

専門用語研究

Journal of the Japan Terminology Association

No.10 1996. 01

座談会 学術用語集作成の問題点 青戸 邦夫氏を囲んで(1)	1
FID 東京大会から	
カテゴリー 2(Information Science of Classification and Terminology)を 中心に発表論文の紹介	11
EDR 電子化辞書と情報科学	横井 俊夫 12
用語データベース設計の基本方針	佐々木由香 17
	ゲオルグ・J.アンカー
抄録による紹介	29
専門用語研究会アンケート 集計結果	32
投稿規定	36
専門用語研究会のご案内	37
編集後記	38

専門用語研究会
Japan Terminology Association

座談会 学術用語集作成の問題点 青戸邦夫氏を囲んで (1)*

出席者 青戸 邦夫 (学術情報センター) 仲本秀四郎 (IRIS)
太田 泰弘 (文教大学, 司会) 中山 亮一
香川 靖雄 (自治医科大学) 牧野 正久 (東京理科大学)
四ノ宮明夫 (大正製薬) 山下 泰弘 (電気通信大学)
戸塚 隆哉 (紀伊國屋書店) 山本 昭 (関東短期大学)

戸塚 本日はみなさまお忙しいなか、ご出席くださいましてありがとうございました。最初に編集委員会を代表して、この座談会を企画、開催しました主旨について簡単に申し上げます。

きっかけは、1994年12月の専門用語研究会シンポジウムにて、井上如氏の講演の中で、一度青戸邦夫氏の話聞く会をもつたらとの提案がなされたことでした。また、今年4月、青戸先生が、『学術月報』に「戦後半世紀の学術用語標準化の歩み」をまとめられました。そこで、先生を囲んで、学術用語集作成にまつわるさまざまな問題点について話し合いを行い、その内容を会報に掲載したいと考え、この座談会を企画した次第です。この座談会の目的として、

- 1)学術用語集作成にまつわる問題点を浮き彫りにすること、
 - 2)できればその解決策についても検討、何らかの方向を出したい、
 - 3)文部省の学術用語集だけでなく、JIS、民間で作成する用語集も同時に取り上げたい、
 - 4)そして、残った問題点については、当会の重要活動の一環として今後継続して検討していくことにする、
- などを挙げました。

本日の司会を編集委員でもおられる太田さんをお願いしました。どうぞよろしくお願ひします。



太田 はい、それでは、どこまでうまく進行できるかわかりませんが、意義のある座談会にしたい、みなさんから、きたんのない問題提起、ご意見、ご批判、ご提言などいただければ幸いです。最初に、青戸さんから、ごあいさつがあるようですので、どうぞ。

青戸 専門用語研究会は、昭和63年5月に発足以来、ターミロジーに関する研究を続け、活発に活動されており、本日、学術用語標準化をテーマとする座談会を企画され、私がこの席にお招き

(1)

- ・「学術用語標準化の手引」について
- ・手続き上の問題
- ・科研費の基準について
- ・後継者の問題

(2) (次号)

- ・造語、表記に関する問題
- ・定義の問題
- ・ナカボチ (ナカマル) の話し

* 1995年6月24日(土)紀伊國屋書店セミナールームにて開催

をいただいて光栄に存じます。これからの学術用語の制定事業のありかたについて有益・適切なご提言をご出席の皆様から賜れば大変ありがたく存じます。

さて、最初に、戦後の学術用語の標準化の歩みについて書いてきたものがありますので、これを読み上げさせていただきます。

◆1945年10月1日に、マッカーサー元帥は、GHQ(連合国軍総司令部)にあいさつにきた幣原喜重郎首相に、日本の民主化を進めるための『当面の五大改革』を示したとのことですが、この中の一つ、「学校教育の自由化・民主化」に関連して、新しい学術用語の制定事業が発足した。

昭和21年11月に「当用漢字表」(1850字)および「現代かなづかい」が内閣告示および内閣訓令で制定・公布された。これらの国語改革は、法令・公用文・報道界など社会一般に広く実施されて、国民の言語生活に深く浸透していった。¹⁾

◆原子爆弾投下による終戦は、科学で破れたとされ、戦後の平和的復興も科学の力にまたねばならず、教育民主化、教育の機械均等の見地からも、科学をすべての人に普及することは重要なことと考えられた。

学界・教育界の要望にこたえ、文部省では、自然科学および人文科学の複雑・難解な学術用語を、平易簡明なものに統一して、新しい学術標準用語を制定・普及する事業に着手した。²⁾

◆戦時中に国家総動員態勢の一環として科学技術全般にわたる標準用語の整備を進めるために、全国科学技術団体連合会(全科技連)と科学動員協会に原案作成の委嘱があって、昭和19年に答申案が作成されたが、終戦とともに断絶した。昭和22年2月に学術研究会議の中に学術文献調査特別委員会学術用語制定科会を設けて、この19年の答申案に戦時中に著しく躍進した米国学術界の新用語を加えて、10専門分野の用語につき調査審議に着手した。³⁾

◆昭和23年7月に、日本学術会議法が公布されたのにもなつて、学術研究会議は廃止され、この事業は文部省の学術用語調査会(昭和24年文部省訓令第1号)に受け継がれ、さらに学術奨励審議会学術用語分科会(昭和24年政令第246号)を経て、現在は学術審議会(昭和42年政令第117号)(より正しくは、その学術用

語分科会)に継承されている。その後、「当用漢字表」「現代かなづかい」「送りがなのつけ方」などの実施の経験から国語施策が再検討された結果、昭和56年に「常用漢字表」が内閣告示された。

専門用語については、「この表を基準として整理することが望ましい」(使用上の注意事項)とする「当用漢字表」は廃止されて、「この表は、科学、技術、芸術その他の各種専門分野や個人人の表記にまで及ぼそうとするものではない」(前書2)とする「常用漢字表」が施行された。⁴⁾

◆学術用語の制定事業は、今日まで一環して継続されており、戦後50年という年を迎えて、戦後の学術用語の標準化事業を概観した拙稿⁵⁾を発表した。この別刷を関係各位に贈呈したが、その歩みを体系的・多面的に一望できる文献としてご好評をいただいている。

◆国内学協会誌の投稿規定調査報告書が最近の「情報管理」⁶⁾に掲載されており、『学術用語集』の支持率ともいうべきものが示されています。和文誌の場合では、学術用語は原則として、文部省編『学術用語集』に準拠すると規定するものが、55誌中の37誌に達すると報告されています。

『学術用語標準化の手引』について

太田 どうもありがとうございました。さて、座談会に入りますが、委員会の方であらかじめ用意した質問事項があるので、これから始めましょう。最初に、『学術用語標準化の手引』が複雑なため、一人の専門家が必要とするほど」という『図書館学編』の担当者からの質問が出ています。これは質問になるのかコメントになるのか良くわかりませんが、青戸さんいかがでしょうか。

青戸 この『手引』の初版は、文部省学術国際局情報図書館課編『学術用語標準化の手引(昭和55年版)』で、関係法令、審査基準、関係資料を整理して載せた資料集でした。昭和56年10月に「常用漢字表」の制定があり、これにともなう学術用語審査基準の改正と、昭和57年11月に学術

1) 「国文学解釈と鑑賞」特集：漢字の新しい書き方・使い方、第225号、1955
 2) 学術用語制定の歴史。日本学。8(4)、1989
 3) 学術用語の標準化28年の展望、学術月報、28(9)、1975
 4) 馬場重徳先生を偲んで－文部省の学術用語制定事

業の回顧。オンライン検索、14(1)、1993

5) 戦後半世紀の学術用語標準化の歩み。学術月報。48(4)、1995

6) 国内学協会誌の投稿規定調査報告書。情報管理。38(3)、1995

用語標準化の促進を図るため、学会などに標準化または制定された学術用語の改訂の計画がある場合に、これに対して科学研究費（特定研究）補助金（以下科研費という）を交付して、学術用語標準化の原案を作成願ひ、この原案を学術審議会学術用語分科会で調整・審査のうえ、『学術用語集』を刊行できることとし、この用語集は文部省と学会などの共編の著作物とするという、学術用語の制定手続きの改善とがあって、これらの新しい資料を編集した文部省学術国際局学術情報課編『学術用語標準化の手引』（平成3年3月）が出されました。

『手引』が複雑だという意見は確かにそうです。科研費というのは自由な研究ですから、勝手にやりやすくなって、『学術用語集』の原案ですと言われても困る。科研費による「学術用語標準化のための調査研究」のための『手引』でもある。この研究成果は、磁気テープの形で提出いただくことになっています。電子印刷で『学術用語集』を刊行するため、学術用語のデータベース化に都合の良いようにというわけです。

牧野 用語を決める委員会が科研費で運営されているから、あまり勝手にやらないようにちゃんとしたルールというか、既存のものを整理したのが『手引』なのだと言戸さんはおっしゃったと思うんです。

青戸 実は、この『手引』の骨子をまとめようということにも科研費が使われています。昭和57年11月に策定された学術用語の制定手続きの改善に基づいて、学術審議会学術用語分科会の意向を受けて、学会などに当該専門分野の学術用語標準化の調査研究に対して科研費を交付して、『学

術用語集』シリーズの統編を刊行する科研費措置は、継続的に毎年2-3件ずつ行われました。⁷⁾

これに並行して、昭和61-62年度科研費（特定研究）「学術用語の標準化のための最適手法の開発とその高度利用に関する研究」（研究代表者井上如学術情報センター研究主幹）によって、(1)学術上の概念を表す用語（学術用語を体系的に整理統一し、その標準化を図ること）、(2)『学術用語集』のデータベース化による高度利用の研究を行い、実用的システムを開発することを目的とした研究開発が行われました。

この研究には、私も研究分担者に一員として参加して、井上さんとともに『学術用語標準化の手引（改定版）』の原案作成にあたりました。これが骨格となって、のちに文部省での検討を経て、現行の『手引（改定版）』が印刷されています。

この研究成果として、『手引』のほかに『学術用語集』21編⁸⁾を入力した『CD-ROM (Sci-Terms) 学術用語集文部省編（予備版）』（学術情報センター、1988、非売品）があります。

牧野 藤原鎮男さんの『文部省学術用語集統合リスト』（1984年）というのは？

青戸 この統合リストは、『学術用語集』23編の用語を、ローマ字のABC順に、用語、英語、出典を挙げてあります（日本語88,975語、英語92,266語）。

この他に、『学術用語集』を集成したものに、自然科学部門24編を電子印刷で和英・英和の冊子体の合本としたものに、『学術用語集集成』（日本科学協会、1984年、非売品）があります。この本は、日本船舶振興会の補助金の交付を受けた

7) 昭和56-57年度『機械工学編（増訂版）』『農学編』
昭和56-58年度『心理学編』
昭和58-60年度『動物学編（増訂版）』『植物学編（増訂版）』
『新言語学編』（後に『言語学編』と改名）
昭和59-61年度『経済学編』『土木工学編（増訂版）』
昭和60年-61年度『物理学論（増訂版）』
昭和60-62年度『電気工学編』（増訂2版）『建

築学編（増訂版）』
昭和61-63年度『歯学編（増訂版）』『遺伝学編（増訂版）』
8) 数学、天文学、気象学、海洋学、地震学、分光学、化学、地学、機械工学、電気工学、船舶工学、航空工学、計測工学、原子力工学、農学、論理学、キリスト教学、心理学、図書館学、地理学、採鉱冶金学

日本科学協会が『総合学術用語集』監修委員会(大塚明郎委員長)を組織し、筑波データベース研究会(中山和彦会長)に業務委託して刊行されました。また、『統合リスト』を作成された藤原鎮男・藤原譲の両先生から提供を受けた『学術用語集』23編の原テープを加工して得られた国立国語研究所の研究成果の野村雅昭・石井正彦『学術用語語基連接表』(1988年)およびこの姉妹編『学術用語語基表』(1989年)があります。これらは、学術用語の研究にとって貴重な資料となっています。

牧野 その後、藤原さんらが1987年に刊行された『総合学術用語集』(発売紀伊國屋書店)は、25部門の用語集を統合して英語のabc順に並べたもので、コンパクトで各部門の日本語の異同がすぐわかり、便利です。CD-ROMにして、調べたいとき自由自在に引けるとよいですね。たとえば「superconductivity」を引くと、電気ではどう決まっています、物理ではどうなっているということが。

青戸 21編の学術用語は、英語とローマ字と日本語とカタカナとの各々から検索できるようにCD-ROMに収めています。その後、学術情報センターでは、「学術用語制定の効率化に関する研究」(平成4・5年度科学研究費補助金)によって、『学術用語集』のデータベース化による高度利用を実現して、学会が科研費措置によって実施する「学術用語標準化の調査研究」を支援する学術用語制定支援システムの開発に取り組んでいます。

仲本 『学術用語標準化の手引』を見ながら各学会に設けられた委員会が責任を持って見るんですか？ それとも誰か専門家がチェックしているのでしょうか？ 文部省の中でその用語集の最終的なチェックは誰がやられるわけですか。

青戸 科研費措置による学術用語標準化の調査研究には、通例で私が研究協力者または研究分担者という形で総括班に参加してきました。文部省に提出された後における研究成果報告書の問題点の

チェックは事実上私がやっています。文部省の事務担当者の手には負えない面がありますので。

仲本 専門分野間の調整は？

青戸 『学術用語集』シリーズの各編に記載する学術用語は、関連専門分野との間で総合調整を経て制定されていて、「学術用語審査基準」細則(9)によってその用語が本来所属すると思われる専門分野の表記を優先的に扱って調整します。この総合調整は、過去に2回行われて、結果が記録されています。⁹⁾

昭和29年2月には、数学、物理、化学、動物、機械、電気、土木、建築、採鉱冶金、農学の10分野の相互間で用語の表記の総合調整が行われ、同年3月に『学術用語集』シリーズの5編(数学、物理、動物、土木、採鉱冶金)が創刊されました。

昭和42年6月に、学術審議会学術用語分科会が発足した時点で、天文、気象、海洋、地震、分光、遺伝、航空、心理、地理の最終案が作成されており、隣接または関連の分野の選定用語の相互間もしくは既刊の『学術用語集』の各編の制定用語との間で表記の総合調整を実施しました。昭和43年3月に調整を要する学術用語を収録した用語集を作成し、物理学・化学系、地学系、生物系、工学系の各調整委員会、そして合同の調整委員会で総合調整を実施し、昭和47年に終わりました。(「総合調整用語集」2冊⁹⁾を出席者に回覧)

仲本 将来総合調整のことをトレースする人がいるかも知れないから、記録をどこかに残すことを考えたほうがいいのではないですか。

青戸 「総合調整用語集」2冊は全部コピーをとって、学術情報センターの学術用語データベースに入力してあります。総合調整で消えた廃語を検索すると電子的に出てくるようになっています。

