地学用かな漢字変換辞書の作成

佐藤岱生・村田泰章・青木光子（地質情報解析室）
Taisei SATO・Yasuaki MURATA・Mitsuko AOKI

はじめに

RIPS（Research Information Processing System：工業技術院共同利用研究情報処理システム）における日本語処理の地学用かな漢字変換辞書を作成した．　その手順や経緯・考え方・問題点等を述べ　今後似たような問題が生じた時に参照していただくために　記録として残しておきたい．

昨今のコンピュータの発展は　すさまじいばかりである．　地質調査所で　コンピュータを使用した標本管理システム GEMS が動きだしたのは　つい12年ほど前の1975年である．　その時は　漢字を主体とした日本語の使用など思いも寄らないことであった．　日本語タイプライタとも言えるワープロが普及しだしたのも　ここ数年のことである．　地質調査所で最初のワープロの報告書原稿が提出されたのは1982年のことで　地域地質研究報告書：信濃池田地域の地質（1983）であろう．

かな漢字変換方式の日本語処理をしていくには　熟語とその読み方を対応させた巨大なかな漢字変換辞書があってこそ可能となる．　特に　地球科学関係では　山や川の名前等を含む地名が　基本的に重要な用語である．　標本産地　地層名　岩体名　構造帯名　異常帯名等々は地名を冠して命名される．　また　岩石名や地質現象に関する用語も一般辞書に期待することはできない．　標本管理や文献情報用のデータベースだけでなく　鉱物資源データベース・岩石分析値データベース・火山データベース・地震データベース等々の開発中あるいは開発の予想されるデータベースのためにも　地学用かな漢字変換辞書を作成することは　重要な課題であることがわかる．

地質情報解析室が　この仕事に取り掛かったのは次のような経緯による．　地質調査所では　資料室が担当して地質文献目録を年単位で発行してきた．　この地質文献目録を計算機にデータをためこむことによって作成することは　単に印刷物を作るだけでなくデータベースからキーワードによって検索が可能になるという点で非常に意義のあることである．　地質情報解析室は　1985年7月の発足早々に資料室より申し入れがあって　地質文献データベース　（GEOLIS；村田泰章　1986：日本地質文献データベース GEOLIS のシステム開発について）の構築について協議した．　入力件数・入力項目・検索のやりかた・インデックスタームの取り方などを検討するなかでかな漢字変換効率によって文献入力のスピードが大きく規制されることが認識された．　地学用かな漢字変換辞書の作成については　「全所的に必要なものだから地質情報解析室が……」ということであった．　地質情報解析室は　研究部門として発足したばかりだったし　予算も人手もなかったが　引き受けることにした．　文献入力数が多くなり　ユーザーにとって便利になることを期待したのである．

この辞書は　当面は RIPS の日本語機能の中で使われ次のような用途に必要である．

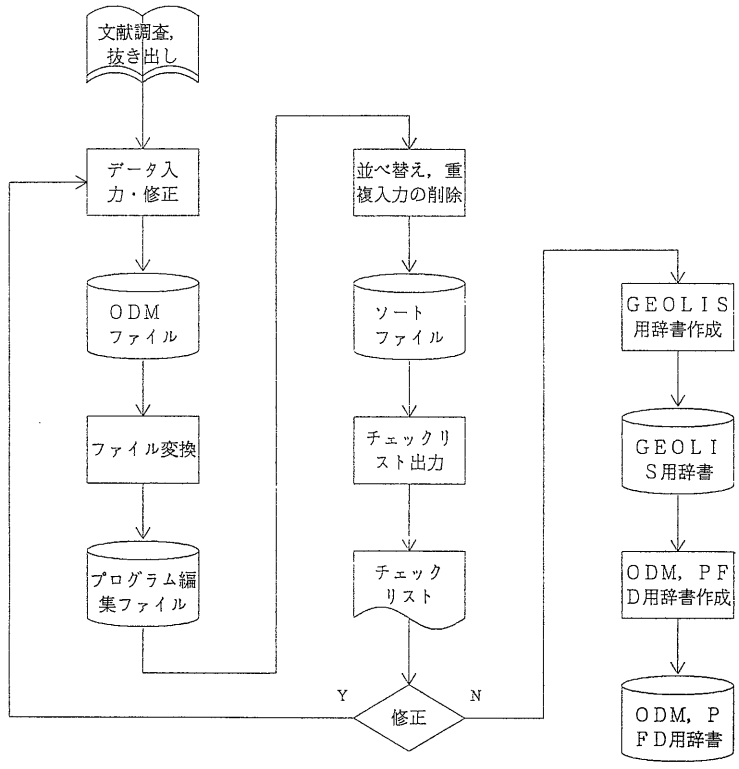
地球科学関係の文献　すなわち GEOLIS 入力
地質標本データベースへの標本産地等の入力
研究論文や地質図幅作成等の ODM 機能使用
PFD を利用した日本語データ作成

