

# 로마자 표기표준의 활용에 관한 연구

## A Study on the Utilization of Romanization

심효정, 이화여자대학교 대학원 문헌정보학과

992LSG11@mm.ewha.ac.kr

Hyo Jeong Sim, Dept. of Lib. and Info. Sci., The Graduate School,  
Ewha Womans Univ.

많은 양의 정보가 실시간으로 생산되는 현재 다양한 정보검색방법이 더욱 필요하게 되었으며 이에 따라 다양한 언어로 된 정보를 신속하고 정확하게 검색해내기 위해서 로마자 표기도 더욱 중요하게 되었다. 이에 본 연구에서는 정보검색 분야에서의 로마자 표기표준의 활용을 중심으로 논하고자 한다.

### 1 서론

다량의 정보가 실시간으로 생산되고 있는 지금 원하는 정보를 신속하고 정확하게 찾아내기란 쉬운 일이 아니다. 따라서 효율적으로 정보를 찾아낼 수 있는 방법 개발에 많은 관심이 집중되게 되었으며 그 중요성은 날로 더해 가고 있다.

또한 세계 도처에서 생산되고 있는 수많은 정보들은 서로 다른 언어로 이루어진 것들이 많아 이런 자료들을 검색하고자 하는 사람들에게 부담감을 안겨 줄 수 있다. 따라서 이러한 각각의 다른 언어로 이루어진 정보들을 효과적으로 처리하여 정보검색에 도움을 주기 위한 방법으로 로마자 표기표준의 활용을 들 수 있다.

본 연구에서는 정보검색 분야에서의 로마자 표기표준의 활용에 관해서 일반적으로 활용되는 경우와 도서관에서 사용되는 MARC에서의 활용으로 나누어 살펴보고자 한다.

### 2 로마자 표기

#### 2.1 로마자 표기의 정의와 필요성

로마자란 옛날 로마제국이 사용하던 라틴어를 표기하던 것으로, 자모음을 합하여 26자가 되는 문자 체계를 뜻하며 라틴 문자라고도 한다.

로마자 표기법이란 로마자를 표기의 수단으로 하지 않는 나라가 자국어에 대한 발음을 로마자로 표기함으로써, 로마자를 표기의 수단으로 하거나 그것을 이해하는 외국인들로 하여금 자국어의 발음을 할 수 있도록 하는 기법을 말한다. 또한 국어의 로마자 표기법이란, 우리말을 로마자로 옮겨적는 일정한 규칙을 말한다.

과거에 국어를 로마자로 표기하는 작업이 필요했던 것은 주로 국어 표기에 사용되는 문자, 즉 한글과 한자를 모르는 외국인들에게 국어의 발음을 전달하기 위한 것이었다. 그러나 오늘날은 한국어의 발음을 전달하는 것은 물론 국내외의 모든 분야에서 정보를 효과적으로 검색

하고 전달하기 위해서 더욱 필요하게 되었다.

## 2.2 한글의 로마자 표기 주요 표준안

한국인이 작성한 최초의 로마자 표기법은 1935년에 발표된 정인섭의 표기법으로써 전자법<sup>1)</sup>의 형태를 띠고 있다. 그 후 1939년에 발표된 머쿤 라이샤워(McCune-Reischauer) 표기법은 전자법<sup>2)</sup>으로, 발표된 이래 오늘날까지 영어권에서는 한글의 로마자 표기법으로 가장 널리 쓰이고 있다. 반달표와 어깨점이 있으며 유성음과 무성음의 차이를 반영하고 있다.

최초의 정부안인 1948년에 발표된 표기법은, 전사와 전자를 합친 절충식 표기법으로 되어 있다. 예를 들어 '값 [갑], 값을 [갑슬]'을 전사법의 원칙에 따라 소리나는 대로 표시하여 'kap', 'kapsúr'으로 나타내었으며, '값만 [갑만]'의 경우는 자음 동화를 표기에 반영하지 않고 전자법의 원칙에 따라 'kap man'으로 나타내었다. 두 번째 정부안인 1959년에 발표된 표기법은 전자법으로, 반달표나 어깨점을 없애 기계 적용에 쉬운 것이 장점이다. 세 번째 정부안인 1984년에 제정된 표기법은 얼마 전까지 사용하던 표기법이며, 전사법을 기본으로 하고 있다. 이는 반달표나 어깨점을 사용하므로 불편하다는 단점을 지니고 있다. 또한 반달표나 어깨점을 경우에 따라 생략해도 좋다는 애매 모호한 규정이 있어 청주(ch'ongju)와 정주(chongju)같이 반달표나 어깨점을 생략하면 구별이 불가능한 경우도 있다.

