

## Cutter-Sanborn저자기호표에 관한 분석적 고찰\*

박 준 식\*\*

목 차	
서 론	3. 기호표의 단점
1. 기호표의 성립배경	1) 표목선정의 범위문제
2. 기호표의 특성과 장점	2) 문자별 번호할당의 비율문제
1) 표의 구조적 특성	3) 표목배분의 간격문제
2) 기호 부여방법상의 특성	결 론
3) 이용상의 특성	

### 서 론

저자기호는 분류를 통하여 이미 유별화된 문헌을 다시 서가상에 배열하고, 문헌의 효과적인 이용과 검색을 위하여 각각의 문헌단위로 個別化시키는 기능을 갖는다. 문헌의 개별화 작업은 1871년 Jacob Schwartz가 분류와 도서기호의 개념이 복합된 조합식 기호법을 고안하여 발표한 이래, 1870~80년대 Charles A. Cutter가 고안한 일련의 저자기호표에서 심진식, 혼합기호법으로서의 기초가 확립되었다.

그후 Cutter-Sanborn Three-figure Author Table<sup>1)</sup>로 개편되고, 이 Sanborn 기호표가 오늘날까지 대부분의 도서관에서 정리업무에 필수적인 도구로 활용되어 왔다.

\* 본 연구는 계명대학교 교수연구비의 지원에 의해 이루어 졌음.

\*\* 계명대학교 문헌정보학과 교수.

1) Sanborn의 기호표는 Cutter-Sanborn Three-figure Author Table이 공식 명칭이나 기호표의 독자성을 인정하여 Sanborn기호표(Sanborn's Table)로 통용되고 있다. 따라서 본고에서도 약칭을 사용하기로 한다.

그러나 Sanborn기호표는 발표된 이래 100년이 지났으나 아직까지 한번도 기호표의 구조와 장, 단점에 대해서 분석적인 평가가 이루어지지 않았다. 저자기호와 더불어 청구기호를 구성하는 DDC는 지금까지 115년간 20번이나 개정되었고, 그 특성과 장, 단점에 대해서 각 나라별로 헤아릴 수 없을 만큼 많은 연구가 있어 왔지만, Sanborn기호표에 대해서는 1890년대 후반과 1900년대 초반 기호표의 사용방법이나 DDC의 적용방법에 관한 단편적인 기사가 발표되었을 따름이고, 평가에 관련된 문헌을 근년의 도서기호법에 관한 Donald J. Lehnus와 John P. Comaromi의 저서에서 활용법을 중심으로 언급되고 있을 따름이다.<sup>2)</sup>

Sanborn 기호표가 분석, 평가되어야 할 필요성은 이 기호표가 표목의 선정순위, 문자별 표목의 점유비율과 표목의 세분화 등 두 가지 관점에서 문제점을 내포하고 있을 것으로 예상되기 때문이다.

표목의 선정범위면에서 예상되는 문제점은 이 기호표가 인명중심의 기호표라는 점이다. 100년전 Sanborn의 기호표가 편찬되었던 당시는 서구의 문헌들이 주로 개인 저자에 의해 발행되는 것이 일반적인 관례였으므로 이 기호표가 주로 인명중심으로 표목을 선정할 것은 당연하였다. 그러나 오늘날에는 단체저자의 문헌과 정부조직이나 정부 산하기관에서 발행하는 문헌이 급속하게 증가하고, 표제가 기호매김의 대상이 되는 학술잡지나 참고도서등이 대량으로 증가하여 Sanborn이 기호표가 만들었던 당시의 상황과 크게 달라졌다. 따라서 기존의 Sanborn기호표로서 이러한 단체명과 연속간행물 및 참고도서의 표제를 개별화하기에는 무리가 많을 수 밖에 없는 것이다.

문자별 표목의 점유율과 표목의 세분화에서 예상되는 문제점은 Sanborn이 기호표의 기준에 의거해서 문자별 점유율을 결정하고, 표목을 세분화했는지 그 근거를 밝히지

2) Lehnus, Donald J. *Book Numbers: History, Principles and Application*. Chicago, ALA., 1970. 이 논문에서는 도서기호법의 발전과 각 기호법의 문자별 빈도할당 비율을 중심으로 논하고 있다.

Comaromi, John P. *Book Numbers: A Historical Study and Practical Guide to Their Use*. Lincoln, Nebraska: Libraries Unlimited Inc., 1981. 이 논문은 도서기호법의 DDC 적용에 대해서 강조하고 있다.

언급하지 않았으며, 실무자들이 실제 사용하는데도 알파벳 문자그룹에 따라, 그리고 표목에 따라 저자번호가 집중되고 분산되는 문제가 있는 것으로 밝히고 있다는 점이다. 열거식 저자기호표는 그 특성상 인명과 단체명, 표제등의 철저한 빈도조사에 근거하여 표를 구성하여야 하는데 Sanborn기호표에서는 이 두 가지 부분에서 문제가 있을 것으로 예상할 수 있다.

이러한 두 가지의 관점에서 본 연구에서 Sanborn기호표의 어떤 요소가 오랫동안 많은 도서관에서 별 비판없이 이 기호표를 사용해 오도록 했으며, 단점이 있다면 어떤 부분에서 어느 정도의 결합이 있는가? 이것은 정리업무에 어떤 영향을 미치는가? 등을 분석적으로 평가하고자 하는 것이다.

### 1. 기호표의 성립배경

19세기의 마지막 4반세기 이후 미국의 도서관계는 괄목할 만한 발전을 이루고 있었다. 즉, 미국도서관협의 창립, *Library Journal* 의 창간, 도서관학교의 설립, 장서량의 확충, 그리고 카네기 재단의 후원등에 의한 새로운 도서관들의 계속적인 설립 등이 이 시기에 집중적으로 이루어졌다.

도서관계의 이러한 의욕적인 발전은 도서관 내부의 환경을 변화시키는데 작용하여 새로운 문헌정리 방법을 필요로 하게 되었다. 특히, 1880년대 말부터 1890년대에 이르러 미국의 대부분의 도서관에서 Cutter의 두 자릿수 기호법이 보편적으로 사용되어 왔지만, 장서량이 많은 도서관에서는 부적합한 것으로 평가되고 있었다. 그래서 Margaret Mann은 “장서가 5,000권이 넘는 도서관에서는 Cutter의 두 자릿수 기호표가 사용되기 어렵다”<sup>3)</sup>고 평가하였다. 또한 Cutter의 두자릿수 기호표는 모음과 자음에 따라 문자의 자리수가 1자리, 2자리 또는 3자리가 되는 일관성 없는 형식을 가지고 있으므로 이용하기 편리한 새로운 기호법이 요구되고 있었다. Kate G. Sanborn이 만든 Cutter-Sanborn 세 자릿수 기호표는 이러한배경에서 출현하게 되었다.

Sanborn은 1891년 메세추세츠 Manchester 공공도서관에서 Cutter 밑에서 목록자

3) Mann, Margaret. *The Classification and Cataloging of Books* Chicago, ALA., 1928. p.115.

로서 근무할 때, 후일 Cutter의 기호표보다 더 보편적으로 사용되었던 2개의 기호표를 만들었다. 첫번째 기호표는 모음과 S를 위한 기호표이다. 이 표에서 그녀는 모음과 자음 구별없이 저자기호에서 채용하는 첫 글자는 모두 한자리 문자로만 결정하였다. 이러한 형식은 그 당시 사서들이 모음과 S자에 대한 Cutter의 두 자리 또는 세 자리 문자를 좋아하지 않았고, 모든 문자에 걸쳐 1문자 기호방식을 더 좋아한다는 사실을 알았기 때문이었다. Cutter자신도 후일 자신의 두 자릿수 기호표에서 채용된 문자와 숫자의 배정이 불합리하다는 점을 인정한 바 있다.<sup>4)</sup> 또한 Sanborn은 A와 S에 대하여는 세 자리 숫자를 사용하였지만 나머지 모음인 E, I, O, U에 대하여는 세 자리 숫자가 불필요하다고 생각하여 두 자리 숫자만을 사용하게 되었다.<sup>5)</sup> Sanborn의 S와 모음을 위한 기호표는 1892년에 정식으로 출판되었다.

Sanborn은 1895년에 모음과 S를 제외한 나머지 자음 20자에 대한 기호표를 완성, 출판하였다. 이 자음 기호표는 모든 이름에 頭文字 1자만을 문자기호로 채용하고 한 알파벳 내에서는 숫자기호의 자리수를 통일하는 등 기본원칙은 첫번째 기호표와 동일하다. 이 자음표는 기호구성에 있어서 J, K, Y, Z는 2자리 숫자, Q, X는 1자리수, 그리고 나머지 자음은 모두 3자리수로 하였다.<sup>6)</sup> Sanborn의 두개의 기호표(모음과 S자, S를 제외한 자음)는 나중에 하나의 기호표로 통합되어 출간되었다.

Sanborn은 첫번째 기호표를 만들 때 Cutter의 요청을 받아서 개정하는 형식을 취하였으나, 두번째의 자음기호표는 Cutter와 관계없이 완전히 다른 별개의 기호표를 만들게 되었다. Sanborn의 기호표는 모든 이름이 母音가 子音의 구분없이 첫 문자만을 문자기호로 채용하여 일관성을 기하도록 하였고, 첨가되는 숫자는 이름의 빈도, 즉 알파벳 문자에 따라서 1자리수에서부터 3자리수까지 다양화시키되, 한 문자내에서는 자리수를 통일시켰다. 이런 형식은 이 기호표가 Cutter의 두 자릿수 기호표와 일치하는 것이 거의 없는 별개의 기호표가 되도록 만들었다.