本地学用かな漢字変換辞書の開発は　富士通の機器を使用したので JEF コードで行われた．　しかし JEF 拡張漢字を除けば JIS コードへの変換も容易なので　将来はパソコンやワープロでの地学用かな漢字変換辞書としての用途も開けるかもしれない．

RIPS における日本語処理

RIPS の富士通M-380を使用した日本語機能としては種々のものがある．　日本語の研究論文を書くなどのためのワープロ機能を持つ ODM（Office Document Manager：文書処理システム）があるほか　日本語データの編集のために日本語PFD（Programming Facility for Display users）がある．　この PFD で　データベースに登録するためのデータを作成することもできる．　また　ユーザプログラムと日本語辞書とのインタフェースも用意されていて　GEOLIS（日本地質文献データベースシステム）等で使用されている．

これらの日本語処理におけるかな漢字変換や文法解析のために　RIPS にはいくつかの辞書が登録されている．



第1図　地学用かな漢字変換辞書作成の流れ

例えばODM・PFDで使用する一般用語辞書　RIPS専用のプログラムの上でのみ使用できる氏名・住所変換辞書　そして漢字一文字の読み方を記憶した辞書などである．　その辞書一つ一つは数万語という内容の膨大なものであるが　一般的な用語が多く　専門用語については私用辞書を作成して各自で対応するという環境にある．専門用語を多用する文書作成には私用辞書あるいは専門用語辞書が不可欠である．

RIPS では　都道府県名・郡名・市町村名までの地名はかな漢字変換されるが　地学関係で頻繁に使われる大字小字名や　山川の名前は辞書に入っていない．

作業の流れ

作業はまず　かなと漢字が対応した適当な文献を探すところから始まった（第1図）．　文献名や用語の選択については　後で述べる．

データの入力作業は　FACOM6658（WDS）をRIPSの端末としてではなくローカルのワープロとして使用した．　かなと対応する漢字を一対にした一覧表をワープロのファイルとして幾つか作成する．

こうして作成したファイルを　RIPSのODMのファイルに転送し　結合してひとつのファイルとする．　このODMファイルを　プログラムで編集可能なプログラム編集ファイルに変換し　さらに読みの順番にソートして重複しているものを削除してソートファイルを作成する．　このファイルからチェックリストを出力して　データのチェックと修正を行った．　修正は元のODMファイルに対して行う．　修正がすんだら再びファイル変換をしてチェックリストを出力して修正すべき所が無いことを確かめる．次に　ソートファイルをODM・PFDの私用辞書として使用するODM・PFD用辞書とGEOLIS用辞書の2つの形式の辞書に変換する．

この方式で分県地図索引から山川湖名と　平凡社地学事典から地質用語の入力をおこなった．　この　約5000語を入力した辞書が出来たところで　地質調査所のおもだったところへ案内を出した．　1986年1月末頃で　開始から約3カ月かかっている．　時間は　ほとんどが文献からの用語の抜出しと入力にかかっており　プログラムの作成に数日を要したほかは　ファイルの変換は1―2時間で終了する．

この段階で　種々のアドバイスによって文部省制定の「学術用語集」からの入力の必要性を感じたとともに日常使っていながら一般辞書には出てこない単語がたくさんあることに気が付いた．　たとえば　「三鉱学会・地質図幅・演旨」のたぐいである．　このような用語は

ワープロ等の私用辞書から拾うことにした.

再び　地学用かな漢字変換辞書作成の流れ（第1図）の最初に戻る.　次に入力した古今書院地学辞典全3巻以降は　ODMの辞書参照コマンド「SHOW」を用いて既入力チェックを行いながらファイルへの書込みを行った.　したがってデータ入力作業はODMのワープロ機能を使用して行われた.

人名については　姓名とその読みが対になり　しかも重複が無いような適当な文献がなかった.　また　学会名簿等では重複が多く　大量のデータになることが予想された.　そこで　工技院筑波研究センターの電話番号簿用のファイルから　姓と名前をそれぞれ一語とし　編集して重複をはぶいた.

1986年11月に　合計25,000語の入力を終わったので第2図の様な案内を所内各部とおもだった使用者に出した.　1987年2月には　国土地理院発行の「図名索引」からの入力も終了した.　蛇足ではあるが　ODMおよびPFDでの本かな漢字変換辞書の使用方法を第3図および第4図に示す.

準拠文献

辞書には　どのような用語が必要かを　広く衆知を集めて検討を行うほうが良かったのであろうが　そのような時間が無かった.　GEOLISの入力開始に間に合せるために　緊急に作成する必要があった.

始めに用語を選ぶ元となった文献を　入力に使用した順に挙げておく.

分県地図索引（人文社　昭和60年度版）
平凡社地学事典（増補改訂版）
古今書院地学辞典（全3巻）
文部省学術用語集地学編（丸善）
ワープロの私用辞書
工技院筑波研究センター電話番号簿
（昭和61年1月版）
図名索引（国土地理院　昭和52・56年度版）

などである.　ただし　工技院筑波研究センター電話番号簿は　RIPS上のデータファイルを編集して使用したので　従来の印刷物として連想される文献とは少々異なるかもしれない.

