

情報探索と質問定式化

—オンライン目録探索を中心に—

Information Seeking and Query Formulation:  
Searching Behavior in an Online Catalog

斎藤泰則  
*Yasunori Saito*

*Résumé*

In this paper, researches on information seeking behavior in a document retrieval system, especially an online catalog are reviewed. The focus of the review is query formulation in information seeking process.

The introduction of an online catalog early in the 1980s in the USA revealed the demand for subject searching of users. As a result, the frequency of subject searching in an online catalog increased in comparison with that of a card catalog. But subject searching caused many problems which were related to the selection of search terms, that is query formulation. So many enhancements and improvements of subject searching have been proposed. Some of them are based on information retrieval researches.

Recently, researches on information seeking behavior in a document retrieval system using the theories and the methodologies of cognitive science appeared. These researches revealed the cognitive process of information seeking and query formulation. The findings implies a new typed user-centered document retrieval system.

The user-centered document retrieval system should be explored and designed, apart from the constraints of current document retrieval system. The application of the theories and the methodologies of cognitive science to the research on information seeking behavior of users is useful in exploring user-centered retrieval system.

I. 序

II. 情報探索の問題点—オンライン目録における主題探索を中心に—

- A. 主題探索の頻度
- B. 主題探索の問題点
- C. 主題探索の改善

---

斎藤泰則：鹿児島経済大学社会学部，鹿児島県鹿児島市下福元町 8850  
Yasunori Saito: Kagoshima Keizai University, 8850, Fukumoto-cho, Kagoshima-shi, Kagoshima-ken.  
1992年5月30日受付

### III. 情報探索過程と質問定式化

#### A. 情報探索過程の全体

#### B. 情報要求と質問定式化

### IV. 文献の構造と質問定式化

#### V. 結

#### I. 序

オンライン目録探索を中心とした 1980 年代以降の情報探索研究は、主題探索の利用頻度調査に代表されるように、情報探索過程の側面、すなわち探索の実行形態を主として対象としてきた。それらの調査結果によって、様々な問題点が指摘され、改善策も提示されている。その改善策には、情報探索研究の成果をオンライン目録の検索方法に取り入れたものもある。目録探索にコンピュータが利用されることによって、雑誌論文の検索を主体とした情報探索システムの検索技法のオンライン目録への導入が技術的には可能になった。しかし、それらの改善策は、オンライン目録探索過程のごく一部を対象とした計量的調査の結果、明らかにされた問題点を踏まえて、目録システム設計者の経験や知識に基づいて提起されたものであって、利用者の情報探索行動に関する実証的研究に依拠したものではない。言い替えば、その改善策は目録探索に関するシステム設計者の考え方から引き出されたものであり、目録情報の内容に関する基本的枠組みを前提としている。目録のもつ基本構造を踏まえた改善は、確かに現実的なものである。しかし、こうした改善の方向は、利用者の情報探索行動の特質を踏まえてシステムを設計すべきところを、利用者の側にシステムへの適応を求めるシステム指向の改善をもたらす結果となる。システム設計者のもつ経験や知識から予測された利用者の情報探索行動が利用者の探索特性を反映しているならば、問題はない。しかし、その予測が利用者の情報探索行動に合致しているかどうかは定かでない。また、目録の構造が利用者の本来の情報探索行動とは異なった行動を利用者に強いていないとも限らない。システム設計者の予測の妥当性を検証するためにも、さらには、利用者の行動を可能な限り反映した文献探索システムを設計するためにも、文献探索システムにおける利用者の情報探索行動を詳細に把握する必要がある。

文献探索システムにおける利用者の情報探索行動については、従来の質問紙やトランザクションログによる計

量的手法とは異なる認知科学の研究方法を用いた調査が実施されるようになった。これらの調査は、情報探索過程全体、あるいはその過程の最初に位置する情報要求から質問定式化に至る部分に焦点をあてている。それらの調査結果から提起され、引き出された改善策は、目録情報の基本的枠組みの再考を示唆するものとなっている。

計量的方法によって明らかにされた問題点の中心は、情報探索における探索語の選択と探索戦略の構築という「質問定式化 (query formulation)」に関わる側面であった。認知科学的手法による調査においても、文献探索システムにおける情報探索の主たる問題点は同じく質問定式化にあり、研究の焦点は定式化をめぐる問題に置かれている。

オンライン目録における主題探索に関する最近のレビュー論文としては、Larson (Larson, 1991) のものがある。これは、主題探索全般を取り上げており、オンライン目録における主題探索の改善を扱った研究について、オンライン目録システムを構成する四つの層 (利用者インタフェース層、データベース管理システムインタフェース層、データベース管理層、データベース層) に分けて論じている。

それに対して本稿は、オンライン目録探索を中心とはしながらも、広く文献探索システムにおける利用者の情報探索行動と質問定式化に関する問題について、主として 1980 年代以降、米国で発表された論文を取り上げ、その研究動向や得られた知見を紹介する。なお、取り上げた論文は、“Online Searching”, “Intermediaries”, “User-System Interface” 等のディスクリプタを使って LISA (Library & Information Science Abstracts) を検索し収集した。

本稿の構成は次の通りである。II 章において、いまなお研究の多いトランザクションログの分析を中心とした計量的調査から明らかにされたオンライン目録における情報探索上の問題点とそれに関連して提起された改善策について述べる。III 章では、オンライン目録を中心とした文献探索システムにおける情報探索行動に関して、

認知科学の方法を用いて行われた研究を取り上げる。IV章では、認知科学的研究から得られた知見を基に、新たな文献検索システムの可能性について検討する。

## II. 情報探索の問題点—オンライン目録における主題探索を中心に—

### A. 主題探索の頻度

米国においては、1970年代末から1980年代初めにかけて、大学図書館を中心にオンライン目録が開発された。オンライン目録の導入を契機に、目録探索に関する多くの調査がこれまでに実施されてきている。その最初のもので、CLR (Council on Library Resources) の後援による質問紙を用いた全国規模の調査であり、Matthewsら (Matthews *et al.*, 1983) によってその結果が報告されている。それによれば、主題探索は、探索全体の59%を占めており、利用者にとって最も難しい探索であり、特に質問の定式化に利用者は難しさを感じていた。その結果を反映するように、オンライン目録に望む改善点は主題探索をめぐるものであり、アクセスポイント数を増やすことや関連語の参照機能が求められた。このように、オンライン目録の導入は、従来のカード目録探索の場合と比べて主題探索の利用頻度を高めたが、同時に主題探索の問題点を浮き彫りにした。

オンライン目録探索の調査には、システムと利用者とのやり取りを記録したトランザクションログが早くから用いられている。Borgman (Borgman, 1982, 1983) は、オハイオ州立大学図書館コントロールシステムから採取したトランザクションログを分析したが、利用者の探索セッションの約35%は主題探索を少なくとも1回含んでいた。

Tolle (Tolle, 1983) は、四つのオンライン目録システム (LC の SCORPIO, シラキュース大学の SULIRS, ダラス公共図書館のオンライン目録, オハイオ州立大学図書館コントロールシステム) から採取したトランザクションログを分析し、主題探索の割合が45%から60%に達することを示している。

Larson & Graham (Larson & Graham, 1983) は、1982年5月中のカリフォルニア大学オンライン目録システム MELVYL の全トランザクションを分析したが、それによれば、主題探索の割合はコマンドモードで51.6%、メニュー方式であるルックアップモード (LOOK UP MODE) で71.7%であったという。

オンライン目録における主題探索の増加については、

2通りの解釈が示されている。その第1は、カード目録上の主題探索の頻度を過小見積っていたというものである。従来、著者/書名探索とされてきた探索の多くが、実は利用者にとっては主題探索を意図して実行されていたのであり、オンライン目録によってそれらが主題探索として顕在化されたという解釈である (Hancock-Beaulieu, 1990; Lipetz & Paulson, 1987)。第2は、オンライン目録はより優れた主題探索機能を提供するとの利用者の期待によるものとする解釈である (Steinberg & Metz, 1983; Dykstra, 1988)。

オンライン目録の導入によって主題探索への要求が新たに生まれたというよりも、キーワード検索など、カード目録では事実上不可能であった主題探索機能が付加されたことにより、著者/書名探索という形態をとっていた探索が、件名標目を利用する本来の探索形態をとるようになったと考えるべきである。一方で、検索機能の付加、拡張が利用者主題探索への期待を抱かせたのも事実である。したがって、それら二つの要因が結びつくことにより、カード目録の場合には潜在していた主題探索への要求がオンライン目録において顕在化され、その結果、主題探索の頻度を押し上げたと思われるべきであろう。