山下 それはオンラインで引けますか？

青戸 オンラインで引けるよう、将来はサービスできると思います。今はセンターの中だけで、うまく動くか試験中です。

香川 文部省の中の担当部局の名前はときどき変わりますが、窓口はどこに相談したらいいんですか？

青戸 学術情報基盤整備推進室でしたが、最近の人事異動で空席です。7月上旬から、文部省学術国際局学術情報課の企画調整係が担当するとのこと。

牧野 二次にわたる「総合調整用語集」というのを今はじめて拝見しましたが、それぞれ千数百とか2千何百もの項目があって大変でしたね。見出し語はだいたい同じ英語のものを一つの項目にまとめて、関連の複合語があれば併記されていますね。

青戸 総合調整に際しては、全専門分野の制定用語と選定原案の用語を1語1枚で「学術用語マスターカード」(B6)に記入したものを、アルバイトの手を借りて英和形式に配列した。この全カー

ドを私がチェックして、引き抜かれたカードから、「調整用語集」を印刷して、これを調整委員会に付議しました。また、『学術用語集』の出版原稿は、「マスターカード」を英和・和英に配列して、英文タイプによってカーボン紙をはさんでコピーしたリストを関係者で校閲して作成しました。今日のような複写機も電子印刷もないため、大変な手間がかけられていました。

手続き上の問題

太田 それでは観点を変えて、手続き上の問題に入ることにしましょう。やはり『図書館学編』の担当の方から、“発行前に何年もかかる、3、4年じゃ済まない”という声が出ています。もっとすごいものもあるんだと思いますが、このために、嫌気がさしているんだが、科学研究費が魅力になっている。この効率化がもっとはかれないだろうか、ということについて、青戸さんにお話いただきましょうか。

青戸 図書館学編の改定は、平成2年に日本図書館学会の「図書館情報学用語標準化の調査研究」がスタート以来、今年(平成7年)6月に、学術審議会学術用語分科会運営委員会が開かれて、改定案が了承されています。5年かかっています。¹⁰⁾

牧野 世の中に聞こえてくるには、委員会側で3、4年努力してまとめたものが、最終的にOKない



- 9) (1) 学術奨励審議会学術用語分科審議会、総合調整用語集 [A-21-2] . 昭和29年2月, 2,428項目)
 (2) 学術審議会学術用語分科会、総合調整用語集 [A-24] . 昭和48年3月, 1,469項目
- 10) 改定の経緯は下記のとおり。
 昭和33年5月 図書館学編刊行
 平成2-4年 日本図書館学会において「図書館情報学用語標準化の調査研究」(科研費補助金交付)
 “ 5年3月 日本図書館学会において研究成果とりまとめ
 “ 6年7月 学術審議会用語分科会運営委員会開催

- 図書館情報学用語専門委員会の設置を決定
 “ “ 12月 図書館情報学用語専門委員会(第1回)開催
 “ “ 7年2月 図書館情報学用語専門委員会(第2回)開催
 “ “ 4月 図書館情報学用語専門委員会(第3回)開催
 “ “ 6月 学術審議会用語分科会運営委員会開催
 学術用語(図書館情報学)の改定案を了承

しは調整ということで出てくるのにさらに2年くらいかかるのでかなわないという意見はよく聞きますね。さっきおっしゃった調整委員会で何千もの項目をチェックして、というような作業もあったからということですか？

青戸 「総合調整用語集」(昭和48年3月)の総合調整の場合は、先にできたところが足踏みしてたんです。この時は、文部省の行政機構改革の巻き添えで、総合調整の態勢が整うまで待ってもらった。growthに対応する用語は、「生長」(植物学編初版)と「成長」(動物学編初版)とがある。これには、「植物学では、生長の問題は非常に重要で、多くの研究があって、さまざまなことが知られているが、動物学では、成長のことを扱うのは少ない。植物学では、時々刻々の生長を研究する方向に進み、いわば微分的である。動物・人体の場合は、成体にまで到達すれば“成長”はやんでしまうので、考え方や研究方法が、いわば積分的であるとも考えられる」(植物学用語専門部会主査)という見解が、学術用語の審査の際に了承された経緯がある。その後、動物、植物、遺伝、農学の生物系の調整委員会で再調整した結果、増訂版では「成長」に表記を統一しています。

牧野 最近刊行された増訂版の動物・植物・遺伝の時に足並みをそろえるために、動物としては決まっても、遺伝や何かが決まるまでは少し待った、ということがあったわけですか？

青戸 そうでした。『動物学編増訂版』(昭和63年)、『植物学編増訂版』(平成2年)、『遺伝学編増訂版』(平成5年)の3編では、90%程度の重複語があり、極力調整を図った。しかし、「コードン」(遺伝学編初版・動物学編増訂版・植物学増訂版)を『遺伝学編増訂版』では関係学会で優勢に使われている「コドン」に改変した。遺伝暗号ということで、ほかの分野からの文句は出ていない。学術用語の他部門との調整には結構ヒマがかかるんです。

牧野 審査のためではなくて調整のためですか？

青戸 科研費措置で作成された『学術用語集』原案は、学術審議会学術用語分科会に付議されて、専門委員会を2-3回開催して、調整の上で審査案を作成し、運営委員会を1回開催してこれを審査する。審査を終えた『学術用語集』の制定案が学術審議会総会に報告されて最終決定となる。これには通例半年から9カ月を要している。

香川 遺伝学は増訂版発行までに7年です。ずいぶん時間がかかりましたね。

青戸 昭和61-62年度に科研費措置として、平成元年3月に作成された「研究成果報告書」は、平成2年1月に学術審議会学術用語分科会の審査を終えました。『遺伝学編増訂版』の出版原稿の作成に際して、いわゆる「差別用語」の再検討を求められた日本遺伝学会では、平成3年8月-9月に該当用語の修正・削除を行って、『増訂版』発行は、平成5年8月になりました。

仲本 後発の分野では、借用語が多いんです。ISO1087のJIS化では、言語学から借用しようと思っているんですが、言語学がまとまらないので、決めかねている。

青戸 『言語学』の場合は、最初に科研費措置をした『新言語学』を土台に成案を得ており、『学術用語集言語学編』の共著者をめぐって当事者間の折り合いがついたので、近く発行されることになっています。

太田 医学会でやってる時に他の学会とのすりあわせ、調整っていうのはもうすませてあれば、文部省に原案が出された時に文部省の委員会でもめごとは少なくなりますけども、その辺のことは？

ISOでもこのごろ何回も何段階も分けて議論するの大変だから合同でやっちゃおうということが行われている。文部省の場合はそういうことは行われていないんですか。

青戸 事実上行われてますね。科研費措置では、合同っていう形式はとれないが、医学、薬学は審

査の段階で一緒にびったりあうようにしたいと思っております。

医歯薬の3分野の科研費措置は、歯学、薬学、医学の順に行われてきました。日本医学会では、さきに『医学用語辞典』を1975年に完成しました。この改訂作業は、最初伊藤隆太東邦大学教授を委員長とする医学用語管理委員会で進めてきましたが、その後、草間悟東大教授を委員長とする新しい委員会に引き継がれ、1991年に『医学用語辞典(英和)』と1994年にその和英編が刊行されました。

『学術用語歯学編増訂版』(平成6年)に収載する医学用語については、科研費による調査研究の段階で、伊藤先生にご協力いただいて表記の整合に当たりました。

薬学についても、科研費による調査研究の段階で、『医学用語辞典(英和)』と『学術用語集歯学編増訂版』の表記を尊重してきました。

医学については、医学会で『学術用語集医学編』の用語選択、記載方法などの基本方針を取りまとめて、平成8年3月に科研費による研究成果報告書を作る運びとなりました。薬学用語の制定は、これとの関連もあり、1年間足踏みさせられてきましたが、このほど学術審議会学術用語分科会に薬学用語専門委員会の設置が決まり、『学術用語集薬学編』の成案の取りまとめが行われる運びとなりました。

私は、目下、「薬学用語標準化の調査研究成果報告書」の問題点を調査中で、医歯薬関連の調整用語について資料を作成しており、薬学用語専門委員会で審議してもらうことになっています。

香川 現在は学際化してきていて、以前は、医学と歯学と大きく分かれていたわけです。内容もずいぶん違っていたし、今は近づいています。そうなりますと、今までの作り方が問題です。用語に対応する学会から委員を選びます。たとえば、動物学編は動物学会に頼むわけです。動物学会が全部作ってから、今度は医学会が検討するのは時間がかかって困ります。生体工学辞典と生化学の時初めて試みましたが、違う学会、違うバックグラウンドの人を、最初からどこかに集める手段が

なかったら早くまとまるものもまとまらない。各学会が完成してから持ってきて調整するのは全部変えなきゃならない。とくに言葉のもとにある語基にふれるものに改正があったら大改訂になる。だから私は、用語集を作る委員会は、学会推薦で一つの学会でやるのではなくて、学際的な見識を持った人が何人か集まって文部省に頼んで、その委員会を医学、薬学、合同のものを先に作り、あるいは医学、薬学、歯学、医療系が合同の委員会をまず作る必要があります。今のように学会が全部作り上げたのを持ってきてそのあと調整しようとしたら、これは大変ですよ。

太田 そりゃあもう始めからですね。科研費を出す時に、科研費は特定の学会に出すんだが、そこでは学際的にやってくださいよと条件はつけてるんですか？

青戸 もちろんつけてますよ。基本的には、「学術用語審査基準」に準拠することになります。『計測工学編』の増訂版では、thermocoupleに対応する用語の「熱電対」は定着したが、thermopileに対応する用語の「熱電対列」はあまり使われないので、この際に外来語の「サーモパイル」に改めたいという強い要望がありました。「熱電対列」は総合調整で決められた経緯があるが、その用語が本来所属する専門分野からの改訂の発議でもあり、関連分野の電気工学、物理学からの反対もなく、すんなり受け入れられました。決められた用語を崩そうというのはむずかしいですね。

仲本 3年立ったから直せばいいという発想にはならないんですか。

青戸 なかなかならないんですね。既刊が28編あって全部がなじがらぬに揃っちゃってる。いまさら「熱電対列」を使わないと言われても、じゃあどうか、というようには。計測あたりからの発議がないとできない。

仲本 基本的な用語はあまり変えたくないんだけど、新しい用語はどんどん変えたほうが、新し

い学問の進歩にフォローアップできると思うんです。今のようにこうぎしり3、4年かけてっていうやり方だと、追隨していくのが大変で、そこでやっぱり用語集の役割ってというのが一つ考え直してもいいのではないんですか。教科書に使う教育関係の用語は基本的にあまり動かさなくてもいいけど、その専門用語っていうのはかなり柔軟性を持っていないと対応できないですね。

太田 JISは3年にいっぺんは見直して確認されるということに対して、『学術用語集』の場合はそういう条件はついてないとすると、変えようって意志はどこから？

青戸 学会から申し出があれば、『化学編(増訂2版)』の場合には、日本化学会がこのままでは放置できないとして、科研費を当てにせず積極的に増訂案を完成して、文部省に持ち込んで刊行されています。

牧野 『化学編』は大変でしたね。初版と増訂版が非常に違うのです。初版は、カタカナの採用などによって、常用漢字以外の難しい字を使わないことを、いちばん忠実にやった、他のあらゆる分野に比べて。ところが元素や化合物もかな書きが多いから、全部かなだらけになり、とても具合が悪いということで、かな書きルールなどを変えて他の用語集よりも割合早く増訂版を刊行した。さらに、新しい言葉が足りないからということで増訂2版を用意したのは、ちょうど常用漢字を使えるようになったときにトップのほうだったので、他の学会はずいぶんたくさん表外漢字を使っているようだけど、化学会は、我慢に我慢をして3つか4つしか提案しなかった。化学は変わらざるをえない内在的な理由があったので早かった。文部省も割合早くそれを認めた。もうまた変えなければならぬ時期にあると思います。増訂2版は、現行のものの中ではいちばん古いのです。

青戸 文部省から関係学会に『学術用語集』の増訂について意向を打診しても、学会が乗り気でないのに無理にやっってくださいとは言えません。学

会によっては、国際学会があって手いっぱい待ってくださという風な学会の内部事情がある。

太田 手続き上のことで、文部省側の審査のために発行まで3、4年かかるということはない。

青戸 科研費措置が終わっても、学会の見直しに1年、文部省での審査に半年、さらに出版までに1年近くかかっている。この間に文部省の担当官の異動、審議会委員の任命(再任)にともなう事務上の遅れも多少はある。

科研費の規準について

山下 学術用語制定の精神を考えますと、学術分野全体を見渡せるように、広範囲に影響を及ぼすような分野を重点的にやっていく必要があると思うんですけど。前に聞いた話では、文部省から提案することもあるけど、その学会からの申請に従って学術用語制定の事業が行われるという例もあるという感じで、あまり一貫していない感じがするんですが、どういう基準、どういう感じで考えているんでしょう？広く浅くばらまくのか、それとも重要な分野だけ重点的に改訂を繰り返していく？

青戸 重点的というか、古くなった、たとえば当用漢字になってるままの、専門分野は採り上げるべく改訂していく。航空工学編を取りあげることにりましたが、航空宇宙も取り込んでいる。航空工学も飛行機だけでなく、ロケット関係も取り込んで、宇宙航空学会の学会の名前をとって『宇宙航空学編』を作ろうってことになります。

香川 ご存知でしょうが、文部省はむしろ学者の自発性を非常に重視するんですね。ですから、さきほどから、研究費がついた、つかないとか話がありましたけど、原則どころか、決まりとして、誰かが申請書を書かない限りつきません。ところが、私は通産省の用語のお世話もしたから知ってますけど、通産は逆なんですね。いろいろな研究のプロジェクトにしても、コンピュータのJISの第一水準にしても、省庁の中に、学識経験者から

多少意見を聞いたら、そしたら積極性があるんですね。あるプロジェクトへ。用語は通産と文部と両方あるわけだけど、通産の方がいろいろな意味で早いつてことがありますね。どうしてもその輸出用のマニュアルを作らなければとか、ネジの規格をちゃんと決めなければ製品が作れないとかつていうことがありますから。通産の方は、今のご質問に答えるためにもう一度はっきり言いますと、文部省は申請者の自発性を待って、申請書が学会なり、研究者のグループから提案されない限り科研費の交付はありえません。だからたとえば、この分野が大事だからやろうじゃないかと、そういう発想を持ってくるのはむしろ通産の方であつて、文部省は違うんですね。

山下 やろうって盛り上がってくる位の勢いがないと、重要じゃないと。

香川 特に医学関係は医師会の武見太郎会長がいて、省庁と喧嘩していましたから、医学会では、用語は統一しなきゃいかんというのは前から言われてきたにもかかわらず、故意に、文部省を避けていた面もある。そういうところに、文部省が金を出して、医学用語を作ろうとしてもできなかったのです。7部に歯学はありましたが、中核の医学さえできればその周辺の歯学や薬学もできるわけです。本来なら文部省や通産省でも必要があればお金を出すわけですよ。それをやらなかったのは申請がないからです。

