

# 国立国語研究所学術情報リポジトリ

## 高校教科書 文脈付き用語索引 解説書

メタデータ	言語: Japanese 出版者: 公開日: 2021-02-26 キーワード (Ja): キーワード (En): 作成者: 国立国語研究所, The National Language Research Institute メールアドレス: 所属:
URL	<a href="https://doi.org/10.15084/00003136">https://doi.org/10.15084/00003136</a>

国立国語研究所 言語処理データ集 1

高校教科書 文脈付き用語索引

解説書

国立国語研究所

1985



## 刊行のことば

これは「高校教科書の語彙調査」の3冊目の報告です。一昨年度に国立国語研究所報告76として刊行した語彙表は、ここに「言語処理データ集1」として報告する「高校教科書 文脈付き用語索引」のデータをもとに作成されたものです。

「文脈付き用語索引」は、単なる「用語索引」とは異なり、ある語がどのような文脈で使われているかを示すものです。語彙表は語の頻度・使用率などを示すものですが、この「文脈付き用語索引」によって、それぞれの語が実際にどのように用いられているかを一覧することができます。また、高校教科書に特徴的に用いられる語彙の分析だけに限らず、日本語全体の語彙・文法研究にも有用であると信じます。

「文脈付き用語索引」の欠点は、ページ数が大量になることです。このデータを本にすると9433ページとなり、1冊500ページとしても19冊になります。これでは、利用するには便利ですが、それを置く場所もとりますし、作成するための費用もたくさんかかります。この欠点を除くために作成費も安く、置く場所もそれほど取らないマイクロフィッシュによる刊行としました。マイクロフィルムとしなかったのは、「文脈付き用語索引」が最初から順に読んでいくものでなく、必要な所を飛び飛びに調べることが多いものだからです。

国立国語研究所に電子計算機を導入してから20年になります。この間に電子計算機を用いた研究の成果をいくつかの報告書や資料集として刊行してきました。しかし、それですべてであったわけではなく、各種の研究に有用な多くのデータが、それが大部になるという理由で報告できませんでした。今回、国立国語研究所の刊行書のなかに「言語処理データ集」というシリーズを新たに設け、マイクロフィッシュのようなこれまでになかった刊行の方法によって、言語研究者、文学研究者および言語情報処理にたずさわる工学研究者の多くの要請にこたえることとしました。

「高校教科書の語彙調査」を進めるにあたって、各方面から多大の御助力・御助言を賜りました。ここに厚く感謝の意を表し、今後も引き続いて御指導くださいますよう、お願い申しあげます。

なお、この調査は、言語計量研究部において、部長斎賀秀夫を中心に、第一研究室、第二研究室、第三研究室の研究員全員で推進しているものです。本データ集の作成には、主として、作業面を第一・第三研究室が担当し、解説書の執筆は、中野洋・齋岡昭夫・山崎誠が担当しました。

昭和60年3月

国立国語研究所長 野元 菊雄

## 目 次

### 刊行のことば

I	高校教科書の語彙調査の概要	1
1.	調査の目的	1
2.	調査の規模	2
3.	調査の手順	4
4.	調査単位について	5
5.	同語異語判別	13
II	文脈付き用語索引について	15
1.	データ量	15
2.	KWICとは	18
3.	文脈付き用語索引の形式	18
4.	用例の配列順序	19
5.	検索の方法	20
6.	語数を知りたい場合	20
	フィッシュ索引	21
	フレーム索引の例	22

## I 高校教科書の語彙調査の概要

### 1. 調査の目的

現代日本語の用語用字の実態を明らかにするために、国立国語研究所では、これまで、新聞、婦人雑誌、総合雑誌、雑誌九十種、新聞三紙を対象として、語彙調査を重ねてきた。とくに、昭和41年の新聞三紙を対象とした調査は、電子計算機を使用した、最初の大規模な調査であった。これらの調査のあとを受けて、国民が一般教養として、各分野の専門知識を身につける時に必要と思われる語彙の実態を明らかにすることを目的として、この高校教科書の語彙調査は企画された。高等学校進学率の増加に伴い、現今では、高等学校教育は、国民大多数の基本的な教養の場となっている。また、大学教育は、この高校教育の基盤に立って進められるものであり、とくに高校の理科と社会は、大学における専門教育の基礎となっていると考えることが出来る。われわれが高校の理科・社会科の全教科を対象とした語彙調査を企画したのは、以上のような理由からである。

なお、本調査は、単に出現した語彙の一覧表を作るのではなく、専門知識体系を記述する語彙・表記・表現の実態を把握・分析することを目標としている。とくに、知識体系の記述を分析するために、この調査では、従来のようなサンプリング法によらず、対象とする文章を限定したのち、その全文を入力するという方法を採った。このやり方では、調査対象の幅をせばめてしまい、高校教科書の全体像を記述するという点では、やや不十分な結果しか得られない面もあるが、文章解析などこれまでの語彙調査では出来なかった数々の分析と記述を可能にしたと言える。

また、この調査では、われわれは、電子計算機を十分に活用し、人力と機械力のそれぞれの長所を最大限に生かすことを心掛けた。恐らく調査の精度は非常に高くなっているだろうし、また語彙表その他の集計表は、電子計算機から出力したものをそのまま使用しているので、誤植も発生しにくいはずである。電子計算機からの出力をそのまま使用するということは、本調査のデータすべてが計算機に納められているということ意味する。今回公開する文脈付き用語索引も計算機出力の一つである。今後、各種語彙表・分析表・用例集の作成は、機械処理によって実現が可能であり、さらに、語彙調査としての成果のほかに現代日本語の分析・記述・実験が可能である。われわれはこの教科書データが日本語研究のデータベースの一部分となり得ることを心掛けた。

## 2. 調査の規模

### 2. 1 調査対象

この調査は昭和49年度に着手したものである。調査対象は、当時使用されていた高等学校教科書のうちの、社会科・理科の全教科の中から、次の9教科とした。

理科…物理I・化学I・生物I・地学I

社会…倫理社会・政治経済・日本史・世界史・地理B

当初は数学も対象とする予定であったが、予算の都合で、取り止めとなつた。どの出版社の教科書を対象とするかについては、教科書協会の協力に負うところが多い。ここに記して感謝する次第である。取り上げた教科書は次の通りである。

(教科名)	(教科書名)	(著者名)
物理 I	標準高等物理 I	大塚明郎ほか
		講談社 昭和49年1月30日発行
化学 I	化学 I	柴田雄次ほか
		大日本図書 昭和49年2月5日発行
生物 I	生物 I	石田寿老ほか
		清水書院 昭和50年2月15日三版発行
地学 I	地学 I	湊正雄ほか
		実教出版 昭和49年1月25日発行
倫理社会	倫理・社会	中村元ほか
		東京書籍 昭和49年2月10日発行
政治経済	政治経済 新訂版	辻清明ほか
		自由書房 昭和49年2月5日発行
日本史	詳説日本史(再訂版)	宝月圭吾ほか
		山川出版 昭和49年3月5日発行
世界史	三省堂新世界史	土井正興ほか
		三省堂 昭和49年3月30日再版発行
地理 B	高校新地理 B	青野壽郎ほか
		二宮書店 昭和49年1月20日発行

上記の教科書の本文部分を全て調査対象とした。本文部分には、表紙・目次・奥付・各ページのノンブル等は当然含まれないが、さらに、次のものも本文とはしなかった。

◎巻末の索引・年表等

◎図表・写真およびその周辺部分にある説明の言葉

◎脚注

さらに、本文に付随する次のものも、本文部分から除いた。

◎人名・地名の上下に付いているアルファベット表記・生没年

◎下付きルビ（外国語や外国音、また別の読みなどを示すのに使われる）

例； アジア・アフリカ会議	四川	かわせ	はくすきのえ	抵抗運動
Asian-African Conference	スーチョワン	為替	白村江	レジスタンス

（これについては「高校教科書の二重ルビ一覧－高校教科書調査雑報4」<言語計量研究部「季報」1980夏号>にまとめてある。）

このようにして規定した本文部分を、極力、忠実に、そのままの形で電子計算機に入力することを心掛けた。ただし、データ処理のつごうでやむをえず手直しをしたものがある。また、わずかだが誤植と判断して改めたところもある。（これについては「教科書本文を改めたもののリスト－高校教科書調査雑報2」<言語計量研究部「季報」1980春号>にまとめてある。）

## 2.2 調査の担当者

この調査は、特別研究「言語計量調査」として、言語計量研究部に属する三研究室が共同で行った。現在の各室の人員は次の通りである。

言語計量研究部長 斎賀秀夫

第一研究室 中野 洋 鶴岡昭夫 山崎 誠

第二研究室 野村雅昭 佐竹秀雄 石井正彦

第三研究室 斎藤秀紀 田中卓史

以上のほかに、調査の企画時より参加し、現在は他に転じたものに、石綿敏雄、田中章夫、土屋信一、米田正人がある。この研究作業を助けたのは、大滝弘美、岡田敏子、小沢（長田）厚子、小高京子、沢村都喜江、科野（白木）千夏、下山いくよ、田島（武田）道子、中俣（堀江）久美子、松浦（小原）美恵子、原田（山口）和子、米田（竹内）純子である。とくに、小高京子、沢村都喜江は、この調査の開始から今日までデータのさん孔・修正を受け持ち、また、米田（竹内）純子は、プログラミングおよびオペレート全般に、終始、力を注いだ。このほかに、アルバイター数十名が、この仕事に参加した。

