分析合成式 著者記號 自動化에 관한 硏究

(한글順圖書記號法 第2表를 中心으로)

韓 啓燮 (동아대학교 경영정보학과 교수)
崔 昌鎬 (동아대학교 경영정보 석사과정)

제1장 序 論

Ⅰ. 硏究의 目的

圖書館에서는 本然의 奉仕 目的을 달성하기 위하여 각종 資料에 書架上의 配列 位置를 指定하는 請求記號를 일정한 法則에 따라 각각 賦與하고 있다.

請求記號는 司書가 資料를 조직적이고 합리적으로 配列 할 수 있도록 할 뿐 아니라 이용자가 스스로 書架에 접근하여 檢索할 수 있도록 하는 連結 기능을 가진 기호이다. 司書들은 이들 資料들이 각각의 독립성을 가지며 서로 중복되지 않도록 "一記號 一資料" 또는 "一位置 一資料" 의 請求記號를 賦與하고 있다.[1)]

請求記號는 제1단계로 資料가 갖는 主題 내용을 分析하여 記號化한 分類記號가 매겨진다. 그러나 分類記號는 동일한 主題의 資料가 많기 때문에 分類記號 자체만으로 완전하게 書架上 위치를 구분해 주지 못하므로, 圖書記號를 竝行하여 記號上으로 個別化시키는 역활을 한다. 즉, 同一 分類 아래 資料를 순서대로 配列하기 위해 圖書記號를 붙여[2)] 書架上에서 각 資料에 대한 配列 위치가 각각으로 유지될 수 있도록 최종적으로 확정해 주는 記號法이다.

分類記號를 일관성 있게 賦與하기 위한 각종 노력이 圖書館 내부에서 계속 시도되었으나,[3)] 著者記號는 分類記號에 비하여 중요성이 결코 떨어 지지 않는데도 一貫性 있는 기호를 賦與하기 위한 연구와 分析 評價가 크게 뒤떨어지고 있다. 왜냐하면 著者記號 매김 방법이 理論上으로는 간단하나 실제 업무에서는 의외로 복잡하기 때문에 司書들이 著者記號를 賦與하는데 비 경제적이라 할 만큼 많은 시간과 노력을 기울이고 있지만 노력에 비하여 만족할 만한 결과를 얻지 못하고 있기 때문이다. 그러므로 圖書館에서 정확하고 일률적인 著者記號를 賦與하기 위해서는 현재의 記號매김 방법외에 次元이 다른 수준의 記號매김 自動化에 대한 연구가 검토되어야 할 상황이다.

따라서 본 硏究의 目的은 첫째, 현재 國內 圖書館중 절반 이상이 사용하고 있는 分析合成式 著者記號表를 自動化시켜, CRT상에서 著者記號와 著作記號가 매겨지도록 하고, 또한 記號의 자리수를 줄이기 위하여 目錄Data Base의 著者들을 통계Package를 이용, 字母 出現 頻度數를 찾아서, 현재 사용중인 記號表가 현대의 著作者 성격을 어느 정도 수용하면서도, 自動化에 적당한 구조인가를 分析하여 自動化되는 記號表를 작성하는 것이다. 또한 둘째 목적은 特定 著者가 分類에 따라 著者記號가 다르게 매겨지던

---

1) 金南碩, 圖書記號, 大邱, 啓明大學校出版部, 1988. p.6.
2) 今まと子, 資料分類法及び"演習, 東京, 水村房, 1985. p.102.
3) 朴俊植, 英美著者記號表硏究, 未刊本博士學位論文, 中央大學校大學院, 1991. p.1.

것을 새로운 記號表의 適用에 의해 區分 能力을 극대화시켜, 그 자리수를 가능한 적게 사용하면서 著者마다 고유한 著者記號를 가질수 있도록 하는 것이다.

셋째, 지금까지는 著者記號를 매기고 나서 書架配列에서 著者記號의 중복 매김을 일일이 확인하던 번거로운 작업방식을, 컴퓨터 CRT상에서 特定 著者와 類似 著者의 기호매김을 한 화면에서 一目瞭然하게 확인하도록 하여 分析合成式 著者記號法의 가장 큰 단점인 기호의 書架配列 확인 과정을 없애려는데 두고 있다.

Ⅱ. 硏究의 方法 및 範圍

본 연구는 국내 圖書館의 절반 이상이 사용하고 있는 分析合成式 著者記號法인 한글順圖書記號法 제2표의 子音과 母音의 記號配列이 적절하게 이루어져 電算化에 적합한 記號表 인지를 確認하기 위하여 D大學圖書館의 目錄Data Base에서 基本記入 (Main entry)인 Tag[4] 1XX와 副出標目(Added entry)인 Tag 7XX에서 기호매김 대상어의 子音과 母音의 출현 빈도를 統計Package SAS로 조사했다.

따라서 종래의 기호매김 방법은 整理業務에서 상당한 시간과 知的 노력을 要하던 것을, 새로운 기호매김 방법인 自動化시스템을 導入함으로써 얻어질 수 있는 著者記號 매김과 記號表의 개선 방안에 연구의 초점을 두었다.

그러나 본 연구에서의 제한점은 한글順圖書記號法에서 기호매김의 예외 사항은 著者記號 自動化시스템에서는 모두 수용할 수 없어서 프로그램상에서 선택적으로 記號를 賦與하도록 하였다. 특히, 現存하는 모든 人名과 團體, 學會등 著作에 관련된 모든 標目을 대상으로 子音과 母音의 모든 頻度調査를 할 수 없으므로 1985년부터 1991년12월까지 D대학圖書館에서 KORMARC Format(KORea MAchine Readable Catalogue = 韓國機械可讀形目錄)으로 入力된 資料중 국내에서 출판된 單行本과 國內 碩博士 學位論文, 中國, 日本등의 漢字文化圈 圖書를 대상으로 制限하였다.

제2장 著者記號法의 機能과 構成

Ⅰ. 著者記號의 機能

著者記號는 동일 分類記號 아래에서 圖書를 二次 配列하기 위하여 著者名을 配列에 직접 사용하지 않고 이에 대응하는 숫자나 문자를 조립하여 記號化 시킨 것이다.

---

4) Tag는 MARC (機械可讀形目錄 = MAchine Readable Catalogue)에서 目錄의 각 요소를 區分하여 컴퓨터에서 식별 할 수 있도록 숫자 記號化 한것.

만약 著者를 어떤 형태로 가공하지 않고 이름을 配列에 직접 사용하게 된다면 藏書가 적을 경우에는 책등의 書誌事項으로 配列이 가능하겠으나, 同一 分類 아래 同一 著者 또는 姓名이 類似한 著者가 많이 모일 경우에는 配列에 혼란이 생기므로 일일이 區分하여 配列할 수 있어야 하기 때문에 著者를 숫자나 기호로 바꾸어서 配列하게 된 것이다. 이러한 기능을 가진 著者記號法은 1880년 Cutter가 體系化된 두자리 著者記號法을 최초로 발표했을때 著者記號의 필요성에 대한 논란이 많이 제기되었다.

1880년대의 圖書館 藏書는 당시의 출판상황이나 사회여건이 著者記號 없이 分類記號만으로 書架 配列과 이용에 지장이 없었기 때문에 이미 著者記號를 채택하여 사용하고 있던 圖書館도 廢棄하는 경우가 있었다.[5] 그러나 文化의 발전으로 階層이 多變化되고 科學 技術이 급격히 발달되어 새로운 硏究者와 著作者가 증가하고 著作物이 폭증했기 때문에, 이용과 관리에 혼란이 생기지 않도록 이를 일일이 개별화 시킬 필요가 제기되었다. 따라서 대다수 圖書館이 사용할 수 있는 理想的인 著者記號表를 작성하기 위한 著者記號表의 作成基準으로 다음과 같은 사항이 필요하게 되었다.

첫째, 區分性과 조기성이 있어 쉽게 기호를 매길 수 있어야 하고,

둘째, 기호는 展開性이 뛰어나 登載되지 않은 새로운 表目이 등장하더라도 새 項目의 追加에 별도 작업이 없어야 하며,

셋째, 기호가 편리하게 사용되기 위해서는 기호의 길이가 짧고, 分類記號의 경제적 限界는 3-4개, 圖書記號는 4개 정도이며,[6]

네째, 記號表를 바꾸게 될 경우에는 비교적 비용이 적게 들고 혼란이 적어 書架配列과 카드 작성이 쉬워야 한다.