地名について

すでに述べたように　地球科学では地名が

1986. 11. 14.
地質情報解析室

PFD／ODM使用の皆さんへ

地学用語辞書の試験公開について

日本語処理のために，昨年から試験的に公開しておりました地学用語辞書を５０００語から約２５０００語に拡充しました．内訳は，地球科学用語約１１０００語，山川名約８５００語，地名約５５００語です．下記によりご利用ください．

地学用語辞書ファイル名
'G0216.CHIGAKU.JISHO'
PFDの場合
日本語ＥＤＩＴに入り，コマンド行でＡＬＴ⇩
出てきたメニューで，上記辞書名を指定する．ＰＦ３（ＥＮＤ）でメニューを抜け，ＥＤＩＴ画面に戻る．
ＯＤＭの場合
ＥＤＩＴ画面直前のメニューで，私用辞書名の所に上記辞書ファイル名を指定する．

この辞書は，資料室で入力中の地学文献データベース（ＧＥＯＬＩＳ）にも使用されております．必要な熟語やお気付きの点は，情報解析室　☎（３６４２）　佐藤岱生まで　ご連絡ください．一応の締め切りを１２月１５日とさせていただきます．

第2図　1986年11月中旬の案内
入力開始から約1年で登録語数は約25,000語である.　ファイル名とPFD・ODMでの辞書指定方法を示している.　入力開始3ヵ月目の最初に作成した辞書に比べると登録語数は約5倍に増えた.

```
--------------------------------<  辞書メニュー  >------
コマンド ===>
辞書のデータセット名を指定して下さい。

単語辞書
データセット名      ===> 'G0216.CHIGAKU.JISHO'
データセット名      ===>
標準辞書を連結      ===> 1          (1-する,2-しない)
漢字辞書
データセット名      ===>
データセット名      ===>
標準辞書を連結      ===> 1          (1-する,2-しない)
```

第3図　PFDでの辞書指定メニュー
データセット名の所に第2図の辞書ファイル名を書く.

非常に重要である．　地層名や岩体名は　地名を頭に付けて命名される．　また　鉱物や化石の産地は　地名が最も重要な情報の一つである．　ここでは　分県地図から自然地名を中心に入力した．

分県地図索引には　自然地名として山岳・峠・山脈・山地・高地・高原・台地・丘陵・平野・盆地・渓谷・峡谷・瀑布・河川・湖沼・半島・岬・鼻・砂浜・砂丘・松原・海岸・湾・港・島・諸島・海・海峡・水道・岩礁・浅瀬・岩のかなと漢字が　県図の位置とともに記載されている．　分県地図には人文社のものと国際地学協会のものがあるが　人文社の方が収録数が多く　「山」や「川」の字にもかなが付いていて都合が良かった．「山」は　ヤマ・サン・ザン・セン等　「川」はカワ・ガワ等と読まれるので　そこまでのかなが必要であるが多くの地図索引ではこの部分を省略している．

用語としては　山川等の付いたものと山川等を付けないものの両方を収録した（第5図）．　「山・川・岳・湖・岬・湾」等の付かない語は　町村や字を表わす行政地名になると考えた．　また　「十和田カルデラ」のように　地学用語として「湖」等を略して使われる場合も多い．　「甲武信ケ岳」「白糸ノ滝」のような場合は「ケ岳」「ノ滝」を削除した．　削除した結果　漢字1文字になるもの　および普通名詞になるものは収録しなかった．槍ケ岳・燕岳・監獄岩・二本岩等がその例である．

