語の発生は終わっています。つまりこれからいえることは、登録語の発生確率の方が高く、未登録語の発生確率は登録語に比べて低いということです。また、図5は、長単位の語になるほどその発生確率は低いんだということを示しています。まあそういう実験を行ったわけです。

図6は、ある単語を含む前後1,000語の間に3造語単位以上のもがどのくらい発生しているかというものをグラフにしたものです。いつも未知語というのは3造語単位以上です。その中で特に未知語、登録語でなかったものがどういふうに発生してくるかという、初めに3造語単位ぐらいの未知語が発生し、それからだんだん4造語単位の語が発生してくる。そういう傾向を示していきまして、長単位の単語ほど発生確率が低くなっている。しかも未登録語の中では特にそういう傾向が目立っているということを示しています。これからいえることは、さっきの四つの条件といいますが、登録語は全部はずさせられるし、しかも高い確率で発生してくれる。それから、2造語単位以下のものは発生しないかです。さっきの四つの条件を満たす造語モデルは確率文法である程度できそうだと考えられます。しかも今の実験は非常に単純な、生起文法で日本語の造語ができるんだと仮定してやりましたけれども、見出し語の木構造を手で丹念に作ってやり、これを基に確立文脈自由文法を作れば、もっと質の良い造語モデルができるでしょう。以上の造語モデルは、複合語の造語モデルにはなりません。複合語の造語モデルでは、造語単位が単語であると考えられます。短単位の単語を造語単位として、同様の手法を用い、

複合語の造語モデルを構成することも可能です。複合語に関しまして最近少し実験をしたことがございますので、それを紹介させていただきたいと思っています。

未登録語の中には派生語というものが多いのです。派生語というのは、一つの単語に語尾が付いている。そういう派生語というのは数がたくさんにのぼりまして、したがって辞書には載ってないわけです。またそれは一つの複合語とも考えられるわけです。その発生する機構を考えてみたいと思います。例えば、その時、語の処理はどうするのだという時に、語の上位下位関係を使うことができないか。例えば「作品」という単語に「展」という派生語尾がついて、「作品展」という複合語になるので、「作品」の下位の語にも「展」が接続して派生語になるものが多いだろうと考えられます。しかし調べてみますと、必ずしも全部の下位語に付くわけではありません。例えば「創作展」というのはありそうですが、「詩展」とか「川柳展」というのはあまり聞いたことはないような感じがするわけです。動物の場合には「動物的」といいますけれども、では「動物」のずっと下位の語に、例えば「キリン的」とか「蚤的」「象的」というかという、あまり派生語らしくない。しかし、「動物」に近い下位の語には付く可能性が高いであろう。派生語の生成にそう

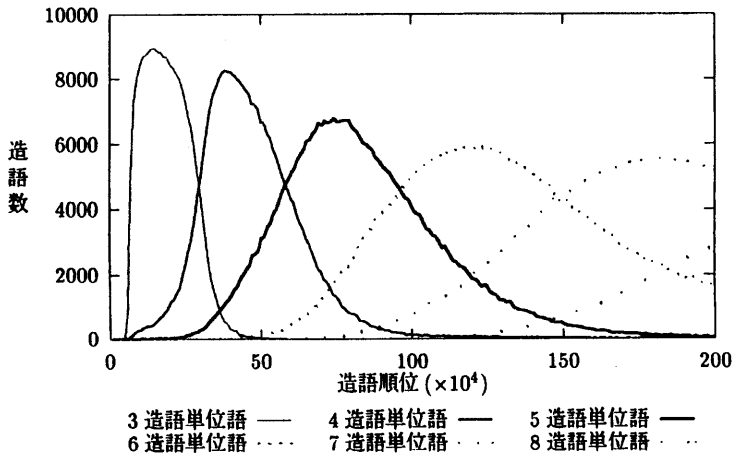


図6. 構成造語単位数別未知語分布

いうことを利用できないでしょうか。これを利用するには、いくつかの派生語サンプルをテキストから集めてくる。もちろん全部集めてくることはできませんから、例えば1万なり2万なり集めてくる。派生語サンプルの中に「作品展」があれば、あとはソーラスを用いまして、「作品」の下位の語にも「展」が付き得るのではないかと考えるわけです。ここで大事なことは、「作品」のずっと遠い下位の語になればなるほど付きにくくする、すなわち発生確率を低く抑えるような造語機能にする。そういう造語モデルを考えただけです。

まず日本UNISYS社の協力を得て集めた、異なり語数9,939語の派生語サンプルと、筑波大の荻野氏が作られた「現代日本語ソーラス」における名詞の上位下位関係を用いて派生語の確立造語モデルを作り、派生語の認識実験をしました。すると、33%程度の認識率しか得られませんでした。これはやはり派生語のデータが少なかつたんだろうというわけで派生語サンプルを30%（約3,500語）増やした造語モデルで実験しました。そうしても、大体38%、わずか5%くらいしか認識率が上がらなかったのです。これでやっていると、いくら派生語サンプルを集めてもうまい造語モデルはできそうにない。そこで何を考えたかといいますと、少し人工的に派生語サンプルを増やしてみようということです。これは、次のようにして作りました。「絵画展」という派生語サンプルがあると、「絵画」のすぐの上位語である「美術」に派生語尾「展」を付けて、「美術展」を人工的な派生語サンプルとします。すべての派生語サンプルについて同様なことをやって、これを基に派生語モデルを拡張しました。拡張された派生語モデルで派生語の認識実験をしたところ、認識率が85%にあがり、予想以上の好結果を出しています。以上の実験から派生語サンプルをさらに充実することにより、派生語認識に充分効果的な造語モデルになるものと考えています。

複合語の造語モデルについても、同様の手法が使えらると思われませんが、複合語サンプルの収集と人手による木構造への解析にも大変な労力を要すると思われま

●柴田

どうも日高先生ありがとうございました。未登録語というのは辞書に載っていない言葉という意味だと思いますが、未知語とおっしゃるのも同じことですね？

●日高

未登録語というのは、単語としてははっきり確立されているんだけど辞書には載っていない語で、未知語はこれから造られるかもしれないがまだ単語ではない。未知語の中には未登録語も含まれています。

●柴田

ああそうですか。未知語の方が広い概念。そういう未知の世界をご研究でございまして、これもいくつか質問したいんですけども、すべてはパネルディスカッションの方へ譲りたいと思います。

講演 3

専門用語の命名における諸問題

On Naming of Some Technical Terms in the Japanese Language

森岡 健二・ MORIOKA, Kenji

私は、昭和24年に国立国語研究所に入りまして、最初に与えられた課題が命名の問題でありました。当時は、戦後間もなくのことで、稲、芋などの品種が農林○号などと呼ばれていて、農家から番号制は親しみがないばかりでなく、注文するとき間違いが起りやすいので、名前らしい名前にもどしたいという要望があったのだそうです。しかし、名前を付けるといっても、何千何万とある品種を名前によって区別するとなると、容易なことではありません。そういうことで、国語研究所がこれを国語の問題として取り上げたのですが、もちろん品種に付いての知識のないものに、具体的な名前の付けられよう筈もなく、結局、私は農作物、商品、あるいは専門用語集などから名前を出るだけ多く採集して、言語にとって名前とは何か、命名にはどんな問題があるかという基礎的な問題を考えることにしました。そこで、今日は先ず、日本語全体の中で専門用語が、どのような位置にあり、どのような性質をもつかということから、話を始めたいと思います。資料を用意してきましたので、資料に沿って説明させていただきます。

