

# 정기간행물기사 색인서비스의 포괄성 측정에 관한 연구\*

-국어학 문헌을 중심으로-

## An Indexing Coverage Study of Periodical Literature in a Subject Area of Korean Language

장혜란 (Hye-Rhan Chang)\*\*

### □ 목 차 □

I. 서론

III. 데이터 분석과 논의

II. 연구방법과 내용

IV. 결론

### 초 록

서지대조법을 활용하여 국내 정기간행물기사 색인서비스의 포괄성을 측정하였다. “국어학논저목록집”으로부터 추출한 국어정책 분야의 문헌 675건을 기본서지로 하여 “정기간행물기사색인”과 “국어국문학논문목록” 두 가지의 색인서비스를 탐색하여 포함여부를 조사하였다. 분석결과 1977년부터 1990년까지 출판된 문헌에 대한 “정기간행물기사색인”의 포괄성은 약 63%, 1986년부터 1990년까지 출판된 문헌에 대한 “국어국문학논문목록”의 포괄성은 약 33%, 두 서비스의 통합포괄성은 약 78%로 나타났다. 그밖에 두 서비스의 중복성, 유일성 등이 측정되었으며 수록된 문헌의 주제별 분석과 제외된 문헌의 수록잡지에 대한 분석도 수행되었다.

### ABSTRACT

This study was performed to measure the indexing coverage of the periodical literature on a subject area in Korean language policy. A bibliography of 675 citations dated 1977-1990 was obtained through “Kugohaknonjoromongnokjip”. Coverage was determined by checking the citations against 2 indexing services, “Chonggiganhangmulgisasakin” and “Kugogunmunhaknonmunmongrok”. For the period 1977-1990, “Chonggiganhangmulgisasakin” consistently provided coverage of about 63%. For the period 1986-1990, “Kugogunmunhaknonmunmongrok” provided coverage of about 33% and the combined coverage of the two services was about 78%. The extent of overlap and uniqueness of both indexing services are measured. The analysis also includes the subject of the indexed articles and the type of periodicals omitted.

\* 본 연구는 1993년도 상명여대 사회과학연구소 학술연구비 지원으로 수행되었음.

\*\* 상명여자대학교 문헌정보학과

## I. 서론

색인, 초록, 서지 등의 이차자료는 출판된 문헌에 대한 주제접근을 가능하게하는 도구로서 전문적인 학술정보를 필요로 하는 연구자나 학생 또는 기타 이용자 모두에게 없어서는 안될 필수적인 것이다. 그 중 특정 주제에 관한 정보의 존재 여부와 그 정보의 소재를 안내해 주는 색인서비스는 서지데이터베이스(인쇄본이나 기계가독형)의 기간을 형성하고 있다.

서지데이터베이스의 주요 평가 기준은 포괄성, 검색성, 예측성, 적시성 등을 들 수 있는데(Lancaster 1991, 116-137), 그 중 포괄성(coverage)이란 일정기간 동안에 출판된 특정 주제에 관한 문헌 중 어느 정도가 평가될 데이터베이스에 수록되어 있는가를 나타내는 수치이다.

포괄성은 이용자 측면에서는 탐색할 데이터베이스를 선정함에 있어서 고려해야할 첫번째 요소가 된다. 이것은 특히 망라적탐색(comprehensive search)을 필요로 하는 이용자에게 더욱 중요한 관심사가 되고 있는데, 왜냐하면 원하는 문헌을 발견할 수 있는 확률은 해당 데이터베이스가 특정주제에 관한 문헌을 얼마나 많이 포함하고 있는가에 달려있기 때문이다.

포괄성은 또한 데이터베이스 제작자 측면에서도 중요한 관심사가 되고 있다. 포괄성은 탐색성능의 평가척도 중 가장 중요하게 간주되는 재현율을 결정하는 근원적인 요소가 되며, 포괄성의 문제는 교육, 훈련, 탐색경력 등 이용자 변인으로 개선될 수 없는 전적으로 시스템에 의존하는 것이기 때문이다. 또한 포괄성의 측정과 평가는 해당 데이터베이스가 이용자요구를 어느정도 만족시켜 줄 수 있는가를 시사하는 동

시에 시스템 개선을 위한 구체적인 정보를 제공한다.

이와같은 문헌검색활동 전반에 걸친 명백한 중요성 때문에 색인서비스의 포괄성은 일찌기 60년대 중반부터 빈번하게 평가의 대상이 되어 왔으며 70년대 이후 주요 색인지와 초록지가 기계가독형으로 변환, 온라인 탐색이 가능해지고 다수의 데이터베이스가 제작 배포됨에 따라 더욱 활성화되었다.

Martyn(1967)의 연구는 이차자료의 포괄성 연구에서 하나의 고전적인 가치를 가지며, 이후 주제면에서나 규모에 있어서 다양한 연구들이 수행되었다. 유사, 관련분야를 다루는 여러가지 데이터베이스가 생산됨에 따라 포괄성뿐아니라 중복성 및 상호 비교 연구도 가능하게 되었는데 그 중 대표적인 것으로는 다음과 같은 연구를 들 수 있다.