分県地図索引に記載されている文字やかなに疑問がある時は　地図の上で場所や文字を確かめ　他の地名辞典等の文献にあたる．　索引と地図上の文字が異なる場合は索引のミスであると考えた．　最終的には5万分の1地形図のふりがなで確かめた例もある．　このやり方では　疑問に思わない物はチェックできない．　特に　川が「カワ」と読まれるのか「ガワ」なのか　「山」が

```
ﾃﾝｲｾｲ                    転位性
ﾃﾝｲｿｳｼｮｳ                転位双晶
ﾃﾝｲﾃﾝ                    転位点
ﾃﾝｲﾃﾝ                    転移点
ﾃﾝｲﾒﾝ                    転位面
ﾃﾝｶｲ                      展開
ﾃﾝｶｲｽﾞ                    展開図
ﾃﾝｶｲﾃﾝ                    転回点
ﾃﾝｶﾐｬｸ                    添加脈
ﾃﾝｶﾝ                      塡間
ﾃﾝｶﾝｼﾞｮｳ                  塡間状
ﾃﾝｶﾝｼﾞｮｳ                  塡間状
ﾃﾝｶﾞｲｻﾞﾝ                  天蓋山
ﾃﾝｶﾞﾓﾘ                    天ケ森
ﾃﾝｶﾞﾝｶﾞﾜ                  天顔川
ﾃﾝｷ                       天気
ﾃﾝｷｭｳ                     天泣
ﾃﾝｷｭｳ                     天球
ﾃﾝｷｮｳｷﾞ                   転鏡儀
ﾃﾝｷｮｸﾀﾞｲ                  点極大
ﾃﾝｷｶｲ                     天気界
ﾃﾝｷｽﾞ                     天気図
ﾃﾝｷﾞﾝ                     天銀
ﾃﾝｷﾞﾝｻﾞﾝ                  天銀山
ﾃﾝｸﾞ                      天狗
ﾃﾝｸﾞｲｼ                    天狗石
ﾃﾝｸﾞｲｼﾔﾏ                  天狗石山
ﾃﾝｸﾞｳｼﾞ                   天狗寺
ﾃﾝｸﾞｳｼﾞｻﾝ                 天狗寺山
ﾃﾝｸﾞｶﾞﾓﾘ                  天狗森
ﾃﾝｸﾞｸﾗ                    天狗倉
ﾃﾝｸﾞｸﾗﾔﾏ                  天狗倉山
ﾃﾝｸﾞｽﾓｳﾄﾘ                 天狗角力取
ﾃﾝｸﾞｽﾓｳﾄﾘﾔﾏ               天狗角力取山
ﾃﾝｸﾞｾｷ                    天狗石
ﾃﾝｸﾞｾｷｻﾞﾝ                 天狗石山
ﾃﾝｸﾞﾀﾞｲﾗ                  天狗平
ﾃﾝｸﾞﾀﾞｹ                   天狗岳
ﾃﾝｸﾞﾀﾞﾅ                   天狗棚
ﾃﾝｸﾞﾀﾞﾅﾔﾏ                 天狗棚山
ﾃﾝｸﾞﾊﾗ                    天狗原
ﾃﾝｸﾞﾊﾗﾔﾏ                  天狗原山
ﾃﾝｸﾞﾊﾅ                    天狗鼻
ﾃﾝｸﾞﾊﾞﾗ                   天狗原
ﾃﾝｸﾞﾊﾞﾗﾔﾏ                 天狗原山
ﾃﾝｸﾞﾓﾘ                    天狗森
```

第5図　地学用かな漢字変換辞書ファイルの一部

ここでは地名が多い．　山・森・川等を付けたものと付けないものの両方を登録している．　「テングイシヤマ・テングセキザン」「テングハラヤマ・テングバラヤマ」等の読みの違いも登録しなければならない．

----------<　環境定義　>----------

```
辞書名      単語（基本）  ===> 'PP1.FDMS.JDICTLIB'
              パスワード  ===>

            単語（私用）  ===> 'GO216.CHIGAKU.JISHO'
              パスワード  ===>

            漢字          ===> 'PP1.FDMS.KDICTLIB'
              パスワード  ===>

テキストマクロ
            ライブラリ名  ===> FDMSSAVE.MACRO
            メンバ名      ===>
              パスワード  ===>

文書データ
    自動退避機能          ===> 1          (1:動作 2:停止)
    退避ファイル名        ===> FDMSSAVE.ALL
    表示形式              ===> 2          (1:編集 2:清書)
    ページ形式            ===> 2          (1:指定あり 2:指定なし)
```

第4図　ODMでの辞書指定メニュー

単語（私用）辞書の欄に第2図の辞書ファイル名を書く．

「ヤマ・サン・ザン・セン」のいずれなのかは　あまり正確ではない.

後に　国土地理院発行の「標準地名集（自然地名）」の存在を知ったが　再チェックはしなかった.　しかし一部比較してみたところでは　分県地図索引のミスと思われる読みがしばしば見られた.　標準地名集は　国土地理院と海上保安庁水路部の間で　地形図や海図で使用する地名を統一するために協議して合意されたものが記載されている.

国土地理院では　5万分の1地形図等の作成のために地名の読みを地元の自治体に問い合わせているが　地元でも呼び方が一定しておらず　変化することがしばしばあるという.　漢字に引かれて呼び方が変わる例は　札幌市の「月寒」が「ツキサップ」から「ツキサム」に変わった事や　白馬村（ハクバムラ）に引かれて　白馬岳（シロウマダケ）まで「ハクバ」と呼ぶ人が増えている事を見るまでもない.

ここに挙げたような例は　同じ漢字列に対して複数のかなの読みを入れておけば良いのかもしれない.　そうすれば　かな漢字変換用辞書としての機能は　より広く果たせる.　しかし　それでは　誤解にもとずく読み方を訂正する機会も失うことになりはしないか.　大いに悩むところである.　もっとも　地名について責任を持っているのは国土地理院で　自然地名　行政地名に分けて計算機ファイルを作成して作業をすすめている.

国土地理院発行の地形図類の図名は　適当な地名から名前を付けている.　しかし　その名前には字名やあまり高くない山名を採用している場合がある.　そこで国土地理院発行の図名索引からも　2.5万分の1・5万分の1・20万分の1の図名を入力することとした.　これまで入力した　山川名や　RIPS　の市町村名とは20—25%が重複しない.

地球科学用語

地球科学用語は　平凡社地学事典・古今書院地学辞典・文部省学術用語集地学編の順に入力したが　用語の選び方は一様ではない.