### B. 主題探索の問題点

主題探索の増加とともに、主題探索の問題点も指摘されるようになった。Kern-Simirenko (Kern-Simirenko, 1983) は、三つのオンライン目録システムのトランザクションログの分析から、1件も検索できないような失敗の事例の占める割合が35%から46%にのぼることを示している。

最近の大規模な調査に、MELVYLを対象としたLarson (Larson, 1991b) のものがある。従来の調査がある特定の時期を扱っていたのに対して、これは6年間に渡って主題探索の頻度を分析している。その結果によれば、主題探索頻度の高さが強調された従来の調査結果とは逆に、主題索引の利用が年2.15%の率で低下していた。その原因としてLarsonは、主題探索がうまくいかないことを利用者が経験したこと、成功しても検索された文献数が多すぎることで、この2点をあげ、これらが利用者の主題探索離れを引き起こしていると分析している。

主題探索が問題となる事例には、Larson (Larson, 1991a) が分類したように、探索語と索引語との不一致による検索文献数が0件の状態と、検索はできては検索

文献数の多さにより、結果を出力することなく検索を終えてしまう状態とがある。こうした問題点は、多くの調査によって報告されており、オンライン目録探索の改善の主たる目標となっている (Kern-Simirenko, 1983; Cochrane & Markey, 1983; Markey, 1984; Frost, 1987; Blazek & Bilal, 1988; Blackshaw & Fischhoff, 1988; Hunter, 1991)。

探索語と索引語との不一致という問題は、質問定式化の問題の典型である。利用者は、情報要求をオンライン目録システムが受け入れ可能な形式で質問に変換しなければならない。受け入れ可能な形式とは、システムが使用する索引語、すなわち件名標目やディスクリプタを使用した質問形式である。このことは、利用者が件名標目やディスクリプタの内容とその使用方法に関する知識をもたなければならないことを意味する。仲介者としてのサーチャーの援助が通常は得られないオンライン目録探索の場合は、利用者(エンドユーザ)にその種の知識が求められる。したがって、当然の結果として、その種の知識を通常はもたないであろう利用者は、適切な探索語の選択はできず、件名標目との一致ができないまま探索を終えることになる。(Bates, 1977a, 1977b; Frost, 1987; Markey, 1984; Blackshaw & Fischhoff 1988; Van Pulis & Ludy 1988)。たとえ、件名標目から語が選択され、検索結果が得られたとしても、それは情報要求を歪めて件名標目に合わせた結果であるかもしれない。その件名標目が情報要求を的確に反映していなければ、検索結果の如何に関わらず、その探索は成功したとはいえないのである (Lancaster, 1977)。

このように、件名標目やディスクリプタのみを使用した質問定式化には、明らかに限界がある。確かに、件名標目へのキーワード検索機能は、標目との正確な一致という負担と難しさを大幅に軽減し、オンライン目録の優れた機能として重要視されているが、件名標目に専ら依存する探索語の選択を抜本的に解決するものではない。それどころか、この機能は、検索文献数の超過という第2の問題を引き起こすことになる。

Blair (Blair, 1980) は、「無用点 (futility point)」と「期待された無用点 (anticipated futility)」という概念を用いて、検索文献数の超過の問題を論じている。この無用点とは、利用者が検索結果をブラウジングできるくらいに抑えられた文献集合のことであり、検索結果の出力の意思を利用者に減退させない程度の文献数である。一方、期待された無用点とは、利用者が進んでブラウ

ジングを始めようとする検索文献数の最大値をいう。Larson (Larson, 1986) の調査によれば、大規模なオンライン目録システムにおける平均検索文献数は 77.5 件であり、書誌事項を表示するのは1探索あたり約 9.1 件であった。また、検索文献数が最も多くなるのは件名標目を利用した探索であり、検索文献数が 100 件以上を超える探索の約 37%、1000 件以上を超える探索の 52% を件名標目からの探索が占めていた。

オンライン目録が導入されてから 10 年が経過しようとしている米国では、主題探索の減少傾向が現れ始めているが、その一因は主題探索に伴う以上の問題点にあることが Larson (Larson, 1991b) によって指摘されている。先述したように、Larson は、MELVYL の 6 年間のトランザクションログを分析し、主題索引の利用パターンを見だしている。それによれば、夏休みと休日において、主題探索が急激に減少している。この期間は利用経験の豊富な図書館員や大学院生による探索が多い時期であることがわかっている。一方で、主題探索は年 2.15% の率で減少しており、代わりに書名のキーワード検索が増加しているという。この調査結果について Larson は、二つの仮説を提示している。第1は、利用者は探索経験を積むことによって、主題探索の失敗や検索結果の超過をもたらすような経験をしてしまうというものである。第2は、経験のある利用者は、より効果的な主題探索を行う方法を学習し、その結果、より少ない主題探索の回数で目標を達成するというものである。今後の検証が待たれるが、書名のキーワード検索が増加している状況を考えるならば、利用者はより簡単な探索方法を採用する傾向にあり、それは件名標目を使用した主題探索の難しさに原因があると予想される。

### C. 主題探索の改善

主題探索をめぐる以上の問題への改善策は、その手法によって二つに分かれる。一つは、目録中の主題情報に関わる改善であり、もう一つは高度な検索技法の導入による改善である。目録中の主題情報は、一般に、件名標目を中心に、分類記号、書名やその他の主題関連情報のキーワードからなる。主題情報の改善とは、これらの情報を豊富にするか、新たに付加するというものである (Mandel, 1985)。

件名標目に関わる改善には、付与する標目数を増やすこと、専門分野のソースを取り入れること、参照構造を含む LCSH を目録システムに組み込むこと、があ

げられている (Mandel, 1985; Markey, 1988)。Bates (Bates, 1986) は、利用者の使用する探索語とできるだけ一致させるために、探索語を統制語へと誘導する語彙を目録システムが備えるべきであると、豊富な参照構造と統制語を補う多くの自然語からなるエンドユーザーソースの構築を提案している。

分類記号を用いた改善策の有効性については、Dewey Decimal Classification と LC Classification の利用を中心に多くの試みが報告されている (Svenonius, 1983; Cochrane & Markey 1985; Markey, 1985, 1986, 1987; Williamson, 1985, 1986, 1989; Chan, 1986, 1989; Huestis, 1988)。その改善策としては、分類本表と相関索引を目録へ組み込むことにより、索引語を豊富にすると同時に、特に DDC の階層性を利用して探索語の文脈を提示する方法が採られている。

書名のキーワード検索の有効性については、Gerhan (Gerhan, 1989) の調査によって検証されている。統制語以外の語彙を利用した主題探索改善の試みとしてよく知られているのが、Atherton (Atherton, 1977) のSAP (Subject Access Project) である。これは図書の日次や索引から抽出された語を付加するものであり、この手法に基づいて改善された目録探索の調査も行われている (Byre & Micco, 1988)。この方法はきわめて簡便な改善策ではあるが、検索文献数の超過という問題を伴うことは否めない。

第4の改善策とでもいうべきものが、自然言語処理の手法や情報検索の研究から開発、提案された手法の応用である。キーワード検索は語単位の一致を前提としているが、そこに語幹処理や、部分一致の手法、スペルチェック機能などを加えることにより、探索語と索引語との一致の可能性を高めている (Walker & Jones, 1987; Lawrence, 1985)。また、レバンスフィールドバック、探索語から統制語への自動マッピング、検索文献の重み付け、分類をもとに件名標目と書名を統合した索引の作成など、様々の検索手法を取り入れたシステムが提案されている (Doszkocs, 1983; Larson, 1989, 1992)。

Hildreth (Hildreth, 1984, 1987) は、オンライン目録を3世代に分けている。第1世代はカード目録をコンピュータ化したもので、探索語と索引語との完全一致を原則とする単純なマッチングに限定された目録である。米国においてもこの第1世代のオンライン目録を提供する大学図書館が現在でも少なくない。第2世代は、キーワード検索機能が提供されるとともに、ブル演算、トラ

ンケーション、ワイルドカードによるマッチングなどの諸機能を備えた目録であり、現行のオンライン目録でも高水準に属するものである。DIALOG などの商用情報検索システムと同様の機能を有するものと考えていい。第3世代のオンライン目録が先ほど指摘した情報検索研究の成果を取り入れたものであり、以下の特徴をもつ。第1に、自然語から統制語へと誘導する仕組みをもつ。第2に、定式化された質問を構成する語の語幹処理、部分マッチングを可能にする。第3に、適合性 (relevance) 確率に基づいて検索文献をランク付けする。第4に、レバンスフィールドバックの利用により、検索拡張を促進する。