Goode(1970), Wilkinson과 Holland(1973)는 의학 데이터베이스의 포괄성을 연구하였다. Smalley(1980)는 심리학과 의학 데이터베이스를 비교 연구하였으며 O'Connor와 Meadows(1968)는 천문학, Robinson과 Hu(1981)는 에너지, Britain과 Roberts(1980)는 범죄학, LaBorie 등(1985)은 문헌정보학, Sekerak(1986)은 심리학 분야에서 색인서비스의 포괄성 및 중복성을 연구하였다.

국내에서의 색인서비스 포괄성에 관한 연구는 찾아보기 어렵다. 그 이유로는 최근까지만 해도 색인지와 같은 이차자료의 생산과 이용 자체가 활발하지 못하였으며, 체계적인 서지서비스가 발전하지 못했던 환경에서는 대조나 비교에 사용할 만한 적합한 서지를 구하기가 쉽지 않았으리라는 점을 들 수 있을 것이다. 또한

선진국과 같이 일찌기 이차자료의 전산화가 되지 못한 상태에서 인쇄본을 수작업으로 일일이 찾아서 포괄성을 측정하는 작업은 막대한 시간과 노력을 필요로 하기 때문인 것으로 생각된다. 따라서 구체적인 데이터에 입각하여 분석한 실증적인 포괄성 연구는 현재까지 이루어지지 못한 실정이다.

본 연구는 국내의 정기간행물 전체를 대상으로 하고 있는 “정기간행물기사색인”과 국내의 국어국문학 논문 전체를 대상으로 하고 있는 “국어국문학논문목록” 두 가지의 색인서비스에 대하여 국어학 중 국어정책에 관한 문헌의 포괄성을 측정코자 하였다. 이러한 연구는 우리나라에서 생산되는 고유의 정보자료를 대상으로 하는 색인서비스에 대하여 포괄성을 수립하는 평가연구로서 가치가 있으며, 이용자의 탐색과 데이터베이스의 질적 개선에 적용될 수 있는 유용한 정보를 산출하고, 포괄성이 측정되어 있는 외국의 각종 데이터베이스나 점차 증가하게 될 국내의 다른 데이터베이스들과 비교할 수 있는 기초자료를 제공하게 될 것이다.

## II. 연구방법과 내용

### 2.1. 연구방법

색인서비스의 포괄성을 측정하는 방법에는 여러가지가 있다.

가장 보편적인 방법은 서지대조법(Bibliography Checking Method)으로 특정 주제에 관한 신뢰할 만한 서지를 선정하여 그 서지에 포함되어 있는 문헌들을 색인으로부터 탐색, 대조하여 수록여부

를 조사하는 것이다. 이 방법의 문제는 기준이 될 적합한 서지를 식별하는 데 있다. 특정 주제에 대한 체계적으로 편성된 서지를 발견하기가 쉽지 않을 뿐 아니라, 기준서지가 편성될 때 평가될 데이터베이스를 참고로 하지 않았어야 한다는 전제조건이 충족되어야 하기 때문이다.

적합한 서지를 식별하지 못하면, 연구에 사용할 목적으로 서지를 편성하여 이것을 기준으로 데이터베이스의 포괄성을 측정한다. 이 경우 서지를 편성하는 방법은 두 가지가 있는데, 하나는 특정 주제에 관한 리뷰 기사를 선택하여 리뷰에 인용된 문헌들을 서지에 대신할 만한 정보원으로 간주하는 것이며, 다른 하나는 특정 주제에 관한 문헌탐색을 수행하여 그 탐색결과로 얻은 서지를 데이터베이스 상호간 비교의 표준으로 삼는 방법이다.

그 밖에 데이터베이스의 포괄성을 평가하는 가장 손쉬운 방법으로는 각 데이터베이스의 수록 간행물리스트를 구하여 조사하는 방법이 있다. 이 방법은 포괄성 연구의 초창기에 많이 시도되어 왔으나, 색인서비스가 수록 간행물에 게재된 논문을 선택적으로 포함하는 것이 가능하기 때문에 개개 논문을 대상으로 조사한 측정치에 비하여 완전치 못하다는 제한점이 있다.

본 연구에서는 포괄성 측정의 가장 전형적인 방법인 서지대조법을 활용하여 색인서비스의 포괄성을 측정코자 하였다.

### 2.2. 서지의 선정과 평가 데이터베이스

본 연구에서 포괄성 측정을 위하여 채택한 서지는 국립국어연구원이 편찬한 “국어학논저 목록집(국어정책)”이다. 이 자료는 국어연구원

이 오랜 기간에 걸쳐 수집해 내려온 목록류와 자체 제작 유지해온 색인카드를 기초로 편찬한 것으로 개화기부터 1990년까지 발표된 관련문헌을 9개의 주제로 분류하고 각 주제 내에서 자료유형별로 나누어 출판년도 아래 표제의 자모순으로 배열되어 있으며 권말에 제목색인과 저자명색인이 있다.

평가될 색인서비스는 국내의 정기간행물기사를 망라적으로 수록하고 있는 국회도서관 편 “정기간행물기사색인”과 국어국문학 분야의 연구성과를 망라적으로 수록하고 있는 한샘국어국문학연구소 편 “국어국문학논문목록”이다. 이밖에도 국어학 분야의 문헌을 망라적으로 포함하고 있는 데이터베이스는 국립도서관 편 “학술논저종합색인(국어국문학)”과 (주)교보문고에서 편찬한 “문학어학관계문헌목록”이 있으나 이 두 자료는 “정기간행물기사색인”을 참고자료로 하여 편성된 것이므로 포괄성 평가의 대상으로 적합치 못하여 연구대상에서 제외하였다.