青戸 そういうことはあります。医学については、平成5～7年度に科研費措置がとられています。医学以外の複数の学会で使われているために、標準化しなきゃいけないものだけに限定して標準化しようという考えかたです。タンパク質をどう書くか。タンパクはいろいろ書き方があるので、どう決まるかわかりません。

牧野 タンパクはみものですね、どう決まるか。

太田 タンパク質の話が出たんですが、タンパク質の場合、常用漢字に入ってませんよね。ところ

がJICSTが先行して作っちゃつた。JICSTシソーラスでいえば88年には「蛋」はもう漢字を使っちゃつてるんですね。そういうふうに先に実績作つてると、用語集の方が後追いになる。つまり、そこでボンボン使われてると、用語集を作るから、いまさらこれひらがなに、カタカナにしようなんてできなくなつちゃうことになりますね。そんなことはありませんか？

香川 たとえば、大阪大学の「蛋白質研究所」という研究所、文部省の機関ですが、名称は漢字です。

太田 あ、蛋白研は漢字ですね。

香川 厚生省はひらがな。学術用語はカタカナ。それでおそらく、タンパク質がどう決まるかですが、厚生省が定めた医師国家試験の受験のための「出題基準」は、漢字の「蛋白質」。私どもが決めた時は、中国と相談しまして、中国は蛋白質は「質」が略字体ですが、「蛋白質」なんですね。だから、「蛋白質」は漢字になる可能性あります。そんなに字画の多い漢字ではないし、いろんな言葉の元になっているんです。たとえば、「リポ蛋白質」の「リポ」はカタカナなんですが、後を全部ずーっとカタカナ書きにすると非常に読みにくくなつたり、全体が長くなる。全部カタカナでその前に外来語の接頭語がいっぱいありますから、そうすると、たいへん使用頻度の多いところにおいては、蛋白質は漢字にしておかないと使いにくくしょうがないという圧力が高まるでしょうね。

太田 JASは全部ひらがなですね。JASは農水産で全部たんぱくと平仮名で、タンパクが出たついでに「デンプン」があるんですね。「デンプン」も「タンパク」が漢字になったと同時にJICSTのシソーラスは、さんずいをつけた「澱」を生かしちゃつた。ところが「沈殿」の「殿」がもとのままなんです。私も文句言ったんですよ。片方を直して片方がそのままなのはどういうことかと。ああそうかと、うっかりしてたなんてね。「沈殿」の「殿」はさんずいなし。こういうちぐはぐなことをやっている。

牧野 化学用語では自らの表記ルールに従って、増訂2版でも「デンブン」です。香川先生、「蛋白質」の「蛋」の字は第2水準漢字に入ってるんですか？

香川 入ってますよ。

太田 だから使えるんですよ。

後継者の問題

太田 いちばん気になるのが、今までの話を聞いてみると、青戸さんに全部おっかぶしちゃってる。青戸さんが引退したら誰がチェックするかというのが最大の関心事。どうなのでしょうねえ。

青戸 やはりこの次は若い人で、電算機の処理が得意な人であり、学術情報センターで構築中の『学術用語集』のデータベース化による高度な利用を実現して、学術用語制定支援システムを動かすために、そういうことに明るく、熱意を持っている適当な方がいらしたら推薦していただけるような形になっていくしかないでしょう。

仲本 専従職がいないとダメですね。これからはネットワークが使えるから、青戸さんの苦労はある程度軽減されるけど、やっぱり一人専従職が必要ですね。

青戸 そういう認識は持ってます。具体的にはどういう人がいいかということになりましょう。

仲本 専門用語研究会から意見を聞いたら。

青戸 そういう意味で後継者を名乗り出る人がいるのが望ましい。

太田 別の機会にそういうこともぜひ、学術用語の将来と仕組みとして、やってもらいたいですね。

香川 いろんな学術用語に共通することですから、

国立国語研究所とかそういうところで、今まで文部省の学術用語の研究がありました。私いつも考えるんですが、定員非補充の原則というのは今後かなり厳しいと思います。だから新しいポストを作るならどこかを減らさなきゃいけない。インターネットが全部の大学にできましたね。あのために定員増は一つもないんです。いろいろと便利になりましたから、逆に定員少なくて済むようになった。国立国語研究所も今後はインターネットに組み込まれていくだろうし、かえって慣れた方がいれば、人手も少なくて済むと思うんです。だから、言語学というか、用語学というか、そういう専門家の中から選んだほうがいいと思いますね。一つのサイエンスのスペシャリティを持った方より、ジェネラリストで、しかも言語学に通じているような方がいいですね。今後はコンピュータを使わないといけな。だからそのためにあと数年いろんな試行があるでしょう。

仲本 国立国語研究所は研究機関であり、実務はなかなかやりたがらない。学術情報センターはサービスも担当しているから、ああいうところに置いたらどうですか。今、人の確保が難しいけども。今日の青戸さんの話を踏まえて、将来いかに考えるかという話が生まれてくると、今日の会合が非常に有意義になるんじゃないですか。そういう人を置くべきだというような提案はあっていいと思うんです。

青戸 専門用語研究会の提言としてちゃんと然るべきところに言ってくれと非常に良いと思います。最後には人の問題になるでしょう。文部省の建物の中にいる人は行政官であって、制度化して予算をつけることはできても、私のしてきた仕事の後継者になることは難しい。『学術用語集』のデータベース化で今までのような手作業はなくても済むことは確かです。しかし、あまりにも新しいことをやってガラガラ変えることには躊躇せざるを得ません。その辺のところで、専門用語研究会の良いご提言をいただけると良いと思われま

す。(つづく)

～FID 東京大会から～

カテゴリー 2 (Information Science of Classification and Terminology)
を中心に発表論文の紹介

1994年10月2日(日)～9日(日), 大宮ソニックシティにて第47回国際情報ドキュメンテーション連盟(International Federation for Information and Documentation, FID)国際会議が開催された。内外の約40ヶ国300人以上の参加者を集めて, 6件の特別講演, 26件の招待講演, 12カテゴリー140件の発表およびさまざまな付設の部会, セミナー, ミーティングがもたれた。ここでは, 配布されたProceedingの中から, 当会に関係したカテゴリー2 (Information Science of Classification and Terminology)を中心として, 発表者から寄稿のあった日本語による7件(フルペーパー2件, 抄録5件)を次ページ以降で紹介する。なお, その他10件については, タイトル・発表者・所属のみ紹介する。

〈フルペーパー〉

- ・EDR 電子化辞書と情報科学(横井 俊夫)
- ・用語データベース設計の基本方針(佐々木由香)

〈抄録〉

- ・安全研究における専門用語データベースの構築: 複数データベースからの安全および危機管理問題のデータの効率的なアクセスを求めて(後藤智範ほか)
- ・TERM IS - 地球規模の用語情報ネットワークのためのデータベースデザイン(佐々木由香)
- ・オンライン情報検索システムに用いられる操作用語の標準化の一つの提案(大前 巖)
- ・用語間の意味関係の抽出法--SS-KWIC(藤原 譲ほか)
- ・自動索引の制限(R.Fugman)

〈その他〉

- ・AZ Method of Characterizing Document Clusters. M.Nakashima, L.J.Yeh, T.Ito (Japan)
- ・Indexing Terms for Computer Documentation:a User-oriented Approach. A. Kukulska-Hulme(U.K)
- ・Effective Multilingual Subject Searching Through Translex. A.Dewar, E.Katsura (Japan)
- ・APTS Code:an Identity Code Applicable to Various Types of Documents. I.Asai (Japan)
- ・Standardization of Chinese Sciencetech Terminology and Information Science. G.Zhang (China)
- ・A New Thesaurus Structure for Semantic Retrieval. G.Rahmstorf (Germany)
- ・EDR Concept Classification and Methodology for Its Development. Y.Nakao, T.Ogino (Japan)
- ・An Approach to Extraction of Semantic Relations from an Ordinary Japanese Language Dictionary. H. Tsurumaru, S.Yoshida (Japan)
- ・An Approach to Application of Systematics to Knowledge Organization. E.Curras (Spain)
- ・Management System for Multilingual Terminology Applications as Vocabularies and Terminology Databases in all European, Russian, Chinese and Japanese Languages. P.Stancikova, M.Smihla (Slovakia)

~FID 東京大会から~

EDR 電子化辞書と情報科学

横井 俊夫*¹ YOKOI, Toshio

言語や言語処理技術は、情報科学の研究を進めるための重要な土台となる。言語処理技術を本格的な安定した技術とするためには大規模な言語データ、すなわち、電子化辞書が必要である。また、電子化辞書の研究自身も言語に関する情報を研究する分野として、情報科学にいろいろな話題を提供する。知識処理という研究分野も情報科学と大きく重なる。知識処理技術を本格技術とするためのテーマである大規模知識ベース構築技術への有望なアプローチとして言語処理と言語知識からのアプローチがある。このような観点から、EDR 電子化辞書と大規模知識ベースを紹介する。

1. 情報科学と言語と知識

EDR 電子化辞書プロジェクトは、コンピュータに自然言語を処理・理解させるための大規模な言語データベース（言語知識ベース）の研究開発を目的とする。このプロジェクトの成果およびプロジェクトの実行中に得られた多くの経験は、今後の情報科学の進展に様々な貢献をするものと思われる。本稿は、情報科学への貢献ということ念頭に置いた上での EDR 電子化辞書とそれに関係した事項の説明である。

情報科学は、情報の仕組み、すなわち、情報の構造や生態を研究する学問である。情報の仕組みの研究は、情報を表現するメディアの仕組みの研究が非常に重要である。メディアの中で、特に、情報の意味を扱う観点から、(自然) 言語が重要となる。したがって、言語の研究、言語処理の研究や技術の開発と情報科学研究とを結びつけることが、今後の情報科学にとって重要な課題となる。

言語や言語処理の研究にとって、最近、多様な言語現象を捕捉するための言語データの重要性が広く認められるようになった²⁾。辞書やコーパスやテキストデータを体系的に大量に集積し、それを利用して言語処理技術を本格的な安定した技術へと成長させようというのである。電子化辞書は高度な内部構造を持つ言語データである。また、電子化辞書は、言語に関する情報の仕組みを研究

する分野であることから、それ自身としても情報科学に新しい話題を提供するものである。

さらに、情報科学にとって情報を知識という観点から取り扱うという分野は、今後、魅力的な分野となる。知識や知識処理の研究にとっても、本格的な技術へと離陸させるために大規模知識ベースの研究開発が重要になってきている。そして、大規模知識ベースの構築技術への有望なアプローチとして、言語処理と言語知識からのアプローチがある。

以上のような観点に基づいて、以下 2 章で EDR 電子化辞書³⁾、3 章で電子化辞書と大規模知識ベースとの結びつきを説明する。

2. 電子化辞書

コンピュータのための辞書

人は不確かな言葉や見知らぬ言葉に出会うと辞書を引く。辞書には、ごく簡単な語義文や用法を示す一二の例文によって意味が説明されている。これらは、意味を理解するきっかけを与えるにすぎない。人間はそれらをきっかけに、膨大な常識を検索し、複雑な推論を行い、納まりの良い言葉の意味を確定していく。人間のための辞書とは、このような人間の能力を前提にして作られているのである。一方、コンピュータは、そのような能力など一切持ち合わせていない。そこで、コンピュータにそのような能力の一部なりとも、まずは与え

* 1 電子技術総合研究所

てやらねばならぬことになる。電子化辞書とは、その役割をはたすコンピュータのための辞書である。もちろん、人間の常識やそれを扱う能力はまだ未解明であるから、コンピュータに与えるのは言葉の周辺にかかわるごく限られた常識に関する能力である。

電子化辞書は、コンピュータに言葉の処理や理解を行わせるコンピュータのための辞書である。とはいえ、人間のための辞書やそのためになされた蓄積が、研究開発の大きな拠り所となることにはかわりない。

プロジェクトの成り立ち

機械翻訳システムの開発費の9割近くを辞書が占めるといわれる。次世代の高度な(自然)言語処理技術の研究開発のためには、大規模で高機能な辞書が不可欠である。基盤技術研究促進センターとコンピュータメーカー8社(富士通、日電、日立、シャープ、東芝、沖、三菱、松下)の共同出資のもとに、1986年に9年間のプロジェクトとしてスタートした。日本電子化辞書研究所(EDR)が設立され、1995年3月に終了した。

目標としたのは、図1に示すような辞書群、辞書の開発や利用を支援するシステム、辞書の実証や評価のための応用システムなどの研究開発である。単語辞書は、単語と概念(語義に相当)との対応とその時の文法特性を与える。コンピュータに形態素処理や構文処理をさせるために用いられる。概念体系辞書は、上位-下位関係(AKO:A-Kind-Of)で全部の概念を体系付けたもので、通常、シソーラスと呼ばれているものに近い辞書である。コンピュータに、同義や類義の概念を見つけ出させたり、概念間の類似度を計算させたりするために用いられる。概念記述辞書は、深層格的なかわり方に関する概念間の意味的な共起の仕方を記述したものである。コンピュータに意味的に正しいか否かを判断させるためのデータとして用いられる。共起辞書は、単語間の表層上の共起の仕方を記述したもので、コンピュータに自然な言い回しの仕方を理解させるものである。対訳辞書は、同じ意味を持つ日本語単語と英語単語との対応を記述したもので、コンピュータに適切な

訳語を見付け出させるために用いられる。EDRコーパスは、文章データから抽出した例文に、形態素、構文、意味、それぞれの解析データを付加した辞書開発用の素材データである。このデータは、言語処理の研究用データとしてさまざまに利用できる。さらに素材データとしては、文章データを大量に集めたEDRテキストベースがある。これは、新聞社や出版社から購入した機械可読テキストデータを整理し、データベース化したものである。

成果の普及

辞書の販売は、1995年4月から開始されている。価格や条件についての基本方針は、次の3つである。

- (1) 小規模な企業を含め、広くさまざまな企業が利用できるようにする。
- (2) 大学や研究機関などの非営利な機関には特別の配慮をする。
- (3) 国内、海外の差は無しとする。

現在、情報・通信の分野では最大の課題として、国家情報基盤の整備の議論が各所で始まりつつある。まずは、土台となるネットワーク整備から話が進められているが、次には、ネットワーク上の情報環境の整備の議論がある。情報環境の骨格となるのが言語、言語処理技術、そして、それらを支える言語データである。本プロジェクトの発端は、機械翻訳システムの開発への熱意からのものであったが、現在は、より広い範囲の技術をより長期的に研究開発していくための共通基盤となる言語データとの位置付けとなっている。このような位置付けのもとに、これからの改良・拡張への対応や国内外の類似の活動との協力がなされることになる。

位置付けと今後

言語処理技術が産業技術へと成長を始めるきっかけを掴んだこと、言語研究が人間の知能解明への確実な手掛かりを与え、着実な蓄積を行ってきたこと、これらは、今、広く認められるところとなりつつある。マルチメディアブームの中に言語を軽視するという片寄った傾向があることや、機