第3世代の目録は、書誌情報、アクセスポイントの種類において根本的な変更を加える必要がないという点で実現可能な改善目標である。しかし、I章で述べたように、これらの方法は、現行の目録システムの基本的な構成要素を維持しつつ、また当面、利用可能な主題関連情報の枠のなかで、研究者が情報探索の改善について理論的、経験的に追求した結果であり、利用者の情報探索行動に関する研究を基礎に提示されているわけではない。

利用者の情報探索行動の特質を反映した情報検索システムについては、現在の検索システム、目録のもつ基本構造からくる制約やその枠組みにとらわれることなく探求する必要がある。次章で取り上げる情報探索の諸研究は、認知科学の研究手法に依拠したものだが、利用者の情報探索行動に新たな知見をもたらすと同時に、新たな方法論の探求にも有効な手がかりを与えるものである。

### III. 情報探索過程と質問定式化

オンライン目録にせよ、雑誌論文の検索を扱う情報検索システムにせよ、文献検索システムを利用した情報探索過程は、第1図に示すような過程を経る。その過程の最初にくるものが、(a) 利用者の情報要求であり、それに基づいて、(b) 探索戦略が構築され、検索可能な質問への定式化が行われる。次に、(c) 探索実行の段階となるが、(d) 出力結果あるいはシステムの応答内容の評価によっては、(e) 探索戦略の見直し、質問の再定式化が必要となる。再定式化された質問を用いて、探索を再実行するが、その結果によっては、再度、質問の定式化が試みられる。この繰り返しの回数は、検索文献数に対するBlairのいう「無用点 (futility point)」と検索文献に対する適合性に関する利用者の判断によって決まることになる。

以下、第1図に示した基本的枠組みにしたがって、認知科学の研究手法を取り入れた情報探索行動の研究を考察する。

#### A. 情報探索過程の全体

ここでは、第1図の a から f に至る情報探索過程の全体を対象とした研究を取り上げ、特に質問定式化という情報探索過程における最も重要な段階に焦点をあてる。

オンライン目録を中心とする文献検索システムの従来利用調査は、既知文献探索と主題探索の利用率、使用された探索語の種類や数というごく限定された範囲を対象としていた。最近になって、文献検索システムにおける情報探索過程全体、さらには文献検索システムの利用を含む情報探索行動をも対象とする調査研究が行われるようになった (Blackshaw & Fischhoff, 1988; Hancock, 1987; Hancock-Beaulieu 1990; Hancock-Beaulieu *et al.*, 1991; Dalrymple, 1990; Chen & Dhar, 1991)。

##### 1. 情報探索過程における目録探索の位置づけ

Hancock (Hancock, 1987) および Hancock-Beaulieu (Hancock-Beaulieu, 1990) は、目録探索は利用者の情報探索行動の一部を構成するものであるとし、書架上の探索をも含めた利用者の探索行動全体を調査の対象とした。この調査の重要な点は、利用者の情報探索行動のある特定の側面のみを切り取るのではなく、情報探索を連続した行動として捉えたことである。こうした視点に立つことで、特定の利用者の目録探索行動は既知文献探索あるいは主題探索のいずれかに分類しうるわけではないこと、さらに、ある時点の探索が既知文献探索であっても、その利用者の情報探索行動全体から見ると、主題探索の一つの方法として試行されたものであることが明らかにされている。

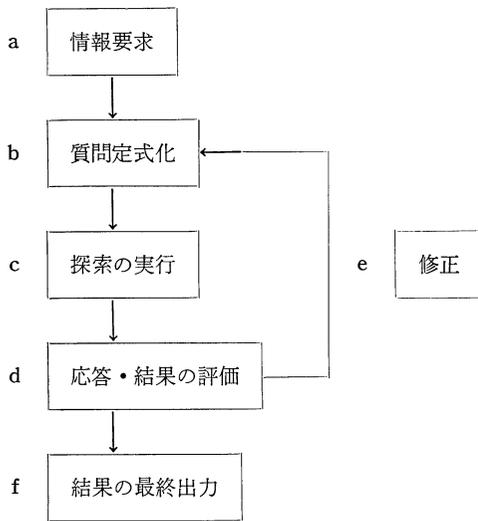
調査はロンドンのシティ大学図書館において、オンライン目録の導入前後の2回に渡って行われた。探索過程のなかで観察された項目は、目録探索の種類 (著者、書名、分類、主題)、アクセスポイント、書架へ行くために利用者が記録した内容、実際に調べた書架番号、書架から取りだした図書の書名である。調査者は、利用者に伴いながら以上の項目について観察、記録するとともに、利用者には探索の際に考えていることを発話するよう求め、そのプロトコルを採取した。

オンライン導入前の探索行動は、著者/書名目録探索、著者/書名中の語を利用した探索、書架上の既知文献の

探索、書架上の主題探索、PRECIS 索引の探索、その他の主題目録探索に分類され、95 の探索事例について、その過程が以上の探索の組み合わせによって記述されている。一方、オンライン導入後の行動は、オンライン目録による著者/書名探索 (既知文献探索)、書架上の既知文献探索、書架上の主題探索、オンライン目録による主題探索に分類され、100 の探索事例について同様の分析が加えられている。その結果によれば、導入前では主題探索を開始する手段として書名の最初の語を利用する事例が 11 あった。既知文献探索のみで終わった事例は導入前が 28%、導入後で 40% であった。既知文献探索として開始され、主題探索へと展開した事例の割合は、導入前では 36%、導入後も 37% と同様の傾向を示しており、そのうちの 74% (導入前)、73% (導入後) は書架上の主題探索であった。

このように、利用者の探索行動は、既知文献探索と主題探索が入り組んで展開されており、特に書架上の主題探索の占める割合が高いことがわかる。このことは、最終的に利用者は原文献にあたるとともに、その関連文献への参照を通じて情報要求を満たす文献を探索、入手することを示している。また、今回の調査で注目される点は、目録における主題探索がオンライン目録導入後に増加していないことである。II 章で述べたように、種々のオンライン目録調査において、主題探索の増加傾向が指摘されてきた。これには2通りの解釈があることは先に指摘したとおりである。Hancock の調査から、オンライン目録導入前においても主題探索への要求は導入後と同程度に存在していたことがわかる。このことから、既知文献探索の形式をとっていた主題探索が、オンライン検索の能力により顕在化したとみるべきで、第1の解釈の妥当性が検証されたといえる。

Hancock の調査でもう一つ注目されるのは、情報要求から質問定式化、再定式化の過程 (第1図の a→b) を詳細に分析した点である。分析枠組みは、(1) 利用者が表現した要求する文献の主題と、最初に使用したアクセスポイント (initial query) との対比 (推移 1)、(2) そのアクセスポイントと、目録の参照を通じて最終的に選択されたアクセスポイントとの対比 (推移 2)、(3) 最初に利用者が表現した文献の主題と、最終的に選択されたアクセスポイントとの対比 (推移 3)、である。対比された語の推移の関係は、一致、上位語への推移、下位語への推移の3通りに分類されている。なお、ここでいう目録とは導入前は冊子体の PRECIS 索引であり、導入後



第1図 情報探索過程

はオンライン目録である。

分析結果によれば、推移1では、オンライン目録探索はPRECIS探索に比べて、求める主題の上位語への推移が少なく、逆に下位語への推移が多かった。その比率をみると、PRECISによる探索では、上位語へと推移した事例が全探索の44%を占め、オンライン目録探索では27%となっている。また、下位語への推移は、PRECIS探索が4%と極端に低いのに対して、オンライン目録探索では15%であった。このように、PRECISによる探索は要求する主題の上位語からのアクセスが多いことが特徴である。一方、オンライン目録探索では、要求する主題の上位語からのアクセスが下位語からのアクセスを上回っているが、その数は3割にも満たない。Hancockは特に注目していないが、要求する主題と探索語が一致した事例は両目録とも半数を超えている。II章で指摘したが、目録探索の大きな問題点の一つは、探索語と索引語とのマッチングの難しさにあることを考えると、この数字は以外な値である。以上の結果は、利用する目録の提供する検索の仕組みが利用者の探索語選択に大きな影響を及ぼすものの、全体的には利用者は求める主題と同じレベルの語またはその上位語からアクセスする傾向があることを示している。

推移2は、質問の再定式化の状況を示すものだが、アクセスポイントを変更した事例は、オンライン目録探索で63%、PRECIS探索では85%であった。その変更の種類は、いずれも下位語への変更であり、それらが約半

数を占めており、一致がほぼ3割、上位語への変更が2割と、きわめて類似した傾向となっており、推移1の結果とは対照的である。利用者は質問の再定式化を、より特定化された語への選択に向けて行うことがわかる。