“정기간행물기사색인”은 국내의 유일한 종합 색인지로서 1945년 이후 국내에서 간행되는 학술잡지, 학회지, 정부기관의 간행물, 사회단체의 기관지 및 대학의 연구논문집 등 정기간행물의 논문 기사를 총망라하고 있으며 1977년도분부터는 전산화되어 온라인 탐색이 가능하다. “국어국문학논문목록”은 한샘국어국문학연구소에서 국어국문학정보센터를 설립하기 위한 기초 자료로서 1986년부터 학회지, 각종 대학간행물, 기념논문집, 잡지 등을 토대로 논문을 수집·정리하여 연간으로 간행하는 것이다. 이 자료에 수록된 문헌은 국어학 국문학 국어교육으로 크게 분류되고 그 밑에 39개의 하위주제로 분류되어 있으며 각 엔트리는 논문의 서지사항과

더불어 가능한 경우 목차 및 논문개요를 포함하고 있다. 권말에는 저자명색인이 있다.

## 2.3 데이터수집 방법과 절차

데이터수집의 첫단계는 포괄성 측정을 위하여 조사될 문헌을 서지로부터 추출하는 것이다. 우선 “국어학논저목록집”에 수록된 문헌 중에서 자료유형별로 단행본과 신문기사를 제외하고 잡지기사로 한정하였다. 잡지기사는 시간적으로 1977년 이후의 문헌으로 한정하였다. 조사될 문헌을 이와같이 제한한 것은 “정기간행물기사색인”이 1977년도분부터 전산화되어 있으므로 수작업보다 완전하고 신속한 탐색이 가능하기 때문이다.

포괄성 연구에서는 평가될 데이터베이스의 정책에 관한 사항에도 유의할 필요가 있다. “정기간행물기사색인”은 일시적인 흥미나 관심을 끄는 학술적 가치가 없는 기사를 제외하고 있는데 그 세부사항을 보면 1) 일반대중잡지기사 2) 순문예지 및 준문예지에 게재된 문예작품 3) 권두어 및 발간사 4) 1면 미만기사 5) 뉴스기사 6) 현지르뽀, 사업보고, 회계보고 등 7) 정기적 통계표, 문헌목록, 제품소개, 광고 등 8) 각종 시험문제자료 9) 법조문 규약문 10) 학보 등에 학부학생이 기고한 기사 11) 참고자료로서 별로 가치가 없다고 생각되는 기사 등이다(이민환 1990, 114-115). 따라서 이에 상응하는 기사들을 서지로부터 제외시켰다.

이와같은 추출작업을 거쳐 서지로부터 포괄성 측정을 위하여 조사될 문헌 총 675건을 선정하여 두 가지 데이터베이스와 대조할 기준서지로 삼았다. 675건에 대하여 서지의 권말색인을

이용하여 저자명을 표목으로하고, 본문에 있는 엔트리를 찾아 첨부하여 조사용 워크시트를 작성하였다. 이 675건의 기준 서지에 포함된 문헌을 “정기간행물기사색인”과 “국어국문학논문목록”으로부터 탐색하여 수록여부를 조사하였다.

“정기간행물기사색인”에 대한 조사작업은 1977년부터 1987년까지의 출판 문헌은 산업기술정보원의 KINITI-IR에 소장된 DIGS(Domestic Information on the General Subjects) 데이터베이스를 온라인으로 탐색하였다. DIGS에 수록되어 있지 않은 1988년부터 1990년까지의 출판 문헌은 국회도서관을 직접 방문하여 NOLIS에 소장된 KPA2 데이터베이스를 열람용 이용자 단말을 통하여 탐색하였다. 탐색은 주로 저자명을 탐색키어로 사용하고 자료가 발견되면 엔트리 전체를 브라우징하여 동일문헌임을 확인기록하였으며 저자명으로 자료가 발견되지 않을 경우에 탐색키어를 논문제목 혹은 잡지명으로 확대하여 탐색을 시도하였다. KINITI-IR의 DIGS탐색은 1993년 7월 중에, 국회도서관 방문탐색은 1993년 8월 중에 수행하였다.

“국어국문학논문목록”에 대한 조사는 수작업으로 수행하였다. 1집(1986. 3-1987. 2)부터 3집(1988. 3-1989. 2)까지는 권말의 필자명색인을 이용하여 본문의 엔트리를 찾아 식별하였으며, 4집(1989. 3-1990. 2)부터 6집(1991. 3-1992. 2)까지는 권말의 필자명논문명색인만으로 조사작업이 가능하였다. 수작업탐색은 1993년 12월 중에 수행하였다.

색인서비스의 평가에서 고려해야할 다른 한 가지 사항은 출판일자이다. 예를들면, 1987년도에 출판된 문헌일지라도 당해년도 이후 몇 년 간에 걸쳐 수록될 가능성이 있기 때문이다. 본

연구에서는 논문 출판일자로부터 데이터베이스 수록까지의 시간적 지연이 “정기간행물기사색인”의 경우 탐색이 수행된 시점까지 약 2년 6개월, “국어국문학논문목록”의 경우 목록집 출판 시점까지의 약 1년 2개월이 허용되었다.