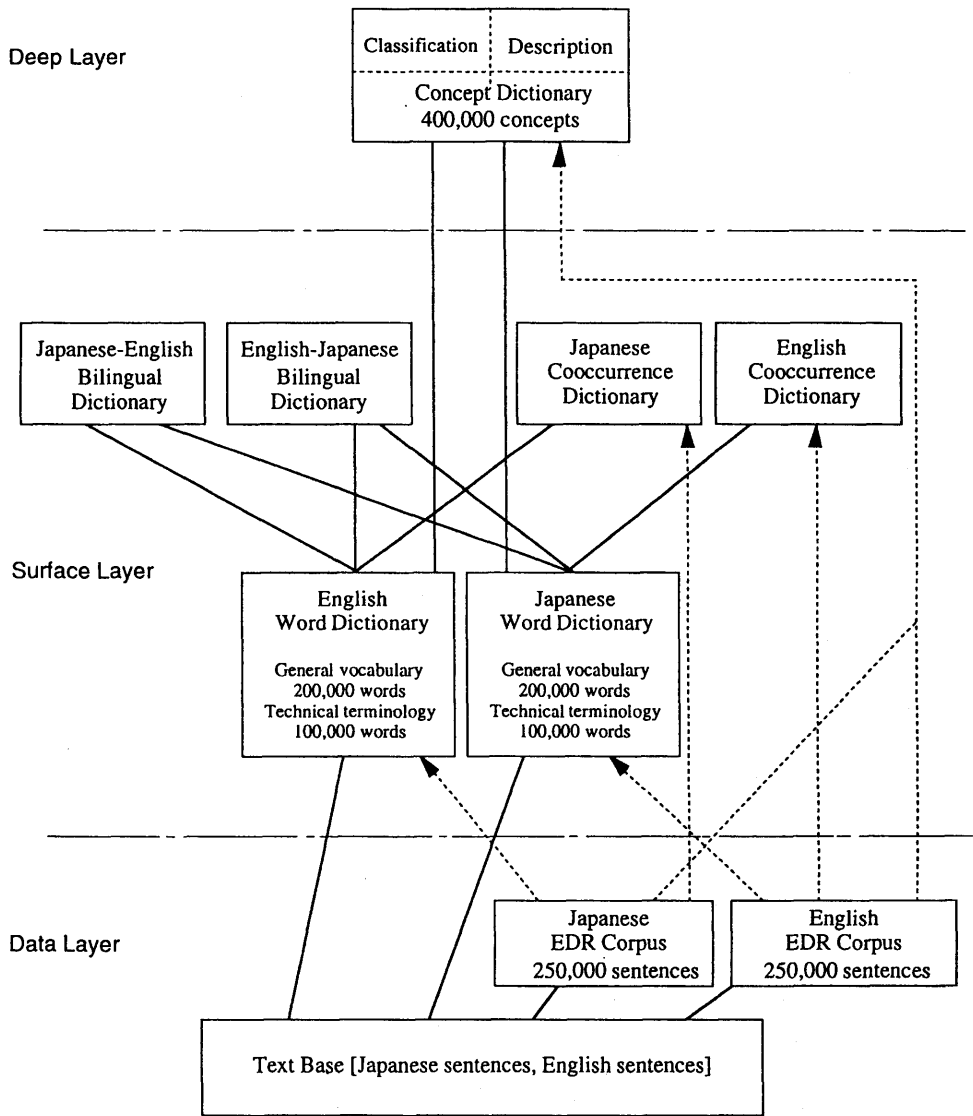


図1. EDR 電子化辞書の構成

械翻訳への過剰な期待に対する過剰な反省があることなどから、少々の曲折はやむをえないであろうが、この共通認識はこれからより確実なものになっていくであろう。そして、これからの言語処理にとって言語データの整備が本質に重要になる。これは、通常いわれる研究開発にとっては客観的なデータが必要であるというような単純な意味からではない。体系的に構造化された言語データが包括的に言語現象をとらえ込み、それ自身が言語処理技術やシステムの仕組に直接結び付くものになるからである。

しかしながら、そのような意味で必要とされる総合的な言語データにとって、EDR 電子化辞書はその第一歩となる部分を仕上げたところである。これから本格的に必要とされる大規模な言語データの研究・開発は、広く関係する方々の同意と協力のもとに、文字通りの情報基盤の整備として取り組まれねばならない。

3. 言語知識から世界知識へ

広く共有化される大規模な知識ベースの構築が、これからの技術開発の重要なテーマになる⁹⁾。いろいろなアプローチが考えられるが、中でも自然言語、電子化辞書からのアプローチが安定した技術を生み出す。電子化辞書の役割を一層明確にするため、このアプローチの概略を以下に説明する。

言語知識の大規模化

電子化辞書は、言葉の使い方に関する常識を集積した大規模な知識ベースである。用いられている知識表現言語は、自然言語と簡単な人工的な言語を組合せたものである。いずれにしても自然言語による表現が多用される。たとえば、言葉の意味の説明も多量の用例データも自然言語の文、文章である。これらは、電子化辞書の改良・拡張(大規模化)のための重要な情報源である。

改良・拡張は、人間と言語知識の獲得機能を持つコンピュータプログラムの協同作業で行われる。電子化辞書をマスター辞書として、さまざまな自然言語処理プログラム用のコンピュータ辞書が生成される。電子化辞書が改良・拡張されるにした

がって、これらの自然言語処理プログラムの精度、性能、ロバスト性がより改善されていく。言語知識の獲得機能は、自然言語処理プログラムによって実現される。したがって、電子化辞書の改良・拡張にともなって、言語知識の獲得機能の能力がより向上していく。

以上のような仕組みによって、言語知識の大規模知識ベースが実現されていくことになる。

世界知識の大規模化

次に取組むのは、世界知識の大規模な知識ベースの実現である。世界知識はそれぞれの分野ごとの知識といろいろな分野に共通の知識(一般常識)に分かれる。用いられる知識表現言語は、自然言語と人工的な言語を組合せたものとする。人工的な言語は、電子化辞書に用いられるものよりより複雑なものであるが、再利用を十分に考慮したものである。自然言語を用いるのは、人間の持つすべての知識をコンピュータに理解可能にする手段は、まだ発見されてはいないからである。まず知識はヒューマンメディア(画像、図形、音声、自然言語等)によって有用な表現がなされると考えるべきである。そして、ヒューマンメディアの中で最も汎用のメディアが自然言語である⁹⁾。自然言語で表現された知識には、教科書や百科事典や用語集などから直接採取してきたものや、新しくこの知識ベースのための作業者が記述したものなどがある。これらは、この知識ベースの改良・拡張(コンピュータ理解可能な部分の規模の拡大)のための重要な知識源(情報源)である。

改良・拡張は、人間と世界知識の獲得機能を持つコンピュータプログラムの協同作業で行われる。自然言語で表現された知識をさまざまに処理することによって、いろいろの形態の知識が抽出される。世界知識の獲得機能は、自然言語処理機能に大きく依存することになる。したがって、言語知識の獲得機能の向上が世界知識の獲得機能の向上に結びつくことになる。

この世界知識に関する知識ベースはそのまま利用されることもあるが、これをマスター知識ベースとしてさまざまな応用向けの知識ベースが生成されることになる。

言語知識と世界知識の関係

言語知識には、自然言語のメディアとしての仕組みに関する知識と自然言語が表現しようとする対象世界に係る知識が含まれる。メディアの仕組みにかかわる知識は、形態論的、構文論的部分にかかわるもので、規則性に富む知識である。対象世界に係る知識は意味論的、語用論的部分がかかわる。規則性や整合性も部分的、局所的である。この部分の多くは言語学でいう言語外知識に対応する。世界知識もこの言語外知識に含まれる。すなわち、言語知識の対象世界に係る知識は世界知識と連続的につながり、最も常識的な部分に対応することになる。

参考文献

- 1) Boguraev, B. and Brisco, T.(Eds.) "Computational Lexicography for Natural Language Processing" Longman, 1989.
- 2) Pustejovsky, J. and Bergler, S.(Eds.) "Lexical Semantic and Knowledge Representation" Lecture Notes in Artificial Intelligence 627, Springer-Verlag, 1992.
- 3) EDR 電子化辞書仕様説明書(第2版), TR-045, 日本電子化辞書研究所, 1995. [http://www.ijnet.or.jp/edr で ftp 可能]
- 4) Fuchi, K. and Yokoi, T.(Eds.) "Knowledge Building and Knowledge Sharing" Ohmsha and IOS Press, 1995.
- 5) Jacobs, P. S.(Ed.) "Text-Based Intelligent Systems : Current Research and Practice in Information Extraction and Retrieval" Lawrence Erlbaum Associates, 1992.

～FID 東京大会から～

用語データベース設計の基本方針

佐々木由香*¹ SASAKI, Yuka

ゲオルク J. アンカー*² ANKER, Georg J.

従来の専門用語データベースや用語管理プログラムの多くは、「概念中心」アプローチによって設計されている。本稿では、これらのシステムが下敷きになっているデータモデルとその欠陥を明らかにした後、用語とは何かを現代言語学、文化理論、情報理論の文脈から捉え直して、用語データベース設計の基本方針を、オープン・アーキテクチャ、情報中心アプローチ、簡略化された現実、複数の現実の共存、言語中心アプローチ、フラクタル・アプローチの6つのキーワードに沿って概括する。

キーワード：用語データベース、データモデル、データベース設計、概念中心の用語整理作業、記述的用語整理作業、用語情報システム

Most of today's terminology databases or programs for the computerized management of terminological data have been designed following the "concept oriented" approach. In this paper we firstly review the data model common to such systems and highlight the problems which may occur due to the deficiency of the underlying model of these systems. After discussing various real world examples, we outline some design principles of a terminology information system, whereby terminology work is understood in the context of linguistics, theory of culture and information theory. The most important principles are: open structure, information approach, reduced and coexisting realities approach, language based approach and fractal approach.

Key words: terminology database, data model, database design, concept oriented terminology work, descriptive terminology work, terminology information system

1. はじめに

本稿のねらいは、多言語用語データベースを設計する際に基盤となる視点を提示することにある。閉じた体系としての用語集や用語辞典と異なり¹⁾、体系の境界が定まらず、用語の意味、用法、体系内での位置づけが時とともに変化してゆく開いた用語体系を扱うデータベース（以下DB）では、（潜在的に）様々な言語の、多様な文化背景を持った、複数の領域とパラダイムの用語が一つのDBの中で扱われることになる²⁾。今日のような広域ネットワークによる国際的な情報流通の時代においては、ターミノロジー分野でも国際的な共同作業の前提としてDB間でのオンラインのデータ交換が求められるが、そのためにも個々のDBが開い

た用語体系を扱えることが必要である。

次章で一般的なDB設計の流れと用語を説明した後、3章と4章で従来の用語DBのデータモデルとその問題点を論じる。5章では現代言語学、文化理論、情報理論の視点から、用語の様態とそれがDBの構成にとって持つ意味を考察し、6章で用語情報システム設計の基本方針をまとめる。別の論文³⁾で、本稿で提示した視点に立ったデータモデルと多言語用語DBのデザインを紹介し、広域ネットワークを利用した国際的な共同作業のための分散型システムの構築に触れているので、興味のある方はご一読願いたい。

2. データベース設計とデータモデル

2.1 対象世界・エンティティ・関連

DB設計は、DB化しようとする対象の世界がどのような実体（エンティティ）で成り立ち、各

* 1 名古屋大学大学院国際開発研究所

* 2 インズブルック大学情報処理センター

エンティティがどのような属性を持ち、エンティティ間にどのような関連があるかを把握することから始まる。例えば企業の販売管理 DB では、製品、顧客、注文などをエンティティとし、住所や電話番号等を「顧客」の属性と考える。こうして DB のデータ項目となる雑多な項目（製品名、製品型番、単価、顧客名、顧客住所、注文日、納期、注文合計など）をグループ化し、各グループがレコードを識別するキー項目とそれに従属する項目のまとまりになるように調整してゆく（データの正規化）。各エンティティ間には 1 対 1 とか 1 対 M 対応など様々なタイプの関連がみとめられるが、何をエンティティや属性ととらえて、そこにどのような関連を認めるかは、対象となる世界のありようや DB の使用目的、つまり対象となる世界をどのような視点から眺めてどうとらえるかによって大きく変わってくる⁴⁾。

2.2 データモデル

続いて対象世界をどう捉えたかを概念的にモデル化する。その際、エンティティ間をその関連で結んだダイアグラム（実体関連モデル・ER 図）がよく使われる。

DB は様々な業務に利用されるが、用途ごとに利用者の視点をモデル化したものを外部モデル（外部スキーマ）、外部モデルを統合して複数の利用者の視点を総括的に抽象化したものを概念モデル（概念スキーマ）と呼ぶ⁵⁾。ある DB を販売管理と在庫管理業務に使用する場合、前者に關係するデータの関連を示したのが販売管理の外部モデル、後者のが在庫管理の外部モデル、両者を統合して DB 全体の姿を表現したのが概念モデルである。なお、概念モデルを DB 管理システムのタイプ（リレーショナル、ネットワーク、階層型など）にあわせて実働可能な表現にモデル変換したものを論理モデルと呼ぶこともある。概念モデルは特定のハードやソフトの制約を考慮せずに対象世界のエンティティ間の本質的な関係を表現したものである。

このあと物理レベルの設計（内部モデル・内部スキーマ設計）を行い、コンピュータの中でデータをどのような構造として蓄えるかを考慮して

DB を実装してゆくが、本稿ではその土台となる概念モデルの設計方針を論じる。最新の DB 技術を用いても、矛盾した概念モデルの上に整合のとれた DB を作り上げることはできない。概念モデル設計は、DB 設計の核となる部分である。

なお、使用するソフトやハードの制約でデータモデルをうまく実現できない場合は、新しい技術が利用できるようになった際に、妥当なモデルがよりよく実現できるように DB を再編成することになる。しかし実際には、新しい技術の導入時にも以前のデータモデルが無批判に踏襲されがちである。次章で見ると、用語 DB においても、電算化以前の紙のカードによる用語整理の方法論を色濃く反映し、データの正規化等のノウハウもあまり考慮されていないデータモデルが踏襲され続けている。

3. 従来の用語データベースのデータモデル

今日稼働している用語 DB や欧米で使用されている用語管理ソフトの多くは、概念中心アプローチによって作られている。これは、言葉は言い替えができたり同じ言葉が複数の意味を持っていたりと混乱しているので、言葉によって「言わんとしていること」や言葉が「指し示していること」を「概念」と呼び、それを中心に据えようとする考え方である。コトバ以前にコトバと無関係に「概念」が存在するという前提に立つもので、「概念」とは思考や知識の単位であり意味素のようなものであって、その「概念」に時には複数の用語が与えられる - つまり、「概念」と「同義語」というエンティティが 1 対 M 対応する - とされている⁶⁾。