4) Cutter, Charles A. *Explanation of the Cutter-Sanborn Author-marks* (3 figure table), 3rd ed. Northampton, Mass. Herald Office, 1899. p. 4, 8.

5) *Library Journal*, 17 (Nov. 1892). p. 468.

6) *Library Journal*, 20 (Oct, 1895). p. 363.

Sanborn기호표는 그후 오랫동안 수정되지 않고 사용되어 오다가 1969년 H. R. Hunting사의 편집자 Ether M. Swift가 Paul K. Swanson의 협력하에 재편하여 출판하게 되었다. 이 새로운 판은 기호표의 구조와 내용을 바꾼 것이 아니라 배열과 활자를 재조정하고 부분적으로 오자를 수정한 것이다. 새로운 판과 안내서의 서지적 사항은 다음과 같다.

*Cutter-Sanborn Three-figure Author Table.* Swanson-Swift Revision.

Chicopee, Mass., Distributed by H. R. Huntting Co., 1969.

Swanson, Paul K. and Swift, Ether M. *Instruction Book for C. A.*

*Cutter-Sanborn Three-figure Author Table.* Swanson-Swift Revision,

1969. Chicopee, Mass., H. R. Huntting Co., 1969.

## 2. 기호표의 특성과 장점

### 1) 표의 구조적 특성

Sanborn기호표는 전형적인 열거식 기호표이다. 저자기호로서 채택가능한 항목을 알파벳순으로 열거해 놓고 여기에 순차적인 숫자를 부여해 주는 형식을 가지고 있다. 열거식 기호는 구조에 따라 계층적(Hierarchical) 기호체계와 序數的(Ordinal) 기호체계로 나눌 수 있다. 계층적 기호체계란 各類와 그 類에 해당하는 하위류에 대하여 새로운 부호를 추가함으로써 기호를 통하여 상위류와 하위류 사이에 종속적 관계를 나타내도록 하는 기호체계를 말한다.<sup>7)</sup>

저자기호법에 있어서 이러한 계층적 구조는 일찍이 Cutter가 처음으로 저자기호법을 고안할 때 채용했다.<sup>8)</sup> 즉 A부터 Z까지 각각 0-9의 번호를 주고, 그 번호내에서 또 다시 알파벳을 세분시켜 두번째 10단위의 자리수를 주고, 필요한 경우 10 자리수의 각 단위별로 또 다시 100단위로 세분시키는 방법이다. 예를 들면 G는 첫

7) Buchanan, Brian. *Theory of Library Classification*. 정필모, 오동근 공역. 문헌 분류이론. 서울, 구미무역, 1989. p.91.

8) Cutter, Charles A. "Another Plan for Numbering Bbooks." *Library Journal*, 3(Sept. 1878). p. 249.

번째 자리수가 3이고, 이를 세분하는 경우 30-39까지 다시 10류로 나누어 두번째 자리수를 삼는다. 만일 한 분류번호아래 많은 문헌이 모이게 되어 세번째 자리수까지 필요로 한다면 30-39까지의 각 10단위마다 다시 10단위씩 세분시켜 300-309, 310-319와 같은 형식으로 100단위의 번호를 부여한다. 이렇게 될 경우 3, 31, 311은 각각 상위의 표목과 종속적 관계를 가지게 된다.

한편, 서수적 기호체계는 표목의 나열과 단순한 순차적인 번호부여 형식을 취한다. 즉, 저자명으로 채택될 수 있는 표목을 순차적(알파벳순)으로 나열하고 여기에 특정 숫자를 계속적으로 부여하는 형식을 취하고 있다. 따라서 서수적 기호체계를 적용할 경우 상위 표목과 하위 표목의 구분이 없으며, 부여되는 번호도 첫째의 표목에서 마지막 표목까지 일련순서로 부여하게 되는 것이다.

Sanborn기호표는 열거식 기호법이면서도 기호의 체계는 서수적 형식을 취하고 있다. 따라서 Sanborn기호표에서는 Cutter가 세 자릿수 기호표에서 채택했던 동일類(한 알파벳 문자내의 목표) 내에서의 3단계식 기호법, 예를 들면 3, 31, 311과 같은 계층적 구조를 가지지 않으며, 따라서 각 번호와 번호사이에는 아무런 因果關係가 없는 단순한 나열에 불과한 것이다.

분류표는 그 목적이 유별화에 있으므로 계층적 구조가 탐색의 범위를 확대하거나 축소하는데 도움을 줄 수 있는 장점을 가지게 된다.<sup>9)</sup> 그러나 저자기호는 근본목적이 동일 분류번호 아래에서의 문헌을 개별화시켜 서가배열 위치를 확정해주는 데 있기 때문에 기호표의 구조가 계층적이어야 할 필요는 없으며, 더욱이 각각의 번호가 특별한 의미를 가질 필요는 없는 것이다. 오히려 십진식 분류표와 같이 계층적 구조를 취할 경우 동일 알파벳내에서 기호가 3, 31, 311등으로 분할되어 배열과 검색에 있어서 혼동을 초래하게 될 것이다. 이러한 경향은 이미 오래전부터 Cutter의 세 자릿수 기호표를 채용하고 있는 도서관에서 문제점으로 지적되어 왔다.<sup>10)</sup>

9) Buchanan, Brian. op. cit., p.93.

10) Barden, Benta R. *Book Numbers: A Manual for Students with a Basic Code of Rules*. Chicago ALA., 1937. pp. 7-8.

Sanborn의 서수적 기호체계는 오늘날 자동화의 관점에서 매우 유리한 조건을 가지고 있다. Cutter의 세 자리수 기호표는 계층적 구조로 되어 있고, 번호부여방식이 동일 알파벳내에서 계층에 따라 한 자리수, 두 자리수, 세 자리수로 분할되어 있어 자동화 시스템을 설계하는데 어려움이 따른다. 한편 LC저자기호법은 번호가 부차적 전개규칙에 따라 두 자리수, 세 자리수로 전개될 때 마다 기존의 장서나 서가목록을 확인해 가면서 번호를 부여해 주어야 하기 때문에 자동 기호 매김에 한계가 있다. 이에 비해서 Sanborn의 기호표는 특정 저자의 저자기호가 고정적이어서 자동화시스템을 설계하는데 LC기호표보다 훨씬 유리한 장점을 가지고 있다.

## 2) 기호부여 방법상의 특성

기호표는 작성에 있어서 기호를 부여하는 방법은 순수기호법과 혼합기호법의 두 가지로 대별된다. 순수기호란 채용되는 기호가 문자만으로 이루어지거나 숫자만으로 이루어지는 경우를 말한다. 전자의 경우는 문자기호라 하고, 후자의 경우는 숫자기호라 한다.<sup>11)</sup>

순수기호는 미국에서 초창기에 고안되었던 저자명을 숫자로 변환하여 기호 매김을 하는 변환시스템<sup>12)</sup>으로서의 숫자기호와, 이 변환시스템이 나타나기 이전 영국에서 성행하였던 저자명의 몇자리 문자만을 따서 기호매김을 하거나, Melvil Dewey가 처음 저자기호법을 채용할 때 저자 성을 그대로 써 넣어주는 방법인 문자기호<sup>13)</sup>의 두 가지 형식이 있다. 문자기호는 Schwartz의 변환시스템으로서의 저자기호법이 고안된 이래 차츰 자취를 감추었으며, 저자를 숫자로 바꾸어 기호 매김을 하는 숫자기호도 John Edmands의 혼합기호법이 고안됨에 따라 오랜 생명력을 유지하지는 못하였다.

한편 혼합기호는 문자와 숫자를 결합하여 기호로 삼는 방법이다. 이 방법은

11) 仙田正雄. 圖書分類 と 圖書記號. 東京, 南書房. 1955. p.165.

12) Jacob Schwartz의 저자기호법과 Charles A. Cutter가 처음 저자기호표를 고안했을때는 저자명을 숫자로 변환시킨 변환시스템을 채용하고 있었다.

13) Comaromi, John P. op. cit., p.36.

1879년 John Edmands가 처음 고안하여 발표한<sup>14)</sup> 이후 오늘날까지도 저자기호법에서 표준적인 방법으로 적용되고 있으며, Sanborn기호표도 혼합기호법을 채용하고 있다.

Sanborn기호표는 문자와 숫자를 결합한 혼합기호법을 택함으로써 구분능력을 더욱 증대시켰다. “문자기호나 숫자기호만 사용하는 경우보다 양자를 다 사용하게 되면 구분능력을 극대화시킬 수 있고, 아울러 기호의 자릿수를 절약하여 기호를 단순화시키는데 기여한다.”<sup>15)</sup> 왜냐하면 26개의 알파벳 문자를 활용함으로써 동일한 숫자 자리수로서 구분능력을 월등히 높일 수 있고, 이러한 가능성은 결국 숫자 자리수를 줄여서 기호를 단순화시킬 수 있기 때문이다. Sanborn의 기호표가 모두 12,330개의 표목을 가지지만 이를 26개의 알파벳 문자별로 분할하였기 때문에 번호는 100단위에서 그칠 수 있게 되는 장점에서 증명된다.