なお、調査は、全員の共同討議で進められたが、特に調査単位については、石綿敏雄、齋岡昭夫、野村雅昭が分担し、機械処理システムについては、秀紀、中野洋、米田正人が分担した。土屋信一は、両者の連絡係をつとめた。この文脈付き用語索引作成までは、主として、第一・第三研究室が作業を担当した。また、解説の執筆は、中野洋、齋岡昭夫、山崎誠が担当した。

### 3. 調査の手順

ここでは、調査の作業手順についての大きい流れについてのみ述べる。単位切り作業については、4に、同語異語判別作業については、5に詳しく述べるので、参照されたい。

調査の全体の流れは、次の通りである。

#### (1) 計画の立案…調査の目的・対象・分析事項の検討と決定。

調査単位・調査システムの決定。作業分担の決定。

なお、調査単位は、文節から助辞を切り出した「W単位」とそれよりも小さく、形態素に近い「M単位」の二種類に決定。

#### (2) 資料の収集

#### (3) 入力データ作成

(ア) 作業台帳の作成、対象とする個所の指定

(イ) 文・段落等の情報の記入

(ウ) 単位切り・その検査

(エ) 清書・その検査…清書の際、読み仮名、代表形、助辞、振り仮名情報を付ける。

(オ) データさん孔（一部分は外部へ発注）

#### (4) 機械処理…データの読み込み、機械的チェック、漢字プリンター出力印字（入力データ形式）

#### (5) 検査・校正…修正データ作成、さん孔、機械処理を含む。

#### (6) 機械処理 修正検査用ミニK W I C（M単位・W単位の二種）および教科書の原文形式の出力

#### (7) ミスデータの検出と修正 修正データ作成、さん孔、修正機械処理を含む。

#### (8) 同語異語判別作業用 K W I C 作成

#### (9) 同語異語判別作業 (8) で機械的に付した判別情報を、全データについて確かめ、判別情報修正データを作成し、機械処理をする。

- (10) 付加情報つけ 語種・人名・地名等の情報、分類語彙表の意味番号、見出し表記形の選択、注記としての情報記入など。意味番号・見出し表記形つけは、機械的に付したものと点検・修正する。
  - (11) 最終ファイル作成
  - (12) 語彙表作成・印字（全体・教科別、五十音順・度数順など）
  - (13) 各種集計・分析表作成
  - (14) 文脈付き用語索引（K W I C）作成
  - (15) (12)～(14) その他を用いた分析・記述、報告書の作成・刊行
  - (16) (11) を用いた各種実験・分析

#### 4. 調査単位について

#### 4. 1 調査単位の種類と長さ

国立国語研究所では、昭和41年から行われた新聞の語彙調査に電子計算機を使用し、長短二種の調査単位を採用した。その理由は、国立国語研究所報告37「電子計算機による新聞の語彙調査」などに述べてあるとおり、電子計算機を用いて処理を行えば手作業ほど労力がかからないこと、語の構造を明らかにできて造語要素の使用状況についての資料が得られること、などである。以後の大規模な調査の多くは、複数の言語単位を用いて行われている。

今回の調査では、それらの各調査の経験をふまえ、また、高校の社会科・理科の教科書という文章の性格を考慮に入れ、新たに決めた長短2種類の単位を用いた。長い単位は、文の構成にあずかる要素（いわゆる文節）にもとづく単位で、wordの頭文字をとってW単位と名付けた。短い単位は、語の構成にあずかる要素（いわゆる最小単位）にもとづく単位で、morphemeの頭文字をとってM単位と名付けた。また、単位切りの段階では、句読点その他の記号類も、一般の言葉と同等に扱っているが、それは単位切り作業と、それに連く一連の機械処理を的確かつ迅速に行うためである。

単位切り作業の手順については「3. 調査の手順」で述べたので、ここで詳しく繰り返すことはしないが、上に示したW単位、M単位の性格から、まず教科書原本の文章をW単位で切り（“／”を入れる）、その後、W単位内をM単位に分ける（“＼”を入れる）、という方式をとった。

なお、W単位、M単位の詳細については、国立国語研究所報告78「高校教科書の語彙調査」および報告81「高校教科書の語彙調査(II)」を参照されたい。

## 4. 2 W単位の切り方（概略）

(0) 単位は“／”に始まり，“／”で終わる。

(1) 記号は1単位とする。

例；／「／不戦条約／」／（／ケロッグブリアン条約／）／

<例外・1>次の記号は無視する。

①数字連続の中に現れる小数点や位取りカンマ。

②それがないときには、全体が、以下に述べるW単位となるものの中に現れる記号。

例；／米・ソ対立／／二・二六事件／／ジョン＝F＝ケネディ／

<例外・2>文中にあって、語と同じような読み方、働きをする記号（アルファベットの略号も含む）およびその連続は、記号とはしないで、一般的の語として扱う。

例；／（　）／に／入れる／／H<sub>2</sub>O／／PTA／

2) 助辞（助詞・助動詞のうち用言の連体形・終止形、または体言につくもの）は1W単位とする。活用する助辞（上記の助動詞）に助詞「て」「で」「ば」、助動詞「う」「た」「ず」などが付いたものは、それを含めた形で1W単位とする。助辞についてはリストを作成した（次ページ）。

例；／外国／で／は／／争い／が／続いて／も／／属国／だった／

<例外・1>「～の～」の形の体言相当句で、「の」を助辞とせずに全体を1W単位としたものが、岩波国語辞典（第2版）により、「天の川」「世の中」など12語あった。

<例外・2>「～が～」の形の体言相当句は、「が」を助辞とせずに全体を1W単位とした。

<例外・3>「こ」「そ」「あ」「ど」につく「の」は助辞とせず、全体を連体詞とした。

<例外・4>人名・地名・国名およびこれらを含む体言的連結の中に、助辞リストの語があっても無視する。

例；／やまのうえの憶良／／高天が原／

<例外・5>助辞のリストにある語を含む語で、その結び付きの強いものは切らない（しかし、その多くはM単位で切る。以下の用例はほんの一部の例にすぎない）。

例；／あくまで／／さらに／／ともに／／まもなく／

(3) 助辞を伴わない文節は1単位とする。文節は、修飾、並列、接続、中止、独立、および述語の、構文機能を有するものをいう。

文節で切る例；／新しい／仏教／／天正11年／完成した／

<付>助辞のリスト（＊のついたものは活用する助辞）

か…合憲か否かの 何回か繰り返す そればかりか 答えはどれか  
 が…人口が減少する 水は通るが溶質は通りにくい  
 かな…ああ、短いかな、人の生命よ  
 から…重力は下向きであるから負の力として示せばよい  
 くらい（ぐらい）…15分くらいたったら 数万年ぐらいまあから  
 けれども…急速な発展を続けているけれども、まだ安定していない  
 こそ…程度の差こそあれ 快楽の状態こそ幸福であり  
 ごとし＊…破ったごときも 大賓を見るがごとし  
 さえ…低い段階のものさえある 正当な手段とさえ認められた  
 し…生けとし生けるもの 過去にもあったし、現在でもある  
 しか…1本しかない 少数の人しか住んでいない  
 しも…今しも戦争が起ころうとしている 必ずしも正しくはない  
 ずつ…2本ずつある 少しずつ出す  
 すら…宗教性すら含んでいる 不平等ですらある  
 ぞ…何の謂ぞやと  
 だ＊…重罪だとされた 立派な行い 主なる神 言えないだろ（う）  
 だけ…自分の都合だけで決めては 日本でだけ通用する  
 たり＊…治者たるにふさわしい 堂々たる態度  
 で…坂下門外で傷ついた 戦争で失う 旧習に固執しないで進める  
 です＊…それは私ですと答えた話 考えてよいでしょ（う）  
 でも…外の入れ物にでも移しておく だれでも考えることができる  
 と…水と食塩をまぜる 下り坂となる 日本とドイツとイタリアの三国  
 とか…水で冷やすとか火で温めるとかすると  
 とて…今さら悔やんだとて仕方がない  
 とも…是非とも実現させようと 板垣死すとも自由は死せじ  
 な…真珠湾を忘れるな  
 ながら…幼いながらしっかりした子ども わずかながら進展が見られる  
 など…イギリス・フランスなどの干渉 兵を進めるなどしたが