1. ことばの町

資 1

言語というものは、いろいろの路地や広場、多くの古い家や新しい家、いろいろな時代に建て増された部分を持つ家が入

り組んで一つの全体をなしている一つの古い町とみなすこともできる。(ウィトゲンシュタイン)

これは、まことに巧みな比喻で、現代の日本語を見渡しても、ここに指摘されるとおりの景観を呈しています。あちこちの路地で種々の方言や位相語、専門語などの仲間うちの言葉が語られているかと思うと、広場では全国どこでも通用する共通の言葉が用いられています。万葉人の愛用した「人 春秋 山川 道波 船花 梅 思 う 聞く 来る 咲く 知る 散る 行く …」などの由緒ある言葉を千数百年を経た今日でも常用している一方で、「オープン クーラー ワープロ CD …」などが家庭に備えられ、「ASA EC IOC ODA OPEC …」などを知らないと生活に不便を来します。中国との交渉はとりわけ古く、時代も出所も異なる大量の漢語を受け入れて、今では漢語は日本語に不可欠の地位を占めています。

このように現代の日本語は、新旧さまざまな言葉から成り立っていますが、このように異質の言葉が乱立して混乱を起こしているかといえば、大局から見ると大した混乱もなく、やはり全体として一つのまとまりをなしています。最近片仮名言葉が著しく増えて日本語が乱れていると嘆く声もありますが、このような現象はどの時代にも大なり小なりあったと思われる。

語彙というのは、その周辺は常に新しい波が押し寄せてきて異分子に侵食されますが、同時に自

* 上智大学名誉教授

然淘汰の力が働いて、時代時代でそれぞれの均衡を保ちながら徐々に動いていきます。ばらばらに見えながら、秩序と統一があり、概してきれいな体系をなしています。

2. 語彙の体系

それでは、語彙が体系をなしているという事実を、日本語における概念の組織という観点から見てみましょう。概念の組織というのは、概念には抽象のレベルに差があり、抽象度の高い概念が抽象度の低い概念をその中に包み、概念相互が含み含まれる関係で成り立っていることをいいます。植物分類学では、この概念の組織を利用して、抽象度が高く、範疇の広いものから順に、植物を「門 綱 科 属 節 種 品種」などと分けています。

資2

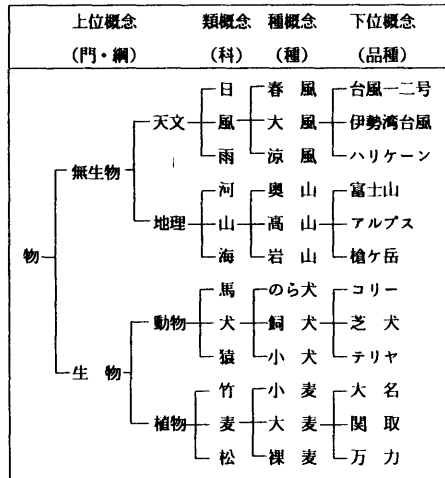
門	綱	亜綱	科	種	品種
緑色植物	種子植物	単子葉類	ラン	キビ	欽傍 神力 八雲
			イネ	イネ	
			ユリ	ヒエ	
		双子葉類	キク	カボチャ	毛馬 北京 地道
			ウリ	キュウリ	
			シソ	カラスウリ	

一般にわれわれは、「神力はイネです」「カボチャはウリだ」などといますが、この判断は、言語的には「神力」は「イネ」に含まれ、「カボチャ」は「ウリ」に属することを表しています。

日常の言葉も同じで、概念は含む含まれる関係で組織されています。複雑で錯綜した日常の言葉を単純化して図示し、植物分類学と対比してみましよう。

日常の言語も植物分類学と同じように、「春風・大風・涼風」が「風」に含まれ、「小麦・大麦・裸麦」が「麦」に含まれ、概念が互いに含み含まれる関係で組織されており。簡単なモデルですが、この表からでも概念のレベルのごとに、言語の質と構造が違っていることが窺われます。各概念のレベルごとに概念の質と語の構造を見てみましょう。

資3



3. 概念の質と語の構造

1 類概念の位置にある語彙 日本語の概念組織の基礎的な枠組みとなるもので、その圧倒的多数が和語単一語基で占められています。和語単一語基は千数百年の間生き続けてきた言葉で、それ以下の小さな要素に分割できない日本語の最小の単位です。

春 秋 日 月 雨 風 山 川 人 牛
馬 鳥 魚 虫 松 竹 梅 桜 木 行 く
咲 く 言 う 見 る 起 き る 過 ぎ る 赤 い 高
い 清 い 寂 し い

これらの和語単一語基は、どの時代・どの地域にも共通して用いられた日本語の基礎語で、日本語の概念の枠組を作り、造語要素となり、日本文法の基礎となっています。

2 種概念の位置にある語彙 種概念を示す言葉は、「種差+類概念」という構造をもち、「春+風」「奥+山」となって、これらが「風」「山」という類概念に含まれる種概念であることを表します。類概念が和語であるため、種概念も和語で作られますが、日本語独特の漢字の音訓の変換によって、和語に当てた漢字を音読することによって、「涼+風 春+風 高+山 火+山 獅+犬 愛+犬 …」などの漢語も和語に基づく種概念になってしまいます。

3 下位概念の位置にある語彙 植物分類学の品種以下のレベルにある下位概念は、すべてが物の名前です。物の名前は、目的・分野・時代によって種々の型と傾向があり、言葉の種類も和語・漢語・外来語などさまざまなものがあります。専門用語は下位概念の名前を特に多く発達させています。

(水稻) 八雲 神力 玉選 秀峰 邦栄 こしひかり (化学) 青口 亜鉛華 アクチノン 亜硝酸エステル (工学) 挟接^{ハツスツグ} 牀板 円縁 歪角 ボギー機

ほんの一部の例ですが、名付け方は専門分野と時代によって傾向が異なります。右のうち、化学用語は最近のもので、工学用語は明治時代のものですが、水稻も含めて、専門外のものには実物と照合することができません。そういう意味で、専門用語は仲間うち言葉であり、類概念や種概念の言葉に対して、日本語の語彙体系の周辺区に位置するといえましょう。要するに、新旧の交替が激しく、短命で、新しい物の誕生や言葉の変化とともに、多くの名前が忘れられていく運命にあります。

4 上位概念の位置にある語彙 このレベルにある語彙は、圧倒的多数が漢語によって占められています。もちろん和語・外来語もあり、特に外来語は現在、急激に増加する傾向を見せています。

抽象語のレベルにある語彙も、a・類概念に基づいている語彙と、b・類概念から遠く離れた難解な周辺区の語彙の二つの層に分けられます。

aは、主として漢字の音訓を変換する働きによって、和語に基づいて作られる抽象語です。

みる { (観) - 察 一 光 拝 一 傍 一
(見) - 物 一 学 発 一 管 一
(看) - 病 一 護 一 視 一 取 一
(視) - 察 正 一 巡 一 疑 一
おもう { (思) - 考 一 想 一 念 意 一
(念) - 願 懸 一 思 一 断 一
(意) - 思 一 向 一 創 一 随 一
(想) - 像 感 一 思 一 回 一