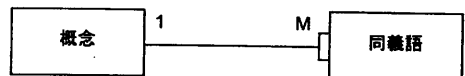


図 1: 概念中心アプローチの ER 図（一言語）

さて、概念中心アプローチをとる人々の多くが主張するように「概念は言語によらず同一」だとすれば、二言語については次の図が成り立つことになる。

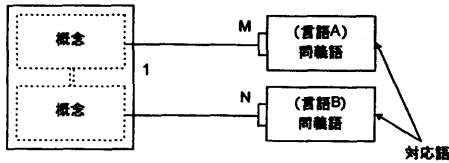


図2：概念中心アプローチのER図（二言語）

この図では、同一言語内の同義語も、異なる言語に属する対応語も一つの「概念」に対応づけられている。

図3は、概念中心アプローチを複数の言語を対象として一般化したER図である。

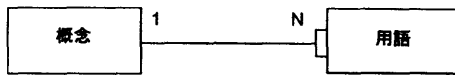


図3：概念中心アプローチのER図（一般）

「語は概念に貼られたラベルないし名前である」⁷⁾という考え方を読みとることができよう⁸⁾。(図4)

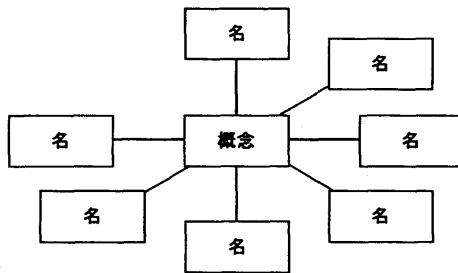


図4：語は概念の名前？

4. 従来モデルの問題点

このモデルに基づく用語DBには、どのような問題があるのだろうか。

4.1 用語間の不適切なコネクションの生成

言語Aのある語と言語Bのある語が対応している、それぞれの同義語は妥当な対応語ないし訳語ではないことが多い。同様に語1（言語A）が

語2（言語B）の訳語で語3（言語C）も語2の訳語でも、語3が語1の妥当な訳語だとは必ずしも言えない。これは各語の「概念」が微妙にずれているとともに、訳語や対応語においては「概念」だけではなく様々なレベルでの文脈や語の用法が問題となるためである。

しかし前章で見たデータモデルでは、すべての同義語や対応語が「概念」を仲立ちとして結びつけられている。このモデルに基づいた用語DBや電子化辞書では、どれだけ細心に各言語間で対応関係を吟味しても、ひとたびDBにデータが収録されると、本来対応していない用語間にも機械的に対応付けがなされてしまう⁹⁾。対象となる言語と専門領域を熟知していれば用語DBや電子化辞書で検索された語彙が不適切だとわかるが、判断のための知識が不足したユーザーは不適切な用語を使ってしまうことになる。機械翻訳システムや翻訳支援システム等のプログラムがDBを利用する場合には、この問題はいっそう深刻である。

4.2 異言語・異パラダイム間の「概念」体系の差異の消失

英語の“water”が日本語では“水”と“湯”となるなど、概念体系は言語ごとに異なる。しかし概念中心アプローチの用語DBでは、各「概念」に対し各言語必ず一つずつ対応語を収録する構造になっているため、データ収録時に次のいずれかを選ばざるを得ない。

- a) 「概念」が細分化されている言語（この場合は日本語）に必要なだけレコード（エントリ）をとり、他方の言語（英語）は、
 - a-i) 複数のレコードに同じ用語を登録する（water(1), water(2), ...）か、
 - a-ii) 例えば cold water と hot water の二レコードとする。
- b) レコード数は「概念」の少ない言語（英語）にあわせて一つとし、他の言語の対応語はいずれか一つ（例えば水）を選ぶ。

英語母語話者なら、a)のように“water”を二つの「概念」としたり、“cold water”と“hot water”という「非論理的」な区分を設けること

には抵抗があろう。また a-i) では、語 water の属性値 (品詞情報など) を二カ所重複して入力・変更しなければならず、一方を忘れることも起こりやすい上¹⁰⁾、英語からの検索の際に water(2) だけを見た場合には、water=湯と誤解してしまう。b) の場合は、湯という用語やその情報を登録も検索もできない。従来のデータモデルに基づいた DB では、このように特定の言語のものさし (バイアス) でデータを構成してゆかざるを得ず、現実に存在している言語間・パラダイム間の概念体系の差異が DB 上は消失してしまう。

4.3 「概念」の電算処理の不可能性

ところで、概念中心アプローチをとった場合でも、DB 上で入力・登録・検索されているのは、実は「概念」ではなく語や語の組み合わせ (コード) である。

情報処理環境で何かをエンティティとして取り扱うためには、それを何らかのコードで、つまり単一ないし複数のコードの要素やその組み合わせで表す必要がある。アルファベットや数字はそれぞれコードであるが、「概念」はその要素や集合が決まっているわけでもなく、それ自身は電算処理可能なコードではない。「概念」を分類して「基本概念」に何桁かの番号をふり、「概念」の集合を明らかにしようとする試みは、「概念」を自然言語から離れた電算処理の対象としようとする試みでもあるが、あらたな言語を対象としたり新しい知見が生まれる度に体系全体の見直しが必要になるわけで、やはり一種の便法であろう。

「概念」はつかみどころがなく、語彙は多義で「概念」を一義的に表せない；概念中心アプローチでは、この問題への解答として「定義」を導入し、「概念」と用語は M 対 N だが「概念」と定義は 1 対 1 対応するとした。しかし、定義とはつまるところ自然言語の語彙からなるテキストである。このアプローチを情報処理環境に持ち込むことは、「概念」を語彙の組み合わせ、正確には、言語に関連するコードの組み合わせによってできる複合体で表して処理しようすることに等しい。言語の多義性を回避することが出発点であった概念中心ア

プローチであるが、「概念」は自然言語から離れて電算処理することはできないというところに行き着く。

5. 用語データモデル設計に先立つ諸考察

コードやコード要素とその関係性しか処理できない情報処理環境では、概念中心モデルは適さない。情報処理環境で把握・処理可能なエンティティを同定し、エンティティ間の関連をできる限り正確かつ柔軟に記述できるモデルが必要である。本章では、記述的な専門用語研究、文化理論、情報理論、現代言語学の知見を踏まえて、我々が扱おうとする対象 (「用語」) とは何なのか考えてゆきたい。

5.1 用語をとりまくさまざまな文脈

自然言語で表された用語の処理には、用語をとりまく様々なレベルの文脈が関連してくる。非常にシンプルな例 (◆) でみてみたい。

従来のアプローチの用語を借りると、読者の頭の中にある「◆」が「概念」ないし「思考の単位」であろう。「◆」は図 5 に示す多様な文脈に依存している。まず、この「思考の単位」らしきものには、多くの言語では - 英語では diamond、日本語ではダイヤないし菱形というように - 最低一つの音声表記と最低一つの文字表記が対応する。(つまり「◆」は各言語ごとに異なる音韻コードや文字コードで表わすことができる¹¹⁾。) 「◆」にはこの他にも (潜在的に) 多くのコードが関与しており、「◆」の導入に用いた絵文字コードの一要素でもあれば、「四角形の集合」コードの一要素でもある。後者のようなコードは、文化、つまり特定の技術ないしパラダイムの使用に強く依存し、「概念」を安定化させるのに大きな役割を果たしている¹²⁾。

抽象概念である「◆」の成立の仕方もこれを取りまく文脈のひとつである。四角形の集合の一要素「◆」は、英語 (diamond) では鉱物の形状に由来し、日本語 (菱形) では植物の葉の形状に由来する。トランプの「◆」は英語では四角形の「◆」と同じ用語であるが、日本語ではダイヤ

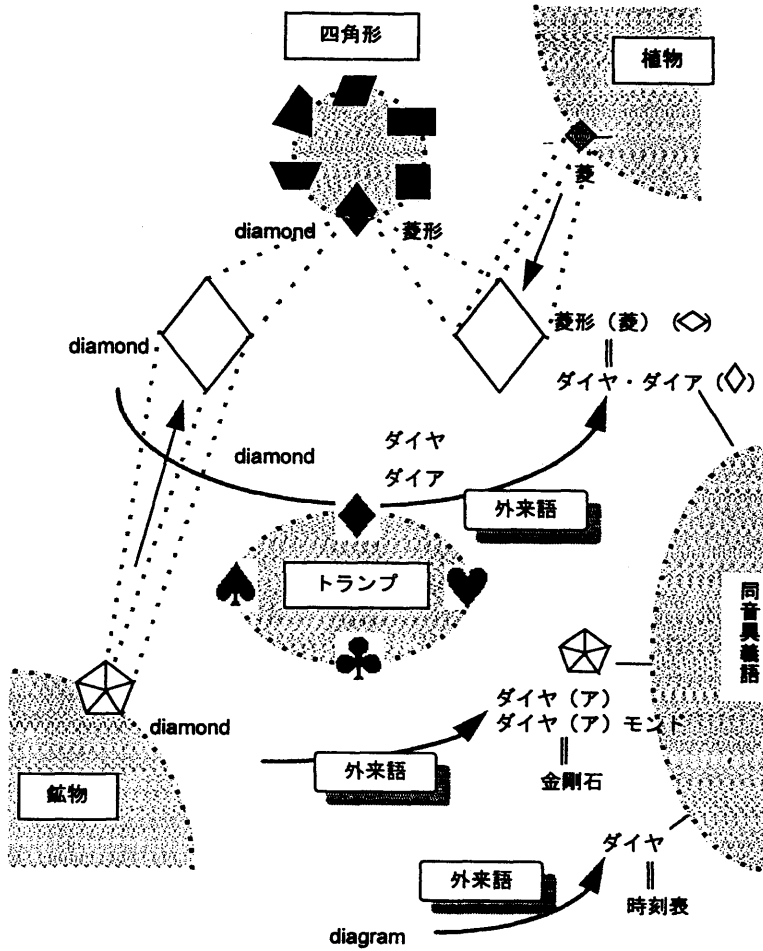


図5: “◆” をとりまく様々な文脈

(diamond から)である。「ダイヤ」という語に焦点を合わせると、公共交通手段のスケジュールもダイヤ (diagram から)である。

対象を三言語に広げると、対応関係は一層複雑になる。例えばドイツ語では、トランプの“◆”(Karo)も幾何学の“◆”(Raute, Rhombus)も鉱物のダイヤとは無関係であり、シンボル“◆”はダイヤを指すのには用いられない。

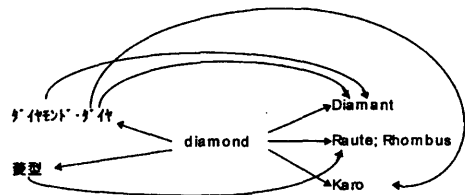


図6: “diamond” の翻訳 (日・英・独)

5.2 データモデル構築の困難性

diamondやwaterの例から、用語DBで対象とする言語を増やすたびに、

- 「概念の細分化」に起因するレコード数の変更や追加
- データ項目の変更や追加
- レコード間の関係の変更や追加

など、既存のデータ構造の見直しが余儀なくされることがわかる。ある文化や言語では重要な「概念」が別の文化や言語ではそれと意識されていないケースは多いし、日本語を対象としたとたんに、欧米語の用語DBでは不要だったデータ項目「読み」(発音)が必要となったりする。

つまり、言語による「現実」の構成(シュミレーション)が複雑で混乱した原理に従ってなされ、相互に関連する複数のコードがその過程に同時に関与してくるため、データモデルの構築は単一言語を対象にした場合でも至難の業である。さらに各種コードの「文法」が言語ごとに異なることから¹⁹⁾、特にDBの設計時点で対象言語が確定していない場合には、エンティティの特定もできず、複数言語を対象としたモデルの構築は多かれ少なかれ不可能となる。

5.3 規範的アプローチと記述的アプローチ

この問題に直面した際にとりうる道の一つは、対象となる言語やデータ項目、DBに記述する用語の文脈のパターンをあらかじめ制限することであろう。その一例が、「ターミノロジーの父」Eugen Wüsterの打ち出した、以下の原則に基づく規範的アプローチである。

- 科学技術分野の用語では、地球上のすべての言語を考慮する必要はなく、少数の特定の言語のみが重要である。
- 「概念」は言語に依存しない。
- 「概念」は階層的に組織化されている。
- 言語間の「概念」のずれは、標準化によって解消されなければならない。
- 用語整理作業では、言語の通時面(時間の経過による変化)は考慮しない。
- 用語整理作業は非言語学的なものである。

E.Wüsterはその博士論文¹⁰⁾の中で、異言語・

異文化間コミュニケーションを円滑にするためには、標準化によって現存する言語の「概念」のずれを解消し¹⁰⁾、将来的には統一言語を導入すべきだとしてそのためのプログラムを展開した。しかし、T.Kuhnのパラダイム論からもわかるように、あるパラダイムの中では意味の確定した用語もパラダイムが異なれば全く別の意味を持つのであり、用語の「意味」や「概念」は、常にそのパラダイムの中でしか捉えられない質の物である。標準化による言語間の差異の解消は(もの見方の一元化を強いるのでないとすれば)、限定された範囲でしか適用できないことは明白であろう。

言葉は使用されることによって常に変化してゆく。用語やその意味の、地域、時代、派閥などによる差異や変化の方が、標準化の努力による差異の解消より大きい。自然科学や技術分野では用語とその意味は一義的に決まるとい説があるが、研究や議論の盛んな分野ほど複数のパラダイムや技術が競合し、新しい理論、パラダイム、技術を説明するために新しい用語が作られ、既存の用語の再定義が行われる。概念体系を静的に捉える規範的アプローチは、記述的な用語収集整理作業においては、技術分野に限っても妥当ではないだろう。この点に関して、カナダの政府機関で長年専門用語の収集と整理に携わっているW.D.Coleは次のように述べている。

この理論の教義を他の領域の研究成果に照らして検討してみると、この教義には言われるほど普遍性はないことがわかる。...[中略]...一つの学問分野が確立してゆく際には、最初は定義されたデータも少なく、理論的枠組みも曖昧で幼稚なのが普通である。研究が進むにつれてデータの定義も進み、信頼のおける物になってゆくのである。...[中略]...こういった展開の特色は、それが一様なものではないという点にある¹⁰⁾。

用語の収集や整理は、標準化等のために規範的な方向付けを持って行われる場合もあるが、ある特定の分野で実際に使用されている用語を学生や翻訳家などに提供するなど記述的な方向性をもつ場合が多い¹⁰⁾。限られた予算、人員、期間内で行われるため、ユーザーのニーズにとって重要でないデータは集められない。従って用語情報システ

ムは、情報が不完全でも曖昧でも機能し、言語による情報やコミュニケーション過程が常にそうであるような混乱した不可知の性格を取り扱えるものでなければならないだろう。

仮に完全な概念ネットワークを描くことが可能だとしても、そんなものはあまり実際の役に立つ物ではないし、大概の場合なくても構わないものである。これまで化学の完璧な概念ネットワークを作ろうなどと試みた人はいない。... [中略]...概念ネットワークなどなくても、多くの信頼に足る用語集がまとめられてきたし、いま現在もまとめられている。... [中略]...記述的な用語整理作業にとっては、該当分野の完全な概念ネットワークが存在しないことなどは、たいてい重要なことではない¹⁰⁾。