なり…本で調べるなり、人に聞くなりする  
 に…水をビーカーに入れる 病気になる 戰争に勝つには勝ったが  
 の…陸軍の将校 命あるものごとき 大きいのに入れる  
 ので…税率を上げたので民衆の不満が高まった  
 のに…戦争に敗れたのにみごとに復興した (c f. 大きいのに入れる)  
 のみ…人はパンのみにて生くるにあらず  
 は…会議には外務大臣が出席した 言うはやすく行うはかたし  
 ばかり…わずかばかりの耕地に 戰いの終わったばかりの国  
 へ…米国へ輸出する 外国への手紙 東へ東へと進出してゆく  
 べき\*…敗れるべくして敗れる 天皇は神聖にして侵すべからず  
 ほど…ビーカーに半分ほど水を入れておく 十年ほど前に  
 まで…外国まで出かけてゆく 言うまでもなく 100°Cまで上げる  
 も…運動が日本でも起こった (c f. だれでも考えられる)  
 もの…そうは言うものの 戰争は始めたものの  
 や…戦争を始めるやいなや 笛や太鼓  
 より…友あり遠方より來たる 案するより生むがやすし  
 らしい\*…大国らしい態度 戰争が始まるらしく思われる  
 を…戦争への道を歩む 平和を望む国民 外国を侵略する

#### 4. 3 M単位の切り方（概略）

- (0) 単位は “＼” に始まり、 “＼” で終わる。ただし、すでに “／” のあるところ、すなわちW単位の切れ目には “＼” を書かない。
- (1) 記号は 1 M 単位とする。W単位で無視した記号もここでは 1 M 単位とする。また、1 W 单位とした記号連続も、分離してそれぞれ 1 M 単位とする。元素記号は 1 元素ごとに 1 M 単位とする。

例 ; / 3 \ . \ 1 \ 5 / / ジョン \ = \ F \ = \ ケネディ /

<例外> 数式、化学式・根号を用いて表した累乗根・整数以外の要素を持つ分数・文章式、などは、全体を 1 M 単位とする（式は “◇○式”，数は “◇○数” として処理をする）。

例 ; / A = 3 / → ◇○式 / √a+b / → ◇○数

- (2) 記号以外の、語 (W 単位語) は次の (ア) ~ (エ) の最小単位をもとにあと A ~ C の規則により M 単位に分割する。なお、数 (アラビア数字、漢数字、ローマ数字とその連続) を含むものはさらに (3) によって処理する。

## (ア) 固有名詞の最小単位

a. 人名・姓・名・通称・雅号・芸名・筆名・あだななどのそれぞれ。

例；／高向／玄理／／三遊亭／円朝／／ジャンヌ／ダルク／

b. 国名・行政区画名・地域名・地形名…「国・王国・共和国・連邦・合衆国・都・道・府・県・州・市・町・村・通り・字・大字・小字・街・山・川・島・海・湾・洋・列島・半島・海峡・岬」などの類概念を除いた部分。

例；日本（国） テネシー（州・川） ヨーロッパ（大陸・州）

## (イ) 和語の最小単位

a. 自立語

例；手〔て・た〕 雨〔あめ・あま〕 物〔もの〕 美しい〔美しいかろ・美しかっ・美しく…〕 行く〔行か・行き・行っ…〕 きっと 最もする〔し・せ・する…〕 こと

b. 付属語

例；（行っ）て （死ん）で （有れ）ば （美しけれ）ど （行こ）う  
（死ん）だ （知ら）ず （知ら）ない （行き）たい （面白）がる  
（迷惑）そう （丸）まる・（丸）める〔「～まる」と「～める」でセットになるもの〕 お（父）さん み（心）

## (ウ) 外来語（和製外来語も含める）の最小単位

a. 一般に原語で1つの最小単位

例；ピーカー ガラス アルミニウム メス＼シリンドー フェノール  
フタレイン

b. 日本で、分解不可能な形に熟合したもの

例；クーデター ヘクタール（ただし h a は h と a とに分ける）

c. 漢字を当てたもの

例；煙草 背広

d. ローマ字、カナで略したもの

例；P T A N H K （塩）ビ ポリ（バケツ）

## (エ) 漢語の最小単位

漢語の最小単位は、字音の漢字1字（和製漢字も含む）とする。

例；一 二 愛 動 信 義 用 的 風 働

## A. 漢語のM単位

漢語要素（上記の（エ）のもの）だけからなるW単位の中では、最小単位の1次結合したものを1M単位とする。W単位の中で、このM単位の前、または後に順次結合してゆく、最小単位および最小単位の1次結合体は、それぞれ1M単位とする。また、W単位の中で1次結合をしていないもの（1字漢語）はそれ自体を1M単位とする。（最小単位、結合について、くわしくは、国立国語研究所報告21『現代雑誌九十種の用語用字』<第一分冊>を参照のこと）。

例；／凝固＼点＼降下／／急進＼的＼民族＼主義＼者／

<例外・1> W単位が、3つ以上の漢語要素（最小単位）の並列からなる場合、それぞれの要素を1M単位とする。

例；／真＼善＼美＼／を／／衣＼食＼住＼／／都＼道＼府＼県＼／（2つの漢語要素の並列は1次結合と見て、全体を1M単位とする。）

例；／町村／）

<例外・2> W単位の中で、2つの漢語要素が並列となって、他の要素、またはその結合したものと結合している場合、次のように処理する。

◎結合の相手が、漢語要素1つの場合 例；／町村長／／原水爆／

◎結合の相手が、漢語要素の1回結合の場合 例；／中小＼企業／

<例外・3> W単位内で、3つ以上の漢語要素が並列となって他の要素（またはその結合したもの）と結合している場合、それを1M単位とする。

例；／市＼町＼村＼長／／都＼道＼府＼県＼知事／

<例外・4> 3つ以上の漢語要素から成る語（3字以上の漢語）でも、現代語として切り離せないと考えられる、次のようなものは切らずに全体を1M単位とする。

例；／阿弥陀／／過不足／／加持子／／勘解由／／下剋上／

<例外・5> 切り方に問題のある、次のようなものは、切らない。

（／〇〇＼〇／か／〇＼〇〇／か決定しがたいもの）

例；／一部分／／海水浴／／加速度／／過半数／／下半部／

## B. 漢語要素を含む混種語のM単位

漢語要素と、他の要素（和語・外来語・固有名詞）とから成る混種語は、まず、Aの方法（例外もあてはまれば適用）で1次結合をもとにしたものに分割し、それをさらに、とのa・b・cのいずれかあてはまるものにしたがって処理する。

a. 2次結合をしていて切り出された部分が、1最小単位であれば、それは1M単位とする。

例；／見本＼品／／町＼工場／（後で、見本、工場はcを適用）

b. 切り出された1次結合が、漢語要素どうしの結合であれば、それは1M単位とする。

例；／運動＼エネルギー／

c. 切り出された1次結合が、漢語要素と他の要素（固有名詞、和語、外来語）との結合、または、漢語要素を含まない結合であれば、それぞれの要素すなわち最小単位に分けて、ともに1M単位とする。外来語「アレルギー」と漢語「性」との混種語「アレルギー性」、和語「古」「屋」と漢語「本」の混種語「古本屋」、固有名詞「スペイン」と和語「風邪」との混種語「スペイン風邪」は、次のようになる。

例；／アレルギー＼性／／古＼本＼屋／／スペイン＼風邪／

<例外・1> 漢語要素と、漢字1字による字音の国名・地域名（略語も含む）との結合は切り離さない。

例；／唐錢／／漢字／／訪米／

<例外・2> 姓などを、臨時に音読したものは、固有名詞要素としない。

例；／平家／／源氏／／藤家／／菅家／

（c f. ／李＼氏／／宗＼氏／…実際の姓である）

<例外・3> 漢字1字の中国人名等に「子」の結合したものは切り離さない。

例；／孟子／／孔子／（c f. ／淮南＼子／）

<例外・4> 漢語要素に、サ変動詞や活用語尾の付いたものは、全体を1M単位とする。

例；／愛する／／信じる／／信する／／課する／／力む／（c f. ／恋＼する／…和語に「スル」が付いたもの）

### C. 漢語を含まない語のM単位

漢語要素を含まない語（W単位）、すなわち、外来語・和語・固有名詞およびそれらの混種語は、1最小単位を1M単位とする。

#### a. 和語のM単位

例；／場＼合／／ま＼ごころ／／子供＼ぽい／／赤＼み／

<例外・1> 最小単位と認められるものを含んでいても、それを切り離すと、後に最小単位と認められないものが残る場合は切り離さない。

例；／けだもの／（c f. ／け＼もの／）／まなこ／／きっかけ／

<例外・2> 次のような、動詞を起源とする副詞・接続詞は、全体で1M単位とする。

例；／あくまで／／あらためて／／かえって／／きわめて／

<例外・3> 形容詞語幹に付く五段活用語尾「む」とその活用形（連用形転成名詞を含む）は切り離さない。

例；／苦しむ／／楽しむ／／慈しみ／／楽しみ／

<例外・4> 助辞リストにある語を構成要素にもち、W単位処理で切られなかったものはM単位で切り離すが、その形式での用法が固定していて、他には用いられないものは切り離さない。