最初に入力した平凡社地学事典では　用語をできるだけ短い単位で選んだ.　すなわち　1）地名　地層名は採用しない.　2）鉱物名　鉱石名は重要度の高いものに限定する.　3）地球科学用語は　原則として漢字2—3文字の熟語を採用する.　などの基準で選択した.　たとえば　「地層・地層圧・地層学・地層面」は選択したが　「地層傾斜検層・地層水・地層単位・地層同定・地層累重・地層累重の法則」は　この段階では選択していない（第6図）.

ﾁｽﾞ	地図
ﾁｽﾞｶﾞｸ	地図学
ﾁｽﾞﾄｳｴｲﾎｳ	地図投影法
ﾁｾｲｾﾝ	地性線
ﾁｾｷｽﾞ	地籍図
ﾁｾﾈｼﾘ	地勢根尻
ﾁｾﾈｼﾘｻﾞﾝ	地勢根尻山
ﾁｿｳ	地層
ﾁｿｳ	地槽
ﾁｿｳｱﾂ	地層圧
ﾁｿｳｶﾞｸ	地層学
ﾁｿｳｹｲｼｬｹﾝｿｳ	地層傾斜検層
ﾁｿｳｽｲ	地層水
ﾁｿｳﾀﾝｲ	地層単位
ﾁｿｳﾄﾞｳﾃｲ	地層同定
ﾁｿｳﾒﾝ	地層面
ﾁｿｳﾙｲｼﾞｭｳ	地層累重
ﾁｿｳﾙｲｼﾞｭｳﾉﾎｳｿｸ	地層累重の法則
ﾁﾀｲｺｳｿﾞｳ	地体構造
ﾁﾀｲｺｳｿﾞｳﾁｼﾂｶﾞｸ	地体構造地質学
ﾁﾀｲｺｳｿﾞｳﾛﾝ	地体構造論
ﾁﾀｲﾘｮｸ	地耐力
ﾁﾀﾊﾝﾄｳ	知多半島
ﾁﾀﾜﾝ	知多湾
ﾁﾀﾞｲ	地台
ﾁﾀﾞﾝｹﾝ	地団研
ﾁﾁｭｳｵﾝﾄﾞ	地中温度
ﾁﾁｭｳｶｲ	地中海
ﾁﾁｭｳｶｲｼｷ	地中海式
ﾁﾁｭｳｽｲ	地中水
ﾁﾁｭｳﾘｮｳｲｷ	地中領域
ﾁﾁｮｳ	地調
ﾁﾁｮｳｹﾞｯﾎﾟｳ	地調月報
ﾁﾁｲ	知々井
ﾁﾁｲﾐｻｷ	知々井岬
ﾁﾁｵﾆ	父鬼
ﾁﾁｵﾆｶﾞﾜ	父鬼川
ﾁﾁｶﾞﾀﾆ	父ケ谷
ﾁﾁｼﾞﾏ	父島
ﾁﾁｼﾞﾏﾚｯﾄｳ	父島列島
ﾁﾁﾌﾞｶﾞｳﾗ	秩父ケ浦

第6図　地学用かな漢字変換辞書ファイルの一部
「地層」の前後.「地層累重の法則」のような長い熟語や「地調・地調月報」のような用語も登録した.

古今書院の地学辞典全3巻では　長い熟語を主体として入力した.　地層名や「塩原動物群・佐川造山運動」のような固有名詞を冠した熟語も入力した.　地層名・火成岩の岩体名等の用語は　後に平凡社地学事典からも入力した.

選択した熟語は事典に直接マークを付け　マークに従って入力する.　この場合　辞書順に並んだ言葉の頭のほうからだけ採用するならば問題は無いのだが　頭に無い熟語も採用しなければならない.　たとえば　「赤石裂線」「有明花粉群」という用語があった場合　「裂線」「花粉群」だけでは　事典に記載がない.　「累重・岩体・軽石層」等も見出し語としては無い.　しかし　これらの熟語もかな漢字変換には必要である.　このように　複合語の二番目以降の辞書順に並んでいない用語を採用した場合は　重複をチェックする必要がある.

入力作業は　上記のようにして選択された用語とその

ページ	本文	訂正
5	圧力形受信器	圧力型受信器
22	外型雌型	外形雌型
27	genti-taisaisekibutu	genti-taisekibutu
33	hanō-kankei	hannō-kankei
56	kai[sai]	kai[sei]
57	kaitei-taisekiati	kaitei-taisekitai
58	kakeki-gun[syū]	kaseki-gun[syū]
64	kasieigan	kaseigan
69	keisitu-keugan	keisitu-ketugan
70	keityōsekisitu	keityōsitu
71	Kenyrū-rui	Kenryū-rui
88	kutyū-hōsyanō-tankō	kūtyū-hōsyanō-tankō
94	menso-kōsi	mensin-kōsi
112	沪過器　沪がＪＩＳコードに無い.	
113	沪速　沪がＪＩＳコードに無い.	
117	ryūsutu-dosyaryō	ryūsyutu-dosyaryō
118	saikkesyō[-sayō]	saikessyō[-sayō]
119	Saiyū-rui	Sairyū-rui
123	青石錦	青石綿
135	自然砂沪井　沪がＪＩＳコードに無い.	
139	層層注入	層々注入
147	syō-sekkaingan	syō-sekkaigan
155	tatuki	tauki
159	tihyō-tyoryoryū	tihyō-tyoryū
162	地質判続	地質判読
168	Tyōzigai-zui	Tyōzigai-rui
175	yonkutu-ōkō	shikutu-ōkō
176	yūkō-ōryokyu	yūkō-ōryoku
180	ziki-kiben	ziki-kiban