Sanborn기호표는 숫자부여 형식에 있어서도 특성이 있다. Sanborn은 저자 성의 첫 문자와 더불어 숫자를 십진식으로 부여함으로써 철저한 알파벳순 배열을 가능케 하는 장점이 있고, 새로운 저자기호가 부여되더라도 기존 장서의 중간 중간에 삽입시킬 수 있는 장점을 갖게 된다.<sup>16)</sup> 이러한 중간 삽입의 이점은 장서량의 증가때마다 서가를 이동하고 재배치시켜야 하는 고정식 기호법의 불편을 개선시킨 것으로 평가할 수 있다.

저자기호법이 출현하게 된 기본적인 배경은 전통적인 수입순 배열법과 같은 고정식 기호법으로부터 탈피하여 문헌을 기존의 장서, 중간 중간에 삽입할 수 있도록 해주는 상대적(상관적) 기호법으로서의 전환에 있다고 할 수 있다. 이렇게 볼 때 알파벳 문자와 십진식 숫자의 결합은 상대적 기호법의 장점을 가장 효과적으로 이룰 수 있도록 해 준다.

Sanborn이 채택한 혼합기호법과 십진식 번호부여방식은 서가배열기준이 대다수의 분류법 즉, DDC, LCC, UDC의 어느 분류법과도 조화롭게 적용할 수 있고, 한 분

14) Edmands, John. “Plan for Numbering, with Especial Reference to Fiction : A Library Symposium.” *Library Journal*, 4(Feb. 1879). pp. 42-44.

15) Buchanan, Brian. op. cit., p. 90.

16) Cutter, Charles A. “Another Plan for Numbering Books.” *Library Journal*, 3 (Sept. 1878) p. 249.



류항목내의 저자별 유취와 서가배열을 용이하게 할 수 있다.

따라서 Sanborn이 채택한 혼합식 기호법, 십진식 기호법, 변환시스템등을 채용한 상대적 기호법의 특성은 저자기호법의 기능을 매우 유효하게 반영하고 있다고 평가할 수 있을 것이다.

### 3) 이용상의 특성

Sanborn기호표의 이용상의 특성은 표의 구조가 단순하고 간결하다는 점과 이에 따른 기호 매김의 시간과 노력이 절감된다는 점으로 요약될 수 있다.

Sanborn기호표는 저자 성의 첫 문자와 저자 성에 해당하는 1-100단위까지의 숫자의 조합으로 이루어져 구조가 매우 단순하다. 특히 기호표의 자리수가 문자, 숫자가 합쳐서 4자리수이므로 표의 경제적 한계에도 적합하다.<sup>17)</sup> 또한 기호표도 저자성의 알파벳순으로 배열되고, 여기에 해당번호를 일련 순서로 배정하고 있기 때문에 매우 손쉽게 검색할 수 있다. 따라서 기호를 작성하는 사서는 기호 매김을 위한 특별한 전문적 지식이 없이도 업무를 수행할 수 있고, 서가배열과 검색을 위해서도 특별한 지식이 요구되지도 않는다.

한편 Sanborn기호표는 표목수가 많아 외울 수 없고, 그래서 기호를 매길 때 일일이 표를 들추어 보아야 하는 불편이 단점으로 지적되기도 한다. 그러나 현재 통용되고 있는 대표적인 네 가지 기호표, 즉, Cutter의 두 자릿수 기호표 및 세 자릿수 기호표, Sanborn기호표, 그리고 LC저자기호표 중에서 기호표를 일일이 들추지 않고 기호 매김을 할 수 있는 기호표는 LC저자기호표 밖에 없다. 그러나 LC저자기호표는 표(전개원칙)는 외울 수 있어도 실제 기호 매김을 하는데 소요되는 시간은 오히려 Sanborn의 기호표보다 더 많이 걸린다.

왜냐하면 LC저자기호표는 항상 기존 장서의 기호 매김을 확인하여 동일한 저자기호 유무를 검토해 가면서 기호의 자리수를 전개해야 하기 때문이다.<sup>18)</sup> 따라서

17) 仙田正雄, op. cit., pp. 164-165.

18) LC저자기호표의 기호매김을 위한 특별한 절차, 즉 기존의 서가목록을 확인하는 과정은 이 기호표가 처음으로 소개된 이후, 1979년의 개정판, 그리고 최근 1986년의 개정판에서도 지속되고 있다. 특히 LC저자기호표는 원래 기호의 기본자리수가 1문자 1숫자로 구성하도록 되어 있었으나, 1986년의 개정판에서는 사용례에서 1문자 2자리수를 기본단위로 삼고 있어 기존장서의 확인절차는 불가결한 요소가 되고 있다.

기호의 첫자리수를 매기는 데는 LC기호표가 분명 빠르지만 둘째 자리수, 셋째 자리수로 전개될 때는 서가목록을 확인하여 기존의 저자기호와 중복되지 않도록 하고, 앞으로 편입될 문헌을 예상하여 번호를 배분하는 등의 절차가 필요하다. 특히, 1986년에 개정된 LC저자기호표의 적용 원칙은 그 예시에서도 나타나 있는 바와 같이 숫자 자리수의 기본단위를 두 자리수로 하고 있기 때문에 기존 장서의 확인 절차는 반드시 필요한 것이다. 이러한 요소를 감안하면 오히려 Sanborn의 기호표가 짧은 시간에 해당번호를 찾아내고 기호를 매길 수 있는 이점을 가지고 있다고 볼 수 있다. 실제로 Sanborn의 기호표로 특정 저자의 표목을 찾는 데 소요되는 시간은 20초 내외에 불과하다.

### 3. 기호표의 단점

#### 1) 표목 선정의 범위문제

“Sanborn기호표의 가장 큰 결함은 저자기호로서 채택될 수 있는 모든 요소, 즉, 인명, 단체명, 표제(주제), 국명, 지명등을 위한 기호 매김에 한계가 있다는 점일 것이다.<sup>19)</sup> 이는 기존의 열거식 기호표들이 가지는 공통적인 결함이기도 하하다.”<sup>20)</sup>

1891년 Sanborn이 기호표를 만들 때 어떤 자료에 근거해서 표목을 채택하였는지 그 근거를 밝히고 있지 않다. 다만 Cutter의 두 자리수 기호표를 확대 전개할 필요성에 따라 Cutter의 제안에 의해 만들었기 때문에<sup>21)</sup> Cutter의 두 자리수 기호표의 전개 원칙을 따랐을 것으로 보인다. 즉, “1878년 Sanborn의 기호표가 각 도서관의 통계로 이루어진 신중한 계산을 바탕으로 표목을 배분했다.”<sup>22)</sup>

19) 정필모. “한국문헌기호법연구 : 현행 열거식 저자기호법에 대한 비판”, 국회도서관보, 9권 5호 (1972. 7), p.10.

20) 리재철. “구조론에 입각한 한국저자기호표의 연구 : 한글구조상의 특성, 기입의 형식, 배열, 표기법의 문제등과 관련한 고찰”, 도서관학, 제13집(1970), p. 10.

21) Lehnus, Donald J. op. cit., p. 38.

22) Schwartz, Jacob. “A combined System for Arranging and Numbering”, *Library Journal*, 3 (Mar. 1878), p. 7.

고 하여 도서관의 목록에 의존하였음을 암시하고 있고, 뒤이어 나온 John Edmands의 기호표도 인명사전과 도서관의 저자목록을 참고하여 표를 구성하였으므로<sup>23)</sup> Sanborn도 전례의 방법을 따랐을 가능성이 높다고 볼 수도 있을 것이다. 그렇다면 인명은 오늘날 그 배분비율에서 큰 편차가 생기지 않았겠지만, 단체저자나 표제가 기본표의 대상이 되는 부분에선 상당한 변화가 있었을 것이다. 왜냐하면 1890년대의 도서출판 경향과 오늘날의 그것이 현저히 다르기 때문이다. 따라서 Sanborn이 어떤 자료에 의해서 표목을 채택하였는가 하는 문제 자체가 분석의 기준으로서 그렇게 중요한 것은 아닌 것으로 보인다. 중요한 것은 저자기호의 매김 대상이 되는 요소가 Sanborn기호표가 만들어 졌던 1890년대와 비교하여 오늘날 얼마나 달라졌는가 하는 점이다.

1890년대의 도서출판 경향은 개인저자에 의해 집필되는 것이 일반적인 관례였던 것으로 보인다. *American Book Publishing Record*의 1876-1949누적판<sup>24)</sup>에서 단체저자로 출판된 1900년 이전의 저작은 별로 발견할 수 없으며, 초기의 저자기호 표들이 주로 인명사전과 도서관의 저자목록을 중심으로 편성하였다는 점이 이러한 경향을 대변한다.

그러나 오늘날에는 학술단체와 연구기관, 정부조직 및 그 산하기관, 지역적, 국가적, 국제적 협회나 기구등 각종 단체에서 발행하는 출판물과 서명이 저자기호의 대상이 되는 연속간행물과 일부의 참고도서의 출판양이 방대하게 증가하였다. 다음에서 단체저자나 서명이 기본표목이 되는 상업출판물, 연구보고서, 정부간행물 및 연속간행물의 출판경향에 대해 살펴 보기로 한다.

