

목차를 이용한 한국어 권말색인 표목의 선정

The Use of Contents in Selecting the Headings of Korean Book Indexes

이 태 영 전북대학교 문헌정보학과

Lee Tae Young Chonbuk National University

색인작성 모형을 구축하는데 필요한 기초적인 자료로서 색인표목 선정
을 목차와 결부시켜 고찰하였다. 색인길이는 본문의 3%이고, 네 개 이
하의 참조 수를 갖는 색인어는 95.8%이며, 목차항목이 색인에 출현한
도서는 100%이었다. 계층표목을 갖는 분석색인에 계층적 목차항목의
50%가 준용되었고, 목차의 고유어는 거의 색인어에 올랐다. 주절과 주
어 등 문법상 주제적인 위치에 있는 단어와 구들이 다수 색인표목으로
선정되어 앞으로 목차항목 등과의 상호관련에 대한 연구가 진행되어야
할 것으로 사려된다.

1. 서론

색인의 기능은 도서, 도서의 장·절, 잡지의
초록과 논문기사, 그리고 제반 기록물의 중요
사항에 대한 안내이다.[Borko]

교과서 등의 단행본 말미에 수록되어 있
는 권말색인은 도서에 대하여 바로 이와같은
안내 기능을 수행한다. 구체적으로 말하면 권
말색인은 도서의 주요 주제개념, 그들간의 관
계, 약어의 전칭, 전문용어에 대한 질문에 해
답을 하고 기대치 않았던 관심사항에 대한
발견의 기쁨을 제공하며 전체적으로는 주제
분야의 포괄적 개요 즉, 어떤 내용을 수록한
것이라는 맥락을 알려준다. 그래서 적절한 식
별자(indentifier)를 선택하고 그들을 상호 연
관시킬 때 색인자는 이용자가 어떠한 탐색에
서도 재현되기를 당위적으로 바라는 모든 사
항이 수록되도록 항상 노력해야 한다. 그리고
이용자가 부적절한 것으로 판정하거나 사소
하게 여기는 어떠한 것도 배제해야 한다. 종
합적이고 접속적인 기능을 수행함에 있어 색
인자는 그자신의 판단을 적용해야만 한다. 때
때로 색인되는 책 또는 사항을 저작한 사람

이 의도한 표현과는 다르게 심지어는 대체
하여 색인하게 되기 때문이다.[Rothman]

이와 같이 훌륭한 색인(사회과학분야의
통계방법론에서 허용하는 오차를 감수한)의
생산을 위해서는 잘 짜여진 작성규칙이 존재
하는 것이 중요하다고 생각한다. 그 규칙을
위하여 권말색인의 개선방안을 제시한 최석
두의 연구와(최석두) 한국어 권말색인의 문제
점과 개선방안을 논하고 작성지침을 제시한
김효열의 연구가 있다.(김효열) 본고는 이러
한 색인작성 모형을 구축하는데 필요한 기초
적인 자료의 성격으로 색인표목선정에 관한
문제를 목차와 결부시켜 고찰하였다. 목차도
일종의 색인으로 볼 수 있어(김효열,p.5) 권말
색인과의 상관관계가 있으리라는 생각에서
였다. 표본으로 인문·사회과학 분야 한국어
도서 24권을 사용되었고 그것들에 나타난 바
를 푸는 과정에서 잡지기사색인의 유형들이
적용하였다.

2. 색인표목의 선정과 기입량

앞에서 말했듯이 색인에는 그 책에서 이

용자가 기대하는 사항 또는 개념들에 대한 충분한 안내가 이루어져야 한다. 반면에 부적절하거나 사소하고 여겨질 수 있는 색인표목은 배제되어야 한다. 색인 후보표목에는 (1) 고유어 : 인명, 지명, 단체명, 상품명, 물체명, 화학명, 등, (2) 전문술어(Nomenclature), (3) 주제구(본고에서 이후 주제 단어와 구 및 절을 지칭한다), (4) 약어 등이 떠오른다.