推移3は、表現された情報要求がどのように変更され、最終的な結果の出力をもたらすアクセスポイントが選ばれるのかを示している。オンライン目録探索では、下位語への変更が44%、上位語への変更が35%を占めていたのに対し、PRECIS探索では逆に上位語への変更が52%、下位語への変更が30%を占めていた。これは、目録の提供する主題探索の仕組みや方法の違いにより、利用者の探索語選択が逆になってしまうことを意味している。理想的には、最後のアクセスポイントは情報要求の記述と同義語であることが望ましい。しかし、一致したのは共に20%程度であり、システムが利用者の要求を表現する語彙を十分に備えていないことを表している。

最後に、第1図のa→bとeとの関係を扱った、結果の出力と実際に書架上で利用者が調べた図書についてみていこう。出力した文献の実際に書架上で調査した図書点数の割合は、オンライン目録探索では情報要求の主題と同じレベルのアクセスポイントをまず選択し、探索を開始した事例では54%を占め、上位語を選んで探索を始めた事例では38%、下位語から探索を開始した事例では8%であった。しかし、調べた図書のなかから実際に選択した図書点数の割合をみると、下位語から探索を始めた事例では89%にのぼるのに対して、一致した語から探索をはじめた事例では26%、上位語から始めた事例では29%にすぎなかった。一方、PRECIS探索では、下位語から開始した探索における図書の選択率は25%にすぎず、オンライン目録での89%という値とは対照的であった。このことは、最初の質問定式化の戦略が同じでも、目録の仕組みが質問の再定式化に影響を及ぼし、その結果、最終的に検索され、選択される文献に違いが出てくることを示している。

Hancockの調査は、図書館での情報探索過程における目録の位置づけと、目録の仕組みや検索方法が探索行動に及ぼす影響を明らかにしようとした点で注目される。特に、情報要求の表現から質問定式化への推移を分析し、上位語、下位語への展開が実際に利用者によって試みられていることを実証した点で評価される。

目録の形態と利用者の情報探索行動との関係について扱った調査には他に、Dalrymple (Dalrymple, 1990) の

ものがある。この調査は、情報探索行動の理論モデルの構築を目指したもので、書誌データベースの探索行動について、人間の想起 (remembering) に関する理論である「再定式化による検索 (retrieval by reformulation)」からの説明を試みている。ただ、分析内容には、その理論による解釈が十分に反映されていない。しかし、情報探索行動をめぐる問題が質問定式化という側面に集約されることに着目し、定式化の問題に正面から取り組もうとした点でこの調査は興味深いものとなっている。

Dalrymple の問題設定は、Hancock のそれと共通している。すなわち、カード目録とオンライン目録という形態の違いが、探索行動、特に質問の定式化に関していかなる影響を及ぼすのかという点である。Hancock の調査では、具体的な要求をもって図書館を訪れた利用者を対象としており、調査にあたっては統制を加えていない。それに対して Dalrymple は、実験というかたちで調査を実施している。被験者はウィスコンシン大学の学生 40 名で、カード目録を探索するグループとオンライン目録を探索するグループとに分けられている。Dalrymple の説明によれば、カード目録とオンライン目録の記述は同一であり、使用可能なアクセスポイントも同じである。違いは、オンライン目録側のもつ、ブール演算等の検索能力とフィードバックの程度にある。オンライン目録は多様な検索が可能であり、入力された語に対してその結果を即座に表示するという点でフィードバックの程度が大きいものとして位置付けられている。

被験者には、五つの問題が与えられ、それぞれの問題を解決するためにはどんな情報が必要かをまず記述させ、指定された目録を使って探索させている。この調査の焦点は質問の再定式化にあるが、ここでいう再定式化とは、目録探索に使用された語の変化であり、その過程を利用者の探索時の認知過程として Dalrymple はみなしている。再定式化の段階を、二つ設定している。一つは、第 1 図の a→b に対応する探索前の再定式化 (pre-search reformulation) であり、これは必要な情報の記述から実際に目録を探索する際に使用された語への変化を指す。次の段階は、第 1 図の e の経路に対応する探索中の再定式化 (search reformulation) であり、最初に使用された語から探索中に変更、修正された語への変化を指す。このように、Dalrymple の分析枠組みは Hancock のそれと共通しており、再定式化のデータとして被験者のプロトコルを使用している点も同じである。しかし、Dalrymple の調査では、再定式化の回数、

すなわち探索語の変更、修正回数に重点がおかれ、それらと目録形態との関係が分析されているだけで、それらの語の意味の関係については残念ながら分析されていない。

この調査の仮説は、オンライン目録を探索する被験者のほうが、(1) より多くの文献を検索し、(2) 探索結果に満足し、(3) 探索により多くの時間を費やし、(4) 再定式化の頻度も高いというものである。この内容は、オンライン目録の検索能力からみて妥当な予測といえる。しかし、検証されたのは、(3) と (4) のみであった。オンライン目録の被験者はその操作を難しく感じ、より多くの文献の検索を試みることはなく、適切な語を見つけることがきわめて難しいと回答していた。再定式化の頻度は、探索中ではオンライン目録の被験者のほうが高かったが、探索前の段階では差異がなかった。このことは、探索中の再定式化の頻度の高さは、オンライン目録検索の難しさに一因があることを示している。全体として、カード目録の被験者のほうが探索結果に満足しており、当初の仮説とは相反する結果が出ている。

Dalrymple は、情報探索における再定式化の性質について今後探求すべき課題をいくつかあげている。

第 1 に、期待との関係である。探索結果への期待の高さは、再定式化の頻度を高めるのか。

第 2 に、再定式化は、探索結果への不満の表れなのか。

第 3 に、再定式化は利用者によって学習される行動なのか。

第 4 に、再定式化には、情報要求の性質や緊急性、サーチャーの個人差が影響を及ぼすのか。

ここでは、第 3、と第 4 点について触れておきたい。

第 3 点については、Fidel (Fidel, 1986, 1991a, 1991b, 1991c) の調査が一部扱っている。Fidel によれば、サーチャーは情報要求の意味を変えないで文献集合の大きさなどの操作的修正を指向する者と、情報要求の意味を変化させる概念的修正を指向する者に半々に分けられるとし、検索修正のタイプはサーチャーの個人的なスタイルであると結論づけている。再定式化の範囲を、操作的修正にまで拡大するかどうかは問題であるが、Fidel の分析は、再定式化は学習される行動というよりも、利用者の個人的特性に属するものであることを示している。ただし、Fidel の調査対象は経験のあるサーチャーであり、必ずしも検索に習熟してはいないオンライン目録の一般利用者にその結果があてはまるかどうかは検討する必要がある。

第4の情報ニーズの性質がなにを指すかは様々である。たとえば、Ingwersen & Wormell (Ingwersen & Wormell, 1988) は、情報要求のタイプを次の三つに分けている。(1) 照合的な (verificative) 情報要求。既知の資料の照合または所在確認。(2) 主題の明確な情報要求。情報要求を表現するのに必要な概念や語が利用者の知識構造の中にあること。(3) 主題の曖昧な情報要求。利用者が情報要求の表現に必要な概念や語に関する知識構造をもっていないこと。情報要求の性質をこのように捉えるならば、それらは再定式化と密接な関係が出てくる。Belkin (Belkin 1980, 1982) は、利用者の知識の変則的な状態が情報要求を引き起こすとする仮説を提示しているが、それによれば、(2) と (3) は変則性の程度の問題である。変則性の程度が小さければ、情報を必要とする問題領域についてある程度の知識を有しており、適切な用語の選択が可能な状態にある。Blair (Blair, 1986) は、主題探索に影響を及ぼす不確定要因として、索引語選択の不確定性 (索引者間の非一貫性) と探索語選択の不確定性をあげている。探索語選択にあたって抄録や書名との照合を想定するならば問題領域の知識がある程度必要であり、このことは検索対象文献の著者の知識や用語を利用者はある程度予測する必要があることを意味する。さらに、索引者が付与する索引語への予測も求められる。このように、質問定式化には、利用者の知識状態が関係すると同時に、著者、索引者の知識や用語への配慮が必要となるため、難しさと複雑さが生じるのである。

質問定式化は、利用者や仲介者としてのサーチャーの作業ではあるが、その過程をいかに支援するかは、検索システム側の問題となる。Dalrymple は質問の再定式化過程は利用者の認知過程であるとしたが、より正確に言えば、システムが設定する範囲においてとりうる認知過程である。利用者は、システム側が用意する定式化の手段の範囲においてのみ、定式化が可能であり、その意味では、システムのもつ制約に常につきあたる。たとえば、キーワード検索やトランケーション機能を提供するシステムとそうではないシステムとでは、定式化の作業は大きく異なる。後者では、LCSH との完全一致が求められることになり、定式化の条件は極端に限定されることになる。そこで、Dalrymple (Dalrymple, 1990) の言葉を借りるならば、利用者による再定式化の自然な行動を取り入れた利用者支援システム (データベースインターフェース) の設計が求められることになる。そのため