먼저 단체저자에 의한 상업출판물의 출판경향을 살펴 보자. *Book in Print, 1990-1991* 년판 Author Series를 통해 조사해 보면, 전체 수록도서 약 864,000종 가운데 단체저자에 의해 출판된 도서는 약 35,000종으로서 전체의 약 4%에 해당된다. 또한 최근 20년 동안 단체저자에 의한 출판물의 증가현상을 *Books in Print*

23) Edmands, John. "Proportion in Initial Letter in Author Catalog". *Library Journal*, 4 (Feb. 1879). p. 56.

24) *American Book Publishing Record, 1876-1949 Cumulation*. New York, Bowker, 1950-1977. 15 vols.

를 통해 조사해 보면 단체저자에 의한 출판활동이 얼마나 가속화되고 있는가를 알 수 있다. 단체명의 첫 단어로 시작되는 대표적인 것이 American, International, National등이므로 이들을 중심으로 증가추세를 조사해 본 결과는 <표 1>과 같다.

<표 1> 단체에 의한 출판물의 증가추세

(단위 : 종수)

구 분	1969	1979	1986	1990
American	900	1200	1800	2800
National	900	900	1850	2800
International	1100	1400	1900	2200
계	2900	3500	5550	7800

이 통계에서 보듯이 70년대 10년 동안 증가량은 약 600종에 불과하였으나 80년대의 10년간 증가량은 4,300종에 달한다. 이는 각종 단체의 지속적인 증가현상과 이들의 활발한 학술활동의 결과에 기인하는 것으로 이런 추세는 앞으로도 계속될 것이다.

둘째, 연구보고서의 경우, Sanborn 기호표가 만들어졌던 당시에는 아직까지 문헌의 형태로 나타나지도 않았던 연구보고서는 2차대전후 급속한 성장을 하여 오늘날 주요한 문헌으로 위치를 구축하고 있다.

과학정보의 체계적 유통을 위한 매체로서의 연구보고서는 1902년 미국 국립표준국(National Bureau of Standard)에서 발행한 *Technologic Papers of National Bureau of Standard* 와 *Professional Papers of the United States Geological Survey* 를 그 시초로 보고 있다.<sup>25)</sup> 연구보고서가 본격적인 문헌의 형태로 나타나기 시작한 것은 2차 대전 이후 미국에서 Publication Board를 설립하고 난 이후였다. 이로부터 미국 정부주관하에 본격적인 연구 개발활동을 전개하여 “1950년대에 이미 정부지원을 받는 연구보고서의 생산량이 연간 75,000-100,000건에 이르

25) Subramanyam, Krishna. *Scientific and Technical Information Resources*. New York, Marcel Dekker, 1981. p. 101.

는 것으로 평가됐다.”<sup>26)</sup>

이런 추세에 따라 “미국 의회도서관은 연평균 10만건의 보고서를 수집하고 있고, NTIS(National Technical Information Services)도 자국의 300여 정부기관과 외국의 기관들로부터 연간 약 7만건 이상의 보고서를 수집하고 있다.”<sup>27)</sup>

보고서를 대규모로 수집하는 도서관에서는 보고서 자체의 번호를 중심으로 정리하는 경향이 있으나<sup>28)</sup> “일반 도서관에서는 보고서 자체가 하나의 주제를 다루고 있고, 보고서의 수집량이 많지 않기 때문에 일반적인 문헌의 정리과정을 거치고 있다. 또한 오늘날 도서관 네트워크의 발달로 인하여 대규모의 보고서를 수집하는 도서관들도 공동 네트워크의 이용을 원하기 때문에 정리방법을 AACR2로 바꿀것을 고려하고 있다.”<sup>29)</sup>

대부분의 보고서는 저자가 다수이고, 기관에 보고하는 형식을 취하기 때문에 단체가 그 문헌에 대한 대표성을 가진다. 또한 이용자들은 검색할 때 개인 저자보다는 단체명으로 접근하기 때문에 저자기호의 대상도 단체명이 된다.<sup>30)</sup>

셋째, 정부간행물의 경우도 유사하다. “미국의 경우 정부간행물은 1859년 설립된 정부인쇄국(Government Printing Office : GPO)이 연방정부의 의회, 사법부 및 행정부의 간행물을 출판, 배포하도록 법률로 규정함으로써 출판이 활성화되었다.”<sup>31)</sup> 그러나 정부간행물이 본격적으로 발행되고 자료로서의 가치가 높게 평가되기 시작한 것은 2차대전 이후의 일이다. 최근 통계에 의하면 “GPO에서만 연간 4만종 이상의 정부간행물을 출판하며, 이는 미국 전역에서 출판되는 정부간행물의 약 절반에 해당한다고 한다.”<sup>32)</sup> 이와 더불어 州政府에서 발행하는 각종

26) Tallman, J. E. “History and Importance of Technical Report Literature.” : *Sci-Tech News*, 15 (Summer 1961). pp. 44-46.

27) Auger, C.P. *Use of Report Literature*, London, Burtworth, 1975. pp. 20-26.

28) Subramanyam, Krishna. op. cit., p. 110.

29) Burriss, Elaine P. “Technical Report : A Comparison Study of Cataloging with AACR 2 and COSATI.” *Special Libraries*, 76 (1985). pp. 187-188.

30) 설문원. 기술보고서 관리에 관한 연구, 서울, 이화여대 대학원, 석사학위논문, 1986. p.48.

31) Levin, Marc A. “Access and Dissemination Issues concerning Federal Government Information.” *Special Libraries*, 74 (1983). pp. 128-129.

32) 최영옥. 대학도서관의 정부간행물 관리에 관한 연구. 서울, 이화여대 대학원, 석사학위논문, 1990. p. 7.

정부간행물을 포함한다면 그 수는 훨씬 더 늘어날 것이다.

이들 정부간행물은 통상 단일 주제를 취급한 독립 출판물로 발행되고 있으며, 자료의 형태도 대부분 장서로 등록될 수 있는 단행본 형태를 취하고 있다. 정부간행물의 자료적 가치가 증대되고 양적으로 확대됨에 따라 관리하는 방법도 일반 도서와 마찬가지로 통합하여 운영하는 방식이 보편화되고 있다.

넷째, 연속간행물은 정기적으로 출판되는 정기간행물과, 1년 단위로 출판되는 年刊物, 그리고 비정기적이지만 동일 지명으로 계속 간행되는 비정기 연속간행물이 포함된다. 이런 종류의 출판물들은 서명이 저자기호의 대상이 된다.

연속간행물의 연차적 증가현상은 과학분야에서 두드러진다. “Derek J. Price는 과학분야의 잡지는 1750년에 10종, 1800년에 100종, 1850년에 1,000종, 1900년에는 10,000종으로 증가하여 50년 단위로 10배씩 증가하는 것으로 보았다.”<sup>33)</sup> 그러나 “20세기에 들어와 연차적 증가현상은 둔화되어 1960년 이래로 연 4%씩 증가하고 있는 것으로 보고되고 있다.”<sup>34)</sup>

1977년 D. W. King이 조사한 바에 따르면 “전세계적으로 생산되는 과학기술분야의 잡지총수는 57,000종으로 집계되고 있다.”<sup>35)</sup> 이 수치에 연 4%씩 증가하는 것으로 계산한다면 1990년에는 약 90,000종에 이른다. 오늘날 학술적 가치가 비교적 높은 연속간행물을 대상으로 편집한 선택서지인 *Ulrich's International Periodicals Directory, 1989-1990* 년 판에 수록된 세계의 연속간행물은 모두 112,000종에 이르고 있다. *Ulrich's* 가 선택자임을 고려한다면 실제 출판량은 이보다 훨씬 더 늘어날 것이다. 이런 상황은 Sanborn 이 저자기호표를 발표하였던 1890년대의 상황과 너무 다르다.

이상에서 살펴 본 바와 같이 단체저자에 의해 출판되는 문헌과, 표제가 저자기호의 대상이 되는 문헌은 방대한 수에 이르고 있다. 이러한 요소를 저자기호표에

33) Price, Derek J. de Solla. *Science since Babylon*. New Haven, Conn., Yale Univ. Pr., 1961. p. 95.

34) Martyn, J. “The Growth of Journals: A Short History”, *Journal of Periodical Communication Studies*, 1 (1979). pp. 259-262.

35) King, D.W. *Scientific Journals in the United States*. Hutchinson Riss, King Research Inc. 1977. p. 12.

반영할 수 있도록 기호표가 만들어져야 하는데 Sanborn의 기호표에서는 취약한 부분이 매우 많다. 먼저 단체저자를 위한 표목의 선정을 중심으로 분석해 보기로 한다.

단체명은 일반적으로 그 첫 단어가 American, Korean등과 같이 국가를 중심으로 한 것, 지역이나 도시명을 중심으로 한 것, 주제명을 중심으로 한 것, 그리고 Association, Institute 등과 같이 단체나 기구를 나타내는 어휘를 중심으로 하여 표현된다.