これまで E. Wüster らによってまとめられてきた規範的な用語の収集整理や用語集作成のノウハウが全く使えないとか、規範的な用語の収集整理が行われてはいけないというわけではない。しかし、これらのノウハウを今一度より現実的な視点から見直し、言語学的基盤を踏まえて、より記述的な方向付けを持った理論の中で捉え直す必要がある。

さて本題に戻るが、コンピュータ上での各用語エントリは、何よりもまず複数の情報の集合である。これらの情報をどのような形で編成するか (DB のデータ構造) によって、情報へのアクセス形態やデータの一貫性制御の難易度が変わってくる。複数のエントリ間の関連づけもデータ構造に左右される。データ構造が単純であれば、複雑なエントリは平面化 (単純化) されざるを得ないが、エントリの複雑さを保つには、一般にデータフィールド数 (リレーショナル DB ではコラム数) が多くなり、妥当なアクセスパスの決定は複雑になる¹¹⁾。言語データベースの場合は - 言語の曖昧性のために - データフィールドの多くが空白になるおそれもある。

市販のシステムや企業内 DB の多くは、登録可能なデータカテゴリ数と収録可能な言語数を比較的少数に限定し、一概念に対し各言語一つしか (見出し) 用語を登録できないようにすることによって、データ構造を単純化する道を選んでいる。

その根拠としてあげられるのが E. Wüster らによる概念中心アプローチであるが、このアプローチがデータの冗長性や情報損失などの深刻な問題の原因となっていることは、すでに見たとおりである。

6. 新世代の用語データベース設計の基本方針

本章では、前章までの考察に基づいて、用語情報システム設計の基本方針をまとめる。これらの方針を実際のシステムにどのように実現させるかは、ANKER, 1994 を参照願いたい。

6.1 オープン・アーキテクチャー

用語情報システムは、複数の DB からなる「知識」を表現するシステムの一部として考えられるべきである。つまり、他の用語 DB や別種の情報システム - 例えば (文法) 形態情報システム、ジオグラフィカル情報システム、DNA データベースなど - との間で情報をやりとりできる規格化されたインターフェースを提供できるものでなければならない。

6.2 情報中心アプローチ

DB や情報システムに収録されるのは、「意味」や「知識」そのものではなく「情報」である。情報は相互に関連する複数のコードによって構成されるが、コード間の相互依存の度合いや情報構成の複雑さは一様ではない¹²⁾。コード自体についての情報も一つの情報として収録できる。

情報の検索は地図を使用するのに似ているが、地図にある建物が記載されていないからといって、その建物が現実に存在していないわけではなく、その地図が間違っているわけでもない。例えば軍事基地は軍事地図では重要であるが観光地図では記載の必然性はうすい。情報システムにおける情報についても同様である。

6.3 簡略化された現実

地図は、土地そのものでも、対象の土地についての全ての情報を記載したものでもなく、ある土

地について特定の視点から、特定の興味に基づいて見た情報をまとめたものである。例えばそれぞれ違う部分が青く塗られた複数の日本地図で、青色は、ある地図では湖沼や河川、別の地図では特定の人口密度、さらに別の地図では特定の年間平均気温といった簡略化された「現実」を表すシグナルである。用語情報システムも地図と同様に、複数の簡略化された現実についての情報を混同せずに組織化できなければならない。

あるシグナル（一つのシグナルによって表現される一つのコード要素）が担う情報量は、コードの要素数によって決まる²⁰。コードの要素数が少ないと一要素の担う情報量が多くなり、逆に要素数が多いと一要素の情報量は少なくなる。複雑な現実（ないし想定された現実²¹）を要素数の少ないコードで表現（より正確にはシュミレート）する際には、複数の事象を同一のコード要素によって表現せざるを得ない。この場合、あるコード要素の適用範囲を別のコードによって限定するなどしてコードを複合化すると、一要素が担う情報量が減少するので、情報をより特定化して処理することができるようになる。専門用語（special purpose vocabulary, special field vocabulary）は、このようなコードの複合化（適用される文脈を「有機化学」に限定するなど）により、一般語彙（general purpose vocabulary）におけるコード要素間の潜在的なリンクの幾つかを切ってしまうことにより成立すると考えられる。つまり用語は、特定の文脈に結びつくことによって潜在的な解釈可能性が減じられた語彙である²²。言語は、このように相互に依存しあう複数のコードを用いて特定の情報に焦点を合わせてゆくため、複雑な「現実」を言語によってシュミレートする際には、その情報処理過程に関与してくるコード数とコード要素数（と信号数）は必然的に大きくなる。

色彩表現といったごくごく単純な語彙群をとってみても、それを組織化してみるとどの言語でもかなり複雑な構造になる。「常識的な現実」の世界に属する「白」とか「黒」などの様々な「色」の用語と光学分野での色の用語とでは、同じ語彙の意味が異なるし、光学の場合は加法混色か減法混色かによって特定の二色の混合色が変わって

るので、色彩用語間の関連もそれに応じて変わる。また色彩表現は多くの言語・国家で政治用語でもあり、「政党」という文脈では「赤」が社会（民主）党、「黒」が保守政党の同義語であるオーストリアのような国もあれば、「黒」と「保守」が結びつかない日本のような国もある。英語圏では熟していない果物を green と形容するのに対し、日本語では「青」となる。語彙集・用語集や DB を編成する際には、このような複数の簡略化された現実の相互関連性を何らかの形でシュミレートする仕組みが必要である。

6.4 複数の現実の共存

情報を検索・収集・提示・利用するコミュニケーション過程²³は、通常何らかの「環境」や状況に埋め込まれている。特定の知識や技術 - 仮にここでは医療 - を学ぶために特定の人物に師事したとすると、そのことによって、あるシグナル（「症状」、「因子」など）の重要性は高まり、他のシグナルの重要性は低くなる。師の選択によって間接的に環境としてのパラダイム（ここでは診断体系）の選択もなされるため、そこでの情報の適用可能な型や解釈可能な型がある程度決まってくる。これも、コード要素間の潜在的リンクの切断（とそれによる専門用語の成立、語義の特定化）の一つの形である。

診断体系が文化・地域・学派によって異なるように、それぞれの体系が了解している「現実」も異なっている。端的な例で西洋医学の用語と東洋医学の用語を同一のシステムに収録・管理できるようにするには、時には互いに矛盾しあう複数の「現実」を扱うしくみが必要となる。具体的には例えば、上位語彙、下位語彙、反義語、可能な訳語などの用語間の関係の網はパラダイムによって大きく異なるので、これを複数平行して扱う仕組みが必要になってこよう。

6.5 言語中心アプローチ

「現実」は、言語と切り放して考えることはできない。言語は我々が特定の社会的文脈の中で現実を「安定化」させるための主要な情報体系であって、他の、例えば数学の体系（一種の高度に形式

化された言語)を導入・説明する際にも、何らかの自然言語が出発点となる。「現実」のシュミレーションが使用する言語によって異なる一例として先に挙げた water の例では、多くの言語で - 英語では ice, water, steam (vapor) というように - 水の三状態を言語上区別しているのに、日本語では水, 水, 湯, 水蒸気(湯気)となり、固体, 液体, 気体という区分とぴったり重なってはいない。

図7は、一つの単純な用語が成り立つ際に関与している諸コードを簡単にスケッチしたものである。ある語彙を用語として安定させるのに関与しているこれらのコードは、用語DBの編成にとっても重要であろう。どんな単純な用語体系も、絡み合った複数のコード体系の束なのである。

用語「water」は、英語の表記コードや音韻コードで表すことができ、またコンピュータ上では5個の1バイトの数字の羅列(文字コード)で表される。日本語のコーディングはより複雑で、「mizu」は一個の漢字(表記)コードで - コンピュータ上では2バイト値(コードにより9085や3F65など)で - 表すこともできるが、音を示す場合などには、ひらがな, カタカナ, ローマ字, 発音記号等の表記コードでコーディングされることもある。コード間の対応も複雑で、一般に漢字

の多くがいくつもの「意味」に対して使用され、複数の読みを持っている。「mizu」と「水」も一対一で対応しているわけではない。

概念中心アプローチに沿って「water」と「水」を、例えばH₂Oを表す別のラベルだとする図式は、化学分野で物質の組成が話題となっているときなどの特定の文脈でしか成立しない。一言語内でも複数言語間でも、このような用語間のリンクは、つねに特定のパラダイムのなかで、特定の文脈でのみ成立する。

概念中心アプローチは、「現実」の構造は一つしかないという思想を背景にしていると考えられるが、仮に「唯一絶対の現実」を論議する場合でも、人は主として既に存在している語彙を使わざるを得ない。どんな語彙も、無数の人々に様々なやり方で(それぞれのものの見方・それぞれの「現実」を反映すべく)構造化されてきた歴史を引きずっている。学問領域でも日常生活でも、特定のパラダイムが優勢になれば、ある語彙の潜在的(ないし歴史的)な文脈や他の語彙とのリンクのうち、特定のリンクの使用頻度が高まり、他のものは忘れられていく。新しい考えやムードを表現するために、既存の語彙をそれまでとは違った文脈で、それまでは組み合わせることのなかった語彙と一緒に使用することもある。「概念体系」

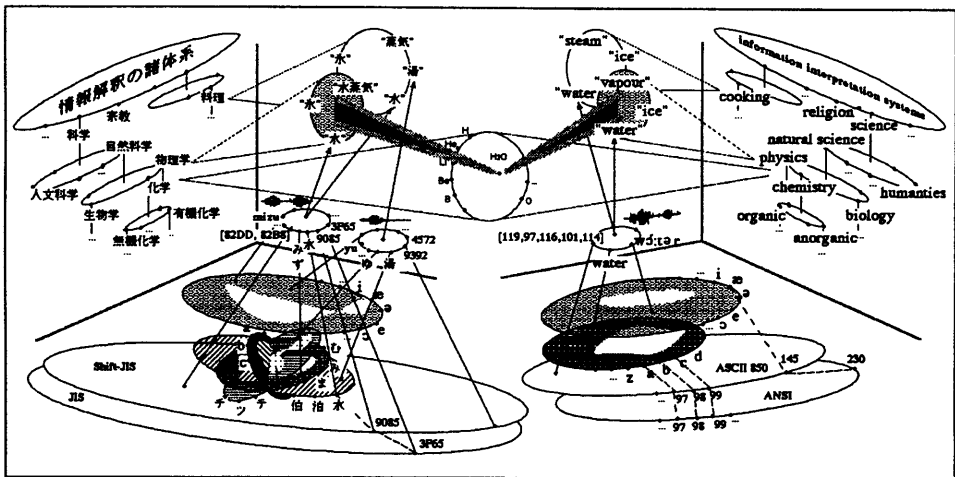


図7: 知の空間

は、言語、文化、地域、サブカルチャー、パラダイムごとに異なり、複数競合しているとともに、時間の経過とともに移り変わってゆく。そのことによって、言葉の形態や発音だけではなく、その意味も変化してゆく。

そういった変化は、もちろんコミュニケーションが成り立たないほど速く進むわけでも、一定の速度で進むわけでもなく、各言語とパラダイムごとに使用される語彙（文脈や他のコードとの潜在的リンク、コーディング）はある程度安定しており、大きな変化は例えばパラダイムの変換に伴って起こる。従って、多言語を対象とした用語整理の方法論としては、まず言語ごとパラダイムごとに用語を整理し、次に言語ペアごとに用語の対応関係を記述するのが妥当であろう。

6.6 フラクタル・アプローチ

常に移り変わってゆくという言語の性格からして、多言語用語DBのデータモデルを細部にわたって確定させることは不可能である。広い意味での「文法」も言語によって千差万別である。しかし、あらゆる自然言語は構造化されており、「語彙」を組み立てるのにそれぞれ一定の数の音素を使用し、複雑な情報を構成するには各々のルールに則って語彙を組み合わせるべく、といった共通の特性がある。言語を用いたコミュニケーション活動は、発話を文脈（常識、テーマ、コミュニケーションの場、話者と聞き手の関係、歴史的・社会的に成立している各語彙の潜在的リンク）に適合させようとする指向性を示し、またこのような明言されない文脈に依存することによって意味を成立させるが、これも各言語共通の特性である。

あらゆる自然言語に共通であろうと考えられる特性を元に核となるデータモデルを立て、個々の言語やパラダイムに依存する要素 - 音韻体系、統語法、語彙のパターン、社会的に成立している語彙間の潜在的リンクなど - はフラクタルと捉えて、収録言語や用途に合わせてDBに採用するというアプローチが採られるべきではないだろうか²⁰。そうすれば、対象とする言語によってデータモデルやデータ構造（リレーショナルDBではコラムの種類と数）自身を変更することなしに、

収録言語の変更や用途の変更に柔軟に対応すること、つまり、データベース編成に必要な論理データの独立性（概念モデルと外部モデルの切り直し）をはかることができよう²⁰。

論理データの独立によって同一の安定したデータモデルを採用しながら多様なアプリケーションを用途に合わせて作成することができれば、複数の組織の異なるニーズにも組織内のニーズの変更にも対応でき、オンラインでインターアクティブなデータ交換も行える、開いた用語体系を取り扱うシステムが実現できよう。遠隔地や外国のパートナーと協力しての用語データの収集整理や、データの幅広いユーザーによる利用、各種アプリケーションでの利用といった、データの高度利用を促すことにもなるだろう。昨今の広域エリアネットワークや分散型DBシステム関連の技術の進歩を見ると、そのための技術的前提はかなり整っていると考えられ、必要なのは方向性の明確化と妥当なデータモデルの確立であろうと思われる。

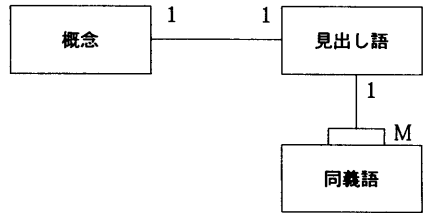
文献

- ANKER, Georg J., and Sasaki, Yuka (1994) "TERM IS - a Database Design for Interactive Global Sharing and Webbing of Terminology" *The 47th FID Conference and Congress*, Tokyo, pp.133-137
- COLE, Wayne D. (1991) "Descriptive Terminology: Some Theoretical Implications" *Meta vol. 36, No.1*, Montréal: Les Presses de l'Université de Montréal, pp. 16-22
- FELBER, Helmut, and Budin, Gerhard (1991, 92) "Terminologie in Theorie und Praxis", Tübingen: Narr
- FELBER, Helmut (1992) "Terminology Manual", Paris: UNESCO
- LAUREN, C., and Picht, H. (1993) "Ausgewählte Texte zur Terminologie", Wien: TermNet, Internat. Network for Terminology
- WÜSTER, Eugen (1970) "Internationale Sprachnormung in der Technik, besonders in der Elektrotechnik: Die Nationale

Sprachnormung und ihre Verallgemeinerung”, Bonn: H. Bourvier u. Co.