第７図　文部省学術用語集地学編の疑問点
これは文部省にお知らせしたリストの写である．　地学愛好者に早く情報が伝わることを願ってあえて掲載する．　日本語がかなのみで書かれた用語はチェックしていない．

読みを対応させた一覧表を　ワープロ（WDS）で作成する．　出来た一覧表は　RIPS　へ転送して辞書に変換する．　山川名と平凡社地学事典の入力が終わった段階で第１図の手順に従って辞書に変換した．　ここで辞書に変換してしまうのは　とりあえず作成した小さな辞書でその後の入力における重複をチェックするために必要だからである．　このようにすると　RIPS　のワープロ機能である ODM 中の SHOW コマンドによって辞書の内容を見ることが出来て　既に登録した熟語を再び入力しないで済む．　また　短い熟語は長い熟語の一部となっていることが多いために　より長い熟語を入力する場合に変換効率が上がることもねらいのひとつである．

古今書院の地学辞典では　当用漢字に無い漢字はかな書きするという原則がある．　また　鉱物名はカタカナで書いている．　しかし　このかな漢字変換辞書ではカナはほとんど漢字に置き換えた．　たとえば　カクセン岩相を角閃岩相　カコウ岩は花崗岩　角レキ岩は角礫岩とした．

文部省学術用語集地学編は　「複雑・難解な学術用語を整理・統一する」目的を持って　昭和24年から検討を始め昭和57年に制定された地学の学術用語を収録したものである．　ここでは　標準的に使用される用語として特に注意深く入力した．「花こう岩・斑れい岩・隆起さんご礁」等のかな文字を含む熟語はもらさず入力した．したがって　このかな漢字変換辞書には「花崗岩」も「花こう岩」も登録されている．

私達が普通に使っている言葉で　標準的な辞書にないものがかなりあると考えられる．　前述の「三鉱学会・地質図幅・演旨」のほかにも「岩鉱学会・地調月報・暗灰色・淡黄色・査読」等々限りがない．　これらの用語は　良く使いこなされているワープロの私用辞書から選択するのが最も良いと考えられた．　そこで　曽屋龍典氏の ODM 私用辞書　および地質部のワープロ私用辞書から用語を選択させていただいた．

人名について

人名は　前述のように適当な文献が見当たらない．一時は　地質学会年会のプログラムから講演者をリストアップしてみたが　読みを付けるのもたいへんな作業で断念せざるを得なかった．　次に思い付いたのが　工技院筑波研究センターの電話帳で　RIPS　上のファイルになっているものを編集して使用することにした．　名前の姓と名を切り離して　JEF コード順に一列に並べて重複を取る．　これで　何人もいた佐藤さんや鈴木さんが　１個の「佐藤・鈴木」という漢字の熟語になる．これに適当に読みがなを付けたが　どうしても分からない読み方十数件については　各所の庶務係に問い合わせた．　地球科学のうえで有名な人名を選択すべきかもしれないが　現時点ではむずかしい．　上記電話帳ファイルによって地質調査所の本所の職員が1985年の時点で網羅されているにすぎない．

原典の疑問点

このかな漢字変換用地学辞書は　過度の厳密性は追及していない．　すなわち　常識的であると言うことで　前述のように名前の読み方も常識に従って付けたにすぎない．　しかし　刊行されている辞典類のことになると針の先ほどのことでも取り上げたくなる．