### III. 데이터의 분석과 논의

조사 결과 얻은 데이터는 수록 데이터베이스별, 문헌의 발행년도별로 구분하여 백분율로 포괄성을 측정하였다. 1977년부터 1985년까지 출판된 문헌은 “정기간행물기사색인” 한 가지 서비스의 포괄성이 측정되었고 1986년부터 1990년까지 출판된 문헌은 “정기간행물기사색인”과 “국어국문학논문목록” 각각의 포괄성이 측정, 비교되었다. 두 가지 서비스의 통합포괄성 및 중복성 그리고 각 서비스의 독점 보유문헌 정도를 나타내는 유일성 등이 측정되었다. 또한 각 서비스에서 포괄하고 있는 문헌의 주제별 분석과 제외되어있는 문헌에 대한 수록간행물 유형 분석도 수행하였다.

#### 3.1 정기간행물기사색인의 포괄성

기준서지의 문헌 중 1977년부터 1985년까지의 문헌은 비교할 만한 다른 데이터베이스가 없으므로 “정기간행물기사색인”과 기준서지만을 대조하여 포괄성을 측정하였다. 1977년부터 1985년까지의 기준서지의 문헌은 총 189건이다. 기준서지의 연도별 문헌 수와 누적문헌 수, 그리고 기준서지의 문헌 중 “정기간행물기사색인”에 수록되어 있는 문헌의 연도별 문헌 수와

누적문헌 수가 <표 1>에 나타나 있다. 이 기간에 대한 “정기간행물기사색인”의 포괄성은 63.5%로 측정되었다.

<표 1>에 나타나 있는 자료에 대하여 시간에 따른 포괄성의 변화를 <그림 1>과 같은 그래프

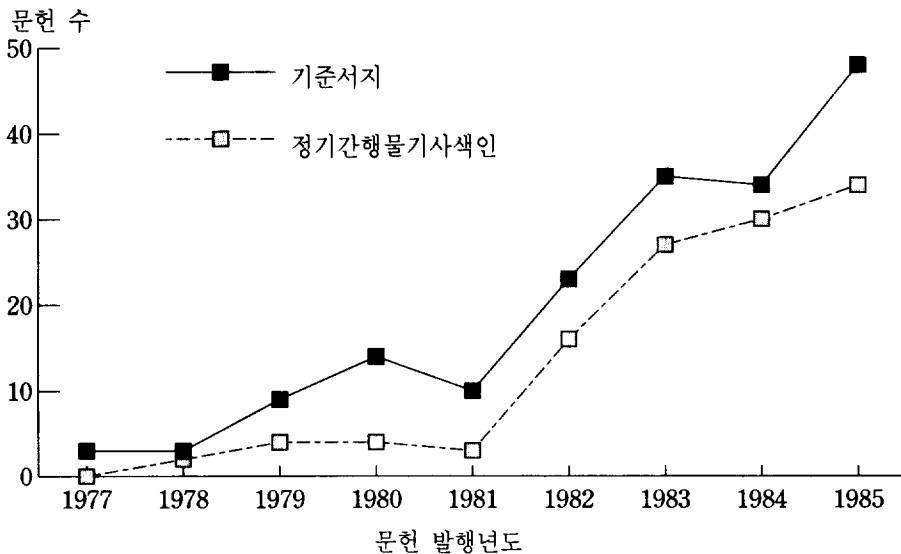
로 나타낼 수 있다.

<그림 1>에 나타나 있는 바와 같이 기준서지의 문헌이 증가함에 따라 “정기간행물기사색인”의 수록문헌도 꾸준한 증가양상을 보이고 있다. 이것으로 이차자료가 일차문헌의 성장을

<표 1> 정기간행물기사색인의 포괄성 (1977-1985)

발행 년도	기 준 서 지		정 기간 행 물 기 사 색 인		
	문헌수	누적문헌수	문헌수	누적문헌수	포괄성(%)
1977	3	3	0	0	$\frac{120}{189} \times 100 = 63.5(\%)$
1978	3	16	2	2	
1979	9	25	4	6	
1980	14	39	4	10	
1981	10	49	3	13	
1982	23	72	16	29	
1983	35	107	27	56	
1984	34	141	30	86	
1985	48	189	34	120	

<그림 1> 정기간행물기사색인 수록문헌의 연도별 변화 (1977-1985)



그대로 수용하고 있음을 알 수 있다. 포괄성은 전반부에서 보다 후반부에서 더욱 향상되었음을 알 수 있다. 시간이 경과됨에 따라 포괄성이 증가되는 양상은 시스템이 질적으로 개선되고 있음을 나타내고 있다.