여기서 색인 후보표목은 첫째, 적당히 반복하여 나타나는 단어나 구, 둘째, 문법적 구조에서 주제적인 것, 예를 들어서 "모기가 호랑이와 같이 사납게 달려들었다"에서 색인어는 '모기'이지 '호랑이'가 아니다. 셋째, 역사적인 배경에 속하지 않는 것, (Cleveland, p.107) 넷째, 그 색인 후보표목에 대한 부연 설명문이 하나 이상 있는 곳에서 추출한다. 그리고 위의 기준과 상호보완하여, 추출되는 대상개념을 보다 확고히 표목의 위치에 올려 놓게 하는 선정방법으로 목차에 출현한 목차 단어나 구 및 절을 단서로 삼아 그것과 관련되거나 또는 그 자체를 색인에 활용하는 것을 제안한다.

저자가 어느 주제분야 또는 관심영역의 지식이나 관찰사항, 그리고 경험과 지각 등에 대한 내용을 체계적으로 정렬하여 지식구조화 시킨 것이 도서이다. 도서는 머리말, 목차, 본문, 참고문헌, 색인, 등이 존재하는데 이들 중에서 도서의 지식구조에 대해 요결적으로 안내해 주는 부분이 목차와 색인이다. 목차는 포괄적이며 넓은 범위를 지시하며, 색인은 세부적이며 좁은 범위를 안내한다. 실제로 한국어 도서 24권을 조사한 바에 의하면 많게는 98%에서 적게는 10%에 이르기 까지 평균 41.3%의 목차 항목(단어·구·절 포함)을 색인에 포함시키고 있다.

따라서 목차에 출현한 단어나 구절은 그대로 색인표목이 될 수 있으며 또 그것의 발췌에 길잡이 역할을 할 수 있다. 아래의 예와 같이 1차 목차항목을 주표목으로 하고 2차 목차항목을 부표목으로 상정한 색인이 있다.

목차항목 : 도서관학 교육의 역사
 독일에서의 도서관학 교육
 미국에서의 도서관학 교육
 도서관학 교육의 문제점
 색인표목 : 도서관학 교육
 독일
 문제점
 미국

위 예처럼 목차항목이 그대로 약간의 수정을 거쳐 색인표목화하는 것을 볼 수 있다. 또 나아가서 목차에는 나타나지 않지만 본문의 장·절에서 항목으로 작용하는 단어들도 훌륭한 색인어 구실을 할 수가 있으며 또한 길잡이 역할을 하게 된다. 더욱이 세분되어 나누어진 목차항목(3차 또는 4차)이면서 목차부위에 출현하지 못하고 본문에만 기재되는 경우의 목

차항목은 색인에서 그 내용과 소재위치를 알려줄 필요가 있다. 또한 목차항목들은 색인표목 발췌에 연관적으로 작용할 수 있다.

아래의 예에서 A가 색인어로 나가면 B도 색인어의 고려대상이 될 수 있고 B가 색인어로 나가면 A가 고려대상이 될 수 있는 것이다.

조직커뮤니케이션 효과 --- A
 생산기능
 유지기능 --- B
 개혁기능

다음으로 색인기입의 양, 즉 색인길이 문제가 되는데 색인길이를 얻기 위하여 한국어 도서 24권을 분석하여 보았다. 평균치로 본문은 페이지 당 250단어가 출현하고, 색인은 페이지 당 72단어가 수록되었다. 이 기준으로 본문 길이에 대한 색인의 길이를 퍼센트 산출하면 3%수준이었다. 길이 문제와 마찬가지로 색인기입에서 소재지시의 수가 너무 많으면 그 페이지를 일일이 찾아 보아야 하는 난점이 생긴다. 이런 경우 부표목을 사용하여 분산시켜야 하는데 그 기준점 즉 색인기입이 갖는 참조의 수는 3-4개가 적당하다고 사려된다. 왜냐면, 3개 이하의 참조수를 갖는 색인어가 전체의 93.8%를 차지하고 4개 이하의 참조수를 갖는 색인어가 전체의 95.8%이기 때문이다.