これらは、訓のある漢字を音読することによ

て、容易に理解することのできる漢語で、日常用いられる広場の言葉といえましょう。また、漢語には和語を仲介としなくても、その造語力によって、組織的に抽象語を作ることができます。たとえば、

(学) 哲 一 科 一 化 一 文 一 歌 一 詩 一
英 一 語 一 農 一 業 一 心理 一
(論) 概 一 本 一 序 一 結 一 俳 一 歌 一
詩 一 空 一 持 一 愚 一 観 念 一
(法) 商 一 民 一 刑 一 公 一 司 一 私 一
六 一 税 一 違 一 合 一 国際 一

など、「一学」「一論」「一法」が類概念となっていて、これらの漢語の示すものが、どの類に含まれるかを組織的に理解することができます。それに対して、外来語の場合は、「エコノミックス ケミストリー サイエンス サイコロジ フィロソフィー リテラチャー …」のように、抽象語が組織的に整えられていないので、個々別々に学の名前を記憶しなければなりません。もちろん漢語にも、「糾弾 勲愨 跋扈 凱旋 邂逅 恍惚 拷問」のように分類しにくいものもありますから、これらの漢語と外来語系の多くの新しい抽象語が、類概念から隔たった周辺区のb・難解な語彙の部類に入ることになります。

以上、ウィトゲンシュタインの言う、新旧さまざまな語彙からなる日本語の町の地図を、試みに描いてみると、次のようになります。

資 4

周辺区	拡張区	中央区	拡張区	周辺区
難解な漢語	○漢語	○和語単一語	○和語複合語	和語
	和語	漢語	○漢語	名前
新外来語	外来語	外来語	外来語	漢語 外来語

町の中央区は、千数百年の伝統をもつ和語単一語で固められ、それに類概念の位置に昇格してきたなにかの漢語(菊 蝶 地…など)や外来語(ペン インキ カメラ…)などが加わっています。中央区の回りには、和語単一語に基づいて拡張された種概念と上位概念の区分が広がっていて、この中央区と上下概念の拡張区は、全国的共

通語圏といってよく、広場の言葉の性格もっています。周辺の新開地域は、外からの移住者が多く、主として小さなグループにだけに通用する物の名前や専門用語が大量に存在し、多くは苛烈な生存競争を展開して、生滅ただならぬ様相を呈しています。下位概念で、しかも大量であるだけに、命名には多くの問題を抱えています。

それでは次に、江戸時代、明治時代、現代の専門用語の命名法を考えることにします。

4. 蘭学時代の術語（資料5）

江戸時代の命名法で、ご覧のようにオランダ語(N.)の語基の一つ一つに漢字を当て、それを音読して術語をこしらえています。昔、日本人が漢字に訓を付けて中国語を読んだのと同じ方法です。上の二段が医学用語、下の二段が化学用語で、このような命名法によって、術語が組織的・体系的になり、現代でもこの術語を用いています。もし、江戸時代に英語(E.)が先に入ってきたら、その術語を見ても解るように、とてもこのような組織

資5

N. Been-vlies	Binde-vlies	Hoorn-vlies	Harde-vlies	Net-vlies	
骨膜	結膜	角膜	鞏膜	網膜	
E. periosteum	conjunctiva	cornea	sclerotic	retina	
N. Bloed-steen	Kleine-hersenen	Moeder-vlek	Opper-huid	Gezichts-zenuw	
血石	小脳	母斑	表皮	視神経	
E. blood stone	cerebellum	mother-spot	epidermis	optic nerve	
N. Grond-stof	Zuur-stof	Warter-stof	Kool-stof	stik-stof	Kleur-stof
元素	酸素	水素	炭素	窒素	色素
E. element	oxygen	hydrogen	carbon	nitrogen	coloring
N. Kool-zuur	Salpeter-zuur	Zwavel-zuur	Phosphor-zuur	Zout-zuur	
炭酸	硝酸	硫酸	磷酸	塩酸	
E. carbonic acid	nitric acid	sulphuric acid	phosphoric acid	hydrochloric acid	

資6

(工学) 収差	無液験圧器	汞和金	電槽	纜	発条
aberration	aneroid barometer	amalgam	battery	cable	spring
(物理) アベラシヨン	アネロイド晴雨計	アマルガメ	バッテリー	ケーブル	バネ
(工学) 交和	分度	擺子	向歪	応力変形	遊間
diffusion	graduation	pendulum	deviation	strain	play
(物理) チラカリ	目盛り	振り子	フレ	ヒヅミ	アソビ

的な命名はできなかつたと思われる。オランダでは、できるだけ自国語で術語を作ろうとしたため、オランダ語に漢字をあてた日本語の術語も組織的になりました。ただ、オランダで自国語で翻訳せず、外来語として受け入れた術語は、日本語でもアルカリ、アンモニア、ウラニウム、カリウムなどと外来語で受け入れています。ここに挙げた術語は、今や専門術語というより一般語(拡張区の語彙)になっていますが、医学や化学その他の学問の基本語が、オランダ語のおかげで組織的になったことは幸いであつたと思います。

5. 明治時代の術語（資料6）

明治になると、日本の洋学は英学全盛の時代になります。医学や化学などの分野では、蘭学時代にできた基礎的な術語とその命名法が受け継がれましたが、明治になって盛んになった学問の分野では、英語はオランダ語とちがって、語基に分解して漢字を当てるのに不適當な言語ですから、術語の翻訳、つまり命名は別の方法が取られるよう

になりました。

人文系の用語は英華辞典などの影響を受けましたが、理工系の用語は参考にするものがなく、大部分は日本人が作ったと考えられます。

ここでは、『工学字彙』（明治21）と『和仏英物理学対訳字書』（明治21）からいくつかの術語を拾いましたが、同じ英語の術語に対する翻訳法がまったく異なっていることがお解りになると思います。工学の方は、蘭学時代の伝統もあって難解な漢字を厭わず漢語で訳そうとしましたし、物理学の方は、無理な造語をしないで、置き換えられるものは日常語で置き換え、置き換えにくいものは外来語として受け入れようとしています。

明治二一年という時代を考えると、工学の術語が普通で、物理学はむしろ特別であったかもしれません。工学の用語から、日常的な用語を拾ってみても、

arch 拱 arm 臂 box 匣 catch 鉤 felt 氈 gravel 礫 nave 敷 lock 闔 asphalt 土瀝青 bolt 繫桿 chain 測鏈 cock 活嘴 cable 鶴頸・打重機

などと、難しい漢字を用いていますが、下位概念を種々の漢語で区別しようとするので、どうしてもこのようになるのでしょう。明治一四年の参謀本部『五国対照兵語字彙』でも、

ring 鑲 rope 索 box 盒 buckle 扣子 noise 喧鬧 alarm bell 号樓

と、日常的な事物の命名に苦労していますが、抽象語はともかく下位の名前には漢語は不向きで、かえって非常に難解になります。私も兵隊のとき、軍隊用語には驚きましたが、兵隊や工事現場で働く人たちの言葉としては、非能率的であったというほかありません。