例；／いかが／／いやしくも／／大いなる／／おのずから／

<例外・5> 現代語で一語と考えられるつぎのようなものは切らない。

例；／あらゆる／／いわゆる／／みずうみ／／玉子（卵）／  
／つまずく／／くちばし／／望ましい／

#### b. 外来語のM単位

例；／ナトリウム＼イオン／／マス＼コミュニケーション／

<例外> 外来語、またはその略語について動詞化する活用語尾は切り離さない。

例；／サボッ＼た／／デモる／／アジる／

#### c. 固有名詞要素のM単位

①人名…姓と名とはそれぞれ1M単位とする。

例；／徳川＼家康／／李＼鴻章／／J＼S＼バッハ／

(注) ルビのみに現れる、姓と名をつなげる「の」は無視する。

例； 源 義経 →／源＼義経／(／みなもと＼よしつね／)

②地名…類概念は外す。また、漢字1字の国名（略語を含む）どうしの並列・結合は切る。

例；／日本＼列島／／東京＼湾／／日＼中／／薩＼長／(～同盟)

<例外> 類概念を外したものが、地名、地形名として独立しないものは切らない。

例；／荒川／／大島／／漢江／(c f. ／揚子＼江／)

#### d. 和語・外来語・固有名詞の混種語のM単位

例；／赤＼ランプ／／ケネディ＼ラウンド／／パリっ＼子／

(3) 以上の規定で分けられたものの中に、数を表す要素が含まれていたら、その部分をさらに切り離し、それぞれの位の数を 1 M 単位とする。数は、アラビア数字、ローマ数字、漢数字（これまで漢字要素としている）のほか、「何十」「幾回」「数人」「百余」などの「何」「幾」「数」「余」なども含める。

例；／第＼一＼回／／四＼分／の＼一／／五＼時／十＼七＼分／  
 ／何＼十＼万＼人／／一＼，＼0＼0＼0＼円／／3＼.＼1＼4＼  
 <例外>その数の場合だけの、特別な意味を表し、数を表すのが主眼でない  
 次のようなものは数の部分を切らない。

例；／一員／／一因／／一円／（／～＼的／）／一応／  
 ／一段／（／～＼と／）／一堂／／一部分／／一瞥／  
 ／一面／（～において）／一門／／一躍／／一様／  
 ／一連／（～の事件）／一貫／（～＼性＼）／一路／  
 ／再三／／十分／（～に調べる）／百姓／／万物／

## 5. 同語異語判別

### 5. 1 同語異語判別作業の概略

この調査システムの特徴のひとつは、同語異語判別作業を、人手と機械力を組み合わせて行ったことにある。同語異語判別とは、一語の範囲を定め、それに属する語に同じ見出し語を付することである。具体的には、異形同語を集め同じ見出し語（代表形）を付することと、同形（同表記）異語を分離しそれぞれに異なる見出し語（代表形）を付することの二作業に分れる。前者は、入力データ作成の際「代表形」を付することにより、ほぼ達成できる。後者は、作業の中間段階で、全ての語に「判別情報」を付すことによって達成される。後者の作業の際「代表形」つけの追加・変更等も行った。また、「判別情報」は、出現した表記形の先頭 1 字を機械的に採用し、それを変更するものについてのみ手を加えた。詳細については、国研報告 78 所収の「6. 機械処理システム」、および国研報告 61 所収の「高校教科書の同語異語判別システム」を参照されたい。

同語とみなすものの範囲は、雑誌九十種調査の  $\beta$  単位の集計単位に準じた。ただし、この調査の M 単位と  $\beta$  単位との単位の長さには違いがあり、また同じ長さの語であっても、 $\beta$  単位とは別の範囲を採ったものもわずかだが存する。例えば、 $\beta$  単位では同語とした「特徴」「特長」は別語とし、 $\beta$  単位で別語とした「食料」「食糧」は、この調査では同語としている。

## 5. 2 代表形のつけ方

「代表形」とは、語彙調査において、同じ語としてまとめられた一群の語の、代表たる語形のことであり、語について使用率がどれほどか、異なり語数が幾らになるかを論ずる際の単体を表す語形のことである。この単体は、現代雑誌九十種の用語用字調査における「集計単位」に相当するものであり、「代表形」は「見出し語」に相当するものの表記形である。「集計単位」については、国研報告21「現代雑誌九十種の用語用字」第一分冊p. 14～p. 20の「集計単位の定め方」を、「見出し語」については、同書p. 22～p. 24の「見出し語の形と決め方・しるし方」を、それぞれ参照されたい。なお、代表形のつけ方の作業規則については、国研報告78「高校教科書の語彙調査」を参照されたい。

(1) この調査での代表形の定め方は、次のような原則による。

- 1 ある語一つを取り上げて、その代表形を立てる一般的な仕方を規定する。この際、その語が活用する語の、ある一つの活用形であるか否かが明らかなら、活用の有無およびその型に関する情報を使ってよい。また資料となつた表現でその語と関係する他の語との、形式上・意味上の相互関係の情報も、明らかな限り使ってよい。
- 2 同じ代表形を持つと認める語の範囲が、一般規則の形か、個々のケースの完全枚挙の形かで規定できるものは、その範囲を規定することによって直ちに、一つの語だけからでもその代表形が決まるようにする。
- 3 同音異義語とするか同じ語の意味の差と考えるかというように、(2)の形で規定しにくいものは、個々の例についてこれを見定める手続きを決め、その手続きを踏んで達した結果によって代表形をつける。このような事例については、国研報告25「現代雑誌九十種の用語用字」第三分冊p. 301～p. 330の「判別の実例一覧表」を参考にした。
- 4 漢字で表記してある語にルビがつけてあれば、そのルビの語形を尊重する。ルビが二通り付されている場合は、主たるもののみをとる。
- 5 作業段階において、出現した表記形が代表形と一致する場合は、作業の能率化を図るために代表形はつけず、出現した表記形を代表形として扱う。また、この原則を生かすために、代表形は最も多く出現すると思われる表記形と一致させる。

### (2) 代表形のしるし方

同じ語の範囲にあると認めた各単位語には、当然同じ代表形を与える。それを書き表す規則は次の通りである。

- 1 活用しない語は、その形を代表形とする。ただし、同じ語の範囲内にある単位語の形に、（活用以外の）差がある場合には、その中の一つに従う。この場合には、おおむね臨時的でないほう、くずれないほうの形を探る。また、その形が優勢でなくとも弁別に便利なときには、それを探ることもある。
- 2 活用する語は、終止形を代表形とする。  
ただし、活用する助辞は、出現した語形を代表形とする。
- 3 代表形を書きしるすには、平仮名を用い、また、ローマ字や特殊記号などを使う方が便利な場合には、それを代表形とする。
- 4 促音を表す「つ」、拗音を表す「や・ゅ・よ」およびこれに準ずる「あ・い・う・え・お」は、小さく書く。  
例；あっさり、きゃく、ファッション