3. 색인과 목차의 연관 분석

3.1 색인의 유형

각 도서에 대한 권말색인들을 살펴 보면 여러가지 유형이 있는 것을 볼 수 있다. 이 유형들은 잡지기사색인에서 출현한 여러 색인의 형식을 빌어서 설명할 수 있다. 잡지기사색인의 종류는 1)계층적 배열 <①분류색인, ②자모순색인>, 2)조합방식 <①전조합색인, ②후조합색인>, 3)주제포함유무 <(1)주제불포함 <①저자명색인, ②요어색인>, (2)주제포함 <①분류색인, ②주제명색인, ③순열표제색인, ④용어열색인, ⑤분절주제색인>>, 4)통제유무 <①통제언어색인, ②자연언어색인, ③자유언어색인>에 따라 <>안과 같이 세부적으로 나눌 수 있다. 요어색인이란 중요한 단어나 구·절로 기입된 색인을 말하고(사공철), 자연언어색인은 문헌에 나온 단어, 구, 절을 그대로 사용하며, 자유언어색인에서는 본문에 등장하지 않는 주제명으로 문헌을 대변하게 한다.

이들 색인형식을 준거로 권말색인의 경우를 아래와 같이 약술한다. (위에서 후조합, 순열표제, 용어열, 분절주제 색인의 형식은 권말색인에서 나타나지 않는다. 그리고 표목들의 배열은 당연히 자모순을 따르고 있다.)

1)분류형식
인명, 영어, 국문색인부와 같이 형식분류가 이루어지며, 주표목과 부표목을 계층적으로 배열한 것이 있다. (계층적 배열은 이후 분석색인으로 칭한다)

표본분석 예 :

- ①단순색인(부표목이 없음) : 70.8%,
- ②분석색인(계층부표목) : 20.8%
- ③분석색인1(한정·보완부표목) : 8.3%,

2)주제명, 전조합, 요어형식
색인된 단어나 복합어 및 구는 그 자체가 주제명이며 전조합형식이다. (전조합형식에 절을 포함시킴).

표본분석 예 :

- ①색인에 단어만 출현 : 33.3%,
- ②색인에 단어와 구 출현 : 29.2%,
- ③색인에 단어, 구, 절출현 : 37.5%

3)저자명형식
기업의 표목이 인물, 기구, 공저자, 정부기관, 대학교명, 등이었을때 이를 저자명색인이라고 하는데 (Cleveland,p.39), 권말색인도 인명, 기관명색인부가 존재한다. 표본에서 이러한 고유어와 주제구가 수록된 방식을 4가지 범주로 나누어 집계하였다. ㉓혼합형 : 주제구와 고유어를 한곳에 혼합배열, ㉔혼합분리형 : 주제구와 고유어를 혼합배열하지만 한글과 영문자를 분리, ㉕분리형 : 주제구와 고유어를 분리배열, ㉖단일형 : 주제구만을 배열하는 색인이 있다.

표본분석 예 :

- ①혼합형1 -- 한글만 사용(외국문자를 한글로 번자) : 25%
- ②혼합형2 -- 한글과 영문자 혼용 : 8.3%
- ③혼합분리형1 -- 한글과 영문자를 분리함 : 16.6%
- ④혼합분리형2 -- 한글(주제어), 영문자(주제어+고유어) : 4.2%
- ⑤분리형1 -- 주제어(한글), 고유어(영문자) 사용 : 16.6%
- ⑥분리형2 -- 주제어(한글,영문자), 고유어(영문자,일본문자) : 8.3%
- ⑦단일형1 -- 한글만 사용 : 12.5%
- ⑧단일형2 -- 한글과 외국문자 분리함 : 8.3%

4)언어형식
통제언어, 자연언어, 자유언어형식이 있다.