*Encyclopedia of Assocoations* 1988년 판을 중심으로 단체명을 나타내는 첫 단어 아래 집합되는 단체의 수를 조사해 보면 American : 약 3,900개, Association : 약 1,900개, National : 약 5,300개, International : 약 4,300개에 이른다. 표목의 선정이 불합리한 것은 이들 단체가 출판한 문헌의 양을 보면 확인할 수 있다. *Bookd in Print, 1990-1991* 년 Author Series를 조사해 보면, American으로 시작되는 단체가 발행한 출판 총량은 약 2,800종에 이른다. 또한 Amer란 표목에 포함될 수 있는 개인저자의 성은 비록 빈도수가 낮기는 하나 Amer, Amerasinghe, Amreinger, Amerio, Amerman, Amermick, Amerong, Amerongen, Amerson, Amery 등 다수의 인명이 있다. 그러나 Sanborn기호표에서는 이를 위해 Amer란 하나의 저자기호밖에 배정되어 있지 않다.

이를 Sanborn기호표의 다른 표목과 비교해 보면 표목선정의 불균형을 알 수 있다. Smith란 성은 Sanborn 기호표에서는 무려 20개의 표목으로 세분되어 있다. *Books in Print, 1990-1991* 에서는 Smith란 성을 가진 사람의 출판 총량이 약 2,800종에 이른다. American으로 시작되는 단체저자의 출판량과 비슷한 수치이다. 그럼에도 불구하고 American은 독립된 표목으로 채택되지도 못하였는데 비해 Smith는 20개의 표목으로 세분되어 있다.

다른 하나의 예를 들어보면, National과 International로 시작되는 단체의 출판 총량은 National 이 2,800종, International 이 약 2,200종에 이른다. 그러나 Sanborn 기호표에서는 독립된 표목이 없이 Nati와 Int만 주어져 있을 따름이다. 이상의 결과를 보건대 Sanborn 기호표는 단체저자를 기호매김하기에 부적합한 것

으로 보여진다.

다음은 주제명을 중심으로 분석해 보기로 한다. 주제는 단체명의 첫 단어로 쓰이는 경우가 많지만 참고도서와 연속간행물의 표제로도 많이 쓰인다. *Ulrich's International Periodicals Directory, 1989-1990* 을 조사해 보면 특정 주제 아래 집합되는 연속간행물의 양을 측정할 수 있다. 연속간행물 표제의 첫 단어로써 많이 쓰이는 몇가지 어휘아래 집합되는 연속간행물의 수는 Journal : 2,600종, National : 1,900종, International : 2,300종, American : 1,900종에 이른다. 그러나 Sanborn 기호표에는 대부분의 주제명들의 독립된 표목을 갖고 있지 못하다. 즉, Amer, Jou, Nati, Int 등의 略字型 표목들이 있을 따름이다.

단체저자의 예에서 언급한 바와 같이 연속간행물도 분류번호에 의해 저자기호의 중복범위를 상당히 줄일 수 있다. 그러나 연속간행물의 첫 단어가 제한된 주제범위를 나타내는 단어로 시작되는 경우에는 저자기호의 중복가능성이 아주 높아진다. 예를 들면 *Ulrich's* 에서 Computer로 시작되는 연속간행물은 무려 400여종에 이른다. Art, Artist, Arts 등으로 시작되는 것은 300여종, 그리고 비교적 좁은 주제인 Library로 시작되는 연속간행물마저도 140여종이나 된다. 그러나 Sanborn 기호표에서는 Art, Compt, Libr 등 불완전한 하나의 표목으로 기호 매김을 할 수 밖에 없다.

표제를 위한 표목의 이러한 불합리성은 참고도서의 경우에도 동일하게 나타난다. 참고도서는 개인 및 단체저자가 기본표목의 대상이 되는 경우도 있으나, 표제가 기본표목이 되는 경우도 많다. 따라서 표제의 첫 단어를 중심으로 하여 기호 매김을 할 수 있어야 하나 Sanborn기호표에서는 주제명을 표목으로 삼지 않았기 때문에 표제를 위한 기호 매김에 한계가 있다.

참고도서에 있어서 표제의 첫 단어를 형성하는 대표적인 어휘들로서 Encyclopedia, Dictionary, Directory, Index, Handbook, Bibliography, Guide 등이 있으나<sup>36)</sup> Sanborn기호표에서는 이들 주제를 대부분 세분하지도, 독립 표목으로 채택

36) 'ALA Rules for Filing Catalog Cards, 규칙 33E, 도치서명'부분을 보면, 도치서명은 사용하지 않는 것으로 권고하고 있다. 따라서 미국 의회도서관은 이미 1940년 12월 30일 부터 도치서명의 사용을 중지해 왔으며, 양서정리업무를 위해 보편적으로 활용되고 있는 *Bibliofile*에서도 일간도서, 정기간행물, 참고도서 등 전부분에서 도치서명이 사용되지 않고 있다.



하지도 않았다.

Sanborn기호표의 표목 'Bibl' 에 해당하는 참고도서는 Bible, Biblio, Bibliografi, Bibliografia, Bibliografiia, Bibliografija, Bibliographia, Bibliographic, Bibliographical, Bibliographie, Bibliogaphisches, Bibliography, Biblioteca, Bibliothek 등의 어휘가 표제의 첫 단어를 형성하고 있다. Sheehy의 *Guide to Reference Books* 10판에서는 여기에 해당되는 서명이 약 1,200종이나 수록되어 있다. 그러나 Sanborn 기호표에서는 이들을 위해 'Bibl' 이라는 하나의 표목만 주고 있을 따름이다. 이러한 예는 International, 국가를 나타내는 American, British등의 어휘, 그리고 주제를 나타내는 많은 어휘에서도 동일하게 나타난다.

이러한 분석결과를 종합해 볼 때 Sanborn은 기호표를 만들 때 단체저자와 주제명(국명, 지명포함)을 별로 중요하게 생각하지 않았음을 알 수 있다.

## 2) 문자별 번호할당 비율의 문제

문자별 번호할당 비율이라 함은 알파벳 26문자의 각 문자별 표목의 점유비율을 말한다. 문자별 번호할당의 비율이 저자기호표에서 특히 중요한 요소로 취급되어야 할 이유는 실제 출판되고 있는(또는 출판되었던) 모든 문헌에서의 저자명의 문자별 점유율과 기호표에서의 배분비율이 유사하여야만 균형있는 기호 매김이 가능하기 때문이다. 만일 이 균형이 무시된다면 어떤 부분에서의 저자기호는 부족하고, 어떤 부분에서는 불필요하게 남아 돌아 기호표로서의 효용성이 반감된다. 따라서 "기호표의 구분능력은 표목수의 많고 적음보다 오히려 문자별 번호할당 비율과 관련이 있다고 보아야 할 것이다."<sup>37)</sup>

각 문자별 번호할당의 비율을 측정하기 위해서는 도서관의 장서목록이 어느 정도 유용할 것이다. 도서관의 저자목록이 개인저자와 단체저자를 포함하고 있기 때문에 개인과 단체가 광범하게 포함된 전화번호부가 오히려 문자별 번호할당비율을 측정하는데 유효한 도구가 될 수 있을 것이다. 이 가정은 1980년 Lehnus가 전화번

37) Lehnus, Donald J. op. cit., p. 50.

호부와 도서관의 저자목록이 밀접한 상관성을 가지고 있다는 조사결과에 기인한다.<sup>38)</sup> 그래서 본 연구에서는 Lehnus가 조사한 전화번호부의 문자별 점유율이 실제 출판상황과 어느 정도 일치하는가를 알아 보기 위해 Lehnus가 조사한 전화번호부에서의 문자별 점유율과 단체저자가 포함된 *Books in Print*의 저자목록에서 나타난 문자별 점유율을 비교해 보았다. *Books in Print*, 1990-1991년 판의 경우 약 500,000명의 저자가 저술한 854,000종의 문헌이 수록되어 있고, 약 20년 정도의 소급능력을 갖고 있기 때문에 문자별 점유율과 빈도수를 조사하는데 매우 적절한 자료가 될 것이다. 조사결과 <표2>에서 보는 바와 같이 최고 차이값 0.84(B)에서 최저 0(X)에 이르기까지 그 편차가 극히 적었으며, 전체 차이값의 합계가 6.24에 불과하였다. 이 결과는 저자기호표에서 문자별 표목의 배분비율을 결정하는데 있어서 이 두 가지 자료중 어떤 자료를 활용하든지 상관없음은 의미한다.

<표2>를 보면 전화번호부와 Sanborn기호표는 V의 최고 차이값 4.57에서 X의 최저 차이값 0.06까지 비교적 광범한 오차를 나타내고 있으며<sup>39)</sup>, 점유율의 차이값의 전체 합계가 40.52에 달한다.<sup>40)</sup> 차이값의 합계가 적으면 적을수록 전화번호부와 저자기호표의 문자별 배분이 일치하는 비율이 높기 때문에 Sanborn기호표의 편차는 대단히 높은 편이다.

그러면 Sanborn기호표는 오늘날의 출판 상황과 얼마나 일치하는가? 이 문제는 *Books in Print*와 Sanborn기호표의 문자별 점유율을 비교해 봄으로써 알 수 있다. 양자의 편차는 전화번호부와 Sanborn 기호표의 비교에서 나타난 것과 거의 유사한 결과를 보여주고 있다. 즉, 최고, 최저 차이값과 분포상황이 아주 유사하며 차이값의 전체 합계도 40.64로서 비슷한 결과를 나타내고 있다. 따라서 Sanborn기호표의 문자별 표목의 점유율은 실제 출판 상황은 정확하게 반영하고 있기 못함을 알

38) Ibid. p. 57.