それに対して、物理学の用語は平易な日常で置き換えたり、無理をしないで外来語を採用したりして、その新しい命名法には感心させられます。現代の専門用語の命名法も、この自由な名付けの影響を受けている所もあるのですが、現代は現代で、この方式ですべての用語をまかなうことが不可能で、現代的な悩みを抱えています。

6. 現代の術語

資7

a 不自然な造語

圧力強さ (compressive strength) 極限強さ (ultimate strength) 主タスク (main task) 運動量厚さ (momentum thickness) 引っかきかたさ (scratch hardness) 曲げこわさ (flexual rigidity) 異常伝わり (anomalous propagation) 一理論段数の相当高さ (height equivalent to a theoretical plate) 一時読み書き誤り (temporary read/write error) 固定ワード長計算器 (fixed word length computer)

b 交ぜ書き

あび法 (brightening) 鋳ばり (fin) えい糸性 (spinnability) 液くん (liquid smoking) えの油 (perilla oil) 解こう (amalgam decomposition) かん出液 (bottoms) せん断 (shearing) 腐れしろ (corrosion allowance) ぎん (grain) 湿ふ (wet gluten) 水素ぜい化 (hydrogen embrittlement) 糖化かん (converter) 水ひ (levigation) ギ酸 (formic acid) ケイ砂 (quartz) ケイソウ土 (silicious earth)

c 意識

出力書き出しプログラム (output writer) 形付け顔色 (raised style printing) 後入れ先出し記憶装置 (pushdown store) 先入れ先出し方式 (pushup) 宛先不明待ち行列 (dead letter) 作業待ち行列項目 (work queue entry) 公衆用市外回線 (message circuit) 瞬間軸線軌跡 (axode) 緩衝域構成単位 (buffer unit)

専門用語は、日本語全体からみるとほとんどが下位概念に位置しますが、学問の発達につれて、下位概念がさらに分化して、現代は長尾先生が言われたように複雑な「言葉の木」ができていきます。しかも、下位の下位に属する概念は、一語では表しにくいので、フレーズの形をとって表されます。幕末・明治期には、これらのフレーズの術

語に対して、

視神経乳嘴 (colliculus nervi optici, 医学)

可変阻汽装置 (adjustable expansion gear, 化学)

平均実効圧力 (mean effective pressure, 工学)

などと、漢字を連ねて一語にするのが普通のやりかたでしたが、物理学では、

静止ノ極角 タシカラシキ誤差 エネルギーノ不滅 見掛ケノ重サ 変則ノ分数

のように、フレーズの形で訳しています。

現代の専門用語は、昔からの伝統もあって、フレーズを一語で訳そうとする傾向が強く、一語で種々の新語を作成しなければならないので、いろいろの困難があると思われます。ここに、私が問題だと思ふ命名法を、a, b, cの三つに分けて拾ってみました。

a 不自然な造語

原語のフレーズを翻訳して、日本語の一語にする方法ですが、ここに挙げた例は、日本語としていかにも不自然に思われます。日常語にも重箱読みや湯桶読みの語があるにはありますが、「圧力強さ」などは無理に一語としないで、明治期の物理学のように「圧力の強さ」とフレーズにするのが自然で、「主タスク」「ワード長」などの漢語と外来語との組み合わせになると異様な感じを受けます。また、「相当高さ」は「相当する高さ」、「一時読み書き誤り」は「一時的な」と文法に即した表現にしてほしい。「引っかけかたさ」「曲げこわさ」になると、なんだか子供か外国人の作った日本語のようです。いくら仲間うちだけの下位概念であっても、非文法的と思える造語はできるだけ避けるべきで、不自然な日本語はいずれは淘汰されていくことでしょう。全体から見ると、この種の造語はそれほど多くはありませんが、目だつので挙げてみました。

b 交ぜ書き

一つの語を表記するのに、漢字と仮名を交ぜて書くのが交ぜ書きですが、これは戦後当用漢字表が制定された際に、学術用語もできるだけ当用漢字表の範囲で表記しようとしたため、このような交ぜ書きが増えてきました。ここに挙げたのは、化学用語ですが、もともと専門用語ですから門外漢に解らないのは致し方ないとしても、漢字で書

けば素人にも解る名前が交ぜ書きにしてあるため、解らなくなったものがあります。

液くん(燻) え(荏)の油 解こう(汞) せん(剪)断 湿ふ(麩) 水菜せい(脆) 化ギ(蟻)酸 ケイ(硅)砂 ケイソウ(珪)土

英和辞書を使って、仮名に漢字を当てみると、事実とはまかくとして、私にも言葉の意味は解ります。術語は符丁のようなもので、専門家が使えばよいのですが、やはり専門外のものにも言葉の意味ぐらいは知らせる必要があります。「解こう」の「こう」などは、始め解らなかつたのですが、明治の用語にアマルガムに「汞」を当てていることに気がついて、やっと理解することができました。また、「ギ酸」「ケイ砂」「ケイソウ」などの漢語を片仮名で表記してあるのも問題で、これでは漢語が外来語のように見えます。専門用語の場合、表記がいかにかつ大切かということを感じた次第です。

c 意識

意識としましたが、原語を敷衍して、丁寧に意味を補足した翻訳で、幕末や明治にはこの種の命名は見あたりません。一語の形をとって、pushupを「先入れ先出し方式」、message circuitを「公衆用市外回線」、dead letterを「宛先不明待ち行列」と訳していますが、辞書のように注釈を加えた命名法です。長尾先生のお話に度々でてきた「行列」などの術語は、私のような専門知識のない者には十分に理解が届きませんが、「宛先不明待ち」とか、「作業待ち」とかの補足語が付いていると、素人ながら「行列」とは何かを考える手がかりは得られます。その点、この種の命名法は私たちには親切ですが、専門家が実際の作業をするときは、一体 pushup と「先入れ先出し方式」のどちらの術語を使うのでしょうか。もし日本名よりも pushup が常用されているとすると、辞書には登録されていても、日本名の専門用語とは何かという問題が生じます。日本語に置き換えるだけでなく、実際に使用されるような術語であってほしいと願うものです。

[注] 現代の術語は、『学術用語に基づく科学技術用語辞典』『工業英語』編集部編のシリーズ、『1. 物理・数学・情報科学・論理学』および

『7. 化学・分光』を用いました。

（お断り）私の話を編集委員の方が克明に文字に起こして下さいましたが、資料なしではお解りいただけないと思って、資料を挿入して書き直しました。資料と話の主旨は当日と変わりませんが、話し言葉では入れられなかった漢語類を説明の中に少々加えました。厄介な文字化をして下さった委員の方のご苦勞に心から感謝いたします。

●柴田

森岡先生、どうもありがとうございました。ご紹介するまでもなく森岡先生のバックグラウンドは人文科学なんですね。しかし、概念の階層構造をお考えになっている点は、長尾先生と全く同じであります。ただ、江戸時代のこと、あるいは明治時代のことまでお触れになりまして、さらに現代の、特に工学、物理、化学、そういう術語に対する批判まで伺えたんですけれども、これは反対のことも覚悟しなければなりません。今文部省では言語学の学術用語を検討中で、今年度中に決まります。おそらく1年、あるいは1年半後には本になって出ます。森岡さんは国語学、私は言語学ですけれども、しかし広い意味での言葉の学問として、その術語集が出た時には、長尾先生、日高先生からの厳しい批判を覚悟しなければいけない。これは、予告だけしておきます。

