

辞書データベースを用いた日本語 シソーラスの構造表示処理

加納豊子 丹羽恵美子 後藤忠彦 安藤一郎*

文献検索，教材資料検索などの情報検索に必要なシソーラスの構成を行うとき，言語間の相互関係を表示しその概念構造を検討することが重要である。このため，本研究では，日本語（漢字）の辞書データベースを用いて，各用語の上位，下位等の関連を調べ，用語間の構造を求める処理系を構成した。本処理の開発により，データベースへ登録された日本語の用語資料を用いて，シソーラスの構造を求めることができ，さらに，各種の教育情報処理に適用することができるようになった。

〈キーワード〉 情報検索，文献検索，教材データベース，辞書データベース，シソーラス，キーワード，教育情報処理

1. はじめに

文献データベース，教材データベース等の情報検索の基礎資料として，各分野で研究開発が進み試作されたシソーラスの用語間の一連の関連を表示し，各研究によりその構造の検討がなされた^{1)~3)}。また，辞書データベースに管理されている用語資料群を用いて，目的に応じた用語間の構造を作成し，それを用いた教育情報処理の管理・検索が文献・教材データベースの基本的処理として要求されている^{4)~6)}。

ところが，これらの用語間の構造表示は，すでに英文字，カナ文字等で処理されているが，辞書データベースの日本語（漢字）を用いたシソーラス構造を表示する処理法は，まだ開発がなされていなかったため，シソーラスの作成に当たって大変な作業を行っていた。

このため，カリキュラム開発研究センターでは教育情報検索用として開発を進めてきた辞書データベースを用いて，各用語間の相互関係を調べ，一連の構造関係の図を作成する処理を開発し，各分野の用語間の関係図を作成し出力できるシステムを構成し実用化に至った。

この日本語（漢字）を用いた関連構造処理の開発により，心理学，教育工学，教科教育研究の分野で辞書データベースを用いて教育用語相

互の関連図を作成し，これを使用したシソーラスの内容構造の検討作業が各分野で進みだした⁷⁾。

2. 処理の方法

辞書データベースに記録されている用語の相互関係から樹状構造を求め，その図表を出力する処理系を次のように構成した。

岐阜大学カリキュラム開発研究センターの教材データベース，文献データベースのもつ辞書データベースの情報は，各用語間の関係として

RT…… Related Term

BT…… Broader Term

NT…… Narrower Term

UF…… Used For

が関連語として記録されていて，これを用いて樹状構造を作成する。（このためには，辞書データベースの整備が必要であり，これらについては別に報告した。）

辞書データベースから用語間の関連とくに樹状構造を求め，それを図表化して出力するためには，次のような処理を行う。

① 主となる用語の設定

構造の主となる用語（樹状構造の先端になる用語）を指示する。（たとえば，物理教育で力学を中心とした構造を求めるには，力学を頂点

として関連図表を作成するため、「力学」を指定する。)

② 分野の指定と該当用語の検索

シソーラスは、同じ用語でも分野により上位、下位の関係、またその意味内容が違う場合がある。このため辞書データベースには、各語についてその用いる分野が記録してあり、この処理を行うに先立ち、求めようとする分野を指示し該当語のみを検索する。

③ 用語間の関連処理

一連の関連図の最初に、①で指示した主となる用語を設定し、次にその下位の語を②で検索した用語群の中から求め第二段に設定する。次に第二段に設定した用語の下位語を前記と同様に求める処理を行う。これを次々とくり返し、全体の樹状構造を計算機の中で構成する。

④ 樹状図表の出力

樹状図表は、用紙の関係で全図を出力することが困難なため、一般には用紙のほぼ全面に印刷できる第五段階まで出力する。(これ以上の下位構造を出力するには、樹状構造の先端に來る用語をあらためて下位の語に設定し、そこから樹状構造として出力し処理する。)

◆ 辞 書 リ ス ト ◆

※ (資料番号)	T 0310130
(見出し語 J)	分数
(属性)	代数 学習指導要領
(関連語)	BT 有理数 NT 分子 NT 分母 NT 約分 NT 通分 NT 逆数 NT 小数 RT 整数 RT 正の数・負の数
(関連資料番号)	113112
※ (資料番号)	T 0310131
(見出し語 J)	分子
(属性)	代数 学習指導要領
(関連語)	BT 分数 RT 分母 RT 約分 RT 通分 RT 逆数 RT 小数
(関連資料番号)	1131121
※ (資料番号)	T 0310132
(見出し語 J)	分母
(属性)	代数 学習指導要領
(関連語)	BT 分数 NT 真分数 NT 仮分数 NT 帯分数 NT 同値 NT 通分数 RT 分子 RT 約分 RT 通分 RT 逆算 RT 小数
(関連資料番号)	1131122