이상과 같이 제시된 記號表 작성 기준을 요약해 보면, 著者記號는 分類記號의 配列原則과 합치되도록 配列性이 있어야 하고, 기호의 구성은 단순하면서도 전개성이 있어야 하며, 論理的인 구조를 가져야 할 것이다.

Ⅱ. 著者記號法의 構成 原理

1. 構成原理

著者名으로서 資料를 個別化 할때, 著者名의 첫 글자 한 자만으로 資料를 개별화하면 區分性이 약하고, 이름 전체를 사용하여 개별화시키면 藏書量이 적을 때에는 문제발생이 적겠지만 藏書가 증가하면 이들을 중복되지 않게 區分해 주기 위하여 整理에

---

5) 박준식, 전게논문, p.45.

6) 仙田正雄, 圖書分類と圖書記號, 東京, 蘭書房, 1955. pp.164-165.

많은 시간이 소요 된다. 또한 著者名이 길면 著者記號의 목적인 簡潔性, 區分性, 單純性이 없어져 二次配列의 목적에 부합되지 않는다. 그러므로 이러한 著者記號法의 목적에 부합되는 記號法은 다음과 같이 크게 2가지로 區分 할 수 있다.

첫째, 圖書記號의 목적에 맞도록 기호를 細分하여 기호의 구성이 단순하면서 區分性과 展開性, 識別性을 좋게 하기 위해서 標目의 대상이 되는 著者名을 미리 列擧하여 表를 만들어 놓고, 그 表에 해당하는 著者名을 숫자화 시키는 列擧式方法(Enumerative System)과,

둘째, 表出하고자 하는 事物과 사상을 일일이 項目을 잡아 늘어 놓는 대신, 문자를 더 이상 分解할 수 없는 最小 元素인 子音과 母音으로 分解하여 숫자를 매겨 놓고, 記號化시키고자 하는 文字의 元素를 어떤 連結原則에 의하여 表의 숫자에 대입하여 이를 다시 組合하여 合成하는 일종의 組立式 또는 組合式 방법인 分析合成式(Analytic Synthetic System)이 있다.[7]

列擧式이나 分析合成式의 記號法에서 사용하는 기호는 區分性을 증대시키기 위하여 주로 문자와 숫자를 混用하는 混合記號法을 사용하고 있다. 混合記號法의 장점은 기호의 區分 能力을 극대화 하는데 있으므로, 2차 區分인 기호값이 동일한 숫자를 가지더라도 1차 區分인 文字記號에 따라 전혀 다른 配列 위치를 가진다.

2. 分析合成式記號表

列擧式 分類表와 著者記號表는 일정 기간이 지나면 表에 登載되지 않은 새로운 項目을 포함하여 더욱 細分하여 列擧된 表가 계속 발표되어야 한다. 현대에는 資料의 주제가 다양화되고, 著者의 성격이 주로 個人에서 團體, 學會, 研究所, 國家機關등 성격이 多樣化되었기 때문이다.

지금까지 著者記號表의 중요성에 비추어 Cutter-Sanborn表나 우리나라의 列擧式 著者記號表가 한번도 개정되지 않았는데, 개정되지 못한 이유가 여러가지 있겠으나 결국 圖書館 司書들은 이 表에 列擧되지 않은 著者에 대한 기호를 賦與하여 資料를 區分시키기 위해서는 별도의 노력이 필요하게 된 것이다.

특히, 특정 문자의 著者가 계속적으로 증가하여 記號表上에 著者가 없을 경우에는 司書가 著者를 記號表에 기록하는것 만으로는 늘어 나는 著者를 다 해결할 수 없기 때문에, 列擧式 記號表는 새로운 著者의 등장에 따라 계속 개정해야 하는 부담을 갖게 된다. 이것을 극복하려고 考案된 것이 分析合成式 방법이다.

---

7) 이재철, "구조론에 입각한 한국인著者記號表에 관한 연구", 圖書館學, 第1輯(1970). p.7.

分析合成式 著者記號表중 대표적인 表는 美國議會圖書館 (LC = Library ofCongress)에서 채택하고 있는 LC著者記號表이며, 한국에서는 이재철 교수의 한글順圖書記號法이 있다. 分析合成式 方法의 理論的 뒷받침은 1957년 영국의 도킹會議(Dorking Conference)에서 分析合成式의 절대적 필요성이 인정되어 동 회의의 6개 중요 합의사항 중의 하나로 채택되어 널리 권장하고 있는 방식으로써 情報檢索의 기초가 된다고 주장하였다.[8]

LC著者記號表의 구성은 대개의 分析合成式 著者記號表와 같이 使用者가 表를 전부 외워서 쉽게 사용할 수 있도록 아주 간단하게 구성 되었는데, 著者記號表를 외울수 있다는 것은 기호를 매기는데 있어서 대상 단어를 일일이 表를 찾아 매겨야 하는 列擧式記號法 보다 짧은 시간에 기호를 매길 수 있는 장점이 있다.

LC著者記號表의 區分性은 記號配定을 위한 기본 원칙만 적용 시킬 경우 불과 188개의 區分 밖에 하지 못하므로, 이론적인 區分 능력에 있어서는 列擧式 著者記號表인 Cutter-Sanborn表의 12,330개보다 區分性이 훨씬 뒤떨어 진다. 그러나 제1차 區分인 LC分類表 자체가 세밀하게 구성되어 分類表에서 거의 區分이 되어지기 때문에 著者記號 자체의 區分性에는 큰 비중을 두지 않았다.

LC著者記號表의 區分性을 계산해보면,

S 다음의 문자 해당 기호수 8개
Qu 다음의 문자 해당 기호수 7개
나머지 子音 19 * 7 = 133개
母音 5개 * 8 = 40개 합계 188개의 區分性이 있다.[9]

우리나라의 分析合成式 記號表인 이재철 교수의 한글順圖書記號法은 8개의 表가 있으나 각 表마다 기호의 배정이 다르고, 區分性이 가장 강한 제2표는 15만8천6백 여개의 區分性이 있다.

산출근거는, 자음 14개 * 모음 9개 = 126개
1259개 문자기호 * 126개 숫자기호 = 158634개 이다.[10]

국내의 列擧式 記號表는 장일세 씨의 東洋書 著者記號表는 4,455개의 區分性이 있고 이춘희 교수의 東書著者記號表는 3,780개[11] 이므로 한글順圖書記號法은 이론상 어느 列擧式 著者記號表보다 방대한 區分性이 있다. 그러나 分析合成式 著者記號表는 출현 빈도가 높은 文字에 대한 기호 配定을 세심하게 配慮하지 못하여, 特定 文字의

---

8) International Study Conference on Classification for Information Retrieval (Dorking,1957), Processing, London, Aslib, 1957. p.82.
9) 박준식, 전게논문, p.19.
10) 이재철, 한글順圖書記號法, 서울, 아세아문화사, 1982. p.17.
11) 이재철, "구조론에 입각한 한국인 著者記號表에 관한 연구", 전게논문. p.18.

출현 빈도가 높은 著者의 경우에는 기호가 중복되지 않도록 기호의 자리수를 임의로 늘여 나가야 한다. 그러므로 著者記號의 자리수가 姓의 첫 문자를 포함하여 최저 2자리 부터 시작하여 계속 늘어나는 단점이 있다.[12] 반대로 잘 사용하지 않는 문자는 기호의 사용이 거의 없는 경우도 있다.

<表-1 한글順圖書記號法 제 2표>[13]

자음기호			모음기호	
1	ㄱ	ㄲ	1	ㅏ
21	ㄴ		2	ㅐ(ㅑ ㅒ)
22	ㄷ	ㄸ	3	ㅓ(ㅖ)
23	ㄹ		4	ㅕ(ㅖ)
3	ㅁ		5	ㅗ
4	ㅂ	ㅃ	6	ㅘ(ㅙ ㅚ ㅛ)
5	ㅅ	ㅆ	7	ㅜ(ㅝ ㅞ ㅟ ㅠ)
6	ㅇ		8	ㅡ (ㅢ)
7	ㅈ	ㅉ	9	ㅣ
81	ㅊ			
82	ㅋ			
83	ㅌ			
84	ㅍ			
9	ㅎ			

3. 列擧式記號表

列擧式 記號表의 대표적인 Cutter-Sanborn 記號表는 알파베트系列 文獻의 著者記號를 매기는데 가장 많이 쓰이고 있는 記號表로써, 1880년 Cutter의 2자리 記號表를 補完하여 1895년 Kate G. Sanborn에 의하여 발표되었다. 그후 Olin記號法, Merrill 記號法등의 列擧式 記號表가 발표되었다. 各國에서는 이를 모방하여 그 나라의 言語에 맞게 작성하여 많이 사용하고 있다. 일본에서는 이를 모방하여 日本人著者記號表를 Cutter식으로 만들어서 사용하고 있으며[14], 우리나라에서는 이춘희 씨의 東書著者記號表, 장일세 씨의 東洋書著者記號表, 정필모 씨의 韓國文獻記號表가 있다.[15]

장일세 씨는 列擧式 記號表의 장점에 대하여,[16]

1) 表가 간단히 구성되어 있다.