それでは、パネルディスカッションの方に移りたいと思いますので、3人の先生方、演壇の方にどうぞ。ここで、お約束どおり、パネルディスカッションの方は横井さんにバトンタッチをいたします。どうもありがとうございました。

パネルディスカッション

司 会：横井 俊夫（日本電子化辞書研究所所長）
パネリスト：長尾 真， 日高 達， 森岡 健二

●横井

司会を司ります日本電子化辞書研究所の横井と申します。日本電子化辞書というのも非常に怪しげな造語であります。ご勘弁下さい。

このシンポジウムの標題ですが、「専門用語の生態学」という、これも非常に怪しげな名前が付いております。これは柴田先生のひんしゅくを覚悟の上でエンジニアらしい強引な怪しげな名づけをしたものですが、長尾先生からは、専門用語に関わる総合的な情報を体系的にどう構成するかというお話、それから日高先生からは、工学的なモデルをベースにしながらか造語のメカニズム、森岡先生からは、万葉時代まで言及されながらの面白いお話を伺い、はからずも何となくシンポジウムの標題に近いという姿になったのではないかと思います。

時間も非常に短いので、先ずお三方の先生方に、それぞれほかの講師のお話を聞かれた上で、最後につけ加えるお話を一言ずついただいて、それから会場の方々からのご質問、コメントを交えてのディスカッションに入りたいと思います。

それでは、ご講演の順番で長尾先生からお一言だけつけ加えることができましたらお願いいたします。

●長尾

最近、辞書に関する関心が世界的に高まってきているように思います。特にヨーロッパとかアメリカでは一般用語の辞書につきましても、そういう傾向があります。機械で使う場合には人間が使う場合と大分様相が違っていて、人間が使う場合は、辞書にミニマムなことが書いてあっても、人間が常識とかいろんな能力でもって読めば大体ど

ういう意味であるというのは分かるわけですが、機械で辞書を使ってやろうとすると、すべてのことを明瞭に書いておかなければだめだと、そういうことになってきております。それで、そういう点から辞書を見直そう。それから、機械でなく人間の場合でも、native language でない言語の辞書を使おうという場合には、コンピュータが辞書を使おうという場合とあまり違わないというか、似たような面がありますので、母国語の人が使う辞書と違って外国人が使う時の辞書はいかにあるべきかということ、機械が使う辞書とがオーバーラップしてくる可能性が非常に高くなってきている。

そういう意味で、辞書の中にどういう情報でどういう形で入れておくべきか、ということが、大いにディスカッションされるようになってきておりまして、これは私どもにとって非常に面白いことだと思っております。

●日高

私がやりました研究に関して言いますと、さっき森岡先生の方から、専門用語の命名法でこういうものはいかがかと思われるおかしなものがある、という話がありました。けれども逆に、では、おかしいとは何か、どうだからおかしんだということを、例えば単語をおかしい単語とおかしくない単語に分けて、それぞれの造語モデルを作ってやる。そして、どういうふうに書換え規則、つまり文法が違って来るかを調べてみるのも一つの方法かと思えます。

●森岡

先ほどちょっと申しましたように、専門用語と

というのは、一般の人たちには分からないことが多いし、分からなくても仕方がないと思っているんですけども、ただその中でも、やはり基礎的な専門用語というのは、どうか広場の言葉をもっていけるように作っていただきたいということです。細かい枝葉末節のことは専門家に任せておいてもよい。江戸時代に作られた酸素とか、窒素とかいうのは非常にうまく作られていると思います。その酸素、窒素をもとにして大きな、長い単語を作りましたけれども、情報科学で一番社会人にも分かって欲しいというような基礎語を先ず教えていただきたい。そう思っております。

●横井

どうもありがとうございます。それでは会場の方からのご意見、ご質問をいただいて議論を進めたいと思います。

●中山 亮一（日本テレビ文字放送センター）

長尾先生にお伺いしたいんですが、具体的な用語をまとめられるということ、文例をもって百何十人の方にお示しになられたという、その辺のノウハウは今後のあらゆる雑誌あるいは辞典を作るのに非常に貴重なノウハウだと思いますが、その辺をちょっとご説明いただけますか。かかった時間とか。いろいろな方が非常に苦心されて、それが一般化されますとあらゆる部門においての辞書を作るのに非常に有効になってくるという気がいたしますので、その重要性をちょっとご説明いただきたい。

●長尾

あんまりご説明するほどのことでもないか思うのですが、「岩波情報科学辞典」の用語については、荒っぽく考えますと三つ、四つの違ったタイプの用語があります。抽象的な機能に関する用語とか、あるいは具体的な対象物に関する名前とか、その他ありまして、それぞれにやっぱり少し違った書き方がどうしても必要になってきます。それに関していろいろ考えました。こういう種類の単語に対しては、こういう順序に、こういう内容を、できるだけ詳しく書いて下さい、というようなこ

とだけでなく、具体的にどういう単語についてはどういう書き方をするのがいいかという、いってみれば規範文例のようなものを何種類か作りました。そして、この単語については、こういうふうに書くのが一つのサンプルですよというのをお見せしました。それ以外に、一度書いた物を読んでもみると、この表現は冗長であるとか、これはもっとこういうふう書き直したほうがいいのか、あるいは普通の人が読んだ場合にこういう順番に読んでいった方が分かりやすいから説明の順序を入れかえるということが出てきます。そこで一度書いたものを修正したらこういうふうになりますよ、といった具体例を示しました。そうすると、最初に書いた文体と次に書いた文体はやはりちょっと違う。さらにもう一度読むとこういうふうになりますよと、まあ自分の恥を曝すようなものですが、文章の推敲を3段階くらいやった実例を見せました。最初に書いたものを refine するとこうなりますよ、necessary and sufficient に書くとすれば、こういうことに注意してこういう情報をつけ加えるとか、こういうものは削除するとかして書いたのがこうですよというわけです。自分で書いたものについてはそういう review and check をして下さいとか。こういうことを色刷りにしまして——まあ受験雑誌の本みたいなもんですけれども——お配りしました。それから、書いていただく時には、この用語については、こういう内容についても頭において、しかも参照しながら書いていただきたいというのを、用語の木の一つの sub-tree を見せるとか、あるいは sub-tree だけでなく、この単語にはこういう大切な連想的な単語があるからこれも忘れずに refer しながら書いて下さいとカードをお渡したとか。こんなふうなことをいくつかやりましたけれども、詳しいことは忘れてしまいました。

●長山 泰介（日本医薬情報センター）

泥臭いことをちょっとお聞きいたします。今、長尾先生がおっしゃったように非常に緻密なサンプルをお渡しして書いていただいたにしても、やっぱり非常にばらばらになると思うんですね。で、それをチェックする体制はどういうふうにされた

のか。長尾先生がほとんど一人でやられたのか、それとも何人かの方がやったのか。あるいは、場合によってはこの点もう一度書き直してくれとか、簡単なことであれば少し直すとか、その辺のことが実際はまあ頭の痛いことでございますけれども、どういうふうにやられたのでしょうか。