표본분석 예 :

- ①통제언어색인(통제언어를 포함한 도서) : 8.3%
- ②자연언어색인(자연언어만 있는 도서) :

58.3%,

- ③자유언어색인(자유언어를 포함한 도서) : 33.3%

3.2 목차의 분석

목차의 분석은 목차항목의 구성체계, 문 구성단위(단어, 구, 절), 고유어, 언어형식, 반복, 문법구조의 범주로 나누어 분석한다.

1)목차항목의 구성체계

목차항목들이 계층적(속/중, 전체/부분, 구성요소적 포함관계)인가를 살피고 그것이 색인에 반영된 여부를 알아보았다. 전체적으로 65%의 항목들이 계층적이며, 분석색인의 형식을 담고 있는 도서들에서는 50%의 계층적 항목이 색인에 반영되었다.

2)문 구성단위

목차의 주제개념이 출현한 양상을 단어, 구, 절로 나누어 아래의 예와 같이 산출하였다.

- ①단어 : 자체(1351/45), 색인(838/62)
- ②구 : 자체(1133/38), 색인(146/13)
- ③절 : 자체(542/18), 색인(9/2)

‘자체’의 왼쪽은 목차에 출현한 개수이고 오른쪽은 전체 합에 대한 퍼센트이다. 색인의 왼쪽은 목차항목 중에서 색인에 출현한 개수이고 오른쪽은 그 목차항목수에 대한 퍼센트이다. 그리고 맨 우측의 괄호안에 있는 숫자는 각각에 참여한 표본도서들의 퍼센트이다. 33%의 목차항목들이 재현되었는데 구나 절의 단어들을 분리하여 계산하면 전술한 바와 같이 41.3%에 이른다.

3)고유어

- ①한국어 고유어 : 출현한 도서에서 100%의 항목 단어들이 색인표목에 올랐다.
- ②영문자 고유어 : 출현한 도서에서 99%의 항목 단어들이 색인표목에 올랐다.

4)언어형식

분석 결과, 목차에 나온 항목 중에서 색인에 수록될 때 통제되는 단어와 구가 있으며 자유언어화하기도 한다. 반면에 절 중에서 절 자체 또는 절 일부가 그대로 색인에 사용된다.

- ①통제언어 : 8.3%의 도서에서 목차항목의 단어나 구를 통제하고 있다.
- ②자유언어 : 33.3%의 도서에서 목차항목의 단어나 구를 변형시키고 있다.
- ③자연언어 : 58.3%의 도서에서 목차항목의 절이나 절의 일부를 그대로 사용한 다.

5) 반복

목차의 단어, 구, 절 또는 절의 일부가 해당 장·절에서 나타나는 문장의 수를 표본 중 6권을 추려서 점검하였다. 평균 28문장에 출현하였다. 이는 그 해당 장·절들의 평균 문장수가 256개 인 것에 비해 11% 가량 출현한 것이 된다.

6) 문법구조

목차에서 출현한 단어, 구, 절이 본문 중의 문장에서 어떤 구조적 위치--주절, 종속절(수식절,비수식절), 주어, 목적어, 보어--를 차지할 때 색인에 등재되는가를 살펴 보았다. 주절과 주어(특히 ~은(는)) 위치에서 색인표목으로 발췌된 경우가 다수이었으나 그래도 목적어, 보어, 수식절에서도 색인표목으로 상정되는데 그 예를 들면 다음과 같다.

수식절 : "자연언어색인을 채택한 시스템은 ~~~"

비수식절 : "~~~에서 정보이론의 정의를 찾아보면 ~~~이라고 써어 있다"

목적어 : "~~~을 사용한 후조항색인을"

보어 : "~~~것을 랑디스크라고 한다"
"~~~이 전조항색인에 속한다"

이상과 같이 색인과 목차에 대한 내역을 분석하였는데 이 분석결과는 이미 진행되었던 연구나, 현재 진행중인 연구의 결과와 유사 또는 상이할 수 있다. 일치한다면 공통의 결과에 도달한 것이 되어 어떠한 도출에 일조를 하게 된 것이며, 상이하다면 그 상이한 이유를 조사하는 동기를 부여하는 것이 된다. 본장에서 이끌어낸 결과적인 사항은 결론에서 말하기로 한다.