네 번째 정부안은 공진청에서 1985년 11월 ISO의 요청을 받아 '기계화를 위한 한글의 로마자 표기법'이라는 이름으로 제정하게 되었다. 1984년 문교부안이 있었음에도 불구하고 다시 제정해야만 했던 이유는 반달표나 어깨점 등의 사용이 기계화 작업에 곤란하다는 점과 표기

체계가 이원화되어 있어 ISO에서 요구하는 한글과 로마자 1대 1 대응 원칙에 위배되기 때문이었다.

다섯 번째 정부안은 문화관광부가 2000년 7월 4일 발표하여 7월 7일부터 시행하게 된 '새 국어로마자표기법'이다.

최종 확정된 '새 국어로마자표기법'은 전자법을 기본으로 하고 있으며, 반달표와 어깨점을 없앴으며, 유성음과 무성음의 차이를 반영하지 않았다는 특징을 가지고 있다. 그러나 기존의 인명이나 회사명, 단체명 등은 그대로 쓸 수 있게 해 로마자 표기에 있어 여전히 논쟁의 여지를 가지고 있다.

## 3 로마자 표기표준의 활용

### 3.1 일반 정보검색에서의 활용

일반 정보검색에서의 문제는 표기가 다양하거나 컴퓨터처리가 불편하다는 데 있다. 이를 해결하기 위해서는 가능한 한 표준표기를 따르게 하고 예외를 인정하는 이형에 대해서는 그 정보를 수집하여 서비스를 제공할 필요가 있다. 기본적으로 표준표기를 채택해야 하는 이유는 다음과 같다.

첫째, 인명 표기의 다양성이다. 로마자 표기표준이 있음에도 불구하고 인명 표기시 상당수가 일관성 없이 표기되고 있어 정보검색 과정에서 많은 어려움이 있었다. 새로 제정된 '새 국어로마자표기법'의 경우에도 인명의 경우에는 관용을 허용한다는 규정이 있으며 특히 성씨의 경우 통일하기 못하고 추후에 결정키로 함에 따라 앞으로는 많은 혼란이 예상된다.

1996년 국립 국어 연구원의 조사에 따르면 표기법을 준수하는 비율은 10%밖에 되지 않는 것으로 밝혀졌다. 예를 들어, 성이 '서' 씨인 경우 'suh', 'seo', 'shuh', 'seu', 'suh', 'seoh' 등으로 다양하게 표기하고 있었는데, 대부분 'suh'로 표기하고 있었고, 당시 표준(84년 문교부표기법)인 'So'는 전혀 사용되지 않고 있었다.

또 외국에서 사용하고 있는 우리 나라의 인

1) 철자를 중시하는 표기 방식을 말하는데, 예를 들어 '한라산'의 경우 'hanrasan' 또는 'hanlasa-n'과 같이 표기하는 것을 말한다.

2) 발음을 중시하는 표기 방식을 말하는데, 예를 들면 '한라산'의 경우 발음은 '할라산'이므로 'hanllasan'과 같이 표기하는 것을 말한다.

명의 표기법을 백과사전<sup>3)</sup> 중심으로 살펴보면 미국, 영국, 독일은 대체적으로 머큐 라이샤워 표기법을 따르고 있으나 개인의 고유한 인명 표기를 따르는 경우도 있고, 프랑스처럼 아예 프랑스의 고유한 표기법<sup>4)</sup>에 따라 표기하는 경우도 있다. 즉, '츠'를 'tch'로 표기하고 '즈'를 'dj'로 표기하는 경우이다.

이처럼 국내에서만 아니라 국외에서도 같은 사람에 대해서 두 가지 이상의 로마자 표기가 가능하다는 것은 정보검색 시 어려움을 가중시킨다.