### 3.2 정기간행물기사색인과 국어국문학논문 목록의 포괄성 비교

1986년부터 1990년까지 국어정책에 관한 연

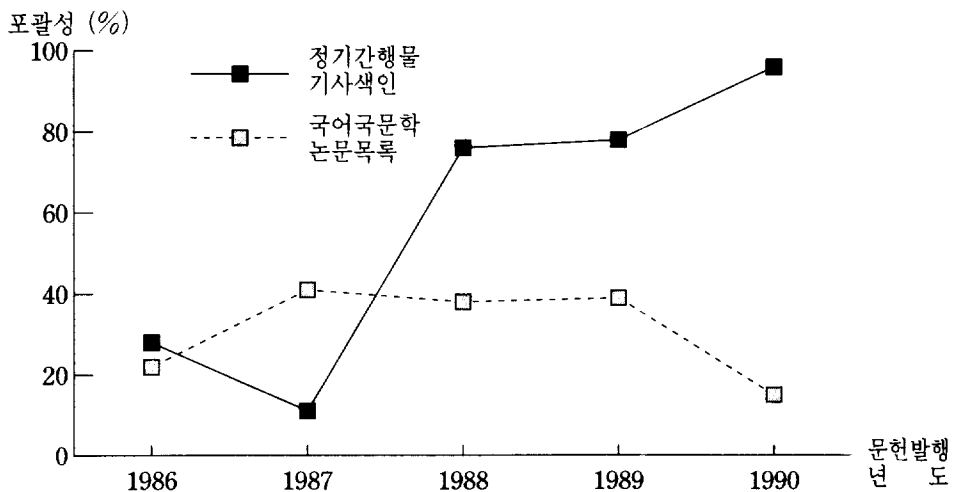
구가 크게 증가하여 이 기간 중 기준서지의 문헌은 모두 486건이다. 이 486건에 대하여 “정기간행물기사색인”과 “국어국문학논문목록” 두 가지의 데이터베이스를 탐색하여 각각의 포괄성을 측정하였다. <표 2>에는 기준서지의 연도별 문헌 수와 누적문헌 수, 그리고 기준서지의 문헌 중 각 데이터베이스에 수록된 문헌의 연도별 수록문헌 수와 수록비율, 누적문헌 수, 그리고 포괄성이 나타나 있다.

이 기간에 대한 “정기간행물기사색인”의 포

<표 2> 정기간행물기사색인과 국어국문학논문목록의 포괄성(1986-1990)

발행년도	기 준 서 지		정 기간 행 물 기 사 색 인			국 어 국 문 학 논 문 목 록		
	문헌수	누적문헌수	문헌수(%)	누적문헌수	포괄성(%)	문헌수(%)	누적문헌수	포괄성(%)
1986	65	65	18(28)	18	$\frac{304}{486} \times 100 = 62.6(\%)$	14(22)	14	$\frac{161}{486} \times 100 = 33.1(\%)$
1987	79	144	9(11)	27		32(41)	46	
1988	133	277	101(76)	128		51(38)	97	
1989	136	413	106(78)	234		53(39)	150	
1990	73	486	70(96)	304		11(15)	161	

<그림 2> 두 가지 서비스의 포괄성 변화 (1986-1990)



괄성은 62.6%이며, “국어국문학논문목록”의 포괄성은 33.1%로 나타났다. “정기간행물기사색인”의 포괄성은 앞서 1977-1985 기간에 대한 것과 거의 비슷한 수치를 나타내고 있다. <그림 2>는 두 가지 색인서비스의 포괄성의 연도별 변화를 그래프로 나타낸 것이다.

“정기간행물기사색인”의 포괄성은 1987년에는 감소하였다(11%) 1988년에 급격히 증가되고(76%), 이후 계속적인 증가양상을 보이다가 1990년에는 최고 96%까지 높아졌다. 이 기간 동안의 전반적인 포괄성은 1977-1986 데이터와 비슷한 수준인 62.6%이다. 그러나 포괄성의 범위는 11%부터 96%까지 다양하여, 일관성이 결핍되어있음을 알 수 있다.

“국어국문학논문목록”의 포괄성은 1987년에 증가되어(41%) 1989년까지 비슷한 수준을(38%, 39%) 유지하였다. 1990년에 와서 포괄성의 급격한 감소를 나타내고 있는데, 이것은 전적으로 데이터베이스 자체의 문제라기보다는 데이터베이스의 최종출판년도와 탐색시점까지의 시간적 지연이 충분히 허용되지 못한 데에

도 기인할 것으로 추정되며, “국어국문학논문목록”의 전체적인 포괄성에 부정적 영향을 미치고 있다.

### 3.3 정기간행물기사색인과 국어국문학논문 목록의 통합포괄성

1986년부터 1990년까지 출판된 문헌을 대상으로 두 가지 데이터베이스의 통합포괄성을 측정하였다. 통합포괄성을 측정하기 위하여 각 데이터베이스가 독점적으로 수록한 문헌 수와 두 가지 데이터베이스가 모두 수록하고 있는 중복 수록 문헌 수가 산출되었다. 통합포괄성 산출을 위한 상세한 데이터가 <표 3>에 나타나 있다.