第8図
「地名地学用語読み方辞典」の一部
地学用かな漢字変換辞書に登録された熟語の漢字からその読みを探す逆引き辞典である．　冊子体で利用する．

凡　凵　刀

14	鳳来湖	ホウライコ
	鳳凰山	ホウオウサﾝ
	鳳凰山花崗岩	ホウオウサﾝカコウカﾞﾝ
	鳳凰山阿部川断層	ホウオウサﾝアベﾙカﾞワタﾞﾝソウ
	凵　部	
5	凹入角	オウニュウカク
	凹地	オウチ
	凹部	オウフﾞ
	出戸	テﾞト
	出戸川	テﾞトﾞカﾞワ
	出水山地	イスﾞミサンチ
	出石川	イスﾞシカﾞワ
	出羽	イスﾞハ
	出羽	テﾞワ
	出羽山地	テﾞワサンチ
	出羽川	イスﾞハカﾞワ
	出来山	テﾞキサﾞﾝ
	出屋敷	テﾞヤシキ
	出屋敷峠	テﾞヤシキトウケﾞ
	出砂	イテﾞスナ
	出砂島	イテﾞスナシﾞマ
	出前沢	テﾞマエサワ
	出島	テﾞシﾞマ
	出島	イスﾞシマ
	出雲岳	イスﾞモタﾞケ
	出雲盆地	イスﾞモホﾞンチ
	出雲層群	イスﾞモソウクﾞﾝ
	出森	イテﾞモリ
	出森山	イテﾞモリヤマ
	凸凹割れ口	テﾞコホﾞコワレクチ
	凸斜面	トツシャメン
	凸部	トツフﾞ
8	函岳	ハコタﾞケ
	函渕層群	ハコフﾞチソウクﾞﾝ
	函館港	ハコタﾞテコウ
	函館湾	ハコタﾞテワン
	刀　部	
2	刀形類	トウケイルイ
3	刃切	ツルキリ
	刃切山	ツルキリヤマ
	刃石器	シﾞンセッキ
	刃状転位	ハシﾞョウテンイ
	刃物ケ崎	ハモンカﾞサキ
	刃物ケ崎山	ハモンカﾞサキヤマ
4	刈込	カリコミ
	刈込湖	カリコミコ
	刈田	カッタ
	刈田	カリタ
	刈田岳	カリタタﾞケ
	刈田岳	カッタタﾞケ
	刈安	カリヤス
	刈安山	カリヤスヤマ
	刈屋	カリヤ
	刈屋川	カリヤカﾞワ
	刈寄	カリヨセ
	刈寄山	カリヨセヤマ
	刈場	カリハﾞ
	刈場山	カリハﾞヤマ
	切甲類	セッコウルイ
	切込	キリコミ
	切込砂利	キリコミシﾞャリ
	切込炭	キリコミタン
	切込湖	キリコミコ
	切出し	キリタﾞシ
	切目	キリメ
	切目川	キリメカﾞワ
	切目崎	キリメサﾞキ
	切羽	キリハ
	切取り	キリトリ
	切取り斜面	キリトリシャメン
	切面	セツメン
	切峰面	セッホﾟウメン
	切峰面図	セッホﾟウメンスﾞ
	切断	セツタﾞﾝ
	切断山脚	セツタﾞンサンキャク
	切雲	キリモ
	切雲谷	キリモタﾞニ
	切頭	セットウ
	分子	フﾞンシ
	分子化合物	フﾞンシカコﾞウフﾞツ
	分子容	フﾞンシヨウ
	分子結合	フﾞンシケツコﾞウ
	分子磁化率	フﾞンシジﾞカリツ
	分子熱	フﾞンシネツ
	分化	フﾞンカ
	分化作用	フﾞンカサヨウ
	分化指数	フﾞンカシスウ
	分化鉱床	フﾞンカコウショウ
	分化説	フﾞンカセツ
	分水界	フﾞンスイカイ
	分圧	フﾞンアツ
	分布	フﾞンフﾟ
	分布様式	フﾞンフﾟヨウシキ
	分光	フﾞンコウ
	分光分析	フﾞンコウフﾞンセキ
	分光光度計	フﾞンコウコウトﾞケイ
	分光器	フﾞンコウキ
	分岐	フﾞンキ
	分岐砂嘴	フﾞンキサシ
	分岐脈	フﾞンキミャク
	分岐進化	フﾞンキシンカ
	分別	フﾞンベツ
	分別作用	フﾞンベツサヨウ
	分別結晶	フﾞンベツケッショウ
	分別結晶作用	フﾞンベツケッショウサヨウ
	分別晶出作用	フﾞンベツショウシュツサヨウ

32

分県地図の自然地名索引については　すでに述べたので省略するとして平凡社の地学事典には気が付くようなミスは見付けられなかった．　ただ「熔岩・熔結凝灰岩」のように「熔」の字を使用している．　ここでは「溶岩・溶結凝灰岩」のように「溶」の字を使うものと両方を登録した．

古今書院の地学辞典には　ごくわずかではあるがミスプリントと思われるものがある．　例えば第Ⅱ巻237ページの「自動酸化」が「自動酵化」に　394ページの「同生団塊」が「同生団魂」になっている．　また「しんせいこうしよう」の様に「ょ」が「よ」になっているものが幾つかある．

文部省学術用語集地学編で気付いた疑問点は　第7図のとおりである．　この図は　すでに丸善経由で文部省にお知らせしたものの控えで　文部省がしかるべき処置を行うものと思う．　にもかかわらず　あえてここに紹介したのは　地学に興味を持っている方々にできるだけ早く情報をお伝えしたいからである．

この学術用語集は　読み方が訓令式のローマ字で示されているので　ミスプリントを発見しにくく　校正ミスと思われる．　また　和英が第1部　英和が第2部となっているが　ミスは両方の対応する所でまったく同じである．　これは　活字を拾った時に生じるミスではなくおそらくどちらか一方を電算機に入力して　他方を計算機処理によって作成したものと考えられる．　はじめに一方を作成した時のミスが　他方に引き継がれたものと想像される．　とはいえ　このような電算機処理の試みは　今後ますます推し進められるべきであろう．