には、システムの制約を可能な限り取り払ったなかでの人間の情報探索行動の調査、分析の可能性を検討する必要がある。システムの制約を考えず、利用者の認知、記憶の特徴から情報探索行動の分析を試みた Allen の研究についてはB節で取り上げる。

ところで、利用者の探索特性を反映したシステム設計の研究は、ややもすれば、探索過程のあらゆる部分の自動化を指向しがちである。こうした研究方向は、利用者指向といいつつ、たとえ利用者の探索行動の研究に依拠しているとはしても、実はその研究で明らかにされた範囲において設計者が想定しうる検索パターンに利用者適合させるやり方である。Bates (Bates, 1990) が指摘するように、利用者支援のシステム設計にあたっては、自動化すべき部分と利用者の決定に委ねる部分を設け、情報探索という作業を人間とシステムとの間の分業かつ協調的作業とする必要がある。

## 2. 問題解決過程としての情報探索

利用者は、探索語の選択に始まり出力の評価にいたるまで、様々な意思決定を迫られている。その意味では、利用者は情報探索という問題に取り組んでいることになる。情報探索は、何らかの問題の処理や解決のために必要な情報を求める行動であるが、情報探索という行動そのものも一つの問題解決行動としてみることができる。利用者がまずもって意思決定を迫られ、解決しなければならぬのは、情報要求を検索質問としていかに定式化し、探索戦略を構築するかということである。探索語の選択と決定に関してサーチャーが用いる意思決定ルールについては、Fidel (Fidel, 1986, 1991a, 1991b, 1991c) の一連の研究が存在する。

Chen (Chen, 1991) は、情報検索の主たる問題点は検索質問を表現する語句と探索戦略の選択における自由度から生じると述べ、利用者が選択する質問定式化の手段の多様性が問題を引き起こすとした。利用者が選択可能な探索戦略については、Bates (Bates, 1979a, 1979b) の研究があり、29の情報探索戦略と、探索時の17の思考戦略 (idea tactics) が提示されている。問題解決の理論からいえば、これらの戦略は、目標状態に近づけるために、現在の問題状態を遷移させるうえで使用される操作子にあたる。情報探索過程においては、情報要求を満たす文献の検索が目標状態である。そこで、利用者は、その目標状態にいたるために、まず、探索語の選択という質問定式化の問題を下位問題として設定し、その解決にあたる。それが解決されたならば、問題状態が目標状態

に一步遷移したことになる。Blackshaw & Fischhoff (Blackshaw & Fischhoff, 1988) は、こうした問題解決過程としての情報探索過程において、利用者がいかなる事柄についてどのような意思決定を行っているかを、意思決定理論に基づいて分析している。先ほどの Bates (Bates, 1979a, 1979b) は、様々な探索戦略を示しているが、それらが実際の探索過程においていかに使用され、また、どのような探索戦略の変更が生じるのかを明らかにしなかった。

第1図の b→e の過程にあたる、利用者が使用する探索戦略とその変更を分析したのが Chen (Chen, 1991) である。Chen は、ニューヨーク大学オンライン目録システム利用時の利用者と図書館員とのやり取り (34 件) を観察し、さらに、図書館員の援助を受けずに自らが選んだ主題に関する文献の探索を行った利用者のプロトコルとトランザクションログファイルを 30 件収集した。このシステムは、書名検索、著者名検索、著者・書名検索の組み合わせ、主題検索、キーワード検索、ブール検索の諸機能を提供している。分析の結果、情報探索過程において実際に使用された探索戦略は次の五つにモデル化されている。

第1は、既知文献利用戦略である。これは、適合文献の探索に既知文献を利用するもので、検索された既知文献に付与された索引語を使って主題探索を行う戦略である。この戦略は、Blair (Blair, 1986) によって推奨されている。Blair は、探索語や索引語の不確定性を扱う伝統的な方法であるレレバンスフィードバックは有効でないとして退けている。というのは、それほど適合していない文献だけが検索された場合、それをもとに急に検索式を修正しても同じような文献しか検索されないという悪循環に陥るからである。そこで、探索語選択の不確定性を小さくするきわめて有効な方法として、非常に有用な文献、すなわち既知文献から探索を開始し、その有用な文献に付与されている索引語を使用して質問の再定式化を行う方法が提案されている。Chen の調査では、利用者は探索する主題領域に関係する著者名を使って検索し、次にその文献に付与された索引語を使用するという戦略をとっていることが明らかにされている。

第2は発見的 (heuristic) 戦略である。これは、統制語による主題探索や、書名検索、統制語および書名のキーワード検索を使用するもので、図書館員やシステムに習熟している利用者がとる戦略である。

第3はソーラースブラウジング戦略である。図書館員

が利用者を援助する場合によく使用する戦略である。

第4は、画面ブラウジング戦略である。これは、探索語とアルファベット順配列上近接する索引語を表示させ、適切な語を選択するものである。

第5は、試行錯誤戦略である。これは、利用者が思いついた語を使用したり、画面に表示された手がかりさえも使用しないやり方で、戦略という表現があてはまらないようなものである。

利用者は以上の戦略のうち平均して二つから三つの戦略を使用し、なかでも第4、第5の戦略をよく使用していた。大多数の図書館員と探索経験のある利用者は、第1、第2、第3の戦略を使用していた。また、質問の再定式化にあたっては、第1、第3の戦略が有効であったという。なお、Chen は分析結果を踏まえた知的検索システムの設計を試みている。

Chen の研究の意義は、調査結果以上に、文献検索システムにおける情報探索行動は問題解決過程として記述することが可能であり、実際に問題解決研究の知見を取り入れて分析した点にある。探索戦略は、問題の初期状態から目標状態に遷移する際に解決者のとるアプローチであり、操作子の系列として位置付けられている。また、使用する探索語をその利用者の知識状態の指標と見なしている。情報を必要としている主題領域の専門家と初心者の違いが、質問の定式化、再定式化の違いとなって現れることに着目しえたのも、問題解決研究の成果を踏まえたからである。

## B. 情報要求と質問定式化

情報探索にあたって、どのような探索語が選択され、どのように質問として定式化されるかは、Chen (Chen, 1991) が示したように、利用者の知識状態に依存する。ここでいう知識には二つある。一つは、探索しようとしている問題領域について利用者がもつ知識である。もう一つは、検索システムに関する知識であり、データベース、探索戦略、件名標目やソーラース、システムの操作に関する知識である。情報探索は利用者の情報要求から生じる。探索の出発点はその情報要求の記述にあるとすれば、そこには前者の知識状態が関係する。情報要求は問題領域に関する利用者の知識の変則性、知識におけるギャップから生まれるため、主題を中心とする情報要求の記述に明確さは期待できない (Belkin, 1980, 1982; Dervin, 1983, 1986)。これは、変則性の程度が情報要求の記述とその明確さを規定することを意味している。質

問定式化はこの情報要求の記述を基に行われるのであるから、利用者の問題領域に関する知識の程度は、質問定式化に対して大きな影響を及ぼすことになる。

しかし、情報要求の記述が明確であっても、少なくとも現行の検索システムのもつ能力を前提にする限り、システムに関する知識がなければ、その記述は質問に反映されない。たとえば、Bates (Bates, 1986) の提案するエンドユーザソーラスをシステムに組み込むことによって、利用者を探索語から統制語へと誘導するか、探索語と統制語を自動的にリンクする仕組みがシステムに備わらない限り、利用者の側に統制語とその使用に関する知識が求められる。仲介者としてのサーチャーは、そうした利用者に求められる知識を有することにより、利用者に代わって探索にあたることになる。なお、そのような知識については、Fidel (Fidel, 1984, 1991a, 1991b, 1991c) の一連の調査が対象としている。