---

12) 이재철, 한글順圖書記號法, 상게서, pp.29 - 31.
13) 이재철, 상게서, p.13.
14) 장일세, "한국인著者記號表의 구성원리", 圖協月報,第3卷(1962), p.43.
15) 박준식, 전게논문, p.14.
16) 장일세, 상게논문, p.43.

2) 보기 쉽다.

3) 누구나 쓸수 있다.

4) 수십년 동안 사용하여 발전되었다.

5) 십진식(十進式)으로 무한히 전개가 가능하다.

6) 라마자(羅馬字)를 사용하고 있는 모든 人名에 적용할 수 있다고 주장하였다.

한편 정필모 교수는 列擧式 記號表의 장점을,기호가 單純하고 簡潔하며 필요한 곳에 細密한 전개가 되어 있다고 하였고,[17] 박준식 교수는 기호의 配列基準이 대다수 分類表의 기준과 합치되므로 配列과 檢索이 편리하다고[18] 주장하였다.

<表-2 全國圖書館 著者記號表 사용현황 >[19]

동 양 서			서 양 서		
저자기호표	도서관수	비율%	저자기호표	도서관수	비율%
이재철식	77	53.1	Cutter표	88	60.7
이춘희	3	2.1	LC	4	2.7
장일세	22	15.1	Merill	1	0.8
정필모	1	0.7	기타	10	6.9
Cutter표	4	2.7	무응답	29	20
기타	8	5.4	수입순	13	8.8
무응답	12	8.3			
수입순	18	12.6			
합계	145	100%	합계	145	100%

이러한 장점 때문에 表 2 에서 보는 바와 같이 국내 圖書館에서 西洋書의 著者記號表를 거의 Cutter-Sanborn表를 사용한다고 가정하면, 동양서의 著者記號表도 역시 列擧式 著者記號表가 사용되어야 하겠으나, 分析合成式 記號表인 한글順圖書記號表를 사용하는 圖書館이 과반수 이상인 것을 보아 圖書記號의 채택은 分類表와의 配列 기준이 합치하는 記號表를 채택하기 보다는, 著者記號表를 채택할 당시에 사용되고 있던

---

17) 정필모, "韓國文獻記號法研究 - 現行列擧式著者記號法에 대한 代案", 國會圖書館報, 9권5호(1972). pp. 6-7.

18) 박준식, 전게논문, p.17

19) 홍옥자, 한국圖書館에서 채택하고 있는 분류법의 실태에 관한 연구, 未刊本碩士學位論文, 梨花女子大學校敎育大學院, 1977. p.67.

著者記號表중 가장 사용하기 쉽고 입수하기 쉬웠던 表가 채택 되었으리라 생각할 수 있다.

列擧式著者記號表의 단점을 살펴보면,

1) 人名만을 表 構成을 위한 頻度 統計의 대상으로 삼고 團體名, 管轄 區名, 標題등을 고려하지 않았기 때문에 標目의 선정이 용이하지 않다.[20]

2) 가나다順 配列이 되지 않는다.

3) 主記號의 수가 부족하다. 한글은 單子音 14자와 單母音 10자 (2중 또는 3중 母音까지 합하면 23자)로 구성되어 있어서, 母音記號는 그 형태상 기호로 사용하기에 부적당하며, 子音記號는 기호로써 그다지 큰 손색은 없으나 그 기호의 수가 적다고 생각된다.[21]

4) 기호가 방대하여 기호를 매길 때 表를 들추는 시간과 노력이 많이 든다.

제3장. 한글順圖書記號法의 特性과 變形 記號表의 作成

Ⅰ. 한글順圖書記號法의 特性과 長點

東洋書 整理를 위해 대다수 국내 圖書館에서 사용하고 있는 分析合成式 著者記號表는 1958년 이재철 교수가 발표한 동서저자기호표가 최초의 記號表였다.

초판에는 제1표와 제2표가 수록 되었으며, 1970년에는 제3,4,5,6表를 수록한 제2판을 발표 한 후 뒤이어 제7,8표등 모두 8종의 記號表를 발표하였고,1982년에는 7표와 8표를 크게 수정하여 한글順圖書記號法을[22] 발표하였다.

이재철 교수는 이 表의 작성을 위하여 전화번호부를 이용하여 人名과 團體名의 記入語에 대한 子音과 母音의 頻度를 조사하였고, 일반어의 表題記入과 外來語記入(인명, 지명등)을 위해서 문교부 조사의 글자잣기 통계를 참고하였으며, 韓,中,日의 人名, 團體名, 標題, 일반 主題名등이 주로 漢字에 의해 구성되기 때문에 漢字사전의 音 索引의 字母頻度 통계를 여기에 가산하여 평균치를 내어[23] 한글의 子音과 母音에 1부터 9까지 숫자를 배정하였다.

分析合成式 記號法인 LC著者記號法과, 表 1에서 보는 한글順圖書記號法의 表 構成은 列擧式 著者記號表와는 달리 表의 構成이 단순하여 외우기가 쉽고, 藏書量에 따라

---

20) 박준식, 전게논문, p.17

21) 정필모, 전게논문, p.7

22) 金明玉, 資料分類論, 서울,歐美貿易, 1986,pp.245 - 249.

23) 이재철, 구조론에 입각한 한국인著者記號表에 관한 연구, 전게논문, p.22.

상대적으로 기호의 자리수를 얼마든지 늘여 나갈 수 있다.

表 1에서 보는 바와 같이 記號表가 子音과 母音으로 區分되어 일련의 순서에 의하여 숫자가 配定되어 있으면 기호매김을 위해서 表를 찾거나 記號의 組合을 위해서 별도의 노력이 필요치 않기 때문에 전문지식이 없어도 짧은 시간에 기호를 매길 수 있고, 새로운 著者記號가 매겨지더라도 書架에 配列되어 있는 기존 藏書의 配列 秩序를 흐트리지 않고 중간 중간에 資料를 挿入 시킬수 있는 이용상의 장점이 있다. 表의 編纂에 있어서도 列擧式보다는 훨씬 용이하게 文獻을 그룹화 할 수 있는 柔軟性이 있고 表의 구성이 간단하기 때문에 表의 작성이 쉽고 새로 발견되는 類를 表에 삽입하기 쉬운 장점이 있다.[24]

Ⅱ. 한글順圖書記號法의 短點

한글順圖書記號法은 기호로 바꾸고자 하는 著者名의 字母를 字素로 잡아 이를 빈도 조사 끝에 숫자 記號化한 것으로 著者記號를 매길때 著者名(記入語)의 첫문자와 더불어 이를 합성하여[25] 기호매김을 하고 있다. 그러나 이 記號法은 문자의 字母를 元素로 分解하여 숫자화한 단순한 기호 구조를 가지므로 기호매김 대상 資料의 복잡성에 대해서는 전혀 對處할 수 없다. 특히, 韓國人은 陰陽五行과 門中의 行列字를 따서 이름을 짓는 특이한 방법이기 때문에 同性同本이 아니더라도 姓이 같으면서 金木水火土의 邊이 들어가는 行列字를 사용하게 되면 거의 비슷한 字母를 가진 글자가 되기 때문에 기호값이 비슷하여 혼란을 加重시키고 있다.

이런 경우에 列擧式 著者記號表는 사전에 細分하여 表를 만들 수 있지만, 分析合成式은 著者名 둘째자의 子音과 母音만으로 기호를 매기기 때문에 둘째자의 字母가 같은 경우에는 둘째자의 終聲이나 셋째자의 字母로 더 區分해 주어야 하지만, 둘째자의 종성 이하는 기호매김 대상에서 제외되므로 字母에 따라 기호를 더 확장할 수 없어 表 3에서 처럼 資料의 區分을 위해 어쩔수 없이 기호의 자리수를 임의로 늘여 나가야 한다. 또한 전 국민의 반수 이상을 차지하는 한국의 5대 姓氏인 金, 李, 朴, 崔, 鄭 氏 姓의 경우에는 같은 姓氏에 本이 다르더라도 대개 行列에 따라 이름이 지어졌기 때문에 기호의 중복이 더 심하게 나타나서 다른 姓氏 보다 자리수가 더 길어 지게 된다.