39) 차이값에 있어서 그것이 +인가, -인가 하는 것은 그렇게 중요하지 않다. 표목을 더 많이 배분하였건 더 적게 배분하였건 간에 오차의 크기 자체가 중요한 것이다.

40) Lehnus도 필자와 동일한 방법으로 분석하였는데, 그 결과 차이값의 합계가 36.02로 나왔는데 이는 Sanborn 기호표에서의 표목의 수와 점유율을 계산할때 몇 부분에서 착오를 일으킨 결과이다. 예를 들면 Sanborn 기호표에서 V의 표목수는 729개이고, 점유율은 5.91이 되어야 하는데, Lehnus는 표목수를 81개를 보고 점유율을 0.7로 계산함으로써 차이값이 0.84에 불과하였다.

수 있다.

본 연구에서는 이에 더하여 이들 세가지 자료의 정확한 편차를 구해 보았다. 그 결과 전화번호부(A)와 *Books in Print* (B)의 표준편차는 0.347인데 비하여 전화번호부와 Sanborn기호표(C)의 편차는 2.086으로서 무려 6배의 편차를 나타내고 있다. 또한 저자명과 단체명의 실제적인 빈도를 나타내는 *Books in Print*와 Sanborn기호표간의 표준편차도 2.098로서, 전화번호부와와의 대비에서와 유사하게 6배의 편차를 나타내고 있다. 이 결과를 보면 Sanborn기호표는 우선 인명의 분포비율과 출판 상황을 정확히 반영하고 있지 못함을 알 수 있다. 따라서 각 문자별 번호할당의 비율이 잘못되었음이 입증된다.

〈표2〉 전화번호부, BIP, Sanborn의 문자별 점유율과 차이값

문 자	점 유 율 (%)			차 이 값		
	전화(A)	BIP(B)	Sanborn(C)	A - B	A - C	B - C
S	10.10	10.04	5.91	+ 0.06	+ 4.19	+ 4.13
M	9.05	8.50	↗	+ 0.55	+ 3.14	+ 2.59
B	8.65	9.49	↘	- 0.84	+ 2.74	+ 3.58
C	7.60	7.32	↗	+ 0.28	+ 1.69	+ 1.41
H	6.55	6.79	↘	- 0.24	+ 0.64	+ 0.88
W	5.80	5.36	↗	+ 0.44	- 0.11	- 0.55
G	5.25	5.29	↘	- 0.04	- 0.66	- 0.62
R	5.00	5.22	↘	- 0.22	- 0.91	- 0.69
P	4.80	4.52	↗	+ 0.28	- 1.11	- 1.39
D	4.65	4.76	↘	- 0.11	- 1.26	- 1.15
L	4.50	4.83	↘	- 0.33	- 1.41	- 1.08
A	4.30	4.18	↗	+ 0.12	- 1.61	- 1.73
F	3.75	3.85	↘	- 0.10	- 2.16	- 2.06
K	3.70	4.35	0.66	- 0.65	+ 3.04	+ 3.69
T	3.50	3.23	5.91	+ 0.27	- 2.41	- 2.68
J	2.75	2.18	0.66	+ 0.57	+ 2.09	+ 1.52
N	2.15	2.31	5.91	- 0.16	- 3.76	- 3.60
E	2.10	2.05	0.66	+ 0.05	+ 1.44	+ 1.39
O	1.42	1.41	↗	+ 0.01	+ 0.76	+ 0.75
V	1.34	1.53	5.91	- 0.19	- 4.57	- 4.38
U	0.90	0.50	0.66	+ 0.40	+ 0.24	- 0.16
I	0.78	0.94	↘	- 0.16	+ 0.12	+ 0.28
Z	0.56	0.62	↘	- 0.06	- 0.10	- 0.04
Y	0.54	0.52	↗	+ 0.02	- 0.12	- 0.14
Q	0.25	0.16	0.07	+ 0.09	+ 0.18	+ 0.09
X	0.01	0.01	↗	0.00	- 0.06	- 0.06
합 계	100.00	100.00	100.00	6.24	40.52	40.64

문자별 번호할당의 비율에 있어서 왜 이러한 문제가 나타났을까? 이미 Schwartz와 Edmands가 알파벳 각 문자로 시작되는 이름의 실제 빈도수에 따라 번호의 정확한 분배를 결정하고자 시도한 바 있다. 즉, Schwartz는 각 도서관의 통제로 이루어진 신중한 계산을 바탕으로 하고,<sup>41)</sup> 아울러 책의 크기별 출판비율을 근거로 기호표를 작성했다.<sup>42)</sup> 또한 Edmands는 문자별 번호할당 비율을 결정하기 위해 인명사전과 5개 주요 도서관의 저자목록을 참고로 했다.<sup>43)</sup> 이 두 가지 방법은 각각 1878, 1879년에 *Library Journal*에 발표되었고, Sanborn의 기호표는 1891-1895년에 만들어 졌으므로 그녀는 이미 이러한 방법론을 알고 있었을 것이다. 그럼에도 Sanborn은 이러한 방법을 응용하지 않았던 것으로 보인다.

Sanborn기호표에 있어서 문자별 번호할당 비율의 부정확성은 각 문자별 자리수의 최대치를 3개의 큰 그룹별로 미리 정해 놓고, 이에 따라 표목수에 배정했기 때문이다. Sanborn은 각 문자별 성의 빈도를 개략적으로 판단하여 세그룹으로 나누었다. 첫째 그룹은 A, B, C, D, F, G, H, L, M, N, P, R, S, T, V, W를 가장 빈도수가 높은 성으로 판단하여 번호의 자리수를 세 자리(999번까지)로 하고, 두번째 그룹은 모음 E, I, O, U와 자음 K, J, Y, Z를 중간 그룹으로 판단하여 두 자리수(99번까지)의 번호를 배정하였다. 그리고 가장 빈도수가 낮다고 생각되는 Q와 X는 한 자리수(9번까지)로 하였다. 그리고 영문 대문자 O와 혼동될 수 있는 숫자 0을 모든 번호배정에서 생략하여 끝자리수가 각각 999, 99, 9가 되도록 하였기 때문에 표목수는 첫번째 그룹이 모두 729개, 두번째 그룹이 81개, 세번째 그룹이 9개로 고정되고 말았다.

이 방법은 엄밀히 말해서 성의 빈도를 정확하게 반영하지 못했고, 자리수를 통일시키지도 못했다. <표2>를 보면 전화번호부에서는 10.10%를 점유하는 S와 1.34%

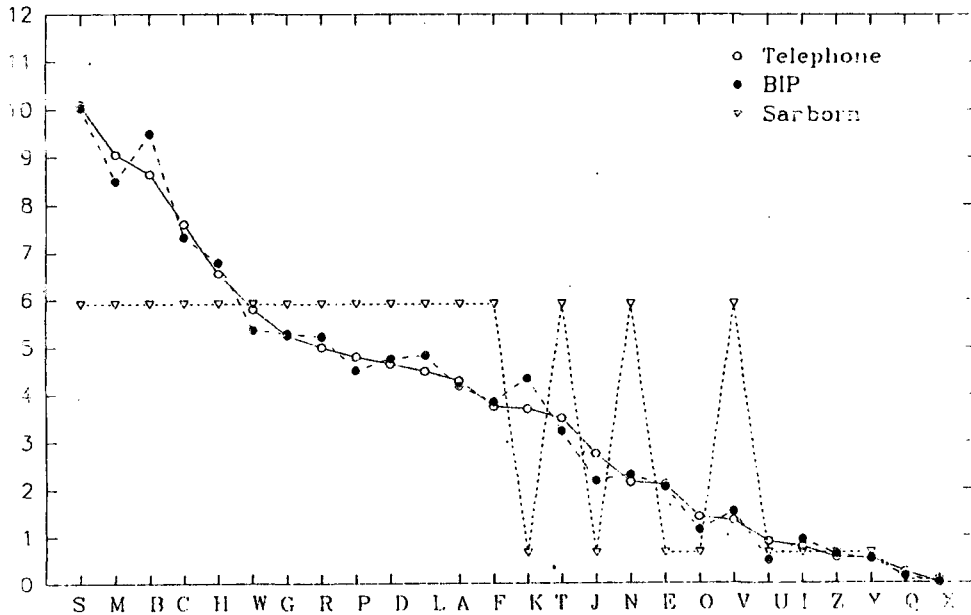
41) Schwartz, Jacob. "A combined System for Arranging and Numbering," *Library Journal*, 3 (Mar. 1878). p. 7.

42) Ibid. pp. 7-8.

43) Edmands, John. "Proportion of Initial Letter in Author Catalog," *Library Journal*, 4 (Feb. 1879). p. 56.

를 차지하는 V가 기호표에서는 공히 5.91%를 점유하고 있다. Sanborn의 기호표에서 문자별 번호할당 비율이 잘못 적용된 더 정확한 예는 K와 V, 그리고 O와 V를 비교해 보면 더욱 분명해 진다. 전화번호부에서는 3.70%를 점유하는 K는 기호표에서 불과 0.66%(81개)를 차지하고, V는 전화번호부에서 1.34%인데도 불구하고 기호표에서는 오히려 K의 9배나 되는 5.9%(729개)를 배정하고 있다. 한편 O와 V는 전화번호부에서 각각 1.42%와 1.34%의 비슷한 점유율을 갖고 있다. 그럼에도 Sanborn의 기호표에서는 O를 위해 81개의 기호(0.66%)를 배정하고 있고, V는 그 9배나 되는 729개(5.91%)를 배정하고 있다. <그림1>은 문자 별로 표목수를 잘못 배정함에 따라 전화번호부 및 BIP와 Sanborn기호표가 얼마나 많은 편차를 나타내는가를 잘 보여 준다.