この節では、利用者に関わる要素と質問定式化との関係を扱った研究を取り上げる。

#### 1. 情報要求の記述と質問定式化

これまで取り上げた諸研究は、いずれも、既存の検索システムを前提としており、そのシステムの枠組みのなかでの利用者の情報探索行動を分析している。ここでいう既存の検索システムとは、索引語およびフリーキーワードによる検索を中心に、ブール演算、キーワード検索トランケーション等の機能を有するものである。また、書誌情報の基本的な要素を変更せず、書誌情報へのアクセスの仕組みも語を単位とするシステムである。たとえば、これまで取り上げてきた調査に使用されたオンライン目録には、収録対象が図書という性格上、抄録は含まれてはいない。また、雑誌論文を対象とする文献検索システムは、その多くが抄録を含んではいないが、抄録中の語への単なるアクセスを提供しているだけである。SAPのプロジェクト (Atherton, 1977) は、図書の目次や巻末索引を索引語として付加しようとするものであったが、これは索引語の種類と数を増やしたにすぎず、アクセスの方法は従来通りの索引語とのマッチングを前提としている。利用者指向といわれる検索システムの設計を考えるならば、既に指摘したように、既存の検索システムのもつ制約条件を取り払ったなかで、利用者の情報探索行動の特性を探求する必要がある。

Allen の一連の研究 (Allen, 1988, 1989, 1990a, 1990b, 1991) は、第1図の a→b、すなわち情報要求の記述と質問定式化との関係を扱っている。Allen は、既存のシス

テムの制約にとらわれることなく、利用者の情報探索行動を人間の認知構造に関する研究を基に探求しており、全く新しい考え方に立った検索システムを示唆している。検索システムの利用者が求めるのは文献である。文献検索という行為は文献中の様々な要素を手がかりにして求める文献を探したものである。利用者が文献中のいかなる要素に着目し、どのようにその内容を捉えているかを明らかにすることは、利用者の最も自然な文献探索行動に適した仕組みを考えるうえできわめて重要である。

では、利用者、すなわち文献の読者は、文献の内容についてどのように記憶し、再生 (recall) するのであろうか。人間は、文献の内容を理解し、再生する際にあるスキーマを用いていることがわかっている。Allen はこのスキーマに着目した。そのスキーマとは、テキストの上部構造 (superstructure) を表したレポートスキーマであり、Kintsch と van Dijk (Kintsch & van Dijk, 1978) によって提示されたものである。Allen が指摘するように、このレポートスキーマは、文献の主題内容を扱っていないために、これまで情報検索システムの研究に応用されることがなかった。Allen の研究は、索引語や統制語という文献の主題内容を表した語を専ら検索の手がかりとしてきた従来の検索システムとは全く異なる発想から出発しているのである。

次のような実験が行われている。被験者 44 名に、心理学関係の論文を読ませ、その後三つのグループに分け、論文の内容について記述させた。第1のグループには、論文について再生できることならすべて記述するよう求めた。第2のグループには、論文の著者、標題、論文の内容を表すキーワードについて記述させた。第3グループには、レポートスキーマから取った次の要素について記述させた。(1) 著者の論文作成の目的、(2) 著者が調査に使用した方法、(3) その方法を通して得られた結果、到達した結論、である。以上の3種類の記述中の語と、論文の標題、抄録中の語 (索引語) との照合が行われた。その結果、索引語との一致数は、第1、第3グループが第2グループの2倍以上であった。次に、論文中の命題と、第1、第3グループの記述中の命題との照合が行われた。それによれば、一致した命題はほぼ同数であったが、一致しない命題数は第3グループが大きく下回った。

以上の結果から、Allen は、レポートスキーマという文献の構造を手がかりにした情報要求の記述とそれに基

づいた質問定式化が、抄録を利用した文献探索にとって有効であると結論づけている。Allen の研究は、抄録がレポートスキーマを基に記述されていることを間接的に明らかにしたともいえ、単に抄録を検索の手がかりとした場合の探索語選択の方略を示したものにすぎないともいえる。また、抄録が常にレポートスキーマを基に記述されているわけではなく、分野によってもその構造が異なることを考えれば、直ちに Allen の研究成果を検索システムに応用することは難しい。しかし、検索手段としての抄録を、現行の検索システムのように単にそこに含まれている語と探索語との照合に利用するのではなく、利用者の文献内容再生の特性を反映した構造にその価値を見いだそうとした点で、Allen の研究の意義は大きい。

また、Allen (Allen, 1991) は、情報探索に使用する探索語の選択、質問定式化に利用者の主題知識が及ぼす影響についても考察している。この研究はイリノイ大学図書館のオンライン目録を使って行われたが、利用者は、主題知識レベルが高いほど、検索結果に満足しない場合に新たな検索戦略を用い、質問の再定式化を多く試みていた。

## 2. 利用者の問題への認識と質問定式化

情報要求の記述、および質問定式化に及ぼす影響は、利用者の主題知識だけではない。より正確に言うならば、主題知識の状態が確かにその基底にあるのだが、様々な要素が質問定式化にあたっては考慮されなければならない。Belkin (Belkin, 1983) は、仲介者としてのサーチャーが、プレサーティンタビューを通して利用者から把握しようとする諸要素を明らかにしている。それらの要素は、利用者の抱える問題状況（情報を必要とする意図や目的）、問題の構造や主題、利用者の個人的属性（利用者モデル）の三つに分けられている。Belkin によれば、主題知識の状態が変則であるために、問題状況が認識され、その変則性によって問題の構造や主題が決まることになる。Belkin のあげた要素とほぼ同じものを、Taylor (Taylor, 1968) は、図書館員と利用者との相互交渉過程において図書館員が質問定式化のために把握することが有用な要素（フィルター）としてあげている。しかし、Belkin らの提示した要素が実際に、質問定式化にどう反映され、さらには情報探索行動全体に及ぼす影響を及ぼすかは明らかにされてはいない。

主題からのアクセスに限定されている現行のシステムでは、質問定式化も情報要求の主題がその対象となるた

め、主題以外の要素をどのように質問定式化に反映させるかは大きな問題となる。そうした限界の中で、Sarcevic ら (Sarcevic, *et al.*, 1988; Sarcevic & Kantor, 1988a, 1988b) は、利用者に関わる変数が情報探索行動に及ぼす影響について定量的調査を試みている。この調査は、第1図の a→f の全過程を扱った大規模なものだが、ここで特に取り上げたいのが、a→b の過程に関する変数が d の応答・結果の評価に及ぼす影響である。Sarcevic らが設定した変数とは、問題の認識の明確さ、情報探索意図の明確さ、知識状態についての評価、公的知識（すなわち文献）に対する期待度の四つである。その結果によれば、利用者が問題を明確に認識している場合、検索結果の適合率は有意に高くなった。また、利用者の公的知識への期待度が高いと、適合率と精度が有意に高くなった。このことは、利用者の問題への認識や、検索システム中の文献から回答が得られるとの期待は適切な質問定式化をもたらすことを意味している。

Sarcevic らの調査は以下の点で問題がある。すなわち、利用者の問題の明確さの基準は何かということである。明確さの判断を利用者自身に委ねているが、複数の被験者を対象とした場合、その判断基準に一貫性が保たれる保証はないので、明確さを判定する項目を規定しておく必要がある。たとえば、そのような項目として、問題の初期状態、目標状態、初期状態から目標状態に至る道筋、とりうる手段などの諸項目をあげることができるだろう。今後の研究課題として重要なことは、基本的に現行の検索システムが主題からのアクセスしか提供しえないなかであって、問題や意図の明確さへの認識の程度や知識状態の変則性の程度が、実際に質問の定式化、再定式化にどのように反映されるのか、具体的には探索語の選択や探索戦略の構築にどのような違いが見られるかを考察することである。同時に、利用者の意図や、利用者の問題状況に関する情報を生かすような検索システムが果たして可能かどうか今後検討すべき重要な課題である。

## IV. 文献の構造と質問定式化

現行の文献検索システムにおける質問定式化は、語およびその組み合わせによって行われている。様々な探索戦略も、語と語の間の上位、下位、関連の諸関係を使用した探索の拡張、限定を基本としたものである。情報検索研究から取り入れたオンライン目録の改善策も、こうした質問定式化の仕組みのうえに成り立っている。

探索対象である文献は、ある一定の構造と枠組みをもった情報として利用者あるいは読者によって記憶、想起されることが、テキスト言語学や心理学の研究から明らかにされている (Kintsch & van Dijk, 1978; Beaugrande & Dressler, 1984)。しかし、その構造を生かした検索システムは現在のところ開発されていない。一定の構造や枠組みをもつ文献内容を、語およびその組み合わせによる表現のレベルに還元することは文献のもつ構造上の特質を捨てることになる。先の Allen の調査が示すように、利用者の側は、求める文献についてその構造に対応した検索の手がかりをシステムに提供できる状態にある。