図1. 辞書データベース(数学)例

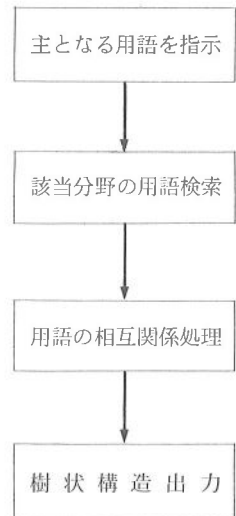
◆ 辞 書 リ ス ト ◆

※ (資料番号)	E3000040
(見出し語 J)	知覚
(見出し語 E)	PERCEPTION
(関連語)	JD 感覚 RT 統覚 RT 注意 RT 視覚 RT 弁別 RT 超感覚的知覚 RT 形の知覚 RT 研究 (知覚) RT 感覚的過程 RT 監視 RT 多知覚 RT 嗅覚 RT 知覚的閉鎖 RT 知覚弁別 RT 知覚の歪み RT 知覚障害 RT 知覚定位 RT 知覚測定 RT 知覚運動過程 RT 知覚調整 RT 知覚スタイル RT 皮刺知覚 RT 信号検出 (知覚) RT 社会的知覚 RT 身体感覚知覚 RT 空間知覚 RT 膈下知覚 RT 触覚 RT 味覚 RT 感觸知覚
(関連語)	UF SENSATION RT APPERCEPTION RT ATTENTION RT AUDITORY PERCEPTION RT DISCRIMINATION RT EXTRASENSORY PERCEPTION RT FORM AND SHAPE PERCEPTION RT ILLUSIONS(PERCEPTION) RT INTERSENSORY PROCESSES RT MONITORING RT NUMEROSITY PERCEPTION RT OLFACTORY PERCEPTION RT PERCEPTIONAL CLOSURE RT PERCEPTIONAL DISCRIMINATION RT PERCEPTIONAL DISTORTION RT PERCEPTIONAL DISTURBANCES RT PERCEPTIONAL LOCALIZATION RT PERCEPTIONAL MEASURES RT PERCEPTIONAL MOTOR PROCESSES RT PERCEPTIONAL STIMULATION RT PERCEPTIONAL STYLE RT ROLE PERCEPTION RT SIGNAL DETECTION(PERCEPTION) RT SOCIAL PERCEPTION RT SOMESTHETIC PERCEPTION RT SPATIAL PERCEPTION RT SUBLIMINAL PERCEPTION RT TACTUAL PERCEPTION RT TASTE PERCEPTION RT TEXTURE PERCEPTION
(SN)	環境についての認知情報あるいは感覚情報を得る過程に関する概念上幅広い用語。できれば、より特定の用語を使用せよ。
(英語出典)	APA

図2. 辞書データベース(心理)例

3. 辞書データベースを用いた処理例

用語間の関係から、樹状構造を求める処理は、辞書データベースの各用語の上位、下位の関係の最初の語(主となる語)を入力し、該当分野の用語を検索し樹状構造を出力する。この処理を用いて、各分野でシソーラスの研究が始まり、次の例で示すように幾つかの試案が出されるに至った。



4. おわりに

教育情報処理の中で、辞書データベースおよびそれを用いた各種の検索処理は、第一にシソーラスの作成が必要であるが、その用語間の関連を調べる作業が大変であった。しかし、この処理により時間的にも速くなった。また、一部の用語の上位、下位の関連語を変更することによって、いつも全体の書き直しが必要であったが、この処理により短時間に樹状の図が出力でき、シソーラスの訂正、修正、用語の追加等の検討を簡単に行うことができるようになった。

このように、作業の簡易化により、日本語の教育研究シソーラスの構成が岐阜大学以外の研究機関、研究者により本処理を用いて進みだし、現在、言語教育、数学教育、物理教育、化学教育、地学教育、生物教育、保健体育、心理学、教育工学等の広い研究分野で試案が作成され始めた。これらの試案の多くは、文献データベースと教材データベースで、それぞれ違ったシソーラスの研究が進んでいる。

現在、これらのシソーラスの試案を用いて、教材データベースでは、登録された資料のキーワードの整備、不足資料の追加・作成等の作業も始められた。また、教科書で用いている用語、説明文、定義文なども辞書データベースに登録し、本処理を用いて学習内容間の関連を表示し、その検討がなされるようになった。

以上のように、本処理は、用語間の関連を検討するための資料作成にも利用されている。

本研究費の一部は、文部省科学研究費試験研究「科学教育における教授・学習資料の管理・検索のための教材データベースの開発」(代表 後藤忠彦)による。また、本研究を進めるにあ

り、大阪大学言語文化部深谷哲教授にシソーラスの構成につき御指導、協力いただき、ここに感謝の意を表します。

文 献

- 1) 永野和男, 西之園晴夫(1981) “教育シソーラス作成のための基礎資料としてのキーワードの選定(I)～電子計算機による用語の抽出方法とその適用結果～” 科学教育研究 5-1
- 2) 西之園晴夫, 永野和男(1981) “教育シソーラス作成のための基礎資料としてのキーワードの選定(II)～教育分野について抽出した用語の問題点～” 科学教育研究 5-3
- 3) 中村治行, 大塚明郎, 八木正明, 磯脇一男, その他(1982) “科学教育語彙の選定とシソーラス化” 日本科学教育学会第6回年会論文集
- 4) 後藤忠彦, 石原正也, 高野卓哉, 興戸律子, 加納豊子(1982) “教育文献検索システムEDMARS—GIFUの構成と利用状況について” 日本科学教育学会第6回年会論文集
- 5) 石原正也, 安藤一郎(1982) “数学教育研究論文・資料に関する文献検索システムの開発について” 岐阜大学カリキュラム開発研究センター研究報告 2-1
- 6) 後藤忠彦, 豊吉律子(1982) “教育文献資料検索システム“EDMARS—GIFU” 岐阜大学カリキュラム開発研究センター研究報告 2-1
- 7) 森幸雄, 石原敏秀, 山田功, 芦葉浪久, 篠原文陽児, 堀口秀嗣, 後藤忠彦(1983) “科学教育(物理分野)の教育用シソーラス試案～日本物理教育学会選定の用語集を用いて～” 岐阜大学CRDCデータレポートNo. 137