이상과 같은 한글順圖書記號法의 단점을 정리해 보면,

첫째, 특정 著者가 경우에 따라 分類가 다르면 동일한 著者記號를 가질 수 없다.

---

24) Buchana Brain, *Theory of Library Classification, London, Clive Bingley,* 1979. 정필모, 吳東根 共譯, 文獻分類理論, 서울, 歐美貿易, 1989, pp. 38-43. 재인용.
25) 이재철, 전게논문, p. 7.

기호매김의 대상이 되는 著者의 성명이 類似하여 姓氏가 같고 성명의 둘째자의 初聲과 中聲이 같을 경우에는 表에 의하여 같은 기호가 매겨지므로 기호의 자리수를 줄이기 위하여 分類마다 기호를 다르게 배정한다.

둘째, 기호가 간결하지 못하다. 장서의 증가에 따라 기호를 계속해서 전개할 수 있도록 하였기 때문에, 同一 分類내에서 同名異人이거나 유사한 이름을 가진 著者의 文獻이 집중될 경우 表 3과 같이 한 단위의 숫자 부가만으로 부족하므로 계속해서 자릿수를 늘여 나가야 한다.

셋째, 한글 字母順 配列이 되지 않는다. 同一 分類 아래에서 著者名이 유사한 경우에는 字母順에 관계없이 작성자 또는 圖書館 내부의 규정에 의하여 임의로 판단하여 숫자를 추가시키기 때문에, 著者나 서명을 配列語로 삼아 字母順으로 配列하는 원칙이 이루어 지지 않는다.[26]

Ⅲ. 變形 한글順著者記號表의 작성

어떤 業務를 電算化하기 위해서는 지금까지 실행해 오던 업무에 대하여 電算化가 가능한가를 분석하여 電算化가 성공할 수 있도록 基礎 資料의 구조와 業務處理方式을

〈 表-3 현대 作家 記號表 〉[27]

작가명	소설	시	수필	희곡
김달진		22		
김대봉		23		
김대현		24		
김덕연	24			
김도양			25	
김도희	25			
김동길			252	
김동리	252		254	
김동립	253			
김동명		25	256	
김동민	254			
김동석		252		
김동수	255			
김동인	256			
김동일		254		
김동진	258			
김동환		258	258	
김두원		26		

26) 이재철, 전게논문, p.5.
27) 이재철, 한글順圖書記號法, 전게서, p.53.

電算化에 적합하도록 바꾼다든지, 또는 組織의 運營체제를 바꾸어서 電算化를 추진한다. 基礎資料의 구조를 바꾼다는 것은 自動化가 가능할 수 있도록 資料의 조직체계를 바꾸기 때문에, 이런 變化가 다른 업무에 어떤 영향을 미치는지 업무 전반에 대한 影響評價를 실시하고 결과에 따라 電算化의 범위가 결정되어야 할 것이다.

1895년 Cutter-Sanborn表를 발표할 당시에는 記號表의 이용도가 낮아 채택마저 불투명하였고, 컴퓨터의 출현은 전혀 예상할 수 없었으므로 장래에 表의 사용 방법이 어떻게 변할지를 예측할 수 없었다. 다만 타자하기 쉽고 發音하기 좋고, 기호를 쉽게 매기되 중복이 일어나지 않는 것이 좋은 著者記號表의 조건이었다. 이러한 관점에서 著者記號表가 만들어졌기 때문에 列擧式이나 分析合成式 모두가 현 체제대로 自動化하기에는 構造的인 어려움을 많이 가지고 있다.

그러나 圖書館 資料의 계속적인 증가는 分類表나 著者記號表가 自動化에 부적합하더라도 自動化를 하지 않을 수 없게 되었고, 컴퓨터와 周邊 技術의 발달은 이런 어려운 문제를 쉽게 해결해 줄 수 있는 단계까지 되었다. 그러므로 著者記號의 自動化는 기존의 기호체계를 근본적으로 바꾸지 않고 새로운 類가 추가되거나 세분되는 정도로 自動化가 되어야만 書架配列에 혼란이 생기지 않을 것이다.

이상적인 著者記號 自動化는 司書가 著者記號를 매기기 위하여 일일이 表를 찾는 노력과 시간을 없앨수 있어야 하고, 特定 著者가 單一記號를 가지며 著作記號도 함께 매길수 있어야 할 것이다. 이러한 목적을 달성하기 위하여 分析合成式 著者記號表를 自動化하는데 있어 著者記號表가 갖추어야 할 이상적인 요건을 살펴보면,

첫째, 分類目錄의 配列 原則과 합치되어야 한다. 著者記號는 속성상 分類의 보완책으로 書架配列의 질서 유지를 위하여 分類記號의 配列 원칙과 합치되어 書架 配列이나 검색에 있어서 혼동을 줄이고 일관성을 갖게 하는 것이[28] 중요하다.

둘째, 기호는 단순하고 간결하며 자리수가 고정되어야 한다. 기호의 자리수가 고정되면 資料를 탐색하거나 配列할 때도 편리하다.[29]

셋째, 特定 著者의 기호가 고정되어야 한다. 기호를 매기는데 있어 시간과 노력이 절감되고 檢索 接近點으로[30] 활용하기 위해서는 一著者,一記號의 원칙이 일관성 있게 유지되어야 할 것이다.

이러한 요건을 갖춘 著者記號表를 만들수 있는 방안을 연구하기 위하여 D대학 圖書館에서 축적된 目錄Data Base에서 다음의 통계를 조사하였다.

1) 기호매김의 대상이되는 標目의 첫째자와 둘째자의 출현 빈도를 조사.

---

28) 이재철, 전게논문, pp.2-4.
29) 박준식, 전게논문, p.12.
30) 박준식, 상게논문, p.29.

2) 둘째자의 初聲,中聲의 출현 빈도 조사.

3) 8대 姓氏별 둘째자의 출현 빈도 조사.

4) 8대 姓氏별 둘째자의 초성,중성의 출현 빈도를 조사하여 한글順圖書記號法 제2표를 다시 작성하였다.

1. 標目의 종류와 선정범위

子音과 母音의 통계대상은 基本記入 標目 부분인 Tag 100 個人名과 110 團體名, 111會議名 및 集會名, 130 統一書名의 資料와, 副出記入 著者인 Tag 700,710,711,730의 資料중 제2指示字(Indicator)가 "0"인 韓國人 또는 韓國人의 한글표기 著者와, 제2指示字가 "1"인 外國人 또는 外國語의 한글표기 著者, 書名 基本記入도 標目의 대상이 되므로 서명의 둘째자를 통계의 대상으로 삼았다.

이것은 한글順圖書記號法의 기호매김 대상 언어는 한글만 대상이 되기 때문에 基本記入 標目과 副出記入 標目이 한글인것만 사용하기 때문이다.

통계를 취합한 결과 韓國人名과 團體名, 한글로 표기된 외국인을 포함하여 총33,749건이 었다. 실제 Data Base에 축적된 資料는 11만여건이 넘었으나 1冊 1著者로 Data Base가 구축되어 있으므로 통계를 위하여 중복되는 著者와 複本 著者와 일부 Data 誤謬는 제외하였다.

2. 字母의 출현 빈도와 記號의 배정

조사된 총 33,749건의 資料는 目錄Data Base에서 전체 이름중 둘째 자의 출현 빈도를 初聲과 中聲으로 字素를 분해하여 표4,5와 같이 조사하고, 8대 姓氏별 둘째자의 初聲과 中聲의 出現 頻度 統計値인 表6,7과 비교하여 한글順圖書記號法 제2표의 變形記號表를 表 8과 같이 작성하였다.

일반적으로 한국의 姓氏는 김,이,박,최,정 氏 순으로 人口가 많은 것으로 알려져 있으나 본 통계에서 특이하게 나타난 사실은 부록 表 11의 전체 통계에서와 같이 "이"자가 제일 많은 6.7%의 빈도를 나타 내었고, 다음은 김,한,박,최,정 氏의 순으로 나타났다.

본 硏究에서는 이,김,한,박,정,오,조,최를 통계의 標本으로 삼고 이 여덟 문자를 편의상 8대 姓氏라 부르고, 初聲과 中聲,終聲의 출현 빈도를 조사 하여 나타난 통계치와, 8대 姓氏에서 사용된 문자의 통계를 기초로 하여 한글順圖書記號法 제2표를 다시 작성 하였다.