4. 결론

색인표목 선정에 목차가 기여할 수 있는 부분을 목차항목(표본도서 전체의 색인에 출현)의 구성체제, 문 구성요소, 고유어, 색인언어 유형, 통계, 문법구조를 통해 살펴보았다. 한국어 도서는 다수(70.8%)가 단순색인 수준에 머물러 있으며 계층부표목을 갖는 분석색인은 20.8%, 한정·보완부표목을 갖는 분석색인은 8.3%로 소수에 불과하였다. 그러나 분석색인은 발달된 형태의 색인이므로 권장되어야 하며, 그 색인을 잘 만들기 위한 시도가 이루어져야 한다. 그 시도의 하나로써 목차항목(실제로 목차의 계층적 항목의 50%가 분석색인에 준용됨)의 계층성을 이용하여 주·부표목을 선정한 후 분석색인을 구축할 수 있다.

표본 중 37.5%의 도서들이 색인에서 단어, 구, 절을, 29.2%의 도서들이 단어와 구를 표목으로 사용한다. 목차(3.2의 2)에서보면 그

중 구나 절이 잘 이용되지 않는 것을 볼 수 있다. 3.1의 3)에서 보면 색인구성이 여러 가지로 발생하는데 이 현상은 바람직하지 않기 때문에 이에 대한 정리가 필요하며, 3.2의 3) 결과는 목차의 고유어는 색인이이자 색인표목 발췌의 단서어 역할을 할 것이란 것을 알려준다.

3.1의 4)로 보아 언어를 통제한다거나 본문에 없는 단어로써 색인하는 방법은 아직 보편화되어 있지 않지만(합계 41.6%) 3.2의 4)에서 33.3%의 도서들이 목차항목을 변형시켰다. 이 목차항목을 변형시켜 색인표목으로 만드는 것은 본문의 구절도 목차의 항목으로 변형될 수 있다는 징표이기 때문에 목차항목이 자유언어생성에 길잡이 역할을 할 수 있리라 생각된다.

색인표목이된 목차의 항목이 보통 그 장·절에서의 총 문장 대비 11%로 출현한 것은 색인표목 선정에 통계치로 참고사항이 될 수 있다. 그리고 색인어 발췌문제에 문법구조를 활용한 결과 주절과 주어에서 다수의 색인표목이 선정된 것을 확인하였지만 또 그밖의 요소에서도 선정이되었다. 그 선정된 것들을 살펴보면 문법상 주체적인 위치를 차지하고 있는 단어, 구절들이었다. 본고에서는 이 주체적인 것과 목차항목과의 관계 파악, 그리고 목차항목과 관련 주제어의 주제개념 표현 정도에 따른 색인표목의 선정방법, 등에 대한 연구는 진행시키지 못하였고 앞으로의 과제로 남겼다.

참고문헌

김효열, 도서관말색인의 작성지침과 자동색인에 관한 연구, 연세대학교 대학원 석사학위논문, 1995.
사공철, 외편, 도서관학·정보학 용어사전, 서울; 한국도서관협회, 1986. p.153.
정영미, 최석두, 서해란, 색인지침, 서울; 문헌정보처리연구회, 1996.
최석두, 도서관말색인에 관한 연구, 이화여자대학교 인문과학대학 제2회 교수학술제 발표문, 1994.
Borko, H., & Berner C. L., *Indexing Concepts and methods*, New York; Academic Press, 1978, p.4.
Cleveland, D. B., & Cleveland, A. D., *Introduction to Indexing and Abstracting*, Englewood; Libraries Unlimited Inc., 1990, p.39.
Knight, G. N., *Training in Indexing, A course of the Society of Indexers*, The M.I.T. Press, Cambridge, 1969. p.11.
Rothman, J., "Index, Indexer, Indexing", in *Encyclopedia of Library and Information Science*, vol.11, New York; Marcel Dekker, 1974, p.288.