## II 文脈付き用語索引について

### 1. データ量

本資料は国研報告76「高校教科書の語彙調査」の調査データを用例集として編集出力したものである。

データ量は下に記すとおりである。詳しくは、表1(p. 16)参照のこと。

マイクロフィッシュ 98枚

1枚のマイクロフィッシュに含まれるフレーム(頁)数

データ部 97

索引部 1

合計フレーム数

データ部 9433

索引部 98

### 語数

	延べ	異なり
自立語	321058	15519
助辞	186589	88
数字	17700	28
記号	68919	27
全体	594266	15662

表1 語種別の語彙量  
括弧内は百分率

	和 語	漢 語	外 米 語	混 種 語	人 名	地 名	自 立 語	助 辞	辞 数	記 字	号	全 体
物 理	15160 (51.00)	13617 (45.81)	685 (2.30)	207 (0.70)	52 (0.17)	3 (0.01)	18111 (52.25)	1908 (31.84)	7144 (3.35)	12.56 (12.57)	7144 (12.56)	56887 (100.00)
化 学	10043 (37.94)	14548 (54.96)	1663 (6.28)	116 (0.44)	96 (0.36)	5 (0.02)	26271 (52.07)	14626 (28.77)	2728 (5.37)	7099 (13.79)	7099 (13.79)	50834 (100.00)
生 物	13035 (42.38)	16154 (53.26)	823 (2.71)	197 (0.65)	115 (0.38)	7 (0.02)	30331 (55.67)	17878 (32.81)	858 (1.57)	5418 (9.94)	5418 (9.94)	54485 (100.00)
地 学	9256 (39.75)	13128 (56.37)	496 (2.13)	131 (0.56)	55 (0.24)	5 (0.05)	23288 (53.69)	2877 (29.34)	1649 (3.80)	5714 (13.17)	5714 (13.17)	43376 (100.00)
理 科	47494 (43.25)	57447 (52.31)	3667 (3.34)	651 (0.59)	318 (0.29)	237 (0.22)	109814 (53.42)	63340 (30.81)	7143 (3.47)	25285 (12.30)	25285 (12.30)	205582 (100.00)
倫理社会	19123 (50.49)	17174 (45.34)	312 (0.82)	309 (0.82)	503 (1.33)	454 (1.20)	3775 (54.63)	23972 (34.57)	392 (0.57)	7095 (10.23)	7095 (10.23)	69334 (100.00)
政治経済	18224 (41.17)	24335 (54.98)	289 (0.65)	412 (0.93)	78 (0.18)	926 (2.09)	44264 (54.38)	2509 (31.82)	2251 (2.78)	8774 (10.82)	8774 (10.82)	81098 (100.00)
日本史	19273 (38.56)	25791 (51.61)	101 (0.20)	420 (0.84)	1824 (0.65)	2668 (0.54)	49977 (53.44)	30860 (33.40)	3007 (3.22)	9679 (10.35)	9679 (10.35)	93523 (100.00)
世界史	13628 (30.20)	24472 (54.23)	611 (1.35)	331 (0.73)	1076 (0.73)	5012 (1.20)	5012 (54.13)	24892 (54.13)	3462 (4.15)	9883 (11.85)	9883 (11.85)	83367 (100.00)
地理	10870 (31.97)	18571 (54.62)	765 (2.25)	172 (0.51)	84 (0.25)	3536 (10.40)	3536 (55.41)	1716 (28.87)	1445 (2.35)	8203 (13.37)	8203 (13.37)	61362 (100.00)
社会科学	81118 (38.40)	110343 (52.23)	2078 (0.98)	1644 (0.78)	3555 (1.69)	12996 (5.92)	21244 (54.35)	12249 (31.71)	10557 (2.77)	43634 (11.23)	43634 (11.23)	388684 (100.00)
全 体	128612 (40.06)	167790 (52.26)	5745 (1.79)	2925 (0.71)	3883 (1.21)	12733 (3.97)	32058 (54.03)	186589 (31.40)	17700 (2.98)	68919 (11.60)	68919 (11.60)	594266 (100.00)
異なる												
	和 語	漢 語	外 米 語	混 種 語	人 名	地 名	自 立 語	助 辞	辞 数	記 字	号	全 体
物 理	657 (36.42)	1012 (56.10)	94 (5.21)	16 (0.89)	23 (1.27)	2 (0.11)	1804 (95.20)	48 (2.53)	22 (1.16)	21 (1.11)	21 (1.11)	1895 (100.00)
化 学	576 (33.07)	99 (56.89)	126 (7.23)	17 (0.98)	27 (1.55)	5 (0.29)	1742 (95.66)	43 (2.36)	18 (0.99)	18 (0.99)	18 (0.99)	18221 (100.00)
生 物	769 (29.89)	1573 (61.33)	152 (5.91)	19 (0.74)	48 (1.87)	7 (0.27)	2573 (97.20)	44 (1.66)	13 (0.49)	17 (0.64)	17 (0.64)	2667 (100.00)
地 学	756 (27.22)	1784 (64.24)	118 (4.25)	17 (0.61)	21 (0.76)	81 (2.92)	2777 (97.30)	44 (1.54)	21 (0.74)	12 (0.42)	12 (0.42)	2854 (100.00)
理 科	1265 (24.30)	3296 (63.32)	415 (7.97)	27 (0.52)	113 (2.13)	91 (1.75)	5205 (98.13)	51 (0.96)	26 (0.49)	22 (0.41)	22 (0.41)	5304 (100.00)
倫理社会	1129 (29.95)	2239 (59.39)	116 (3.08)	45 (1.19)	176 (4.67)	65 (1.72)	3770 (97.22)	83 (2.14)	10 (0.26)	15 (0.39)	15 (0.39)	3878 (100.00)
政治経済	892 (23.54)	2573 (67.89)	137 (3.61)	48 (1.27)	50 (1.32)	90 (2.37)	3790 (97.66)	54 (1.39)	23 (0.59)	14 (0.36)	14 (0.36)	3881 (100.00)
日本史	1249 (17.61)	4366 (61.60)	60 (0.85)	66 (0.93)	911 (12.85)	27 (6.16)	437 (98.91)	54 (0.75)	10 (0.14)	14 (0.26)	14 (0.26)	7169 (100.00)
世界史	772 (15.27)	2961 (58.61)	185 (3.66)	45 (0.89)	542 (1.07)	545 (10.86)	5057 (98.60)	42 (0.82)	14 (0.27)	16 (0.31)	16 (0.31)	5129 (100.00)
地理	688 (17.05)	2121 (52.57)	234 (5.80)	29 (0.72)	47 (1.16)	47 (2.37)	916 (98.29)	44 (1.07)	10 (0.24)	16 (0.39)	16 (0.39)	4105 (100.00)
社会科学	1961 (15.04)	7447 (57.10)	585 (4.49)	105 (0.81)	1491 (11.43)	1452 (11.13)	13041 (98.98)	88 (0.67)	23 (0.17)	23 (0.17)	23 (0.17)	13175 (100.00)
全 体	2279 (14.69)	9121 (58.77)	941 (6.06)	107 (0.69)	1586 (10.22)	1485 (9.57)	15519 (99.09)	88 (0.56)	28 (0.18)	27 (0.17)	27 (0.17)	15662 (100.00)

代表形	(注記)	C 高校教科書 判 J 科 貞 段落	K W I C )	M単位 出現形
あおき		人 青 日 217 02	けとなつた。ついで吉宗時代に、	青木 昆陽・野呂元丈などがオランダ語を
		青 日 280 00	清輝や、ロマン的画風を代表する	青木 繁らがで、近代西洋画が確立した。
あおぐ		和 あ 政 077 00	資金のかなりの部分を財政資金に	仰い でいる。こうして、国は金融の面で
		あ 日 224 01	させ、条約調印の勅許を朝廷に	あおい だ。しかし朝廷の情勢はこれを許
		あ 日 100 01	中央の権門寺社を本所・領家と	あおい たり、國司の支配を受けたりして
		あ 世 093 04	れ、また東ローマの皇帝を君主と	あおい でいた。これらのゲルマン人のう
		あ 日 065 00	されていった。大武士団の棟梁と	あおが れたのは、中央で思うような官職
		あ 日 062 02	進し、貴族や寺社を本家・領家と	あおぐ とともに、みずからは預所・下司
あおさ	(植)	和 ア 生 141 01	である。図3のように、緑藻類の	アオサ では胞子体と雌雄の配偶体が同じ
あおる	(畠)	和 あ 日 261 01	三国干涉によって国民の敵愾心は	あおら れ、政府も日露の開戦を予想して
あか		和 赤 生 123 01	ばしにある点の色によって違ひ、	赤 あるいは黒に強く反応し、点の色が重
		赤 生 198 04	遺伝子に複雰に支配されている。	赤 毛は特別の遺伝子による、短指・多指
		赤 政 080 03	収入より多くなって、国際収支が	赤 字になると、一国の保有する外貨が減
		赤 政 094 03	ショーンの大きな原因となっていた	赤 字財政を抑え、復興金融金庫の貸出し
		赤 政 193 01	58年から年々国際収支の大幅な	赤 字に悩まされるようになつた。そのた
		赤 世 250 01	進者として、長期的な国際収支の	赤 字とドルの価値低下にならすことにな
		赤 政 096 01	り、輸入の増加による国際収支の	赤 字と、金融の引締めによる景気後退を
		赤 政 117 02	財政による食糧管理特別会計の	赤 字補填額は、増加の傾向をたどってい
		赤 政 177 01	る役割をもっており、国際収支の	赤 字を生じた園にドル資金などの外貨を
		赤 政 187 03	制限政策を強化して、国際収支の	赤 字を防止しなければならなかつた。対
		赤 政 086 00	めには、経常収支の黒字、または	赤 字がこれに加減される。すなわち黒字
		赤 地 147 02	力資源の輸入が多く、貿易収支は	赤 字であるが、観光収入と海外出稼せざ
		赤 政 086 00	あるから、これを加えるし、また	赤 字のときは、その分だけ外国から借り
		赤 日 009 00	宿の関東ローム層(洪積世末期の	赤 土層)から打製石器が発見された、こ
		赤 学 191 08	備海岸の砂および土(火山灰土(	赤 土)がよい)。偏光顕微鏡・スライド
		赤 学 182 08	方眼紙に、気圧は黒で、気温は	赤 で記入する。ただし、気圧・風向・風
		赤 生 099 01	きによる、色素胞には、黒・黄・	赤 などの色素果粒がふくまれていて、光
		赤 地 113 00	赤地113 00 達した市街は、クレムリン宮殿・	赤 の広場を中心、主として放射同心円
		赤 生 178 02	すように、F2は紫花・長花粉と	赤 花・丸花粉とが3:1に分離して、紫
		赤 生 171 02	遺伝学では、洋種オシロイバナの	赤 花と白花との交配による雑種第1代(
		赤 生 178 02	赤生178 02 :1に分離して、紫花・丸花粉、	赤 赤花・長花粉のものは全然あらわれず、
		赤 生 179 00	示すようにF2に紫花・丸花粉、	赤 赤花・長花粉のものが比較的少数である
		赤 生 123 01	赤生123 01 ラ、トゲウオの闘争はおすの腹の	赤 を信号刺激として行動がひきおこされ
		赤 学 201 08	赤学201 08 明るい星については、星の色を	赤 ・だいだい・黄・白・青白というよう
		赤 倫 102 02	赤倫102 02 限な輪郭をきわだてる白・黒・	赤 赤・黄などの色彩が多く使われていたの
		赤 生 178 01	赤生178 01 スイートピーの花色(紫-AAと	赤 -a a)と花粉の形(長-BBと丸-
		赤 生 055 02	赤生055 02 イ・ヒルなどの環形動物のほか、	赤 アカガイ・ユスリカの幼虫などにも見ら
		赤 生 195 02	赤生195 02 遺伝子)も多い。ビードルたちが	赤 アカパンカビでつくった多くの栄養要求
		赤 生 190 01	赤生190 01 伝子にもとづくタンパク質の合成	赤 アカパンカビの実験からわかるように、
		赤 生 189 01	赤生189 01 株の研究から、つぎに記すように	赤 アカパンカビの菌糸内におけるアルギニ
		赤 生 187 04	赤生187 04 いて、化学変化の差異を調べた。	赤 アカパンカビを用いたのは、その菌糸が
		赤 生 187 04	赤生187 04 いない。そこでビードルたちは、	赤 アカパンカビの種々な突然変異株を用い
		赤 倫 055 02	赤倫055 02 りるのである。ところで、どの	赤 ぼうも。多少のずれはあるものの、あ
あかい		和 化 110 04	和化110 04 にしていくと、△○式ぐらいから	赤 みを帯びはじめ、pHの値が大きくな
		赤 化 111 01	赤化111 01 くと、△○式ぐらいからしだいに	赤 みが加わり、△○式ぐらいで黄みがな
		赤 生 122 01	赤生122 01 づくと闘争をしかけるが、これは	赤 娼媚色をもった腹が信号刺激として
		赤 生 123 01	赤生123 01 グロカモメのくちばしの先端には	赤 斑点があり、ヒナがそれをついてむ
		赤 化 110 04	赤化110 04 あるが、アルカリ性の溶液中では	赤 それを含む酸性の溶液にアルカリ
		赤 化 165 03	赤化165 03 また、湿った青色リトマス紙を	赤 した後漂白したり、ヨウ化カリウム
		赤 化 102 01	赤化102 01 し、フェノールフタレイン溶液を	赤 する。これらの水溶液を酸の水溶液
		赤 生 123 01	赤生123 01 図のように形はおすに似て、腹は	赤 ない模型(N)にはまったく反応を
		赤 化 111 01	赤化111 01 △○式ぐらいで黄みがなくなり	赤 なる。しかしそれ以上pHが小さく
		赤 化 119 01	赤化119 01 Hの値は急に大きくなり、溶液は	赤 なる。この変色によって、中和点を
		赤 生 123 01	赤生123 01 いのに、形は違っていても腹側を	赤 ねった模型(R)にはすべて闘争を見
		赤 生 199 02	赤生199 02 に色素がなく、血液の色がすけて	赤 見える。皮膚は白、毛髪は白または
		赤 学 154 01	赤学154 01 既日食のときに観測され、彩唇は	赤 コロナは真珠色に輝いている。光
あかい		和 明 倫 136 01	が必要であると考えられた。滑き	明き 心仏教や儒学が伝来する前の日本人
		明 倫 136 01	し、その心のもちかたとして滑き	明き 心を重んじた。清明さとは、底まで
		明 倫 153 00	るすなおさであり、一面で「滑き	明き 心の伝統を受けつぎながら、心情
あかぎ	(正)	地 赤 学 015 07	いろいろ地方名でよばれる局地風(	赤城 おろしなど)や陸風・海風・雷雨性
あかし		和 あ 倫 106 03	が神に選ばれたものであることの	あかし にはなりうるかもしれないという
あかす		和 明 物 217 01	さらに原子の内部の問題を解き	明かし ていくためには、17世紀以来1
		明 倫 077 00	から疑いが生じ、この疑いを解き	明かす ために眞の知を得ようとする探究