여기에 集計된 여덟자가 전체 통계에서 23.7%를 차지하고 있고, 첫 글자에서 使用頻度가 높은 순서대로 33번째 字 까지가 19,249회로 50.1%를 차지하고 있어 1차 區分은 약50자 이내에서 거의 區分이 되기 때문에 8대 姓氏의 통계를 중요 자료로 삼았다.

8대 姓氏에서는 부록 表 12와 같이 "이"자가 26.53%로 가장 많은 빈도를 나타내었다. "이"자가 가장 많이 나타난 이유는 日本 人名 "이시-", "이와-", "이노-", "이토-"

〈表-4 둘째자 子音 出現 頻度 統計 〉

SAS

JUNG1	FREQUENCY	PERCENT	CUMULATIVE FREQUENCY	CUMULATIVE PERCENT
	551			
ㄱ	3548	10.8	3648	10.8
ㄲ	622	1.8	4270	12.7
ㄴ	1028	3.0	5298	15.7
ㄷ	1930	5.7	7228	21.4
ㄸ	97	0.3	7325	21.7
ㄹ	2643	7.8	9968	29.5
ㅁ	1766	5.2	11734	34.8
ㅂ	1848	5.5	13582	40.2
ㅃ	6	0.0	13588	40.3
ㅅ	4530	13.4	18118	53.7
ㅆ	129	0.4	18247	54.1
ㅇ	5977	17.7	24224	71.8
ㅈ	3633	10.8	27857	82.5
ㅉ	231	0.7	28088	83.2
ㅊ	1295	3.8	29383	87.1
ㅋ	754	2.2	30137	89.3
ㅌ	975	2.9	31112	92.2
ㅍ	548	1.6	31660	93.8
ㅎ	2089	6.2	33749	100

〈表-5 둘째자 母音 出現 頻度 〉

JUNG1	FREQUENCY	PERCENT	CUMULATIVE FREQUENCY	CUMULATIVE PERCENT
	551			
ㅏ	5673	16.8	5673	16.8
ㅐ	1395	4.1	7068	20.9
ㅑ	4918	1.5	7559	22.4
ㅒ	3353	9.9	10912	32.3
ㅓ	1526	4.5	12438	36.9
ㅔ	2206	6.5	14644	43.4
ㅕ	205	0.6	14849	44.0
ㅖ	4021	11.9	18870	55.9
ㅗ	679	2.0	19549	57.9
ㅘ	65	0.2	19614	58.1
ㅙ	570	1.7	20184	59.8
ㅚ	3981	11.8	24165	71.6
ㅛ	393	1.2	24558	72.8
ㅜ	36	0.1	24594	72.9
ㅝ	148	0.4	24742	73.3
ㅞ	519	1.6	25261	74.5
ㅟ	2795	8.3	27932	83.1
ㅠ	289	0.9	28056	84.0
ㅡ	5404	16.1	33749	100
ㅢ				
ㅣ				

"이마-", "이노-", "이둥-", "이이-"등의 人名이 포함되어 있고 일부 英美系列의 著者도 포함되어 있기 때문이다.

기호를 매길 言語인 한글은 單子音은 14자이고 (雙子音은 單子音에 포함시킴), 母音은 單母音이 10자이다. 기호를 공평하게 배정하기 위해서는 각 子音에 한개씩 숫자

<表-6 8대 姓氏 둘째자 子音 出現 頻度 >

SAS

JUNG1	FREQUENCY	PERCENT	CUMULATIVE FREQUENCY	CUMULATIVE PERCENT
	37			
ㄱ	1617	17.4	1617	17.4
ㄲ	84	0.9	1701	18.4
ㄴ	207	2.2	1908	20.6
ㄷ	492	5.3	2400	25.9
ㄸ	15	0.2	2415	26.1
ㄹ	76	0.8	2491	26.9
ㅁ	431	4.6	2922	31.5
ㅂ	427	4.6	3349	36.1
ㅃ	1	0.0	3350	36.1
ㅅ	1443	15.6	4793	51.7
ㅆ	8	0.1	4801	51.8
ㅇ	1710	18.4	6511	70.2
ㅈ	1194	12.9	7705	83.1
ㅉ	24	0.3	7729	83.4
ㅊ	407	4.4	8136	87.8
ㅋ	64	0.7	8200	88.5
ㅌ	195	2.1	8395	90.6
ㅍ	50	0.5	8445	91.1
ㅎ	824	8.9	9269	100

<表-7 8대 姓氏 둘째자 母音 出現 頻度 >

JUNG1	FREQUENCY	PERCENT	CUMULATIVE FREQUENCY	CUMULATIVE PERCENT
	37			
ㅏ	1151	12.4	1151	12.4
ㅐ	564	6.1	1715	18.5
ㅑ	109	1.2	1824	19.7
ㅓ	989	10.7	2813	30.3
ㅔ	156	1.7	2969	32.0
ㅕ	1037	11.2	4006	43.2
ㅖ	79	0.9	4085	44.1
ㅗ	1103	11.9	5188	56.0
ㅘ	294	3.2	5482	59.1
ㅙ	9	0.1	5491	59.2
ㅚ	228	2.5	5719	61.7
ㅛ	1508	16.3	7227	78.0
ㅜ	127	1.4	7354	79.3
ㅝ	5	0.1	7359	79.4
ㅞ	33	0.4	7392	79.7
ㅟ	232	2.5	7624	82.2
ㅠ	468	5.1	8086	87.3
ㅡ	122	1.3	8092	88.6
ㅣ	1055	11.4	9269	100

를 배정하여야 하지만, 여기에 사용할 숫자기호 "0"는 혼란의 여지가 있으므로 생략하면 사용 가능한 숫자는 "1"부터 "9"까지 아홉개 밖에 없으므로, 14개의 子音을 9개 그룹으로 나누고 한 그룹당 약11%정도의 出現 頻度가 되도록 子音을 배정한다. 그러나 일정하게 11%로 구분하게 되면 실제 기호를 매길 때는 어떤 그룹은 기호를 매기는데 여유가 있고, 어떤 그룹은 기호가 부족한 경우가 있기 때문에 出現 頻度가 높은 문자라도 姓氏에 따라 出現 頻度가 다른 점을 감안하여 表를 구성하였다.

특히 기호매김의 대상인 둘째자에서 보면 30여자가 과반수 이상을 차지하므로 表의 구성을 통계에 나타난 수치에 의하여 일정하게 기호를 배정할 수 없으므로 8대 姓氏를 중심으로 기호를 배정하였다.

이러한 원칙 아래 기호의 子音을 먼저 9개 그룹으로 나누면, 먼저 숫자 "1"은 전체 통계에서 각각 10.9%와 1.8%, 3%의 頻度를 나타내는 ㄱ과 ㄲ, ㄴ으로 정하였다.

ㄴ은 전체 3%로 1024회 출현 빈도를 나타내지만 韓國人名에서는 208회 출현한다. 208회중 "나"가 31회, "남"이 67회, "노"가 58회 나타난다. "나"와 "노"는 이씨 姓에서, "남"은 김씨 姓에 가장 많이 나타나므로 ㄴ은 한국인 보다 外國人에 많이 나타나는 것으로 볼 수 있어 ㄱ의 그룹에 넣어 ㄷ과 分離하였다. 그러나 ㄴ은 ㄱ의 19와 記號가 重複되므로 ㄴ에만 별도의 母音記號表를 만들어 기호의 중복을 막았다.

ㄷ은 5.7%로 ㄴ보다 빈도가 높고 특히 韓國 人名에서 많이 사용하므로 한자리 기호를 배정하였다. 특히 ㄷ은 "도,독,돈,동"의 글자가 통계에 의하면 총234회 출현하여 ㄷ 전체의 반 정도를 차지하므로 상당히 밀집되어 중복이 많이 일어나서 分類에 관계없이 一著者 一記號를 賦與하게 되면 기호의 자리수는 4 - 6자리 이상 증가할 수 밖에 없으므로 이들 著者의 著者記號 자리수를 줄이기 위해서는 ㄷ을 한자리 기호 "2"를 配定하였다.

ㄹ은 전체적으로 7.8%의 높은 빈도를 나타내지만 8대 姓氏에서는 67회 밖에 출현하지 않고 그중 가장 빈도가 높은 글자는 "리"와 "림"으로 각각 14회 였다. 그러므로 ㄹ은 ㄷ에 포함시켜도 기호의 중복이 발생되지 않기 때문에 ㄷ에 포함시켰다.