기준서지와 대조한 결과 “정기간행물기사색인”이 독점적으로 수록하고 있는 문헌 수는 217건으로 유일성은 44.7%(217/486×100)이며 “국어국문학논문목록”이 독점적으로 수록하고 있는 문헌 수는 74건으로 유일성은 15.2%(74/486×100)로 판명되었다. 기준서지의 문헌 중 두 데이터베이스에서 모두 탐색된 문헌의 수는 모

<표 3> 두 가지 서비스의 유일성, 중복성, 통합포괄성

발행년도	기준 서지		독점 수록문헌				중복수록 문헌		통합포괄 문헌	
			정기간행물 기사색인		국어국문학 논문목록					
	문헌수	누적문헌수	문헌수	누적문헌수	문헌수	누적문헌수	문헌수	누적문헌수	문헌수	누적문헌수
1986	65	65	13	13	9	9	5	5	27(42)	27
1987	79	144	5	18	28	37	4	9	37(47)	64
1988	133	277	65	83	14	51	37	46	116(87)	180
1989	136	413	75	158	23	74	30	76	128(94)	308
1990	73	486	59	217	.	74	11	87	70(96)	378



두 87건으로 이는 기준서지 486건의 17.9%에 해당된다. Goldstein공식을 (LaBorie et al. 1985, 186-189) 적용하여 산출된 두 서비스간의 중복성은  $23\%(87/(304 + 161 - 87) \times 100)$ 이다.

〈표 3〉의 데이터로부터 기준서지의 문헌 중 두 가지 데이터베이스에 포함되어 있는 문헌의 총수는 모두 378건으로 산출되었다. 통합포괄성은  $77.8\%(378/486 \times 100)$ 으로 판명되었다. 이것은 국어정책 분야의 주제에 관하여 1986년부터 1990년까지 출판된 문헌을 찾을 때 두 가지 서비스를 모두 탐색한다면 최대 77.8%까지의 재현율을 달성할 수 있음을 암시한다.

〈그림 3〉은 두 가지 서비스 각각의 포괄성을 나타내고 있는 〈그림 2〉에 통합포괄성을 부가하여 그 연도별 변화를 그래프로 나타낸 것이다.

위의 그림을 보면, 두 가지 서비스의 상호보완적인 역할이 잘 나타나 있다. 1988년을 기점으로 이전에는 “국어국문학논문목록”이 “정기

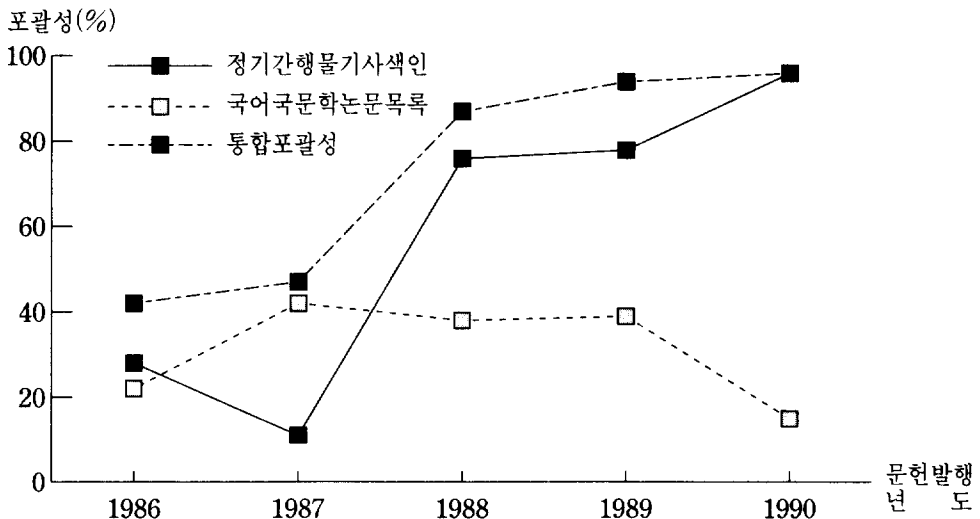
간행물기사색인”을 보완하며 이후에는 “정기간행물기사색인”이 “국어국문학논문목록”을 보완하고 있다.

### 3.4 정기간행물기사색인과 국어국문학논문 목록의 주제별 포괄성

1986년부터 1990년까지 출판된 기준서지의 문헌에 대하여 각 색인서비스에 수록된 문헌을 주제별로 분석하였다. 주제별 구분은 “국어학논저목록집”에서 사용된 분류를 그대로 적용하였다. 〈표 4〉는 두 가지 서비스 각각의 주제별 탐색 문헌 수와 그 포괄성을 나타내고 있다.

“정기간행물기사색인”의 경우 서비스 전체의 포괄성보다 낮은 포괄성이 산출된 주제 분야는 맞춤법 표준어 분야와 국어순화 국어운동 분야이다. “국어국문학논문목록”의 경우에는 로마표 기법과 외래어표기법 등 법규관련의 문헌이 하

〈그림 3〉 두 가지 서비스의 통합포괄성 변화 (1986-1990)



〈표 4〉 두 가지 서비스의 주제별 포괄성 (1986-1990)

	기준서지 문헌 수	정기간행물기사색인 수록문헌 수(%)	국어국문학논문목록 수록문헌 수(%)
글자꼴 기계화	26	17 (65.4)	5 (19.2)
맞춤법 표준어	76	35 (46.1)	29 (38.2)
국어순화/운동	111	63 (56.8)	37 (33.3)
국어정책일반	68	49 (72.1)	15 (22.1)
한 자	49	33 (67.3)	14 (28.6)
로마자표기법	5	4 (80.0)	0
북한어	80	51 (63.8)	37 (46.3)
외래어표기법	11	8 (72.7)	0
기타	60	44 (73.3)	24 (40.0)
계	486	304	161

나도 없었으며, 그 밖에 글자꼴 기계화, 국어정책일반, 한자 분야에서도 색인서비스 전체의 포괄성보다 낮은 포괄성이 산출되었다.