●長尾

詳しいことは忘れてしまったんですが、一番大きなフィールドが6つか何かあって、そのそれぞれのフィールドに2人ずつの編集委員を貼り付けてまして、その編集委員2人ずつの下にそれぞれ7、8人ずつの方を配しました。一つのフィールドに10名くらいがおられて、その方が最初に書かれて出てくる説明をチェックして下さいました。学術的に間違っていないかどうか、学術的に抜けのある書き方になっていないかどうかをチェックしてもらい、著者に戻して書き直してもらおうということが何回かありまして、その後、一応大体できたという時に計算機に打ち込んでゲラを出し、それを一番上から次のレベルの30人くらいの方が読んで下さいました。その後、それを用語の木にしたがって打ち出したものを、私が全部3度ほど読んで、この文章はやっぱり分かりにくいんじゃないかと、これは何だとか、相当やりまして、書いて下さった方には大分失礼なことになりました。

●濱 豊彦(横川・ヒューレット・パカード)

辞書の見出しといいますか、用語を抽出する時に、こういうことは起こらなかったのでしょうか。同じ意味なだけけれども、ちょっと微妙に表現が違うというような、例えば、デジタル処理というのがありますが、これはデジタルと表現しろと決めてある全国紙もあるわけでございます。こんなことはなかったのでしょうか。あったならば、どんな具合に処理したか教えていただきたいんですが。

●長尾

それはものすごくたくさんありまして、用語を集めた段階ではいろいろなものが出ました。「岩波情報科学辞典」としては何を standard とし考えるかというのは、6人か7人の一番上の編

集委員の方々の会合でいろいろディスカッションして、まあとにかくこれに決めようというふうにやったんです。特別にこれに関する基本的な理論というか、principle はなくて、結局いろいろな使い方を想定してみて、やっぱりこれが一番抵抗感が少ないなという、そんな程度でしか決められませんでした。

●細野 公男(慶應義塾大学)

専門用語を聞いてこれは何かおかしい、あるいは用語としてどうもしっくりしないというようなことは、確かにあると思うんです。この件に関しまして森岡先生からいろいろ例を出していただいておりまして、非常に納得するんですけども、先ほど日高さんもおっしゃいましたように、なぜおかしいと思うんだろう、どういうところがおかしいんだろうかということに関して、ちょっとお教えいただければと思います。

それから日高さんにお尋ねしたいんですが、いわゆる推移確率でもって造語を作っていくわけですけども、これは今までの辞書の中の語に基づいて一つのルールをお決めになったと思うんですね。ですから、いわばボトムアップ方式だろうと思いますけれども、例えば、おかしいということが何かルール化できたならば、トップダウン方式に何かそういうふうなルールを入れられるのかどうか。例えば、これはどなたに聞いたか覚えていないんですけども、駅で券を買われますね、自動販売機で。あれを「売券機」という。「売券機」とはいわない。中国では「売券機」というんですか。何かそのようなお話を聞いたことがあるんですけども——。日本語の文法からそういうルールが出てくるのなら、そういう情報が入れば造語も変わってくるかもしれない。それから、音韻ということもあって、こういう音と音がつながると何となくきたないとか、そのようなことも、何か造語には関係するのかもしれないのか。この点どのようにお考えになるか、お聞かせいただきたいと思います。

●森岡

確かに、おかしいかおかしくないかということ

は主観的なように見えますけれども、幕末・明治からずーっと専門語を作ってきた経路をたどってみますと、先ほど申したような例は、何か非文法的なような気がして仕方がないんです。つまりこの頃、「見れる」「食べれる」といった「ら抜き言葉」がはやっているように。「一時読み書き誤り」という述語は、何のことかよく分からなかったんです。専門の方はお分かりになると思いますけれども、これは「機械の一時的な読み書きの誤り」と「的」を入れて下されば文法的になります。それから「永続読み書き誤り」も同じですね。何か「ら抜き」「的抜き」などは舌足らずで非文法的と言いたくなります。

それから、漢語と漢語の組み合わせとというのは、日本人は割に自然に受け入れられますけれども、外来語と漢語との組み合わせというのは、もちろん「アメリカ合衆国」などとたくさんありますけれども、どうも「ワード長」というのは音節の具合から見ても何か感覚的に非日本語みtainな気がしてしょうがない。これは主観でしょうか、どうなのでしょう。慣れればいいのでしょうか。やっぱり「語の長さ」と訳してほしい。「固定ワード長計算機」なんかもうちょっと何とかかなりそうな気がします。こんなことをいっては大変失礼なんですけれども、造語能力の衰弱だといいたくなるんです。

●日高

言葉の並びの順番で、こうも言えるし、こうも言えるんだけど、どちらが日本人として自然に感じるかということに関して、「券売機」と「売券機」というお話がありました。「券売」にしても、「売」が主な意味(head)になるのですね。headに当たる語が後ろにあった方が安心するといえますか、日本人には自然に聞こえるんですね。ところが「売券機」といいますと、漢語の影響だろうと思うんですけど、「売券」と目的語が後ろにくるわけですね。だけど、「売券」でも、「機」に掛かっている意味はやっぱり「売る機械」という意味で掛かるわけで、「売券」の主たる意味は「売」です。確かに単語としてはあるけれども、現在の日本人が——漢語の知識が多い時代の

日本人と現在の日本人とは違うので——どちらが自然かというとおそらく「券売機」の方ではないかと思います。

それからもう一つ、意味的なつながりだけでなく音もあるのではないかと。確かにそうでありまして、私も今日お話するのを忘れていたのですけれども、単語は、書き言葉であると同時にしゃべり言葉です。だから、書いている場合は意味だけでよろしいんですけども、しゃべり言葉として相手に情報を伝えるという時には、前の言葉と後ろの言葉の音のつながりがしゃべりやすいことが望ましい。だから造語モデルなんかを考える時には、音の結びつきと意味の結びつきと、両方を考えなければいけない。そういう意味では、造語モデルの造語単位というのは、今日は漢字だけを見せましたけれども、実は漢字と読みのペアになっておりまして、読みの方が音を示しており、漢字の方が意味を表している。そういう単位で造語単位を考えているわけです。そうすると、この場合は読みがスムーズにつながらないから単語になりにくいなあという機構もある程度つかめるのではないかと考えております。

●井上 孝(翻訳業)

日高先生にお伺いします。素人なのでとんちんかんな質問かもしれませんが、一つは、このモデルをお作りになるに当たって、仮定1にあるように出発点としては漢字云々という字種に依存した命題をおいていらっしゃると思うんですが、できあがったシステムは字種とは独立な形で動くといえますか、字種のことは考えなくても一つの体系になっているのかどうか。と申しますのは、先ほどの森岡先生のお話にもありましたけれども、交ぜ書きの術語とかで頭が平仮名で始まっているとかいうのがいろいろあるわけですね。最初のコーパスとしてはそういうものは扱いにくいから落としたけれども、システムが上がってしまえば、そういうものが出てきてももう対処できるという状況なのかどうかということなんですけれども。

それからもう一つは、同じようなことなんですけれども、それが何らかの用途をもったツールとしてインプリメントされた時にはそのテキストの

文字種はもう考えなくてもいいということなのかどうか。

●日高

造語モデルは意味と音のペアで、意味は漢字といましたけれども、実は意味をもった最小単位であればよろしいんですね。そういう意味では表記法とは一応別のものであるということです。ですから、造語モデルで生成された語のある部分を漢字で書くか平仮名で書くかという、表記法については、別の機構であると考えています。表記法について言えば、例えば「繰り返す」という単語のいろいろな表記を作る時に、「繰り返す」とやさしい漢字を仮名で書いて難しい漢字を漢字で書くということは人間の場合は普通やらないわけですから、表記に直す時にある種の優先ルールがあるかと思う。こうも表せるし、こうも書けるんだけれども、こういう表記で現れることは少ないよ、ということはあるかと思う。