한글順圖書記號法 제2표에서는 ㅊㅋㅌ표을 각각 81,82,83,84를 주었으나, 출현 빈도 통계에서 ㅊ은 404회의 높은 출현 빈도를 보였고, ㅋ은 65회 ㅌ은 196회 ㅍ은49회 출현하였다. 그러므로 ㅊ은 자주 사용하는 글자이기 때문에 독립된 한자리 기호 "8"을 가지도록 하여 자주 등장하는 人名에 대해서는 著者記號의 자리수를 줄여서 전체 著者記號의 자리수가 안정되게 되었다.

ㅋ은 "오카-", "이케-", "오쿠-"에만 거의 출현하므로 ㅋㅌㅍ중 8대 姓氏에 가장 많이 쓰이는 ㅊ에 한자리 기호를 배정하기 위하여 ㅊ그룹에 ㅋ을 포함 시켜 한글順圖書記號法 變形 2表를 작성하였다.

子音에서 두자리 숫자를 가지는 ㄴ과 ㄹㅋㅌㅍ에서 적은 숫자인 1부터 사용하지 않은 것은 대부분의 문자는 子音 다음에 붙는 母音이 ㅏ, ㅑ, ㅗ, ㅜ등의 앞모音을 많이 사용하기 때문에 出現 頻度가 낮은 뒷 번호 7,8,9등을 사용하였다.

지금까지 中聲의 기호인 母音에는 숫자"0"을 한글 子音 "ㅇ"과 알파베트 "O"자와 구분이 곤란하다고 사용 하지 않았으나, 실제 表의 사용에 있어서 혼란을 초래할 이유가 적기 때문에 變形 2表에서는 숫자 "0"을 포함시켜 母音은 10區分으로 하였다.

이상과 같은 原則에 의하여 기호를 배정하면 表 8과 같다.

< 表-8 한글順圖書記號表 제2표의 變形 記號表 >

	자음	모음	ㄴ에 붙는 모음
1	ㄱㄲ	0 ㅏ	0 ㅏ
19	ㄴ	1 ㅐㅑㅒ	1 ㅐㅑㅒ
2	ㄷㄸ	2 ㅓㅔ	2 ㅓㅔㅕㅖ
29	ㄹ	3 ㅕㅖ	3 ㅗ
3	ㅁ	4 ㅗ	4 ㅘㅙㅚㅛ
4	ㅂㅃ	5 ㅘㅙㅚㅛ	5 ㅜㅝㅞㅟㅠ
5	ㅅㅆ	6 ㅜ	6 ㅡ ㅢ
6	ㅇ	7 ㅝㅞㅟㅠ	7 ㅣ
7	ㅈㅉ	8 ㅡ ㅢ	
8	ㅊ	9 ㅣ	
87	ㅋ		
88	ㅌ		
89	ㅍ		
9	ㅎ		

3. 變形 한글順記號法의 區分性 檢證

表 8의 한글順圖書記號法 變形 2表에서 ㄱ과 ㄴ에서 重複되는 부분인 19를 살펴보면, ㄱ에서 19를 가질수 있는 文字를 變形 記號表에 의하여 初聲과 中聲을 조합하여 記號를 매기면 다음과 같다. 단, 받침은 ㄱㄴㄷㄹㅁㅂㅅㅇ만 출현하였으므로 字母順으로 배열하기 위하여 편의상 순서대로 번호를 賦與한다.

<表-9 ㄱ과 ㄴ의 重複記號 매김표>

	ㄱ열	ㄴ열
191	각	내 냐 냬
192	긴	너 네 녀 녜
193	귿	노
194	길	놔 놰 뇌 뇨
195	김	누 눠 눼 뉘 뉴
196	깁	느 늬
197	깃	니
198	깅	

ㄱ에서는 19를 사용하는 문자는 8대 姓氏에는 기, 길, 김만 출현하였다. "길"에 대응하는 ㄴ의 문자는 놔 놰 뇌 뇨가 있으나 8대 姓氏에는 한번도 出現하지 않았고, "김"에는 누 눠 눼 뉘 뉴가 있으나 "누"가 이씨에서 4회 출현하였으나 "김"은 김氏에서만 1회 출현하여 전혀 중복이 되지 않았다.

긱, 긴, 긷등의 문자는 이름에 잘 사용되지 않는 자이며, 놔 뇌 뇨 누 눠 눼 뉘 뉴등의 글자는 韓國 人名에서는 거의 쓰일수 없는 문자들이다.

ㄷ과 ㄹ에서 ㄹ은 韓國 人名에서는 자주 사용되지 않기 때문에 ㄷ과 중복되지 않았다. ㄷ에서 29를 사용한 문자는 "디"한자 뿐이고 "다"는 8대 姓氏중 이氏에서 2번 오氏에서 1번 밖에 출현하지 않았으므로 記號가 重複되는 일은 없었다.

ㅊ과 ㅋ, ㅌ, ㅍ은 8, 87, 88, 89로 8을 사용하는 子音 4자가 밀집되어 있어 記號가 중복 될 수 있는 소지가 많으나, ㅊ에서 87을 사용하는 문자는 춰, 췌, 취, 츄중 8대 姓氏에서 "취" 1자만 1번 출현하였고, 88은 츠, 츼로서 8대 姓氏에서는 출현할 수 없는 문자이다.

〈表-10 ㅊ과ㅋ, ㅍ의 重複記號 매김표〉

	ㅊ열	ㅋ열		ㅊ열	ㅍ열
871	칙첵칙츅	캐캬	891	칙	패 퍄 퍠
872	췬췐친츈	커케	892	친	퍼 페
873	춴췐친츋	켜켸	893	칟	펴 폐
874	췰쳴칠출	코	894	칠	포
875	침쳼침츔	콰쾨쿄	895	침	퐈 퐤 푀 표
876	칩쳽칩츕	쿠	896	칩	푸
877	칫쳇칫츗	쿼퀘퀴큐	897	칫	풔 풰 퓌 퓨
878	칭쳉칭츙	크킈	898	칭	프 픠

89는 치, 칙, 칠, 칭으로 8대 姓氏에는 출현하기 어려우나 ㅋ에 대비하여 記號를 매겨보면 表 18과 같다. ㅋ과 대응하는 ㅊ에서는 전혀 출현할 수 없는 문자가 만들어지고 ㅍ에서는 "칠"이 김씨에 2회 박씨에 1회 출현하였으나 이에 대응하는 ㅍ의 "포"는 한번도 출현하지 않았다.

제4장. 著者記號 自動化시스템의 開發 및 適用

Ⅰ. 자동화시스템의 概要

表 2에서 보는 바와 같이 국내 圖書館의 절반 이상이 分析合成式 著者記號表인 한글順圖書記號法을 사용하고 있으나 記號를 매기는데 새로운 기법이 개발되지 않아

많은 分類 司書들은 著者記號의 중복방지를 위해 記號를 매길때 마다 일일이 書架配列을 확인하고 記號를 매기는 불편한 방법을 사용하고 있다.

본 시스템은 分類 司書가 標目을 CRT에 입력하면 記號매김의 대상인 저자명의 두번째 문자가 子音과 母音으로 분해되고, 분해된 字母는 變形 한글順圖書記號法 제2표를 對照하여 값을 취한 다음, 다시 조합하여 1차적인 著者記號를 매기도록 하였다.

著者記號의 매김을 자동화하면 記號매김 대상어의 子音과 母音이 같으면 동일한 記號가 매겨지므로 이를 방지하기 위하여 標目과 著者記號를 檢索接近點으로 삼아 目錄 Data Base에서 해당 資料를 檢索하여 이미 賦與된 著者記號를 확인하면서 최종적으로 著者記號와 著作記號가 중복되지 않도록 매긴다.

본 시스템에서는 開發 言語를 ANSI COBOL을 사용하였고, 데이타베이스는 MV의 DG SQL 5.1을 사용하였다. 하드웨어는 Data General사의 MV 15000 MODEL 20을 사용하였다.

Ⅱ. 시스템의 구성과 특성

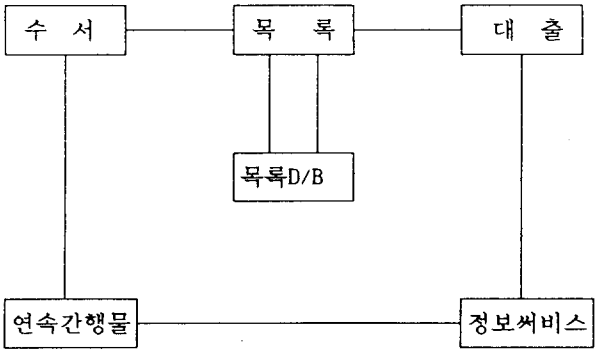
첫째, 본 시스템은 書誌中核型시스템중 目錄業務의 일부로 구성되어 있다.