文献の構造とは、論文であればその論文の構成要素とそこに実際に含まれる情報とを切り離すことなく結び付けることによって構築される全体のことを指している。文献の全文 (フルテキスト) を今述べたようにして再構成することが文献の構造化になる。抄録は文献の構造が圧縮されたものとして、文献の全文の代替としての利用が可能であり、雑誌論文を対象とする現行の検索システムの多くに既に含まれている。そこで、文献の構造を抄録が反映していると仮定したうえで、文献の構造を生かした質問定式化の可能性について以下検討する。図書を中心としたオンライン目録システムの場合、図書の全文はもちろんのこと、その抄録を含むようなシステムはこれまで存在しないが、今後の可能性として、抄録レベルの情報を含む検索システムを想定することにしよう。

文献の構造を反映した検索システムと質問定式化の可能性を示唆した研究に、Liddy (Liddy, 1991) の抄録に関する研究がある。この研究は、抄録作成者がメンタルモデルとして構築している論文の抄録構造の抽出を試みたものである。対象は実証研究を扱った論文であるが、分析の結果、抄録の階層構造と七つの要素からなる基本構造が明らかにされている。また、各要素に頻出する語彙についても集計、分析している。これらの構造と構成要素は、抄録作成者が頭のなかに持っているものであり、それが実際の抄録の構造をよく反映していることが見いだされている。文献に対して利用者が抱く構造と抄録作成者が抱く構造とが類似するものと仮定すれば、抄録構造は利用者のもつ文献構造を反映しているとみなすことができる。

Liddy は、抄録の構成要素をフレームとする抄録に関するフレームシステムを示している。文献の構造がこのフレームシステムにしたがって記述されるならば、抄録

中の語彙とそれが含まれている要素とを組み合わせる検索の手がかりとするシステムが可能となる。情報を求めている主題領域にあまり精通していない利用者にとって探索語の選択と質問定式化は難しい。利用者が関心のある主題領域の概念のもつ役割を質問として定式化できない場合、そのような検索システムはその概念が含まれる抄録中の種々の構成要素と各要素の抄録構造上の位置づけを示すことによって、利用者の質問再定式化を支援することができる。文献の構造を取り入れたシステムは、選択した探索語が文献中においてもつ意味を明示できるという点で、従来のシステムに比べて、適切な質問定式化を可能にし、検索の精度を高めることが期待される。

## V. 結

文献検索システム、特にオンライン目録システムにおける情報探索行動とそこでの質問定式化をめぐる諸問題を扱った論文を取り上げ考察してきた。研究の方向は、検索システムがいかに利用者の情報探索行動を規定し、質問の定式化の方法を決定するのかを探求する段階から、利用者の情報探索行動の特質を取り入れ、利用者による選択の幅を広げるシステムを探求する段階へと展開している。

現在の検索システムは、利用者のすべてに一律の検索機能を提供することを前提に構築されている。もちろん、そのなかでどの機能を選択、活用するかは利用者にかまされてはいるが、利用者の個別性を考慮したインタフェースの設計は行われていない。オンラインネットワークの発達により、現在では、個々の利用者が端末を独占的に使用し、探索することが可能である。例えば、大学図書館という場でのみ端末が利用可能であった時代は過ぎ去ろうとしている。こうした環境の変化に伴い、個々の利用者の要求や嗜好さらには探索特性を学習するユーザモデリング (user modeling) の能力を備えたシステムの設計が求められる (Croft & Thompson, 1987)。

文献検索システムにおける利用者の情報探索行動については、認知科学の手法と成果を取り入れた研究が開始され、新たな知見が得られつつある。利用者の情報探索の特性を十分に取り入れたシステムを設計するためにも、認知科学の手法と成果に基づいた情報探索行動に関する本格的な研究が必要である。

Allen, Bryce. Text structures and the user-intermediary interaction. RQ. Vol. 27, No. 4, p. 335-

- 341 (1988).
- Allen, Bryce. Recall cues in known-item retrieval. *Journal of the American Society for Information Science*. Vol. 40, No. 4, p. 246-252 (1989).
- Allen, Bryce. Knowledge organization in an information retrieval task. *Information Processing & Management*. Vol. 26, No. 4, p. 535-542 (1990a).
- Allen, Bryce. The effects of academic background on statements of information need. *Library Quarterly*. Vol. 60, No. 2, p. 120-138 (1990b).
- Allen, Bryce. Topic knowledge and online catalog search formulation. *Library Quarterly*, Vol. 61, No. 2, p. 188-213 (1991).
- Atherton, P. Books are for use: final report of the Subject Access Project to the Council on Library Resources. Syracuse University School of Information Studies. 1977. (ERIC Document Reproduction Service No. ED156 131). 190 p.
- Bates, Marcia J. Factors affecting subject catalog search success. *Journal of the American Society for Information Science*. Vol. 28, No. 3, p. 161-169 (1977a).
- Bates, Marcia J. System meets user: problems in matching subject search terms. *Information Processing & Management*. Vol. 13, No. 6, p. 367-375 (1977b).
- Bates, Marcia J. Information search tactics. *Journal of the American Society for Information Science*. Vol. 30, No. 4, p. 205-214 (1979a).
- Bates, Marcia J. Idea tactics. *Journal of the American Society for Information Science*. Vol. 30, No. 5, p. 280-289 (1979b).
- Bates, Marcia J. Subject access in online catalogs: a design model. *Journal of the American Society for Information Science*. Vol. 37, No. 6, p. 357-376 (1986).
- Bates, Marcia J. Where should the person stop and the information search interface start? *Information Processing & Management*. Vol. 26, No. 5, p. 575-591 (1990).
- Beaugrande. Robert-Alain de and Dressler, W.U. テクスト言語学入門. [Introduction to Text Linguistics.] 池上嘉彦 ほか共訳. 紀ノ国屋書店. 1984. 348 p.
- Belkin, N.J. Anomalous state of knowledge as a basis for information retrieval. *Canadian Journal of Information Science*. Vol. 5, p. 133-143 (1980).
- Belkin, N.J. ASK for information retrieval: Part 1. background and theory. *Journal of Documentation*. Vol. 38, No. 2, p. 61-71 (1982).
- Belkin, N.J. Cognitive models and information transfer. *Social Science Information Studies*. Vol. 4, Nos. 2/3, p. 111-130 (1984).
- Blackshaw, L. and Fischhoff, B. Decision making in online searching. *Journal of the American Society for Information Science*. Vol. 39, No. 6, p. 369-389 (1988).
- Blair, D.C. Searching biases in large interactive document retrieval systems. *Journal of the American Society for Information Science*. Vol. 31, No. 4, p. 271-277 (1980).
- Blair, D.C. Indeterminacy in the subject access to documents. *Information Processing & Management*. Vol. 22, No. 2, p. 229-241 (1986).
- Blazek, R. and Bilal, D. Problems with OPAC: a case study of an academic research library. *RQ*. Vol. 28, No. 2, p. 169-178 (1988).
- Borgman, C.L. Online monitoring of users of the Ohio State University online catalog: methodological issues and results. In: A.E. Petrarca et al ed. *Information Interaction: Proceedings of the 45th ASIS Annual Meeting, Columbus, Ohio October 17-21, 1982*, p. 35-36. New York, Knowledge Industry Publications, 1982.
- Borgman, C.L. End user behavior in an online information retrieval system: a computer monitoring study. In: J.J. Kuehn ed. *Research and development in information retrieval: Proceedings of the Sixth Annual International ACM SIGIR Conference*, p. 162-176. Baltimore. Association for Computing Machinery, 1983.
- Byrne, A. and Micco, M. Improving OPAC subject access: the ADFA experiment. *College and Research Libraries*. Vol. 49, No. 5, p. 432-441 (1988).
- Chan, L.M. Library of Congress Classification as an online retrieval tool: Potentials and Limitations. *Information Technology and Libraries*. Vol. 5, No. 3, p. 181-192 (1986).
- Chan, L.M. Library of Congress class numbers in online catalog searching. *RQ*. Vol. 28, No. 4, p. 530-536 (1989).
- Chen, H. and Dhar, V. Cognitive process as a basis for intelligent retrieval systems design. *Information Processing & Management*. Vol. 27, No. 5, p. 405-432 (1991).
- Cochrane, P.A. and Markey, K. Catalog use studies—since the introduction of online interactive catalog: impact on design for subject access. *Library and Information Science Research*. Vol. 5, No. 4, p. 337-363 (1983).
- Cochrane, P.A. and Markey, K. Preparing for the use of classification in online cataloging systems and online catalogs. *Information Technology and Libraries*. Vol. 4, p. 91-111 (1985).
- Croft, W.B. and Thompson, R.H. I<sup>3</sup>R: a new approach to design of document retrieval systems.