## 2. KWICとは

KWICとは、Key Word In Contextの略で文脈の中にキーワードを配置したものである。文脈付き用語索引とも呼ばれる。本索引はKWICの形式で編集している。図に示すごとく、一つの語（用例の中央部一字分の空白で囲まれた文字列）がどのような文脈でどのように使われているかが、たちどころにわかる。本来KWICは図の右半分を指す。左側はキーワードに関する情報である。

## 3. 文脈付き用語索引の形式

各項目の説明を下に記す。

### ①代表形 「代表形」欄

語彙調査において、同じ語としてまとめられた一群の語の代表たる語形。p. 14 参照。

### ②注記 「注記」欄

代表形および判別情報だけでは語を区別しにくい場合、〔植〕〔鳥〕などのように意味を示したり、「かよう〔通〕」「かよう〔斯様〕」のように漢字表記を示したりしている。また、和語、漢語、外来語、混種語、人名、地名はそれぞれ「和」・（省略）・「外」・「混」・「人」・「地」の略号で示した。

### ③判別情報 「判」欄

代表形が同形である異なる語を機械で区別するための情報。主に漢字表記の頭一字を用いる。それで区別できない場合は他の文字を利用している。

### ④助辞情報 「J」欄

規則に示した助辞情報を「J」で示す。規則に示した語以外に若干助辞としたものがある。これは、国研報告76「高校教科書の語彙調査」「五十音順M単位語彙表 2付表」に示した。

### ⑤教科書 「科」欄

教科書を示す。略号は以下の通り。

物…物理 I	倫…倫理 社会
化…化学 I	政…政治 経済
生…生物 I	日…日本史
学…地学 I	世…世界史
	地…地理 B

### ⑥頁 「頁」欄

その語がどの教科書の何頁にあるか、その頁数を示す。一語が二頁にわたる場合は前の頁に属する。

⑦段落 「段落」欄

その頁の何段落目にあるかを示す。段落が前の頁から続いている場合は 0 段落とする。

⑧出現形 「出現形」欄

用例中、空白で囲んで示す。前後の用例は前 15 字分、後はその語および 1 字の空白を含め 19 字分を示す。

用例中に用いた記号は次のとおり。

◇○式 化学式、数式など

◇○数  $10^2$ ,  $\sqrt{a}$ , 3. 8440 × 10<sup>5</sup>,  $\frac{b}{a}$  など数値をあらわすもの

◇○○ 。、、など特殊記号

◇○1, ◇○2 … ①, ②など○つき数字

◇A 1, ◇A 2 … I, II などローマ数字

◇A D, … Δ などギリシャ文字の大文字

◇B A, ◇B B … α, β などギリシャ文字の小文字

#### 4 用例の配列順序

用例の配列順序は、代表形の 50 音順である。同じ代表形の中では出現形の 50 音順、それも同じ場合には用例の直後の語・直前の語の 50 音順に並んでいる。

詳しくは下に示すソートキーによる。

##### ソートキー

1. 代表形の 50 音順
2. 判別情報のコード順
3. 助辞情報
4. 出現形の 50 音順
5. 出現形のコード順
6. 直後の語の 50 音順
7. 直後の語のコード順
8. 直前の語の 50 音順
9. 教科書名のコード順
10. ページ数
11. 段落数

注) コード順　ここでいうコードとは、一つの文字に与えられた電子計算機の中で用いる記号をいう。この記号(コード)は順序性を持っている。国立国語研究所特有のコードがあり、これを国研コードと呼ぶ。

詳しくは松本昭「国研用漢字テレタイプと同機利用の言語情報処理」(国研報告31「電子計算機による国語研究」)参照。

## 5. 検索の方法

この文脈付き用語索引は代表形によって配列されている。したがって、調べたい語の代表形を求め、これによって検索する。本索引はM単位の索引である。比較的短い単位であるから、いわゆる複合語の語構成要素で語中、語尾に用いられている語を検索することも容易である。用例は一行30字で示されている。用例をもっと長くほしい場合には、用例中の先頭、あるいは最後の語を求め、その語の用例を調べればよい。

調べたい語が何枚目のフィッシュに入っているかを調べるために後掲のフィッシュ索引を用いる。また、1枚のフィッシュの中のどのフレームに目的の語があるかを調べるために、1枚のフィッシュの最後のフレーム(H-14)にあるフレーム索引を用いる。

次に検索の例を示す。

### ①検索したい語 「書く」

代表形 「書く」の代表形は「かく」

フィッシュ索引 「11 がいこく」により、11枚目のフィッシュを見る。

フレーム索引 「C-07かく」により、C-07のフレームを見る。H-07のフレームに目的の語「かく」の用例が始まる。

### ②検索したい語 「研究所」の「所」

代表形 「所」の代表形は「しょ」(語彙表で確かめると早い)

フィッシュ索引 「25 しゅちょう」により、25枚目のフィッシュを見る。

フレーム索引 「E-03 じゅんれい」により、E-03のフレームを見る。E-03のフレームに目的の語「所」の用例がはじまる。

## 6. 語数を知りたい場合

この用例集は国研報告76「高校教科書の語彙調査」の調査データを編集したものである。したがって、各語の語数、順位、使用率、教科書別語数、語種別語彙量などはこの国研報告76を見ればわかる。