둘째, 입력된 標目은 1차적으로 著者記號表에서 記號를 조합하고, 이 記號와 標目을 검색접근점으로 삼아 目錄Data Base를 검색하여 관련 書誌事項을 검색할 수 있다.

셋째, 分類記號와 著者記號를 매기기 위하여 검색된 데이타를 브라우징하면서 記號의 중복을 확인하고 複本調査도 동시에 실시할 수 있다.

넷째, 入力司書가 標目의 선정을 위하여 별도 작업을 하지 않도록 分類 司書가 입력한 標目으로 著者典據파일(Author Authority File)에 연결하여 標目으로 채택할 著者를 작업지시표에 자동 출력시킨다.

〈圖-1 書誌中核型시스템 구성도 〉



Ⅲ. 著者記號自動매김 과정

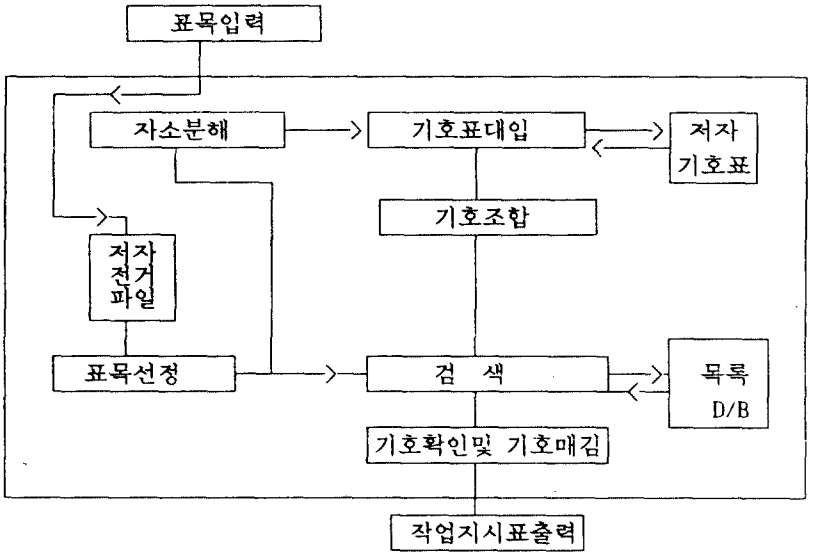
記號매김의 基本 要素를 入力할 때 다음 여섯가지를 유의해서 입력해야 한다.

첫째, 標目은 가능하면 전부 입력하여야 한다. 基本記入標目은 전부 한글로 사용하기 때문에 外國 人名의 경우 여러가지로 표기되는 人名을 著者典據파일에서 標目으로 채택할 著者名으로 記號를 매기기 때문에 완전한 이름이 있어야 한다.

둘째, 성과 이름 사이에는 컴마를 사용하지 않는다.

셋째, 입력된 標目으로 目錄 데이타베이스에서 사용된 著者記號를 검색해야 하므로 검색시간을 줄이고 精度率(Precision ratio)을 높이기 위하여 저자의 주제적 성격을 파악하여 分類를 類,綱,目 또는 細目까지 입력 한다.

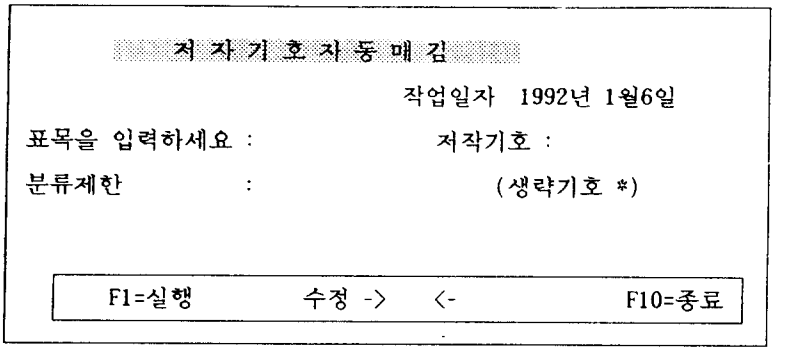
〈圖-2 著者記號자동화시스템 개요도〉



네째, 완전한 圖書記號를 매기기 위하여 著作記號인 서명의 첫글자를 입력한다. 圖 3의 기본화면에서 標目을 "김동길"로 입력하고 多作家이므로 分類는 81* 또는 814.6으로 제한하고 著作記號는 서명의 첫자인 "감"을 입력하면 圖 4와 같다. 分類의 제한은 저자의 主題的 성격을 사전에 제한하므로써 검색된 資料에서 雜音(Noise)을 제거하고 精度率을 높이기 위하여 사용된다.

記號매김 대상어의 첫자와 둘째자의 子音과 母音이 이미 사용되었을 경우에는 記號의 자리수를 한자리 더 附加하여 주어야 하는데, 부가되는 記號는 입수되는 순서대로 1,2,3순으로 賦與하지 않고 어느 정도의 字母順 配列을 유지하기 위하여 5,2,7,4,9,3

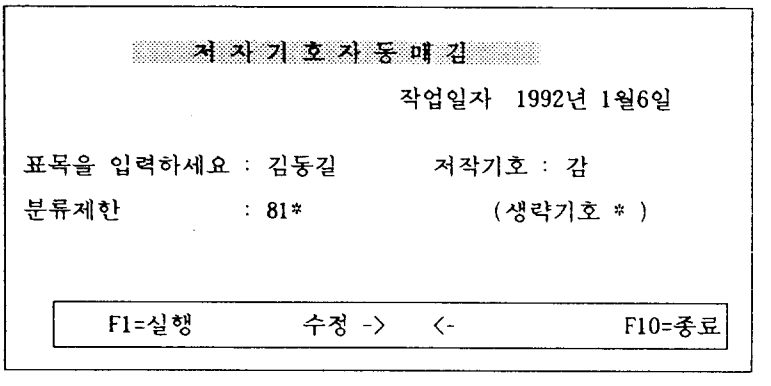
〈圖-3 저자기호자동매김 기본 화면〉



6,1,8순으로 부가 한다.

예를 들면 子音 "ㄷ"과 母音 " ㅗ"로 만들어지는 문자의 경우에 "도"에서 부터 "동" 까지는 基本記號가 모두 24이므로 처음 賦與되는記號는 24를 사용하고, 그 다음부터 입수 되는 資料는 245, 242 ,247, 244, 249, 243, 246, 241, 248순으로 自動 賦與 되며 "도"에서 "동"까지 10개 이상의 資料가 입수되면 역시 같은 순으로 2455, 2452, 2457..... 순으로 부가한다.

〈圖-4 입력 화면 〉



著作記號의 매김은 특정 저자의 分類아래 사용된 著作記號와 記號를 매기기 위하여 입력한 記號와 비교하여, 記號가 중복되는 경우에는 사용된 著作記號의 終聲보다 1자리 뒤의 終聲인 子音을 사용한다. 예를 들면 입력된 標目이 "김동길"이고 著作記號가 "감"이면 검색결과 "김244감"은 이미 사용 되었으므로 프로그램내에서 자동으로 "감"의 Hexa Code값에 "1"을 더하면 바로 뒤 문자인 "갑"이 오게 되어 사용할 著作記號는

"김244갑"으로 매겨진다.

同名異人의 경우에는 基本 記號가 중복되므로 F9 Key를 사용하여 著者記號가 중복되지 않도록 임의로 記號를 매겨주어야 한다.

Ⅳ. 著者記號 自動化시스템의 長點과 實行 效果

手作業에서는 著者記號를 매기기 위해서는 특정 저자가 分類에 따라 어떤 著者記號가 매겨졌는지 書架配列에서 확인하고 記號를 매기는데, 가능하면 字母順이 유지되도록 著作 記號를 매기는 방법으로 지금까지 작업을 해 왔다.

만약, 書架카드의 배열이 잘못되거나 카드가 유실 되었을 경우, 또는 정리를 위해서 카드가 제자리에 있지 않을 경우에는 정리가 끝나고 書庫에 入庫된 다음에, 또는 貸出되었던 도서가 返納되어 書架에 배열되었을때 誤謬를 확인할 수 있으므로 번거롭기 짝이 없는 再整理 작업을 계속해야 한다.