●長尾

私どもの分野の、特に翻訳語について、英語の単語を日本語でどういうふうしているかという現象を見ていますと、英語である単語が出てきた時に、日本国内でそういうことに関心をもっていろいろ研究をしている人たちが個人個人で勝手に好きなように翻訳して使いだすわけです。そうするとほかの人も自分で考えるのは面倒くさいからその単語を使う。そうすると、ある分野の人たち間でその日本語表現が割とポピュラーになる。そのポピュラーになった段階で、ほかの人がその分野に入ってきた時に、はたと考えて、この訳語はちょっとよくないなあ、こういうような訳語の方がいいんじゃないか、といってももう手遅れで、それが流布しちゃってるからいまさら変えるのも面倒くさいなというんで、妙な訳語がそのまま定着するという現象もあるわけです。誰かが最初に翻訳して使った時に本当に慎重に考えてやったかという、最近忙しい時代ですから、適当に考えてやるという程度なのです。何か用語に関するコミッティーでも作って、新しい用語が出てきたら慎重審議してこの訳語にしましょうとかいって

決めれば、森岡先生がおっしゃっているような非常に quality の高い訳語が揃うんでしょうが、実際上そんなことはやってられないんで、おかしなことがいっぱい起こることじゃないかと思うんですね。

●森岡

一つだけ申します。明治14年に「哲学字彙」という哲学関係の用語辞典が出ましたけれども、それを見ておきますと、西周とういう方の「観念」とか「理論」とか「絶対」とか「相対」とか、現在でもこの用語を使わないと文章が書けないくらい基礎的な言葉が出ています。それでもその「哲学字彙」中の半分以上は現在捨てられています。不自然な訳語というのは多分淘汰されてなくなっていくから、あまり心配することはないだろうと思います。

●藤原 譲（筑波大学）

各先生方一つずつお伺いしたいのですけれども、まず長尾先生。用語の体系的構成ということで、語の収集から全体のバランスという意味で非常にご苦心なさったと思います。その意味では大変立派なものができたと思って、感心して拝見させていただいていますし、使わせていただいています。電子出版もお考えということでございましたが、電子出版となりますとどうしてもメディアの容量からいまして、量的にということと、それから使える機能が大幅違ってきますから、いろんなことをお考えだと思うんです。機械での処理という観点で申しますと、先ほど外国人向け辞書をどうするかということについてもお触れになられたけれども、何かこう、比較的完全な系ということもちょっといい過ぎかもしれませんけれども、そういうふうな形である使いやすいと思うんです。そういうふうな、ある程度大量でかつ機械向きということに対して、こういう辞書を作られる時の内容の範囲・量といったことに関して何かお考えがございましたらお聞かせいただきたいと思えます。

日高先生の造語のメカニズムというのは、私全部はフォローしきれなかったんですけれども、基

本的にマルコフ過程でというのは大変自然なお考えというかモデルだと思うんですが、ちょっと気になりましたのは多義性の問題とか、それから実際に作っていきますと言葉に省略を入れるとかいうようなことがあるんですけども、この点についてもう少し説明いただきたいと思うんです。

それから森岡先生、本日の全体のテーマとも関連するんですけども、専門用語も、先生の例を拝見させていただきましても、時代とともに次第に変わっているというような感じがいたします。いい言葉を作るといことと、ある意味で時代に流されてしまうといこととの関連でございませけれども、標準用語があって、みんながそれを守ってくれたらいいんですが、どうしても守りきれなくなってしまう。そういうことについて結局はどうあるべきなのか、お考えをお聞かせいただければと思います。

●長尾

的確な答になるかどうか分からないのですが、電子出版、例えばCD-ROMのような格好で出すという場合には、もっともっと単語数を増やす。今13,000とありますが、多分30,000語くらいにはしないと駄目なんじゃないかと思ひます。ところが、そうするにはどうするかといひますと、用語の自動収集システムを作つて関連のテキストを常にそのシステムを通過させ、新しい用語とか面白い単語とかを自動的に収集するようにしなければいけないんじゃないかというふうに思ひていひます。それから、CD-ROMにした場合の検索の方式については、ハイパーテキストシステムみたいなものをやるのがいいと思ひますが、それをどこまで自動的に作れるかといひことと、それからハイパーテキストのアクセスの仕方についてはいろいろなことが考えられますが、これはそれぞれ森岡先生がおっしゃいましたように自然淘汰的に一番便利なものが最後に残つていかざるを得なくて、今できるだけのことを考えても本当にそれでいいのかどうかといひのはちょっと分からない。そういう感じでおひります。

●日高

造語単位で、文字単位にしましても単語単位にしましても多義があるのではないか、その扱ひをどうしているのか、というご質問ですが、文字単位でも厳密には多義があります。だけどここでは、人手を省くためにそれを区別していないといひこととです。もう一つ屁理屈をいひますと、文字の場合は語源的な意味は一つで、他の意味はそれから派生したものです。その単位から単語を作り上げていく過程を追う時には、その文字単位の語義は語源で充分だと思ひます。ただしこれが複合語になりますと、単語単位で——単語と単語を合わせて——造語していきまますね。そうすると相当言語文化が上がつている。この場合、多義の選択を厳しくやらなければいけない。といひのは、シソーラス自体が多義をちゃんと扱ひていて、この単語の何番目の語義はどの単語の何番目の下位であるといひことを使っているのだから、単語に対してはそれを厳しくやっていかなければいけないと思ひておひります。複合語に関しましてもさっきの派生語に関しましても、実は多義の選択をして造語モデルを作る必要があります。

省略に関しましては、センテンスの場合にも文法処理の場合にも大変難しくてよく分かりません。文章の場合にはほかの文脈とも関係するわけですが、単語の中で何が省略されるかといひことは、よく分かりません。

●大塚

音と訓との混ぜこぜのやつを自由勝手に、ambiguity を避けて通用さえすればいいといひいう実用的な立場に立つことを許すといひうような考えを取り入れるのも一つの方法ではないか。まったく実的にうまく通用すればいい。間違ひが起らないようにambiguity を少なくして通用させればいいといひいう、一つの割り切つた方針を常に採るといひのはいかがでございませう。

●森岡

専門用語の変化の問題でございませうね。明治時代に作られた専門用語と現代の専門用語とで変わつていひるものを見ますと、先ほど申しましたように基本的なものはあまり変わらないんですけども、