- Journal of the American Society for Information Science. Vol. 38, No. 6, p. 389-404 (1987).
- Dalrymple, P. W. Retrieval by reformulation in two library catalogs: toward a cognitive model of searching behavior. *Journal of the American Society for Information Science*. Vol. 41, No. 4, p. 272-281 (1990).
- Dervin B. An overview of sense-making: concepts, and results to date. Seattle, School of Communication. University of Washington, 1983, methods 72 p.
- Dervin B. and Dewdney, P. Neutral questioning: a new approach to the reference interview. *RQ*. Vol. 25, No. 4, p. 506-513 (1986).
- Doszko, J. CITE NLM: natural language searching in an online catalog. *Information Technology and Libraries*. Vol. 2, p. 364-380 (1983).
- Dykstra, M. LC subject headings disguised as a thesaurus. *Library Journal*. Vol. 113, No. 4, p. 42-46 (1988).
- Fidel, R. Towards expert systems for the selection of search keys. *Journal of the American Society for Information Science*. Vol. 37, No. 1, p. 37-44 (1986).
- Fidel, R. Searchers' selection of search keys: I, the selection routine. *Journal of the American Society for Information Science*. Vol. 42, No. 7, p. 490-500 (1991a).
- Fidel, R. Searchers' selection of search keys: II, controlled vocabulary or free-text searching. *Journal of the American Society for Information Science*. Vol. 42, No. 7, p. 501-514 (1991b).
- Fidel, R. Searchers' selection of search keys: III, searching styles. *Journal of the American Society for Information Science*. Vol. 42, No. 7, p. 515-527 (1991c).
- Frost, C. O. Faculty use of subject searching in card and online catalogs. *Journal of Academic Librarianship*. Vol. 13, No. 2, p. 86-92 (1987).
- Gerhan, D. R. LCSH in vivo: subject searching performance and strategy in the OPAC era. *Journal of Academic Librarianship*. Vol. 15, No. 2, p. 83-89 (1989).
- Hancock, M. Subject searching behaviour at the library catalogue and at the shelves: implications for online interactive catalogues. *Journal of Documentation*. Vol. 43, No. 4, p. 303-321 (1987).
- Hancock-Beaulieu, M. Evaluating the impact of an online library catalogue on subject searching behaviour at the catalogue and at the shelves. *Journal of Documentation*. Vol. 46, No. 4, p. 318-338 (1990).
- Hancock-Beaulieu, M. et al. Evaluation of online catalogues: eliciting information from the user. *Information Processing & Management*. Vol. 27, No. 5, p. 523-532 (1991).
- Hildreth, C. R. Pursuing the ideal: generations of online catalogs. In: B. Aveney and B. Butler ed. *Online catalogs, online reference: covering trends*. Chicago, American Library Association, 1984, p. 31-57.
- Hildreth, C. R. Beyond Boolean: designing the next generation of online catalogs. *Library Trends*. Vol. 35, No. 4, p. 647-667 (1987).
- Huestis, J. C. Clustering LC classification numbers in an online catalog for improved browsability. *Information Technology and Libraries*. Vol. 7, p. 381-393 (1988).
- Hunter, R. N. Successes and failures of patrons searching the online catalog at a large academic library: a transaction log analysis. *RQ*. Vol. 30, No. 3, p. 395-403 (1991).
- Ingwersen, P. and Wormell, I. Means to improved subject access and representation in modern information retrieval. *Libri*. Vol. 38, No. 2, p. 94-119 (1988).
- Kern-Simirenko, C. OPAC user logs: implications for bibliographic instruction. *Library Hi Tech*. Vol. 1, p. 27-35 (1983).
- Kintsch, W. and van Dijk, T. A. Toward a model of text comprehension and production. *Psychological Review*. Vol. 85, p. 363-394 (1978).
- Lancaster, F. W. *The measurement and evaluation of library services*. Washington, D. C., Information Resources Press, 1977. 395 p.
- Larson, R. R. *Workload characteristics and computer system utilization in online library catalogs*. Doctoral Dissertation. University of California at Berkeley, 1986. 417 p.
- Larson, R. R. Managing information overload in online catalog subject searching. In: *Managing information and technology: Proceedings of the 52nd Annual Meeting of the American Society for Information Science*, p. 129-135. Medford. Learned Information, 1989.
- Larson, R. R. Between Scylla and Charybdis: subject searching in the online catalog. *Advances in Librarianship*. Vol. 15, p. 175-236 (1991a).
- Larson, R. R. The decline of subject searching: long-term trends and patterns of index use in an online catalog. *Journal of the American Society for Information Science*. Vol. 42, No. 3, p. 197-215 (1991b).
- Larson, R. R. Evaluation of advanced retrieval techniques in an experimental online catalog. *Journal of the American Society for Information*

- Science. Vol. 43, No. 1, p. 34-53 (1992).
- Larson, R.R. and Graham, V. Monitoring and evaluating MELVYL. *Information Technology and Libraries*. Vol. 2, p. 93-104 (1983).
- Lawrence, G.S. System features for subject access in the online catalog. *Library Resources & Technical Services*. Vol. 29, No. 5, p. 16-33 (1985).
- Lipetz, B. and Paulson, P.J. A study of the impact of introducing an online subject catalog at the New York State Library. *Library Trends*. Vol. 35, No. 4, p. 597-617 (1987).
- Liddy, E.D. The discourse-level structure of empirical abstracts: an exploratory study. *Information Processing & Management*. Vol. 27, No. 1, p. 55-81 (1991).
- Mandel, C.A. Enriching the library catalog record for subject access. *Library Resources & Technical Services*. Vol. 29, No. 5, p. 5-15 (1985).
- Markey, K. Subject searching in library catalogs: before and after the introduction of online catalogs. Dublin. OCLC. 1984, 176 p.
- Markey, K. Subject searching experiences and needs of online catalog users: implications for library classification. *Library Resources & Technical Services*. Vol. 29, No. 5, p. 34-51 (1985).
- Markey, K. Class number searching in an experimental online catalog. *International Classification*. Vol. 13, No. 3, p. 142-150 (1986).
- Markey, K. Searching and browsing the Dewey Decimal Classification in an online catalog. *Cataloging and Classification Quarterly*. Vol. 7, No. 3, p. 37-68 (1987).
- Markey, K. Integrating the machine readable LCSH into online catalogs. *Information Technology and Libraries*. Vol. 7, p. 299-312 (1988).
- Matthews, J.R. et al. Using online catalogs: a nationwide survey: a report of a study sponsored by the Council on Library Resources. New York. Neal-Schuman Publishers, 1983. 255 p.
- Saracevic, T. et al. A study of information seeking and retrieving: I. background and methodology. *Journal of the American Society for Information Science*. Vol. 39, No. 3, p. 161-176 (1988).
- Saracevic, T. and Kantor, P. A study of information seeking and retrieving: II. users, questions, and effectiveness. *Journal of the American Society for Information Science*. Vol. 39, No. 3, p. 177-196 (1988a).
- Saracevic, T. and Kantor, P. A study of information seeking and retrieving: III. searchers, searches, and overlap. *Journal of the American Society for Information Science*. Vol. 39, No. 3, p. 197-216 (1988b).
- Steinberg, D. and Metz, P. User response to and knowledge about an online catalog. *College and Research Libraries*. Vol. 45, No. 1, p. 66-70 (1984).
- Svenonius, E. Use of classification in online retrieval. *Library Resources & Technical Services*. Vol. 27, No. 3, No. 1, p. 76-80 (1983).
- Taylor, R.S. Question-negotiations and information seeking in libraries. *College and Research Libraries*. Vol. 29, p. 178-194 (1968).
- Tolle, J.E. Transaction log analysis: online catalogs. In: J.J. Kuehn ed. *Research and development in information retrieval: Proceedings of the Sixth Annual International ACM SIGIR Conference*, p. 147-460. Baltimore. Association for Computing Machinery, 1983.
- Van Pulis, N. and Ludy, L.E. Subject searching in an online catalog with authority control. *College and Research Libraries*. Vol. 49, No. 6, p. 523-533 (1988).
- Walker, S. and Jones, R.M. Improving subject retrieval in online catalogues: I. stemming, automatic spelling correction and cross-reference tables. *British Library Research Paper 24*. London, British Library Research and Development Department, 1987.
- Williamson, N.J. Classification in online systems: research and the North American perspective. *International Cataloguing*. Vol. 14, p. 29-30 (1985).
- Williamson, N.J. The Library of Congress Classification: problems and prospects in online retrieval. *International Cataloguing*. Vol. 15, p. 45-48 (1986).
- Williamson, N.J. The role of classification in online systems. *Cataloging and Classification Quarterly*. Vol. 10, nos. 1/2, p. 95-104 (1989).