田中久美子, 梅村恭司 (1993) “和英・英仏辞典からの和仏・仏和辞典の抽出” ソフトウェア科学会自然言語処理シンポジウム

長尾真 (1992) “辞典形式での専門分野の知識の体系的構成法” 人工知能学会誌 vol.7, No.2, pp.320-328



脚注

- 1 冊子体であれ電子媒体によるものであれ、用語集や辞書の作成では、対象となる言語、分野、収録用語範囲、データ項目、用途、索引項目等をあらかじめ決める。発行後は改版しない限りデータの変更や追加は行われない。このような用語集は、特定の視点から特定の時点でまとめた閉じた用語体系である。
- 2 用語 DB では随時データが追加入力ないし削除されるとともに、必要に応じて収録する言語や分野、データ項目が追加・変更されてゆく。用語 DB に収録されている用語体系は実際の用語体系そのものではないが、それと同様に開いた体系である。
- 3 ANKER, 1994
- 4 例えば、製品単価が常に一定であれば、製品番号から単価が一義的に決まるので、単価は製品番号をキーとしたエンティティ「製品」の属性であると考えられる。これに対してアカデミックプライスなどを設定している場合は、製品番号と単価が M 対 N 対応するため、先のグループ化では役にたたず、単価は製品番号と顧客タイプのペアに従属すると捉える必要がでてくる。
- 5 本稿で「データモデル」や「モデル」は、概念モデルをさしている。
- 6 なお、「専門用語は一般語彙と異なり、同一言語内では概念と用語が 1 対 1 で対応する(べきである)」というテーゼの下、同義語は一つの正式な用語の付随情報として登録させる DB が多い。この場合は、図 1 を変形させた以下の図となる。
- 7 用語管理ソフトのマニュアルには、英語では

「chair」ドイツ語では「Stuhl」と用語は違っても、どちらも同じひとつの概念に貼られた別のラベルだ、といった類の説明がよくみられる。

- 8 西欧哲学史・言語学史から見ると、概念中心アプローチは、プラトンや聖書以来の西欧の伝統的な言語観であった言語命名論や言語名称目録観に連なるものである。E.ヴュスターがこれを用語管理の哲学的基礎として提示する(1932)のに先立ち、このような言語観を学問上否定する F.ソシュールの「講義」や L.ヴィットゲンシュタインの「論理哲学論考」がでていることに注意したい。

しかし哲学や言語学ではともかく、応用分野ではこの言語観が今日も繰り返し議論の対象となっており、用語 DB や機械翻訳システムなどの多くの言語情報システムに採用されている。

- 9 田中, 1993 なども参照。
- 10 これを DB 設計の観点からみると、不必要なデータの繰り返し(冗長性)があり、データが正規化されていないために、DB の一貫性制御が妨げられることを意味している。
- 11 コードの要素や要素数ばかりではなく、コードのあり方も言語ごとにかなり異なる。例えば日本語は、英語とは違い表音文字コードと表意文字コードを合わせたハイブリッドな文字コードを用いている。
- 12 この種のコードはしばしば分類法(クラシフィケーション)と呼ばれる。特定の分類法が適用するか否かは、慣習の問題であって論理の問題ではない。普遍的な分類法はないことに注意したい。

- 13 ここでは、統語面の文法だけではなく、音韻、意味、語用など様々なアスペクトの各コードのルールという意味で「文法」という語を用いている。統語面に限っても、現存するすべての自然言語をカバーする「メタ文法」といったものはまとめられておらず、言語の意味や語用面については言わずもがなである。
- 14 WÜSTER, 1970
- 15 FELBER, 1991, 92, FELBER, 1992, LAUREN, 1993 にも同様の主張が見られる。
- 16 COLE, 1991, p16
- 17 機械翻訳システムや言語コードによる知識を扱う他のシステム（エキスパートシステムなど）でも、今後このアスペクトが重要となつてこよう。
- 18 COLE, 1991, p19
- 19 例えば見出し語と同義語のフィールド（コラム）を別にすると、検索の際に調べたい語彙が入っていそうなフィールドをすべて対象にしなければならない。
- 20 「コード」の定義については情報理論の文献を参照。
- 21 この文脈では「情報」と「意味」を混同しないように注意。一般に一コードの情報量が多いと、そのコードは多義（意味が曖昧）であり、逆に情報量が少ないと、そのコードの意味は明確になる。
- 22 情報システムにとっては、「現実」が真に存在するか否かはさして重要ではない。
- 23 専門用語をこのようにとらえるならば、E. Wüster の唱えた「概念」を想定する必要はない。なんらかの共通のコードを学び（教育し）、使用し、それによってコードを再構成していくコミュニティがある限り、そのコミュニティの各成員は、他の成員が了解（使用）しているコード要素やコード要素間の慣習的関連性などを、最低限は学び使用するようになる。他の成員との意志疎通を図るためには、そうせざるを得ないとも言えよう。コード要素間の潜在的なリンクの切断は、ある特定のリンクだけを明示的に示そうとすることによって（それによってその他のすべての可能性を除外することによって）行われることが多い。このプロセスが失敗すると「理解」できなかったり「誤解」されたりするが、成功すればコミュニケーションの「パフォーマンス」が高まる。
- 24 これらのコミュニケーション過程は、常に一種の社会的相互作用である。
- 25 ANKER, 1994 に詳しく述べる。
- 26 データの追加や削除、データ項目の変更、データ量の増加に伴うハードウェアの変更、検索効率改善のための新しい DB 管理システムを用いた DB の再編成とデータの移しかえなど、DB はたえず成長し変化してゆく。そのような際にも既成の応用プログラムをそのまま使えるようにする必要がある。DB の変化に応用プログラムが影響されないですむ性質はデータ独立性と呼ばれ、論理的データ独立性と物理的データ独立性の二つのレベルに分けられる。論理的データ独立性とは、他の応用プログラムの変更などの論理的に無縁なデータの変化によって応用プログラムが影響を受けないこと、つまり概念スキーマの変更が外部スキーマに影響を与えないことであり、物理的データ独立性とは、内部データの構造変化やハードウェアの変更等のデータの物理的変化が応用プログラムに影響を与えないこと、つまり内部スキーマの変更が概念スキーマに（従って外部スキーマに）影響を与えないことを指す。

～FID 東京大会から～

抄録による紹介

安全研究における専門用語データベースの構築:複数データベースからの安全および 危機管理問題のデータの効率的なアクセスを求めて

Construction of Terminology Database in Safety Research:For efficient accessing
safety information from numerous databases

後藤 智範, 山本 晴彦, 高木 伸二, 河村 正一,
小幡 行雄 (神奈川大学理学部), 藤原 鎮男 (国文学研究資料館)

抄録 安全問題は、労働、運輸、環境、化学、薬学などの様々な分野に見られる。そのため、これらの分野から安全に関する情報を収集するためには、複数のデータベースの検索が必要である。書誌データベースは、個別の索引語体系を持っており、同一の概念に対しても、データベースにより索引語が異なることがあるため、複数データベースから安全情報を包括的に収集することは、事実上不可能と言ってよい。この状況の改善、すなわち、複数データベースからの包括的な検索において検索効率を著しく改善するために、著者らによる安全情報についての用語データベースが開発中である。

当研究では、現段階において約4,000用語を収集したが、当論文は、当用語データベースの構築に際し、下記の項目について報告したものである。

- (1) 安全問題に関する専門用語の収集方法
- (2) これらの用語における基本用語の収集方法
- (3) この基本用語の分類カテゴリーおよび、約4,000語の同定方法
- (4) 上記の用語を用語データベース化するためのデータ項目

TERM IS — 地球規模の用語情報ネットワークのためのデータベースデザイン

TERM IS--A Database Design for Interactive Global Sharing and Webbing of Terminology

佐々木由香 (名古屋大学大学院 国際開発研究科)

抄録 従来、ほとんどの専門用語データベースや用語管理プログラムが、特定の OS やプラットフォーム上で稼働することを前提に設計されてきた。多くの場合、その設計規準とされてきたのは、概念中心アプローチであった。これからの専門用語データベースには、インターアクティブなデータ交換、多言語・多文化対応、他機能、マルチ環境への対応といった要素が求められるが、概念中心アプローチでは、データベース構造の一貫性を実現することができず、これらのニーズに対応できない。

これに代わるアプローチとして、筆者らが提案する TERM IS では、異なる言語を対象とし、異なるニーズを持ったユーザーグループが、それぞれのニーズに合わせたユーザーインターフェースや異なるアクセスパスを用いながらも、同一の基本データベースデザインを採用して、今日の標準的なデータベース技術による柔軟な用語情報システムを構築することをねらっている。

本論文では、TERM IS 構造を技術面から概括し、ユーザーインターフェースの一例を紹介する。

キーワード: 用語データベース, 用語情報システム, 用語の共同整理作業, 用語データのネットワーク, インターネット, リレーションアルデータベース

オンライン情報検索システムに用いられる操作用語の標準化の一つの提案

One Proposal for Standardization of Operating Terms on
Using the Online Information Retrieval Systems

大前 巖 (帝人技術情報)

抄録 今日、多くのオンライン情報検索システムが、インフォメーションスペシャリストおよび研究者などによって便利に利用されている。しかしながら、これらのシステムのそれぞれの操作用語は、それぞれのディストリビュータによって、それぞれに定められている。それゆえ、オンラインユーザーは、それぞれのシステムのディストリビュータが定めた操作用語を用いなければならない。したがって、オンラインユーザーが、二つ以上のシステムを用いると、しばしば混乱することになる。多くのユーザーは、DIALOG が世界でもっとも早くから用いられるようになったことと、もっとも多くのオンラインユーザーを有するシステムであることから、標準的な用語として DIALOG の用語が用いられている。しかし、DIALOG の用語は、多くのオンライン情報検索システムの中で DIALOG にだけ用いることのできる用語であり、DIALOG の用語は、必ずしも、情報検索の個々の操作の意味をもっとも適切に示しているとは言えない。そこで、筆者は、標準的な操作用語として、DIALOG の操作用語の次に示すもの：Descriptor, Identifier, Posting, Basic Index, Additional Index, Search Session, Set Number および Search Save については、それぞれ、Controlled Term, Uncontrolled Term, Number of Hits, Index File, Index File, Session, Search Number および Save と用いるように提案する。これらの新しい操作用語は、Table2 端に Operatin Term として示した。また、これらの日本語の標準的な操作用語が Table2 の項目に提案した。

キーワード：操作用語、オンライン情報検索、標準化、ディストリビュータ、オンラインユーザー、DIALOG、オンラインシステム、標準操作用語

用語間の意味関係の抽出法 — SS-KWIC

Extraction of Semantic Relationships Among Terms—SS-KWIC

頼 静娟, 陳 漢雄, 藤原 譲 (筑波大学 電子情報工学系)

抄録 情報構造化の重要性および必要性が、情報や知識の意味処理の観点から除々に認められてきた。概念を表現する単位としての用語の組織化は、情報を構造化するために有用である。

本研究で報告する SS-KWIC の方法は、用語間の意味関係を抽出するために開発した手法の一つであり、用語の生成規則に基づいて回想的な関係および関連的な関係を自動的に抽出するために役に立つ。これは、シソーラスの自動構築とも言える。専門用語の表現モデルとして新たに開発したスーパーグラフモデルに基づく SS-KWIC の方法の原理および実装法が示され、実験の結果の例を示す。原理は簡単であるが、詳細な階層および豊富な関連関係が得られ、その結果は、方法の信頼性および実用性を示した。

キーワード：SS-KWIC、自己組織化、意味解析、意味関係、情報構造化、階層関係、関連関係、自動シソーラス、スーパーグラフモデル、ターミノロジー

自動索引の制限

The Limitation of Automatic Indexing

R.Fugman (International Society for Knowledge Organization, Germany)

抄録 コンピュータ化を考える時、常に「決定論のプロセス」と「非決定論のプロセス」の差を認識していなければならない。非決定論のプロセスとは、認知可能なパターン、法則に従わず、予見不可能な進行をするものである。どんなに精緻なメカニズムを持ってしても、人間のアプリオリな指示で制御できるものではない。

無限の可能性の中から、ある時のある一つのできごとが予見、定式化できたとしても解決にはならない。そのよ

うな例に惑わされて、すべてのできごとが適切な指示のもとに予見しようと考えるてはならない。

同様に、初期状態が不可知なプロセスも非決定論的といえる。そのプロセスが成功に向かうことができるのは、アポストロリに決定をくださる人間の、知的介入をもってのみである。

人間が考えや概念を自然言語で表現するのは、典型的な非決定論的プロセスである。全自動翻訳が成功しないのは、予見不可能性すなわち、非決定論性が原因である。

インデクシングには、(索引言語への) 翻訳と、抽象化という非決定論的段階が含まれる。したがって、インデクシングの全自動化は不可能である。反論のある人は、自然言語のテキストの全自動翻訳を理解すべきである。

(人手による) 知的インデクシングは、しばしば自動インデクシングと比較され、後者は同等あるいは優位とされる。しかし、これらの実験で行われている知的インデクシングは、その性能を発揮しているとは言いがたい。この原因は、インデクシングの信頼性の低さや、語彙体系の分類構造の欠如、文法の欠如といった、索引言語の不備によるものである。

効果の面で、自動インデクシングが、知的インデクシングを代替できることがあつたら、自然言語による表現が極端にあいまいな場合か、情報検索に「個別概念」(「一般概念」の反対語として) の取扱いしか期待しない場合だけである。高度な索引言語、およびその信頼性の高い機械支援の開発にもっと力を注ぐべきである。

専門用語研究会アンケート 集計結果

「専門用語研究」「ニュースレター」
編集委員会

編集委員会では、会員の方々が、本研究会とターミノロジーをどのように考えておられるのかを調査するため、1995年8月、全会員に対しアンケートを実施しました。その結果、33会員から回答をいただきました。下記に、集計データおよび寄せられたコメントを紹介して、今後の会の活動の参考に供します。

はじめに

当会の案内パンフレットから、設立の趣旨を引用、紹介します。

「情報化社会の進展にともない、情報や知識の表現と交換の基盤である言語の持つ役割はきわめて重要になってきました。とりわけ、自然科学、人文・社会科学、さらにそれらの応用である技術やビジネスの世界において、語の持つ概念を明確にして確実な伝達を行うことは、学術の進歩や社会の発展のためにぜひとも必要なことであります。

これまでも、専門領域ごとに用語集の作成などに大きな努力が払われてきましたが、学際的な活動が増大するにつれて、学際的かつ国際的な観点にたつて、学術用語を扱う必要性が一層高まってきたと思われまふ。ターミノロジーは、学術用語を対象とする科学および技術であり、Infoterm、ISO/TC37などの国際機関はその推進のために意欲的に活動しております。日本においては、最近に至つてこの分野への関心が高まり、1987年に開催された「専門用語と知識工学に関する国際会議」には、日本からも多数の専門家が参加しましたが、これを機会に日本にターミノロジー活動のノードを作り、Infotermなどの国際的活動への参画を強化してほしいとの要望が出されました。このような情勢に対処するため、専門用語に関心を持つ人達の情報交換と相互協力の場として当会を設立しました」