### 3.5 누락된 정기간행물의 유형별 분석

색인서비스의 평가에서 마지막으로 살펴봐야 할 부분은 각 서비스가 어떤 유형의 출판물을 포함하지 못하고 있는가를 식별하는 것이며 이러한 사항은 시스템을 개선할 구체적인 정보를 제공한다.

1977년부터 1990년까지의 출판문헌 중에서 “정기간행물기사색인”에 누락되어 있는 간행물은 총 68종이며 1986년부터 1990년까지의 출판문헌 중 “국어국문학논문목록”에 누락되어 있는 간행물은 총 92종이다. 누락된 간행물을 살펴보면 전체적으로 뚜렷한 자료의 분산현상을 볼 수 있다. “정기간행물기사색인”의 경우 누락

간행물 68종 중 1개의 기사만 실고 있는 것이 45종, “국어국문학논문목록”의 누락 간행물 92종 중 1개의 기사만 실고 있는 것이 62종으로 나타났다. 각 색인서비스에서 누락되어 있는 정기간행물을 유형별로 구분하여 분석한 데이터가 〈표 5〉에 나타나 있다.

〈표 5〉에 나타난 데이터를 보면, 전 분야를 망라적으로 색인하고 있는 “정기간행물기사색인”에서도 상당수의 누락 간행물이 있음을 알 수 있다. 유형별로는 학회지, 대학논문집 등이 두드러지게 누락되어있다. 이것은 이 유형자료의 체계적 수집에 문제점이 있음을 암시한다. “국어국문학논문목록”은 유형별로 대학논문집과 기타잡지 부문에서 두드러지게 누락되어있다. “국어국문학논문목록”은 국어학과 국문학 문헌을 대상으로 하는 학문지향적 색인서비스이다. 따라서 색인정책상 타학문이나 학문일반을 다루는 문헌은 제외될 수 밖에 없다. “국어

〈표 5〉 누락된 정기간행물의 유형별 분석

자료유형	정기간행물기사색인 (1977-1990)	국어국문학논문목록 (1986-1990)
학 회 지	23종 (36.5 %)	15종 (16.3 %)
대 학 논 문 집	19종 (30.2 %)	23종 (25 %)
연구소간행물	8종 (12.7 %)	16종 (17.4 %)
정 부 간 행 물	2종 ( 3.2 %)	5종 ( 5.4 %)
협 회/단 체 간 행 물	7종 (11.1 %)	15종 (16.3 %)
기 타 잡 지	9종 (14.3 %)	18종 (19.6 %)
계	63종 (100 %)	92종 (100 %)

국문학논문목록”에 누락되어 있는 92종의 간행물을 학문분야별로 분석해본 결과 국문학 국어학 어학 문학분야의 간행물이 30종, 학문 전분야를 망라하는 간행물이 29종, 타학문 분야의 간행물이 33종으로 나타났다.

#### IV. 결 론

본 연구에서는 서지대조법을 적용하여 색인서비스의 포괄성을 측정함으로써 우리나라의 서지데이터베이스에 대한 평가를 시도하였다. 구체적으로 “국어학논저목록집”으로부터 추출한 675건의 문헌을 기본서지로 하여 “정기간행물기사색인”과 “국어국문학논문목록” 두 가지 색인서비스의 포괄성을 측정하였다.

분석결과 “정기간행물기사색인”(1977-1990)의 포괄성은 약 63%, “국어국문학논문목록”(1986-1990)의 포괄성은 약 33%, 두 가지 색인서비스(1986-1990)의 통합포괄성은 77.8%로 나타났다. 색인서비스의 독점 수록문헌 정도를 나

타내는 유일성은 “정기간행물기사색인”의 경우 44.7%로 “국어국문학논문목록”의 경우 15.2%로 나타났으며, 두 가지의 서비스의 중복률은 23%로 나타났다. 각 서비스에 대한 수록문헌의 주제별 분석결과를 보면, “정기간행물기사색인”은 맞춤법 표준어 국어순화/운동 분야에서 “국어국문학논문목록”은 표기법 기계화 국어정책일반 한자분야에서 비교적 포괄성이 낮았다. 누락된 간행물에 대한 유형별 분석도 수행되었는데 그 결과를 보면, “정기간행물기사색인”은 학회지와 대학논문집 부문에서, “국어국문학논문목록”은 대학논문집과 기타잡지 부문에서 크게 결핍되어있음을 알 수 있다.

“국어국문학논문목록”의 경우 비교적 낮은 포괄성이 나타난 것은 국어정책이라는 주제 자체가 순수학문지향적이라기 보다는 어느 정도 다양성을 가지고 있는 업무지향적 주제라는 점, 앞서 논의한 바와 같이 1990년 출판문헌에 대한 시간적 지연이 충분히 고려되지 못했던 점, 그리고 수작업 탐색의 불완전성 등 여러가지 요인이 작용했을 것으로 보인다.

이상과 같은 평가연구는 색인서비스의 포괄성에 대한 흥미로운 결과를 산출하였을 뿐 아니라 데이터베이스 탐색자와 데이터베이스 제작자를 위한 다음과 같은 실제적인 정보를 제공하고 있다.