漢字 JIS コードについて

本州最東端の岬はどこですか？　小中学生むけのクイズだが　答えは岩手県宮古市の「魹ヶ崎（とどがさき）」である．　このことは　広辞苑にも記載されているが　この「魹」の字が JIS コードに無い．

文部省学術用語集では　当用漢字以外の漢字の使用をごくわずかに限って認めている．　その中に「沪」の字が含まれており　「沪過器・沪速・自然砂沪井」に使われている．　ただし　文部省学術用語集の化学編では「ろ過〔沪過〕」のように書かれている．

JIS コードあるいは　その富士通企画版である JEF コードには「炉・爐・濾」とも在るのに「沪」はないが「爐」を「炉」と書くなら「濾」を「沪」と書いても良いような気がする．　ちなみに　「爐」の字の「盧」は「ろ」という音を表わしているが　「戸」には「ろ」の音はなく単なる記号としての役割しか無いようにおもう．　「炉」は JIS 第１水準であり　JIS 第２水準には「枦・鑪・鑪」の文字がある．　どのような理由で「沪」が採用されなかったのであろうか．

もっと困ったことは　鉱山や鉱床学で使われる「鎚（ひ）」と鉱物学で使われる「槌（せつ）」の字が無いことである．　「鎚」は「鎚押・鎚先・鎚の内・鎚肌」等に「槌」は「槌石・槌面・両槌体」のように使われる．

JIS を決めた国語学者の視野が狭いのか　変な文字を使う地質屋がへそ曲りなのか　いずれにしても興味深いことである．

また　斑糲岩の糲の字も　「栃・励」が許されるならば「粟」と書いて良いような気がする．　漢字は　画数の少ない方が覚え易く　書き易く　読み易い．　さらに活字や計算機のプリンタ用のフォントも作り易い．

JIS コードは　なぜ画数の多い字が多くて　画数の少ない文字が少ないのだろうか．　不思議に思う．　「花崗岩」の「崗」も意味が同じで発音も同じなのだから「岡」で良いのではないかと言う人もいる．　画数の少ない方が何かにつけて良いと思うのだが　「花岡岩」では感じが出ないだろうか．

日本語に起因する問題

日常生活で遭遇する熟語で読み方の分からないものはそれほど多くはないだろう．　また　辞書を手に入れることもむずかしくはない．　ところが　地名の類はそうではなく　同じ漢字がまったく予想もつかない読み方をされることもまれではない．

本地学用かな漢字変換辞書は　熟語の読み方が分からなければ使いようがない．　使う人の思い浮かぶ読み方で　たまたま目的の漢字に変換されたとき　それが正しい読み方だと誤解される可能性は高いであろう．　また正しい読みが入っていたとしても間違いは起こり得る．本辞書は　このような誤解から生じたり　辞書の間違いによる日本語の変化に対して責任を負うものではない．使う人の責任にかかわる部分の方が大きいように思われる．　しかし　本かな漢字変換辞書を使用して間違いを見付られた方は　すぐにお知らせ頂ければ幸いである．

また　漢字から熟語の読みを探せるような「地名地学用語読み方辞典」を作成した（地質調査所研究資料集登録 No. 31　第８図）．これは　地学用かな漢字変換辞書のために作成したファイルの並べ替えを行って　熟語を漢字の部首および画数順に配列して　漢字からその読みを引くようにしたものである．　同じ漢字列に対して複数の読み方があるかどうか確かめることができる．　地名・人名等の　難読熟語の読みを探すのにも有効であろう．GEOLIS 入力中に地名の読み方が分からないので苦労することが多いのである．　また　冊子体となっているので一覧性があり　この地学用かな漢字変換辞書にどのような熟語が登録されているかを読んで確かめることもできるだろう．　熟語をどの程度の長さで切ればよいのかも分かって　漢字変換効率を上げることもできよう．

ちなみに　本地学用かな漢字変換辞書に登録されている最も長い熟語は「八方尾根超塩基性岩体」で　漢字10文字である．

おわりに

地学用かな漢字変換辞書を作成した途中経過とその間に感じた問題点を書いた．　この辞書に対して　お気づきの点や要望などがあったらどんどん知らせていただきたい．　実現の努力を惜しまない．　また　当面 GEOLIS に組込んで使用するほか　PFD・ODM の私用辞書として RIPS の利用者は 断わりなしに 使用することができる．　また　ODM では私用辞書を一つしか使えないので　時期をみて RIPS の標準辞書に組み込む方策を考えたい．

謝辞：この辞書の作成にあたって　環境地質部の曽屋龍典課長・地質部山田直利課長ほかから有益なアドバイスをいただいた．　データ入力は青木の他に　鈴木みゆきさん・広瀬真知子さんの手をわずらわせた．　地質情報解析室の花岡尚之室長および野呂春文氏には　辞書作成過程で種々の協力を頂いたが　本報告作成にあたっても粗稿を読んで有益なアドバイスをいただいた．　以上の方々に深く感謝いたします．