そして、会の目的として、「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つて、専門用語に関する研究や技術の向上に貢献すること」とあります。

これらの趣旨・目的のもと、現在約130機関・個人が参加、活動を行つていますが、層は広く、専門主題を持つ専門家、図書館情報学、言語学、情報システムに関係する人達、そして実務者として翻訳者、編集者、辞典編纂者、図書館員、教育者、言語産業関係者など多様です。そのために、本会の活動もどこに焦点をあてたら良いのか、議論を重ねてまいりました。

このたび、編集委員会では、会員の方々が、本研究会とターミノロジーをどのように考えておられるのかを調査するため、会員に対しアンケートを実施しました。その結果、33会員からの回答をいただきましたので、下記に、集計データおよび寄せられたコメントを紹介して、今後の会の活動の参考にさせていただければ幸いです。

1. 質問1：専門用語研究会に何を求めて入会したか？

1.1 集計結果（複数回答あり）

- | | |
|-------------------------------------|----|
| ▷ 広く言語、記号、意味などに興味があり、これらと関連した研究をしたい | 15 |
| ▷ 専門用語の標準化、統一化を研究したい | 14 |
| ▷ 専門用語が持つ特殊性を研究したい | 10 |

▷ 専門用語集, 専門用語データベースの作成ノウハウを得たい	12
▷ 知識, 認識など, 人間の思考活動に興味があり, これらと関連した研究をしたい	11
▷ ターミノロジーに興味があり, 知的好奇心を満たしたい	10
▷ ターミノロジーの応用 (学際領域への活用, 研究者交流のツールなど) を研究したい	8
▷ その他	6
• エスペラントの専門用語と日本語との接点を求める	
• 印刷と用語の標準化	
• データベースと用語とに関連する問題全般	
• 自然言語処理 (用語の計算機処理, 機械翻訳など)	
• データベース検索ツールとしての用語の研究	
• 専門用語の適正化	

1.2 コメント：専門用語研究会について

ほかの機関との連携・交流

◆専門用語の作成過程が, 戦後一時期までは, その関係分野 (学問) 全体で充分議論されていたが, ある時期以後は一部の人間だけで用意されるようになったと思われる。専門用語の適・不適について相互批判盛んになり, 違分野の人間同士の考え方が良く交流されることを期待したい。即効は無理な領域だと心得ているので, 永続し充実した交流が形成されることを願う。

◆何らかのプロジェクトを国などから受けて, ワーキング・グループで作業してもよいと思うが, 現状の研究会のパワーでは無理か。

◆専門分野別に関連する学会を巻き込み, 用語研究を強力に進めることが重要である。

データベース関係の立場から

◆データベースを作成しているため, 単語を登録する際, どのような形で, どのような情報を持たせればよいか, ヒントが得られると良いと思う。

◆分類, シソーラスにより indexing 作業をしているが, 分類, シソーラスの用語概念と, indexing 対象概念との落差に悩む。単に用語を統一, 標準化すれば良いわけではない。専門用語研究会の発表内容には, 参考になることが多い。

◆データベース提供業にかかわることから, 検索用語としてのシソーラス作成に興味がある。現在の会は, 図書館, 言語学, 情報システム, 翻訳者, 辞書編纂者, 教育者, 言語産業関係者など, 多様な会員構成となっており, 検索言語としての用語研究だけというわけにも行かないだろう。すべての会員の問題意識にどこまで応えられるか。

会誌・ニュースレターについて

◆定期的にニューズレターを発行して欲しい。シンポジウムの焼き直しでない会誌を発行して欲しい。シンポジウムだけでなく, 研究会・セミナーなども開催して欲しい。

◆ニューズレター, 会誌, シンポジウムが弱い。

その他

- ◆言語の発生と変化。それと、日本の現状に対応し、主導的立場をとり、専門用語研究会を発展させること。文部省の学術用語集の経緯を見ると発展の必要性を痛感する。
- ◆専門用語にかかわるすべての分野間の情報交換の場として期待する。学問の立場から、世界的、現実的に、多くの分野で活躍するとともに、これを総合的にまとめ、名実ともに日本において官僚的セクショナリズムを越えた指導力、影響力を持つ団体に発展させたい。
- ◆「情報や知識の表現と交換の基盤」の勉強に魅力を感じ入会した。
- ◆学問としての研究にもう少しウェイトを増すことが望ましい。研究成果を基にした高度な応用も期待している。
- ◆文科系分野の用語標準化についても会で扱ったらおもしろいのではないか。概して、文科系分野には用語標準化の要求が少ないのかもしれないが、法律のように用語が厳密に定義される必要のある分野もあるので。派閥争いも多いようなので、多少泥仕合のようになるかもしれない。

2. 質問2：今後、ターミノロジーに関して、どのようなことを行いたいのか？

2.1 集計結果（複数回答）

▷ 専門用語集，専門用語データベース作成を考える	15
▷ ターミノロジーの応用を考える	14
▷ 専門用語の標準化を考える	13
▷ 他の学問分野の視点からターミノロジーをとらえる	8
▷ 知的欲求が満たされればよく，特にこれといったことは行わない	1
▷ わからない	1
▷ その他	5
• 研究会を持ちたい	
• 専門用語の歴史	
• 専門用語の創成	
• 日常言語との違い	
• データベースと用語とに関連する問題全般	
• 自然言語処理（用語の計算機処理，機械翻訳など）	
• 良い日本語の一部として専門用語を考える	

2.2 コメント：ターミノロジーのどのあたりに興味があるか

多様な標準化を考える

- ◆特に医学用語に興味があるが、医学の中でもさらに細分化された領域で用語が統一されていない。ここから、専門用語の標準化に興味を持った。活字や写植にはない文字が統一用語の字体として定められているケースがあり、出版の際に困ることがある。
- ◆機械翻訳のように多国語間の変換には、似た概念の用語の統一化と、似た概念を定義し、あまり間違いない変換方法と、を検討したい。
- ◆専門用語の単なるデータベースはほとんど意味がない。意味を含めた情報源としては価値がある。標

準化すべきでないもの、標準化できないものと、その対処法に興味がある。

◆フルテキスト DB から、いかに効率よく目的とする情報を検索するか。この役割を担うツールが用語と思う。表記の不統一や、意味概念の相違（言語間も含む）により数多くの問題がある。どのような相違があり、どのように整理、標準化したらよいか、背景を含めて研究してみたい。

専門用語データベースを考える

◆専門用語のデータベースの feasibility 研究でも行ったらよい。

◆現在ターミノロジー・データベースの作成をしている。その関連で、技術・理論両面からのアプローチに興味がある。

◆検索機能を高める要因を、システム側の機能に依存するか、検索用語（統制語）に求めるか、それらの相互関係について関心がある。

用語集のあり方考える

◆カナ書きの用語が氾濫する傾向をふせぐ手だてを考えたい。日常語と専門用語の間に溝がありすぎる。専門用語が名詞ばかりなものもおかしい。専門用語をピックアップする時の姿勢が問題ではないか。

◆索引に過ぎない用語集の作成から抜け出し、学の体系をタームで示し、その索引が用語集になるようにしたい。

◆用語集は、新しい用語を取り入れる点において、かなり時代に遅れていると思う。誰を対象としているのか不明であるが、用語の選択、同義語、関連語の指示、などにさらに配慮の必要がある。

その他

◆ある学術分野について、当該分野の用語体系ができていなければ、その分野は科学とはいえないと強調してきたし、これからも強調してゆきたい。学際活動が頻繁になるほど、INTERNET の利用が増えるほど、用語の使い方に慎重さが要求される。

◆技術翻訳を通じ、専門用語（日本語）を各技術分野で独自に作り出しており、その基本的な考え方、概念がどうあるのか究明するところに興味がある。同じ概念に対し、いろいろ異なる用語があり、これが一層専門用語の複雑さを支えている。

◆用語は、

- 1) 情報科学の発展
- 2) 学問分野の学際化
- 3) 学問の高度化、国際化

の時代に入る。これらに対応する必要がある。

◆ターミノロジーに関する国際標準化の情報・動向について。

◆用語とするための努力（概念の明確化、他の用語との識別）はきわめて重要な仕事と心得ている。

◆ターミノロジーの国際的なハーモナイゼーションに大きな関心がある。

◆情報を一般大衆に伝えるための用語の選択に苦労している。

◆言語（専門用語）と社会の関係に興味を持っている。ことば（用語）は、それが使用されている社会の仕組みと密接に関係しているので、ことばを知ることを通じて社会を知ることが、ある程度可能なのではないか、と考えている。

「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、下記の内容に関する論文・記事を掲載する。

- ・ターミノロジーの理論と応用
- ・専門用語集の作成技術
- ・その他、専門用語に関するもの

2. 会員は、会誌に自由に投稿することができる。編集委員会からの依頼により執筆することもできる。

3. 原稿の書き方

3.1 原稿用紙

原稿は、通常のA4サイズ横書き原稿用紙(20字×20字)か、ワードプロセッサを使用する。ワードプロセッサを使用する場合は、A4用紙に1行20字、20行で作成し、印刷する。また、可能であれば、MS-DOSテキスト形式でフロッピーに保存し、印刷物と一緒に送付する。

3.2 原稿の長さ

全体で図表ほかを含めて、原稿用紙16枚から32枚とする。原稿用紙4枚で刷り上がり1ページとなる。執筆依頼時に別途指定ある場合はそれに従う。

3.3 原稿の仕様

原稿には、以下の内容を記入する。

- ・和文と英文の、表題、著者名、所属
- ・和文の、抄録(250字前後)とキーワード(5から10語)
(可能ならば、英文の抄録(150語前後)とキーワード(5から10語)も)
- ・本文(ページをつける)
- ・図表など(番号と表題をつけ、朱筆で文中に挿入位置を指定する)
- ・引用文献(本文中に肩付き数字¹⁾²⁾・・・を記入する)
- ・参考文献、参考図書(本文を読む上で参考になるものがあれば)

4. 原稿の受理、査読

投稿原稿は、当研究会事務局が受け付けた日を受付日とし、会誌編集委員会で査読を行なう。査読結果をもとに、会誌編集委員会で掲載の可否を決定する。委員会で内容・表現などについて修正が必要と認めた場合、執筆者に修正依頼する。

5. 校正依頼

執筆者に初校を依頼する。この際、大幅な修正・加筆は行なわないこと。なお、論旨に差し支えない範囲で、編集委員会が内容の変更を求めることがある。

6. 掲載原稿の扱い

会誌に掲載された原稿、フロッピーは返却しない。

7. 謝礼

執筆者には、掲載された会誌10冊を無料贈呈する。これ以上および抜刷を希望する場合は、有料となる。校正時に申し込むこと。

8. 著作権

本誌に掲載された論文、記事の著作権は、当研究会に帰属する。

9. 原稿提出先

専門用語研究会会誌編集委員会

専門用語研究会 (Japan Terminology Association)

設立まで

情報化社会の進展にともない、情報や知識の表現と交換の基盤である言語の持つ役割はきわめて重要になってまいりました。とりわけ、自然科学、人文・社会科学、さらにそれらの応用である技術やビジネスの世界において、語の持つ概念を明確にして確実な伝達を行うことは、学術の進歩や社会の発展のためにぜひとも必要なことです。

これまでにも、専門領域ごとに用語集の作成などに大きな努力が払われてきましたが、学際的な活動が増大するにつれて、学際的かつ国際的な観点に立って学術用語を扱う必要性が一層高まってきたと思われます。ターミノロジーは、学術用語を対象とする科学および技術であり、Infoterm、ISO/TC37などの国際機関はその推進のために意欲的に活動しております。日本においては、最近に至ってこの分野への関心が高まり、1987年に開催された「専門用語と知識工学に関する国際会議」には、日本からも多数の専門家が参加しましたが、これを機会に日本にターミノロジー活動のノードをつくり、Infotermなどの国際的活動への参画を強化してほしいとの要望が出されました。

このような情勢に対処するため、専門用語に関心を持つ人達の情報交換と相互協力の場として、「専門用語研究会」を1988年5月に設立しました。

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いたします。

活動内容

1. シンポジウム、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『専門用語研究』および『ニュースレター』を年数回発行します。
3. Infoterm、ISO/TC37などとの連絡を密にし、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す。

運営

役員として、会長、理事、監事を置きます。当面は、理事が世話役となって運営いたします。企画委員会、編集委員会、国際委員会が活動の主体となっております。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

編集後記

◆第10号をお届けします。座談会のまとめに手間取り、年を越してしまい、発行が遅れましたことをおわびいたします。また、座談会のご報告を、ページ数の関係で2回にわけざるを得ませんでした。次号をできるだけ早い時期に発行すべく努力いたします。次号には、昨年12月のシンポジウム発表内容も掲載の予定です。

◆昨年8月にアンケートをとらせていただきました会員の意識調査報告は、今回は、アンケート結果だけのご紹介とさせていただきます。次号または、近いニュースレターで、この結果に対する会員皆様からのご意見・感想などをご紹介の予定です。今後の会の活動に反映させたく、本号報告をお読みのうえ、ぜひご意見をお寄せいただければ幸いです。
(戸塚)

◆昨年、青戸先生を囲む座談会は、青戸先生の論文には書かれていない、普段なかなか聞くことのできない話を直接うかがうことができ、大変嬉しい企画でした。しかし、自分なりに整理して質問したつもりでも、文字になると支離滅裂で、普段こんな話し方をしている、周りの人は本当に理解してくれているのか心配になります。その場に居合わせた人にわけのわからない発言を理解させてしまう、文字にならない周囲の雰囲気的重要さに、校正をしながら妙に感心してしまいました。
(山下)

◆字幕を耳の悪い方々のために書いてますと、読みやすく、わかりやすく、おもしろく、かつ内容が豊富な表記表現の文とは、と常に考えさせられます。手話にも、伝統型と日本語対応型とがあり、勉強に限りがありません。二つの型の発生発展の経過は、言語体系としても興味を持てます。一方、電子情報通信学会の用語辞典の編集が本格化する年でもあり、語の選択、説明文の表記は、有効性商品価値を左右するので、改めて、言葉の難しさ、面白さを痛感し、同時に英語との対応の問題にも直面しています。すべては、自己の個の表現でもあるので、客観性の点で、大変責任を感じています。専門用語研究会の重要さを認識し、発展を祈り、編集に努力する決意の新年です。

(中山)

編集委員会委員

太田 泰弘 文教大学
四ノ宮明夫 大正製薬
戸塚 隆哉 紀伊國屋書店(委員長)
中山 亮一
牧野 正久 東京理科大学
山下 泰弘 電気通信大学
山本 昭 関東短期大学

専門用語研究 第10号
(1996年1月31日発行)

発行所 専門用語研究会
〒102 東京都千代田区一番町4-6
一番町中央ビル2F
日本総合研究所(JIST)内
Tel. 03-3262-8956
Fax. 03-3262-8960