이러한 결과를 보면 Sanborn기호표는 번호할당에 있어서 크게 보아 두 가지의 문제점을 가지고 있는 것으로 보인다. 그 하나는 B, C, E, H, J, K, M O, S의 경



<그림1> 전화번호부, BIP, Sanborn기호표의 문자별 점유율 비교

우 처럼 번호의 할당비율이 낮게 책정된 경우와, 다른 하나는 F, R, T, V처럼 기준치 보다 높게 배정된 경우에서 나타난다.

번호의 배분비율을 적정치보다 낮게 책정한 경우, 즉, B, C, H, M, S처럼 번호의 자리수를 100단위(999번 까지)로 제한함으로써 더 이상 전개가 불가능한 경우는 어쩔 수 없는 것으로 보인다. 더 많이 전개할 경우 숫자 자리수가 1,000단위로 가기 때문에 불가피하게 729개의 표목을 배정할 수 밖에 없었을 것이다. 그러나 숫자 자리수를 더 많이 배정할 수 있는데도 불구하고 10단위 자리수(81개 표목)로 묶어 놓은 E, J, K, O의 경우는 문자별 점유율을 전혀 고려하지 않았거나 잘못 판단했기 때문에 생긴 결과이다. 그래서 이 4개의 문자그룹에서는 표목마다 저자가 집중되는 양상을 보이게 된다.

다른 한가지 문제는 F, N, T, V와 같이 번호할당 비율을 불필요하게 높게 책정함으로써 다른 문자그룹과 불균형 현상을 보이고 있는 경우이다. 그러나 이 문제는 세분화 능력을 저해하지 않는다는 점에서 볼 때 전술한 문제점과는 비교가 되지 않는다.

이렇게 각 문자 별로 적절한 배분비율을 적용하지 않으면 각 頭文字 그룹내에서 나열된 표목과 표목간의 간격과 번호배정이 균형을 이루지 못해 기호표의 효용성이 반감된다. 예를 들면 V나 N처럼 인명 및 단체명의 빈도가 적은 문자그룹에 번호배정의 최대치를 높게 설정해 놓으면 그 문자그룹의 표목은 불필요하게 세분될 것이며, 반대로 K나 J처럼 비교적 빈도가 높은 문자그룹에 번호배정의 최대치를 낮게 설정해 놓으면 그 반대현상이 나타나 저자가 밀집되는 현상이 나타나게 된다. 이 문제는 전화번호부에서 각 문자별 점유율과 Sanborn 기호표에서의 점유율이 유사한 경계는 전화번호부에서 각 문자별 점유율과 Sanborn 기호표에서의 점유율이 유사한 경우를 적정치로 보았을 때 나머지는 그 차이값이 +, -로 커질수록 더욱 부적합하게 된다. Sanborn은 기호표를 만들 때 문자별 번호배정의 최대치를 '100 무 큰 그룹별'로 미리 정해 놓고 표목을 배정했기 때문에 이런 현상이 나타나게 되는 것이다.

<표3>의 예시는 전화번호부나 Sanborn 기호표에서 유사한 점유율 분포를 가진

는 0와 V, 그리고 두 자료에서 각각 4.35%를 차지하는 K부분이 기호표에서 번호할당 비율을 잘못 책정함으로써 각 문자 별로 저자가 얼마나 집중되고 분산되는가를 보여 준다. 이 데이터는 *Books in Print, 1990-1991* 에서 각 표목에 해당되는 저자의 수를 조사한 것이다.

〈표3〉의 예시는 번호할당 비율이 잘못 책정되었을 경우 기호표의 효용성에 어떤 영향을 미치는가를 극명하게 보여 준다. 0와 K는 번호배정의 최대치를 낮게 책정함으로써 매 표목에서 저자가 집중되는 양상을 보인다. 특히 K부분은 밀집되는 양상이 심각하다. 반대로 V는 번호배정의 최대치를 너무 높게 책정함으로써 대부분의 표목에서 저자가 매우 분산되는 경향을 보이고 있다. 이런 양상은 차이값이 큰 문자의 전 부분에서 나타난다. 즉, 차이값이 -쪽으로 큰 F, N, T, V 등에서는 위의 V와 같은 분산현상이, +쪽으로 큰 B, C, J, K, M, S 등에서는 0 및 K에서와 같은 밀집현상이 나타나게 된다.

〈표 3〉 Sanborn기호표의 각 표목에 집합되는 저자의 수

Oa	160	Va	8	Kas	825
Ob	212	Vac	4	Kau	480
Obr	205	Vacc	9	Kaw	65
Obs	21	Vaccar	5	Kay	320
Oc	52	Vacch	0	Ke	145
Och	47	Vacchi	1	Keat	190
Oco	7	Vacco	1	Kee	450
Oconn	264	Vach	6	Kei	135
Ocor	7	Vacho	4	Keith	95
Oct	10	Vacq	7	Kel	1125
Od	71	Vad	5	Kem	60
Ode	120	Vade	3	Kemp	150

Sanborn이 기호표를 만들 때는 이런 문제를 예상하지 못했던 것으로 보인다. Sanborn은 당시까지 사용되던 Cutter의 두 자릿수 기호표가 표목수가 적어 구분능력이 취약하다는 요구를 반영할 목적으로 자리수를 늘리는 데는 관심이 있었지만 표목을 문자별로 적절하게 배분하는 데는 큰 의미를 부여하지 않았음이 확실하다.

### 3) 표목 배분의 간격문제

표목 배분의 간격이란 한 알파벳 문자내에서 표목으로 나열된 항목과 항목사이의 細分된 상태를 말한다. 문자별 번호 할당비율이 적합치 않을 때 표목 배분의 전체적 불균형 상황이 초래되는 것과 같이, 표목간의 배분간격이 일정치 않으면 어떤 번호에는 너무 많은 저자가 모이게 되어 구분능력이 떨어질 것이며, 반대로 그 간격이 너무 세분되어 있으면 기호가 불필요하게 복잡해져 작업능률이 떨어지게 된다. 따라서 표목의 배분간격은 저자기호표의 유용성을 결정하는 요소로서 문자별 번호할당 비율에 못지않게 중요하다.

Sanborn기호표에서는 표목간의 배분간격이 균형을 이루지 못하는 부분이 너무 많이 나타나 기호표로서의 가치를 떨어뜨리고 있다. 원래 열거식 저자기호표는 저자 성의 빈도에 따라 세밀하게 전개함으로써 구분능력을 높일 수 있는 장점을 가지고 있으나,<sup>44)</sup> Sanborn기호표에서는 이 장점을 잘 활용하지 못하고 있는 것이다. 먼저 전화번호부와 문자별 번호할당 비율이 상대적으로 균형을 이루고 있는 기호표의 A부분을 중심으로 표목 배분간격의 문제점을 분석해 보기로 한다.

〈표 4〉는 표목이 부분적으로 너무 세분된 상태를 나타낸 것으로서, *Books in Print*, 1990-1991년판과 1986-1987년 판에서 Sanborn기호표의 각 표목아래에 집합되는 저자의 수를 조사한 것이다.

*BIP*, 1990-91년 판을 보면, Arin에서 Ariu까지 모두 8개의 표목에 저자명이 39개, Ayr에서 Aza까지는 모두 7개의 표목에 85개의 저자 성이 나열되어 있다. *Books in Print* 에서의 A부분 전체 저자 약 21,000명을 Sanborn기호표에서의 해당 표목 729개를 나누면 평균적으로 한 표목에 약 30명의 저자가 배분되는 것이 정상이다. 따라서 Arin에서 Ariu까지는 2개의 표목만으로 충분할 것이며, Ary에서 Aza

44) 정필모. “한국 문헌기호법연구 : 현행 열거식 저자기호법에 대한 대안” 국회도서관보 9,5 (1972, 7). pp. 6-7.



〈표 4〉 Sanborn 기호표의 표목 세분화 양상

	90-91	86-87		90-91	86-87
Arin	4	4	Ayr	2	3
Ario	6	6	Ayre	45	36
Arip	0	1	Ayrt	1	1
Aris	11	10	Ays	5	6
Aristi	1	1	Ayt	1	1
Aristo	1	2	Ayton	18	14
Aristop	10	6	Aza	13	7
Ariu	6	3			
계	39	33		85	68

까지도 3개 표목으로 충분할 것이다. 이러한 현상은 26개 문자에서 공통적으로 나타난다.