〈圖-5 1차 검색화면 〉

```
                    저 자 기 호 매 김

                                        작업일자  1992년 1월06일

입력저자 : 김동길                      분류제한 : 81*
표목저자 : 김동길                      저작기호 : 감

제2표기호 : 김 24                      기호매김 : 김244갑

사용된 기호
   (1) 810.8 : 김244사  김244새  김244여  김244홍
   (2) 814.6 : 김244감  김244그  김244극  김244래  김244삭

유사저자 : 김24 -

   (3) 김도영 : 813.6 김247소  김247영
   (4) 김동리 : 810.8 김242노  김242뇌  김242생  김242조
                813.6 김242사  김242상  김242여
   (5) 김동인 : 813.6 김245감  김245노  김245저

F2=개별간략확인 F4=개별상세확인 F5=전체확인 F6=작업지시표작성
F9=기호매김      PgUp  PgDn                    F10 = 작업종료

확인할 번호는 ?: _____  F1= 실행

F9=기호매김 저자기호:                  분류기호:
```

그러나 著者記號 自動化System은 著者記號를 매기기 위해서 입력된 標目으로 CRT화면에서 圖 5와 같이 書架 配列을 확인하고 複本調査까지 동시에 실시할 수 있어, 서가에 배열되고 난 다음 잘못 分類된 資料가 발생되는 경우가 적어졌기 때문에 整理業務의 시간단축은 물론 分類記號와 著者記號를 일관성 있게 매길수 있는 장점이 있다. 또한 CRT화면에서 특정 저자와 類似 著者의 記號를 동시에 확인이 가능하므로 記號매김을 分類에 관계없이 특정 저자가 하나의 記號를 가지도록 조정이 가능하다.

특정 저자가 分類에 관계없이 동일한 著者記號를 가지게 되면 記號매김 대상 문자의 字母가 같을 경우에는 자리수를 늘여야 하는 단점이 있다. 그러나 列擧式 記號表와는 달리 記號를 사용하는 圖書館에서 記號의 자리수를 한자리씩 임의로 늘여 나갈 수 있기 때문에 새로운 표목이 등장하더라도 자동으로 區分이 가능하다.

記號를 區分할 수 있는 자리수는, 한자리 記號를 가지는 子音의 경우에는 전체 기호의 자리수를 다섯자리로 사용하면 1,001개 까지 區分할 수 있으므로 충분히 區分이 되고, 한자리 더 늘리면 10,001개의 區分性이 있어 記號매김의 대상인 둘째자의 子音과 母音이 같은 저자가 밀집되더라도 특정 저자를 分類에 관계없이 列擧式 記號表보다 더 상세하게 독립된 著者記號를 매길수 있어 區分性이 뛰어나다고 볼 수 있다.

分析合成式 著者記號表의 제한점은 列擧式에 비하여 著者記號의 자리수를 일정하게 고정시키지 못하는 短點이 있으나 記號의 자리수를 1 - 2자리 늘여서 사용하면 列擧式보다 엄청난 區分性을 가질수 있는 장점이 있다.

본 연구의 著者記號 자동화시스템은 列擧式 記號表의 장점인 저자와 記號를 즉시 화면상에 일일이 列擧하여 확인 할 수 있는 점과, 分析合成式의 장점인 새로운 項目의 등장을 손쉽게 처리할 수 있는 列擧式과 分析合成式의 두가지 장점을 동시에 가지게 되었다. 또한 著者記號를 검색 접근점으로 이용하여 資料를 검색하게 되므로 著作記號의 자동매김은 물론 複本調査도 동시에 실행할 수 있고, CRT화면상에서 著者記號가 자동으로 매겨지기 때문에 著者記號를 중복되지 않고 정확하게 관리할 수 있는 효과와, 圖書記號를 매기는 데 소요되는 시간은 물론 整理業務에 소요되는 시간을 단축시킬 수 있는 효과가 있다.

제5장. 結 論

한글은 初聲,中聲,終聲인 音節式으로 구성되어 있다. 記號매김 대상을 文字 構成 要素 전부를 대상으로 하게 되면 기본적으로 記號의 자리수가 늘어나기 때문에 初聲과 中聲만으로 記號를 매기도록 한글順圖書記號法에 지시되어 있다. 그러나 이 규정은 처음 意圖와는 달리 특정 문자에서 資料의 계속적인 증가로 記號의 자리수는 4 - 5자리

이상으로 확대되어 簡潔性이 없어지고 書架配列을 확인하면서 記號를 확정해야 하는 불편함을 해소 하지 못했다.

본 硏究에서는 이런 문제들을 해소하기 위하여 著者記號매김을 自動化하기에 앞서 정확한 子音과 母音의 出現 頻度를 조사하여 전체적으로 記號의 자리수를 균형있게 사용하기 위하여 많이 사용되면서 두자리 記號를 가지는 子音은 한자리 記號를 가지도록 재 구성하여 著者記號 自動化 시스템을 구축하는 내용이다.

지금까지 手作業으로 著者記號를 매길때 지적된 문제점들이 記號매김을 자동화하므로서 얻어지는 결과로 한글順圖書記號法의 문제점들을 일부 해결할 수 있었다. 그러나 著者記號 自動化시스템의 단점은 KORMARC Format으로 目錄Data Base를 構築한 圖書館이거나 또는 請求記號만이라도 Data Base를 구축한 圖書館만이 사용이 가능하기 때문에 일반 圖書館들이 당장 적용하기에는 현실적으로 무리가 따른다.

본 硏究에서 문제로 제기된 사항은,

첫째, 記號의 정확한 字母順配列은 명확하게 확정하지 못하였고,

둘째, 記號의 자리수를 일정하게 고정시키지 못하였으며,

셋째, 記號의 자리수를 확장할 때 字母順에 의한 완전 자동화를 하지 못하고 일부 人爲的으로 하였으며,

넷째, 현재 目錄業務가 전산화 되어 있는 圖書館만이 사용이 가능한 점이다.

그러나 본 硏究에서 나타난 사실은,

첫째, 一著者,一記號 원칙을 준수할 수 있고,

둘째, 記號의 區分 能力을 극대화 시켰고,

셋째, 記號의 重複 방지를 위하여 별도로 記號表를 만들지 않아도 일관성 있게 記號를 매길 수 있으며,

네째, CRT화면에서 書架配列을 직접 확인하여 整理에 필요한 書誌事項을 별도 檢索을 하지 않더라도 확인할 수 있었기 때문에 整理 시간을 단축하였고,

다섯째, 著者記號를 檢索接近點으로 삼아 複本調査와 標目과 分類를 통일할 수 있었다.

이와 같은 결과를 볼 때 著者記號를 자동화함으로써 目錄業務의 書誌事項 入力만이 산적한 圖書館업무의 문제해결이 아니라, 타 업무에 비해서 아직까지 原始的인 방법으로 처리하는 請求記號 매김작업의 전산화가 시급한 문제이기 때문에, 본 연구에서 현재의 目錄 Data Base에 著者記號 自動化System을 추가하면 업무의 效率이 크게 증가될 것으로 본다. 특히 신설되는 圖書館이나 藏書의 규모가 적은 圖書館에서는 본 연구에서 나타난 著者記號 自動化시스템을 적용하면 그 업무 효과가 기존 圖書館의 System하에서 보다 급진적으로 향상 되리라고 본다.

參考文獻

1. 金南碩, 圖書記號, 大邱, 啓明大學校出版部, 1988.
2. 金明玉, 資料分類論, 서울, 歐美貿易, 1986.
3. 朴俊植, 英美著者記號表硏究, 未刊本博士學位論文, 中央大學校大學院, 1991.
4. 장일세, "한국인著者記號表의 구성원리", 圖協月報, 第3卷(1962).
5. 정필모, "韓國文獻記號法硏究 - 現行列擧式著者記號法에 대한 代案", 國會圖書館報, 9권5호(1972).
6. 이재철, "구조론에 입각한 한국인著者記號表에 관한 연구", 圖書館學, 第1輯(1970).
7. 이재철, 한글順圖書記號法, 서울, 아세아문화사, 1982.
8. 홍옥자, 한국圖書館에서 채택하고 있는 분류법의 실태에 관한 연구, 未刊本碩士學位論文, 梨花女子大學校敎育大學院, 1977.
9. 今まと子, 資料分類法及び"演習, 東京, 水村房, 1985.
10. 仙田正雄, 圖書分類と圖書記號, 東京, 蘭書房, 1955.
11. Buchana Brain, Theory of Library Classification, London, Clive Bingley, 1979.
12. International Study Conference on Classification for Information Retrieval (Dorking, 1957), Processing, London, Aslib, 1957.