## フィッシュ索引

番号	代表形	番号	代表形	番号	代表形
0 1	あ～	3 4	た～	6 7	はげしい～
0 2	あらわす～	3 5	た～	6 8	はんらん～
0 3	ある～	3 6	だい～	6 9	ふうちょう～
0 4	いすばにあ～	3 7	たしかめる～	7 0	ぶんか～
0 5	いる～	3 8	ちいき～	7 1	ほうけん～
0 6	うち～	3 9	ちょうぜい～	7 2	また～
0 7	おおい～	4 0	つよい～	7 3	みんしゅ～
0 8	おりる～	4 1	て～	7 4	も～
0 9	が～	4 2	て～	7 5	もの～
1 0	が～	4 3	で～	7 6	ゆく～
1 1	がいくく～	4 4	で～	7 7	よくねん～
1 2	かそく～	4 5	てき～	7 8	られる～
1 3	から～	4 6	と～	7 9	れる～
1 4	かんたく～	4 7	と～	8 0	わたる～
1 5	きょう～	4 8	ど～	8 1	を～
1 6	くも～	4 9	どくりつ～	8 2	を～
1 7	けん～	5 0	な～	8 3	を～
1 8	こうぎょう～	5 1	ながい～	8 4	k～
1 9	こくみん～	5 2	なる～	8 5	1～
2 0	この～	5 3	に～	8 6	3～
2 1	さいか～	5 4	に～	8 7	9～
2 2	さんち～	5 5	に～	8 8	・～
2 3	しだい～	5 6	に～	8 9	(～
2 4	しゃ～	5 7	ねつりょう～	9 0	.～
2 5	しゅちょう～	5 8	の～	9 1	.～
2 6	しょくぶつ～	5 9	の～	9 2	.～
2 7	ず～	6 0	の～	9 3	,～
2 8	する(さ)～	6 1	の～	9 4	,～
2 9	する(し)～	6 2	の～	9 5	,～
3 0	する(する)～	6 3	の～	9 6	,～
3 1	せいじ～	6 4	は～	9 7	,～
3 2	せる～	6 5	は～	9 8	,～
3 3	そさい～	6 6	は～		

## フレーム索引の例

## INDEX - FRAME

FRAME	FROM	TO
B 0 1 . . . .	あ	[記～あーとまん]
C 0 1 . . . .	あーとまん	～
D 0 1 . . . .	あい	～
E 0 1 . . . .	あい	～
F 0 1 . . . .	あい	[会～]
G 0 1 . . . .	あい	[会～]
H 0 1 . . . .	あい	[会～]
 B 0 2 . . . .	あい	[会～]
C 0 2 . . . .	あい	[会～]
D 0 2 . . . .	あい	[会～]
E 0 2 . . . .	あい	[会～]
F 0 2 . . . .	あい	[会～]
G 0 2 . . . .	あい	[会～]
H 0 2 . . . .	あい	[会～]
 B 0 3 . . . .	あい	[会～]
C 0 3 . . . .	あいする	～
D 0 3 . . . .	あいだ	～
E 0 3 . . . .	あいだ	～
F 0 3 . . . .	あいだ	～
G 0 3 . . . .	あいだ	～

国立国語研究所 言語処理データ集1

高校教科書 文脈付き用語索引

昭和60年3月

國立国語研究所

〒115 東京都北区西が丘3-9-14

電話東京(900)3111(代表)

UDC 371.671(083.87)

NDC 375.9031

本書の市販品発行所

(〒113) 東京都文京区本郷1-10-11 (03)815-1231(代)

日本マイクロ写真株式会社

国立国語研究所刊行書一覧

国立国語研究所報告

1	八丈島の言語調査	秀英出版刊	品切れ
2	言語生活の実態 —白河市および付近の農村における—	〃	〃
3	現代語の助詞・助動詞 —用法と実例—	〃	2,000円
4	婦人雑誌の用語 —現代語の語彙調査—	〃	品切れ
5	地域社会の言語生活 —鶴岡における実態調査—	〃	〃
6	少年と新聞 —小学生・中学生の新聞への接近と理解—	〃	〃
7	入門期の言語能力	〃	〃
8	談話語の実態	〃	〃
9	読みの実験的研究 —音読にあらわれた読みあやまりの分析—	〃	〃
10	低学年の読み書き能力	〃	〃
11	敬語と敬語意識	〃	〃
12	総合雑誌の用語(前編) —現代語の語彙調査—	〃	〃
13	総合雑誌の用語(後編) —現代語の語彙調査—	〃	〃
14	中学年の読み書き能力	〃	〃
15	明治初期の新聞の用語	〃	〃
16	日本方言の記述的研究	明治書院刊	〃
17	高学年の読み書き能力	秀英出版刊	〃
18	話しことばの文型(1) —対話資料による研究—	〃	2,000円
19	総合雑誌の用字	〃	品切れ
20	同音語の研究	〃	〃
21	現代雑誌九十種の用語用字(1) —総記および語彙表—	〃	〃
22	現代雑誌九十種の用語用字(2) —漢字表—	〃	〃
23	話しことばの文型(2) —独話資料による研究—	〃	2,000円
24	横組みの字形に関する研究	〃	品切れ
25	現代雑誌九十種の用語用字(3) —分析—	〃	〃
26	小学生の言語能力の発達	明治図書刊	2,100円
27	共通語化の過程 —北海道における親子三代のことば—	秀英出版刊	品切れ
28	類義語の研究	〃	〃
29	戦後の国民各層の文字生活	〃	400円
30-1	日本言語地図(1)	大蔵省印刷局刊	品切れ

	日本言語地図 (1) <縮刷版>	"	17,000円
30-2	日本言語地図 (2)	"	品切れ
	日本言語地図 (2) <縮刷版>	"	17,000円
30-3	日本言語地図 (3) 大蔵省印刷局刊	品切れ	
	日本言語地図 (3) <縮刷版>	"	17,000円
30-4	日本言語地図 (4)	"	品切れ
	日本言語地図 (4) <縮刷版>	"	17,000円
30-5	日本言語地図 (5)	"	品切れ
	日本言語地図 (5) <縮刷版>	"	17,000円
30-6	日本言語地図 (6)	"	品切れ
	日本言語地図 (6) <縮刷版>	"	17,000円
31	電子計算機による国語研究	秀英出版刊	品切れ
32	社会構造と言語の関係についての基礎的研究(1) ——親族語彙と社会構造——	"	"
33	家庭における子どものコミュニケーション意識	"	350円
34	電子計算機による国語研究 (II) ——新聞の用語用字調査の処理組織——	"	品切れ
35	社会構造と言語の関係についての基礎的研究(2) ——マキ・マケと親族呼称——	"	"
36	中学生の漢字習得に関する研究	"	"
37	電子計算機による新聞の語彙調査	"	"
38	電子計算機による新聞の語彙調査 (II)	"	"
39	電子計算機による国語研究 (III)	"	"
40	送りがな意識の調査	"	1,500円
41	待遇表現の実態 ——松江24時間調査資料から——	"	900円
42	電子計算機による新聞の語彙調査 (III)	"	1,200円
43	動詞の意味・用法の記述的研究	"	6,000円
44	形容詞の意味・用法の記述的研究	"	4,000円
45	幼児の読み書き能力	東京書籍刊	4,500円
46	電子計算機による国語研究 (IV)	秀英出版刊	700円
47	社会構造と言語の関係についての基礎的研究(3) ——性向語彙と価値観——	"	700円
48	電子計算機による新聞の語彙調査 (IV)	"	3,000円
49	電子計算機による国語研究 (V)	"	900円
50	幼児の文構造の発達 ——3歳～6歳児の場合——	"	品切れ
51	電子計算機による国語研究 (VI)	"	1,000円
52	地域社会の言語生活 ——鶴岡における20年前との比較——	"	1,800円
53	言語使用の変遷 (1) ——福島県北部地域の面接調査——	"	2,500円
54	電子計算機による国語研究 (VII)	"	1,000円
55	幼児語の形態論的な分析 ——動詞・形容詞・述語名詞——	"	品切れ

56	現代新聞の漢字	"	6,000円
57	比喩表現の理論と分類	"	6,000円
58	児童の文法能力	東京書籍刊	5,500円
59	電子計算機による国語研究(Ⅷ)	秀英出版刊	1,300円
60	X線映画資料による母音の発音の研究 —フォネーム研究序説—	"	2,500円
61	電子計算機による国語研究(Ⅸ)	秀英出版刊	品切れ
62	研究報告集(1)	"	1,700円
63	児童の表現力と作文	東京書籍刊	6,000円
64	各地方言親族語彙の言語社会学的研究(1)	秀英出版刊	2,000円
65	研究報告集(2)	"	3,000円
66	児童の語彙能力	東京書籍刊	8,000円
67	電子計算機による国語研究(Ⅹ)	秀英出版刊	1,500円
68	専門語の諸問題	"	4,000円
69	児童の連想語彙表	東京書籍刊	6,800円
70-1	大都市の言語生活(分析編)	三省堂刊	7,800円
70-2	大都市の言語生活(資料編)	"	12,000円
71	研究報告集(3)	秀英出版刊	4,800円
72	児童の概念形成と言語	東京書籍刊	6,800円
73	企業の中の敬語	三省堂刊	9,500円
74	研究報告集(4)	秀英出版刊	4,200円
75	現代表記のゆれ	"	2,700円
76	高校教科書の語彙調査	"	5,000円
77	敬語と敬言意識 —岡崎における20年前との比較—	三省堂刊	8,000円
78	日本語教育のための基本語彙調査	秀英出版刊	6,000円
79	研究報告集(5)	"	4,200円
80	言語行動の日独比較	三省堂刊	
81	高校教科書の語彙調査(2)	秀英出版刊	5,000円