이러한 상황과 반대되는, 표목간의 간격이 밀집되는 경우의 한가지 예를 보면, Sanborn기호표에서 Ad1과 Adm은 *Books in Print*, 1986-1987, 1990-1991을 조사해 보면, 이 표목 사이에 들어가는 저자의 성이 Adlam, Adland, Adleman, Adlen, Adler, Adlerbum, Adlercreutz, Adler-Golden, Adler-Nissen, Adlerny, Adlington, Adlmann, Adolff등 14개의 성이 나열되어 있다. 이러한 성은 대개 흔치 않은 성이어서 출판량이 많지 않다. 따라서 Ad1이란 하나의 표목만으로 가능할지 모르나, Adler란 성의 경우는 다르다. *Books in Print*의 Adler란 성아래 모여 있는 저자의 수는 1986-1987년 판에서 175명, 1990-1991년 판에서는 무려 240명에 이른다. 앞에서 언급한 두개 부분의 성을 위한 표목 15개를 모두 합친 것 보다 많다. 그럼에도 불구하고 이들을 위해 하나의 독립된 표목도 배정되어 있지 않다. 다른 표목들과의 균형을 고려한다면 Adler란 성만을 위해서 최소한 Adler, Adler, C, Adler, F, Adler, J, Adler, M, Adler, S, Adler, W의 7개 표목은 배정되어야 한다.

표목간의 배분간격이 불균형을 이루는 문제는 거의 전 문자그룹에서 광범하게 나타난다. 이를 구체적으로 살펴보기 위해 문자별 배분비율을 과다하게 높게 설정한 V, 비교적 적절하게 배분한 W, 비교적 적게 배분한 O, 그리고 최소한으로 배정한 K의 4가지 문자그룹에서 저자가 밀집되고, 분산되는 경향을 <표5>의 예시를 중심으로 살펴 보기로 한다.

먼저 V부분을 살펴 보기로 하자. V부분은 표목이 너무 세분화되어 있기 때문에 대부분의 표목에서 10개 미만의 저자 성이 집합되는 경향이 있으나,<sup>45)</sup> <표5>의 예시에서 보는 바와 같이 어떤 표목에는 200명 이상의 저자가 모여 있는가 하면, 1명의 저자도 해당되지 않는 표목도 많이 있다.

W부분은 전화번호부와 기호표의 문자별 점유율이 비교적 균형을 이루고 있는 문자 그룹이기 때문에 불균형 현상이 나타난다. 특히 Wolffe에서 Wolffe, P까지 3

<표 5> Sanborn 기호표의 표목별 저자의 밀집 및 분산경향

Voltu	0	Wolcott	12	Omo	32	Kas	825
Voltz	2	Wolcott,M	6	Omu	8	Kau	480
Volu	5	Wold	39	On	265	Kaw	69
Volv	19	Wolf	102	Ons	63	Kay	320
Von	115	Wolf, J	65	Op	82	Ke	145
Vond	244	Wolf, P	350	Opp	136	Keat	190
Vonk	131	Wolffe	0	Or	38	Kee	450
Vono	184	Wolffe, J	0	Orb	53	Kei	135
Voo	6	Wolffe, P	0	Ord	41	Keith	95
Voor	35	Wolfg	0	Ore	95	Kel	1125
Vop	4	Wolfr	44	Orf	20	Kem	60
Vor	37	Wolk	76	Org	48	Kemp	150

45) *Books in Print*의 V부분은 총 저자수가 약 7,800명 정도이므로 이를 표목 729개와 나누면 한 표목당 평균 10명 정도 배분된다.

개의 표목은 한명의 저자도 해당되지 않는다. 그런데도 바로 위의 표목인 Wolf, P 에는 350명의 저자가 밀집되어 있다. 또한 Wolf를 3개의 표목으로 세분시켜 놓았는데 각각 102명, 65명, 350명씩 배분되어 하나의 인명내에서도 균형을 이루지 못하고 있다.

전화번호부의 문자별 점유율보다 조금 적게 번호가 할당된 0부분도 문자간 배분 간격의 편차가 매우 심하여 표목에 따라 수십 배씩 편차를 나타내고 있다. 또한 번호배분비율이 매우 적게 할당된 K부분에서는 거의 대부분의 표목에서 저자의 밀집경향이 나타나고, 표목간의 불균형 현상이 심하여 문헌의 개별화가 어렵도록 되어 있다.

표목간 간격의 불균형 현상은 한 문자내에서 저자가 밀집되고 분산되는 경우를 비교해 보면 더욱 뚜렷하게 나타난다. <표6>은 문자 M에서 저자가 밀집되고 분산되는 상반된 경향을 보여 준다.

M부분은 *Books in Print* 에서 약 44,000명의 저자를 포함하고 있다. Sanborn기

<표6> Sanborn 기호표의 문자 M에서 나타난 표목의 간격

저자번호	표 목	저 자 수	저자번호	표 목	저 자 수
141	Mace	174	792	Montig	0
142	Macer	71	793	Montl	0
143	Macf	197	794	Montlu	1
144	Macfi	5	795	Montm	1
145	Macg	399	796	Montmi	0
146	Macgo	126	797	Montmo	4
147	Macgr	237	798	Monto	12
148	Macgu	154	799	Montp	2
149	Mach	155	811	Montr	6
150	Macho	72	812	Montri	0
151	Maci	310	813	Montro	17

호표의 M부분 표목 729개를 나누면 한 표목당 평균 60명 정도의 저자가 배정되는 것이 정상이다. 그러나 141번부터 151번까지는 대부분 표목에서 저자가 밀집되어 개별화가 어려운 반면, 792번부터 813번까지는 표목을 불필요하게 너무 세분화시켜 놓았다.

이상 표목간격의 분석을 통해서 볼 때, Sanborn은 표목의 세분화를 위해 신뢰할 만한 참고자료를 활용하지 않고 편자의 개략적인 판단에 따라 나누었음을 추정할 수 있다.

## 결 론

이상에서 Sanborn기호표의 성립배경, 구조적 특성, 그리고 장점 및 단점에 대해서 분석적으로 평가해 보았다. 논의된 내용을 요약하면 다음과 같다.

Sanborn 기호표의 특성과 장점은 :

- 1) 기호법의 기본 구조는 열거식이면서 기호의 체계는 서수적 형식을 취하고 있다. 이러한 구조적 특성은 기호의 체계를 손쉽게 이해하도록 하고, 각 표목마다 고유한 번호가 주어져 있어 기호표의 자동 검색시스템을 설계하기에 용이하다.
- 2) 저자 성의 첫 문자와 숫자를 결합한 혼합기호법을 택함으로써 구분능력을 극대화시키고, 기호의 자릿수를 절약하여 경제적 한계에 적합하도록 했다. 아울러 변환시스템을 채용함으로써 저자의 철저한 알파벳순 배열을 가능케 하며, 배열 변환시스템을 채용함으로써 저자의 철저한 알파벳순 배열을 가능케 하며, 배열시 기존 장서의 중간 삽입이 가능하도록 하였다.
- 3) 기호의 구조가 단순하고 간결하여 기호배정의 시간과 노력을 절감할 수 있으며, 배열과 검색을 용이하게 해 준다.

한편 단점으로는 :

- 1) Sanborn기호표는 인명을 중심으로 표목을 선정하였으므로 저자기호의 매김대상인 모든 요소 즉, 인명, 단체명, 표적(주제), 국명, 지역등을 광범히

게 기호 매김하는데는 한계가 있다. 따라서 단체 저자와 연속간행물 및 참고 도서의 표제를 기호화하기 어려운 단점이 있다.

- 2) 알파벳 26개의 각 문자 그룹별로 번호를 할당할 때 그 비율을 기준없이 적용함에 따라 기호표의 문자별 점유율에서 표준적인 참고자료와 심한 편차를 나타내고 있다. 이러한 현상은 문자 그룹별로 저자가 밀집되거나 극도로 분산되어 저자기호표로서의 효용성을 반감시키는 요인이 된다.
- 3) 한 문자그룹내의 표목과 표목간의 간격을 결정하는 데 있어서 실제적인 인명 및 단체명이나 출판의 빈도와 일치시키지 않았기 때문에 표목별로 저자가 집중되고 분산되는 경향이 심하다. 이러한 현상은 특정 표목(저자번호)에서 너무 많은 문헌이 집중되므로 개별화를 불가능하게 하고, 어떤 표목에서는 해당되는 문헌이 전혀 없어 저자번호를 쓸 수 없도록 한다.

이러한 평가를 통해 볼 때 Sanborn기호표는 구조적 특성면에서 장점도 많으나 표목을 선정하여 번호를 할당하고, 문자별로 번호할당의 배분비율을 책정하며, 각 표목간의 간격을 조정하는데 있어서 신뢰할 만한 참고자료를 활용하지 않았음을 알 수 있다. 이 결과 Sanborn기호표는 오늘날의 관점에서 새롭게 개편될 필요성이 있는 것이다.

---

## On the Cutter-Sanborn Thtee-figure Author Table

### Abstract

Park, Joon-Shik

The purpose of this is to evaluate the contents through the analysis of the merits and demerits, and its special quality of the *Cutter-Sanborn Three-figure Author Table*. The abstract of this study is as follows : With the results of the analysis of *Cutter-Sanborn's Tabls*, several advent-ages were found in the area of enumerative-ordinal numbering system, mixed notation, and information retrieval. A defect is that each corporate author and serials and reference book's titles are difficult to be numbered, and there is serious diviation in percentages of the numbers assigned to each letter because each letter apportion rate was applied without any criteria. Such phenomenon decrease the effectiveness as author notation in several letters. The separation of headings are seriously unbalanced because the separation of heading is not consistent with actual publishing trends. This phenomenon made difficult individualization of materials in several letters.