どちらかという、用言系といえますか動詞の言葉は非常に変わりやすい。例えば、昔は「摩剥する」といったのを今は「摩耗する」といってみました。皆さんが英和辞典をお引きになると、お分かりになると思いますが、名詞は、漢語訳が非常に多いのですが、動詞は、まず和語で訳して、それから漢語訳が出てきます。つまり、動詞はもともと抽象語で、その多くは和語で間に合い、専門用語として漢語で言い表してみても、必ずしもその漢語に定着しないで、他に置き換えられてしまいます。そういう点で、動詞系の単語は、比較的自由に改められる場合が多いと思います。それに対して、名詞系の述語は、物の名前ですから、物と固く結びついていて簡単には変えられないので、古くから定着するものが多いのですが、現在では、漢字・漢語などの不的確なものがあって、変える場合には、片仮名言葉にしてしまう率が多いと思います。それから、明治時代でも不安定な専門用語といえますか、一つの訳語に対していろんな訳語がついているものは、相変わらず現在でも不安定で勝手に作られている。例えば「absurd」とか「accuracy」とか「analyser」とか昔から訳語の一定しなかったものは非常に変わりやすく、いつまでも揺れているということですね。なお、もし改める場合には、やはり明治時代の語基の中の一つの漢字を入れ換えるだけの修正とか、すでにできている用語の一部を改訂するというのも少なくありません。これは傾向で、統計を取ったわけではありませんけれども、全体としてそんな感じを受けております。

●吉村 朱美(富士ゼロックス情報システム)

実は今新しい専門用語の流れが三つあると思うんです。一つは今日高先生が話された意味を含めた派生だとか造語。これは明治からずっと続いてきている。もう一つは単なる文字の長さによって決定されて随分不自然な長さにもなってしまうものをどうするんだろうか。これは今いくつかお話が出て、だんだんある形に淘汰されていくだろうと。それから、時代の流れということ。今若い方たちは、いわゆる意味だけでもない、長さでもない、いわゆる音のリズムだけで言葉を作ってしまう

うというところがありますよね。C言語というのがありますが、聞いていると「Cプラプラをやっている」という。「Cプラプラってそれ何よ」というと、「C+言語に+が付いたもの」なんだという。「じゃあ元はC言語というものがあって、それが拡張されてC+になって、それをもうちょっとレベルを上げたものか」というと、まあそんなもんだという。そういう形になると、実はその分野ではもう「拡張」ではなくて「プラス」という言い方をしてしまい、私みたいな、専門から少しずれると何で「拡張C言語」といわないのかと思います。「拡張C言語」なんてまだるっこしいというんです。そうすると言葉に2本立てあるいは3本立てができてしまって、「拡張」という言葉が「プラス」という言葉に取って代わられるということも考え得るわけなんです。そんなふうにくっつかの流れが複合して造語に働きかけていると思うんです。そういうことでいえば専門用語なんかでも一種類には絶対収まってこないだろうと。で、外からみる、さっき先生がお話になられた根の部分はやっぱり分かるように「酸素」は「酸素」で、「拡張」は「拡張」で置いておいて欲しいなというイメージがあるんですが、「Cプラプラ」とやってしまう。そういう傾向が自然に消えていくものか、やっぱり何本立てかで今後も残っていくものなんでしょうか。

●森岡

言葉に関心を持つ人は、この頃日本語が駄目になってカタカナがいっぱいあるとって嘆いておりますが、私はあまり心配していないんです。と申しますのは、フランスというところは言葉を非常に大事にするというので有名ですが、フランスの60くらいのおじいさんで、若い娘の言うことがこの頃さっぱり分からないといって、嘆いているという話を聞いたことがあります。言葉は、みんな変わっていくわけですね。日本でも若者の言葉が心配になるようなテレビ番組もありますけれども、その若者たちがちゃんとした場所ではそんなくずれた言葉を使わないで敬語も使える。敬語はいらないなんてといいながら、やはり日本語では敬語を使わざるを得ないわけですね。明治時

代の人は、漢字というのは非常に能率が悪いからもう漢字はやめようと一所懸命になったんですけども、漢字はやめることはできない。敬語も無駄だからやめろといっても、絶対敬語はなくならない。これは日本語そのものなんですから、なくならないと思います。専門語にからめていいますと、外来語というのは日本人にも苦手ですが、日本語を勉強している外国人にも日本語でいちばん分かりにくいのは外来語だそうですね。元の自分たちの言葉に直せないらしいのです。私の書きました「言葉の町」の地図に、日本語のコアになっている共通語と、中心から離れた周辺という言葉がありました。周辺という言葉というのは泡のように浮いて泡のように消えていく運命にありますので、そんな言葉がどんなに乱れていようと日本語の中心の部分がきちんとしている以上心配ない。私は、その周辺言葉はバブルであって、どうせ消滅する運命のものが多いと思います。

●横井

長い間3人の先生方のお話を交えましてシンポジウムを続けてまいりました。確かに用語が乱れてきておりますし、たくさん用語が行き交う時代になりました。したがって明治の時代のように少数の権威者が用語をコントロールできるという時代ではなくなった。そういう意味からしましても用語の作り方——正しい用語の作り方といいますか——であるとか、その用語の定義の仕方というものに対しまして、簡単なルールであるとか何かそういうものを作らなければならない時代になったのかもしれない。まあ、大塚先生がいわれたように簡単なルールを当てはめてしまえというのも一つの行き方なのかもしれません。今日の議論を一つの出発点にそういう議論が起きればと思います。

専門用語研究会

目的

専門用語研究会は、専門用語全般に関して研究を促進し、情報交換の場を提供し、会員相互の交流を図って、専門用語に関する研究や技術の向上に貢献することを目的としています。

活動内容

1. 研究会、講演会、分科会などを随時開催します。
2. 刊行物として、会誌を年2回、ニューズレターを随時、発行します。
3. Infoterm、ISO/TC37等との連絡を密にし、国際交流活動を推進します。
4. このほかこの会の目的達成に必要な活動を行いません。

運営

役員として、会長、理事、監事を置きます。

専門用語に関心を持つ個人または機関は、どなたでも入会できます。

年会費

個人会員 5,000円
機関会員 30,000円

投稿規定

「専門用語研究」投稿規定

1. 「専門用語研究」（以下「会誌」という）には、ターミロジーの理論と応用、専門用語集の作成技術、その他、専門用語に関する記事を掲載する。
2. 会員は、「会誌」に投稿することができる。投稿にあたっては、「会誌」投稿の手引きにもとづいて原稿を作成する。編集委員会からの依頼によって執筆する原稿も同様とする。
3. 学術論文の原稿が提出された時は、編集委員会は査読を経て、その採否を決定する。
4. 「会誌」に掲載された論文の著作権は、原則として専門用語研究会に帰属する。

編集後記

新編集委員会のもの第1号は、昨年12月に開催されましたシンポジウムの報告です。各ご発表の先生方には、大幅な修正・加筆あるいは書き直しをしていただきました。当日ご参加の方々にも新鮮な目でお読みいただけることと思います。

新しい編集委員は、下記のとおりです。今後、会誌およびニューズレターの企画・編集を担当してまいります。

次号の会誌は、年内発行を予定しております。会員各位からの投稿をお待ちしております。(戸塚)

編集委員会委員

太田 泰弘 文教大学
加藤 信哉 東京大学附属図書館
後藤 智範 神奈川大学理学部
四ノ宮明夫 大正製薬研究開発計画部
戸塚 隆哉 紀伊國屋書店電子情報部
(委員長)
中山 亮一

専門用語研究 第5号(1993年7月)

発行人 大塚 明郎
発行所 専門用語研究会
〒141 東京都品川区西五反田7-1-9
HSビル
Tel.03-3495-4511
Fax.03-3495-2995