첫째, 국어학 분야의 문헌에 있어서 “정기간행물기사색인”은 비교적 높은 포괄성을 유지하고 있지만 망라적탐색을 위하여는 “정기간행물기사색인” 한 가지 색인서비스로는 충분치 못하며, 특히 학회지나 대학간행물 등에 수록된 핵심논문을 발견하기 위하여는 “국어국문학논문목록”의 탐색이 병행되어야 한다.

둘째, 두 가지 색인서비스의 존속이 합리화되었다. 이차자료는 레코드의 작성 축적 유지에 막대한 비용이 소요되므로, 그 중복성의 문제는 중복투자 및 자원낭비의 견지에서 논의되고 있다. 본 연구에서 평가된 두 가지 색인서비스는 어느 정도의 중복성을 가지고 있지만 상당수준의 유일성 또한 측정되었으며 연도별, 주제별, 수록간행물 유형별로도 포괄의 차이를 나타내고 있다. 또한 색인서비스의 존재여부는 목표이용자집단과 직접적인 관계가 있는데, 두 서비스는 성격이 다른 이용자집단을 대상으로 하고 있으며 앞으로도 “정기간행물기사색인”은 이용자일반을 “국어국문학논문목록”은 전문인을 위한 서비스로 계속 발전되어야 할 것이다.

셋째, 데이터베이스 제작자들은 각 색인서비스에 포함시킬 문헌의 범위를 확대시킴으로서 서비스 개선을 도모할 수 있다. 각 서비스는 특히 결핍된 부분에 대한 보완책이 필요한데, 여기서 확대 수록할 논문의 간행물 유형에 관한 사항이 고려되어야 한다. 특정주제분야의 문헌을 모두 수록할 수 있는 완전한 색인서비스란

존재할 수 없다. 다만 위의 두 가지 서비스간에는 분업 내지는 상호협력의 가능성을 추구할 수 있을 것이다. 두 가지 서비스는 각각의 장점을 살려서 “국어국문학논문목록”은 학회지와 대학논문집을 중심으로 “정기간행물기사색인”은 기타 간행물을 중심으로 수집을 확대하여 색인하는 것이 바람직한 것으로 보인다.

## 참고 문헌

- 고영진. 1992. "국회정보검색시스템 개선에 관한 고찰". 국회도서관보26(6): 35-48.
- 국립국어연구원. 1991. 국어학논저목록집(국어정책). 서울:국립국어연구원.
- 국회도서관. 정기간행물기사색인. 1977-.
- 산업기술정보원. 1992. 정보검색:KINITI-IR 이 사용자 지침서. 서울:산업기술정보원.
- 이민환. 1990. "정기간행물기사색인 작성방법에 대한 소고 - 국회도서관을 중심으로", 국회도서관보 27(1):112-123.
- 한샘국어국문학연구소. 1987-1993. 국어국문학논문목록.(1집-6집) 서울: 한샘출판(주).
- Britain, J. M. and Roberts, S. A. 1980. "Rationalization of Secondary Services: Measurement of Coverage of Primary Journals and Overlap Between Services." Journal of the American Society for Information Science (31):131-142.
- Edwards, T. A. 1976. "Comparative Analysis of the Major Abstracting and Indexing Services for Library and Information Science." Unesco Bulletin for Libraries

- (31):18-25.
- Goode, D. J. 1970. "Comparative Analysis of Epilepsy Abstracts and a MEDLARS Bibliography." *Bulletin of the Medical Library Association* (58):44-50.
- LaBorie, T. et al. 1985. "Library and Information Science Abstracting and Indexing Services: Coverage, Overlap, and Context." *Library and Information Science Research* (7):183-195.
- Lancaster, F. W. 1991. *Indexing and Abstracting in Theory and Practice*. Champaign: University of Illinois Press.
- Martyn, J. 1967. "Tests on Abstracts Journals: Coverage, Overlap, and Indexing." *Journal of Documentation* (23):45-70.
- Montgomery, R. R. 1973. "An Indexing Coverage Study of Toxicological Literature." *Journal of Chemical Documentation* (13):41-44.
- O'Connor, J. G. and Meadows, A. J. 1968. "Physics Abstracts as a Source of Abstracts in Astronomy," *Journal of Documentation* (24):107-112.
- Oppenheim, C. 1974. "The Patents Coverage of Chemical Abstracts." *Information Scientists* (8):133-138.
- Robinson, J. and Hu, M. 1981. "DOE's Energy Database versus Other Energy Related Database: a Comparative Analysis." *Database* (4):10-27.
- Sekerak, R. J. 1986. "A Comparison of Journal Coverage in Psychological Abstracts and the Primary Health Science Indexes: Implications for Cooperative Serials Acquisition and Retention." *Bulletin of the Medical Library Association* (74):231-233.
- Smalley, T. N. 1980. "Comparing Psychological Abstracts and Index Medicus for Coverage of the Journal Literature in a Subject Area in Psychology." *Journal of the American Society for Information Science* (31):143-146.
- Tenopir, C. 1982. "Evaluation of Database Coverage: A Comparison of Two Methodologies." *Online Review* (6):423-441.
- Wilkinson, D. and Hollander, S. A. 1973. "Comparison of Drug Literature Coverage by Index Medicus and Drug Literature Index." *Bulletin of the Medical Library Association* (61):431-432.