#### 国立国語研究所資料集

1	国語関係刊行書目(昭和17~24年)	秀英出版刊	品切れ
2	語彙調査—現代新聞用語の一例—	"	"
3	送り仮名法資料集	"	"
4	明治以降国語学関係刊行書目	"	"
5	沖縄語辞典	大蔵省印刷局刊	4,300円
6	分類語彙表	秀英出版刊	1,800円
7	動詞・形容詞問題語用例集	"	1,700円
8	現代新聞の漢字調査(中間報告)	"	品切れ
9	牛店雑談 安愚樂鍋用語索引	"	1,500円

10	方言談話資料(1) ——山形・群馬・長野——	"	6,000円
10-2	方言談話資料(2) ——奈良・高知・長崎——	"	6,000円
10-3	方言談話資料(3) ——青森・新潟・愛知——	"	6,000円
10-4	方言談話資料(4) ——福井・京都・島根——	"	6,000円
10-5	方言談話資料(5) ——岩手・宮城・千葉・静岡——	"	6,000円
10-6	方言談話資料(6) ——鳥取・愛媛・宮崎・沖縄——	"	6,000円
10-7	方言談話資料(7) ——老年層と若年層との対話——	"	6,000円
11	日本言語地図語形索引	大蔵省印刷局刊	1,500円

国立国語研究所研究部資料

1	幼児のことば資料(1) ——2歳・3歳誕生日のことばの記録—	秀英出版刊	3,800円
1-2	幼児のことば資料(2) ——4歳誕生日のことばの記録—	"	3,800円
1-3	幼児のことば資料(3) ——1歳児のことばの記録—	"	6,000円
1-4	幼児のことば資料(4) ——2歳児のことばの記録—	"	6,000円
1-5	幼児のことば資料(5) ——3歳前半のことばの記録—	"	6,000円
1-6	幼児のことば資料(6) ——3歳後半のことばの記録—	"	6,000円

国立国語研究所論集

1	ことばの研究	秀英出版刊	品切れ
2	ことばの研究 第2集	"	"
3	ことばの研究 第3集	"	"
4	ことばの研究 第4集	"	1,300円
5	ことばの研究 第5集	"	1,300円

国立国語研究所年報

1	昭和24年度	品切れ	13	昭和36年度	160円	25	昭和48年度	品切れ
2	昭和25年度	"	14	昭和37年度	220円	26	昭和49年度	600円
3	昭和26年度	160円	15	昭和38年度	250円	27	昭和50年度	700円
4	昭和27年度	160円	16	昭和39年度	品切れ	28	昭和51年度	非売品
5	昭和28年度	品切れ	17	昭和40年度	"	29	昭和52年度	"
6	昭和29年度	200円	18	昭和41年度	300円	30	昭和53年度	800円
7	昭和30年度	品切れ	19	昭和42年度	300円	31	昭和54年度	1,200円
8	昭和31年度	"	20	昭和43年度	品切れ	32	昭和55年度	1,300円
9	昭和32年度	"	21	昭和44年度	"	33	昭和56年度	1,300円
10	昭和33年度	"	22	昭和45年度	"	34	昭和57年度	2,000円
11	昭和34年度	"	23	昭和46年度	450円			
12	昭和35年度	"	24	昭和47年度	品切れ			

国語年鑑

昭和29年版	品切れ	昭和34年版	品切れ	昭和39年版	品切れ
昭和30年版	"	昭和35年版	"	昭和40年版	"
昭和31年版	"	昭和36年版	"	昭和41年版	"
昭和32年版	"	昭和37年版	"	昭和42年版	"
昭和33年版	"	昭和38年版	"	昭和43年版	"

昭和44年版 品切れ	昭和49年版 3,800円	昭和54年版 品切れ
昭和45年版 1,500円	昭和50年版 3,800円	昭和55年版 "
昭和46年版 2,000円	昭和51年版 4,000円	昭和56年版 "
昭和47年版 2,200円	昭和52年版 品切れ	昭和57年版 5,500円
昭和48年版 2,700円	昭和53年版 "	昭和58年版 5,500円

高 校 生 と 新 聞 国立国語研究所 共編 日本新聞協会	秀英出版刊 280円
青年とマス・コミュニケーション 日本新聞協会 共編 国立国語研究所	金沢書店刊 品切れ
国立国語研究所三十年のあゆみ —研究業績の紹介—	秀英出版刊 1,500円

#### 日本語教育教材

1 日 本 語 と 日 本 語 教 育 国立国語研究所 文 化 庁 共編 ——発音・表現編——	大蔵省印刷局刊	650円
2 日 本 語 と 日 本 語 教 育 ——文字・表現編——	"	850円
3 日 本 語 の 文 法 (上) ——日本語教育指導参考書4——	"	450円
4 日 本 語 の 文 法 (下) ——	" 5 —— "	550円
5 日 本 語 教 育 の 評 価 法 ——	" 6 —— "	700円
6 中・上級の教 授 法 ——	" 7 —— "	500円
7 日 本 語 の 指 示 詞 ——	" 8 —— "	500円
8 日 本 語 教育基本語彙比較対照表 ——	" 9 —— "	1,000円
9 日 本 語 教育参考文献一覧 ——	" 10 —— "	1,400円
10 談話の研究と教育 I ——	" 11 —— "	550円

#### 日本語教育教材映画一覧 (各巻16ミリカラー、5分、日本シネラル社販売)

〈卷 題 名〉	〈プリント価格〉
第1巻* これは かえるです 一「こそあど」+「は～です」一	30,000円
第2巻* さいふは どこにありますか 一「こそあど」+「が～ある」一	"
第3巻* やすくないです、たかいです 一形容詞とその活用導入一	"
第4巻* なにを しましたか 一動詞一	"
第5巻* しづかなこうえんで 一形容動詞一	"
第6巻* さあ、かぞえましょう 一助数詞一	"
第7巻* うつくしいさらに なりました 一「なる」「する」一	"
第8巻* きりんは どこにいますか 一「いる」「ある」一	"
第9巻* かまくらを あるきます 一移動の表現一	"
第10巻 おかねを とられました 一受身の表現1一	"
第11巻* どちらが すきですか 一比較・程度の表現一	"
第12巻* もみじが とてもきれいでした 一「です」「でした」「でしょう」一	"
第13巻* きょうは あめがふっています 一「して」「している」「していた」一	"

- 第14巻\* そうじは してありますか 一「してある」「しておく」「してしまう」一 "
- 第15巻\* おみまいに いきませんか 一依頼・勧誘の表現一 "
- 第16巻\* なみのおとが きこえてきます 一「いく」「くる」一 "
- 第17巻 みずうみのえを かいたことが ありますか 一経験・予定の表現一 "
- 第18巻\* あのいわまで およげますか 一可能の表現一 "
- 第19巻 よみせを みに いきたいです 一意志・希望の表現一 "
- 第20巻\* てんきが いいから さんばを しましょう 一原因・理由の表現一 "
- 第21巻\* さくらが きれいだそうです 一伝聞・様態の表現一 "
- 第22巻 あめに ふられて こまりました 一受身の表現2一 "
- 第23巻 おけいこを みにいっても いいですか 一許可・禁止の表現一 "
- 第24巻 あそこに のぼえば うみが みえます 一条件の表現1一 "
- 第25巻 いえが たくさんあるのに とてもしづかです 一条件の表現2一 "
- 第26巻 このきっぷを あげます 一やり・もらいの表現1一 "
- 第27巻 にもつを もってもらいました 一やり・もらいの表現2一 "
- 第28巻 てつだいを させました 一使役の表現一 "

第1巻～第3巻は、文化庁との共同企画

VTR価格1/2インチオープンリール21,000円、3/4インチカセット 20,000円

\*印については日本語教材映画解説の冊子がある。

A CONCORDANCE  
OF SENIOR HIGH SCHOOL TEXTBOOKS

CONTENTS

DATA are 9 text books which were used in 1974. There are 4 science books (physics, chemistry, biology, geology) and 5 social-studies books (morals, political economics, Japanese history, descriptive geography).

CONCORDANCE with context printed on 98 sheets of microfiche. There are 15662 different words and 594266 running words.

THE NATIONAL LANGUAGE RESEARCH INSTITUTE  
TOKYO

1985