둘째, 지명 표기의 다양성이다. 지명 표기 또한 인명과 마찬가지로 표기표준이 잘 지켜지지 않고 있다. 예를 들어 '독도' 같은 경우 표기표준으로 나타낸다면 'dogdo'로 나타내야 하지만 현재 많은 표기는 'Tokto'로 하고 있어 '토크토'라는 이상한 발음으로 읽힐 수가 있다. 또 '안압지'의 경우도 표기표준에 따르면 'Anapji'로 나타내야 하지만 현재 많은 표기는 'Anapchi'로 나타내고 있어 '아나포치', '애남치' 등의 이상한 발음으로 읽힐 수 있다.

또한 지명의 경우도 인명에서와 마찬가지로 외국에서 사용하고 있는 우리 나라의 지명 표기법을 지도<sup>5)</sup>나 백과사전 중심으로 살펴보면 미국, 영국, 독일, 프랑스에서는 대부분 머큐 라이샤워 표기법을 따르고 있다는 것을 알 수 있다.

셋째, 회사, 단체, 기관 및 상품명 표기의 다양성이다. 인명, 지명과 마찬가지로 체계적인 표기를 택하지 않은 예가 많다. 예를 들어, '삼성', '현대'라는 회사명을 개정안에 따르면 당연히 'Samseong', 'Hyundae'로 표기해야 하지만 현재 'Samsung', 'Hyundai'로 표기하고 있어 얼마나 편의적으로 표기하고 있는지 잘 보여 준다. 또한 현 로마자 표기법에 따르면 그 동

안 사용해 온 회사명에 대해선 기존처럼 계속 사용할 수 있다는 규정이 있어 새로 만들어진 회사명과 기존에 사용하던 회사명이 같은 철자를 가지고 있다고 하더라도 다르게 표기되는 경우가 생기게 된다.

넷째, 앞에서 살펴본 인명, 회사명, 지명 등의 문제뿐만이 아니라 관광 안내 책자 속의 명승고적 지명이나 사찰 명칭, 문화재 안내판, 최근 들어서는 인터넷 도메인에 이르기까지 로마자 표기에 있어 많은 혼란이 야기되고 있는데, 이는 앞에서 언급했던 것과 마찬가지로 표기표준이 있음에도 불구하고 표기표준을 지키지 않기 때문에 발생하는 문제들이다.

### 3.2 MARC에서의 활용

도서관에서는 목록데이터 구축 시 MARC 형식을 사용하고 있다. MARC 형식의 특정 필드에 로마자를 부기한 예로는 일본과 대만을 들 수 있다. Japan MARC은 표목 전체에, Chinese MARC은 표목의 일부<sup>6)</sup>에 로마자를 부기 하는 정책을 사용함으로써 검색의 효율성을 높이고 있다.

로마자 표기표준을 목록데이터 구축 시 활용한다면 여러 가지 이점이 생길 수 있는데, 예를 들면 자료를 검색할 수 있는 액세스 포인트가 한 가지 더 늘어나게 된다는 점, 각 국가의 이용자가 현지어 음의 로마자로 된 액세스 포인트를 이용하여 데이터를 검색할 수 있으므로 정확률이 높아진다는 점, 한글의 입력이 어려운 이용자가 쉽게 국어데이터에 접근할 수 있다는 점, 해외의 관련 기관이 국내 작성 목록 데이터를 이용할 수 있다는 점 등이다.

위와 같은 이점들을 생각할 때 국내 도서관에서 표준 로마자번자<sup>7)</sup> 필드는 반드시 필요하

3) 영어는 *Encyclopaedia Britannica*(1995), 독일어는 *Brockhaus-Enzyklopädie*(1988), 프랑스어는 *Quid*(1996).

4) *Dictionnaire des oeuvres de tous les temps et de tous les pays*(1987).

5) 영어는 *Philip's Atlas of the world*(1995), 프랑스어는 *Grand atlas*(1995), *Le grand atlas de géographie*(1986).

6) 저자사항에서 로마자형을 기술할 수 있도록 하고 있다. 저자사항은 개인저자, 단체저자, 가족저자로 구분하며 각각 주저자와 공저자, 보조저자의 필드를 설정하고 있다.

7) 번자란 도서관계에서만 사용하는 용어로 한 언어의 단어, 인명, 지명, 명칭 등을 나타내느라 쓰여진 문자를 같은 음을 나타내는 다른 문자로 바꾸어 나타내는 것을 말한다. 일반 대중에

다고 할 수 있다.

전화 면담 조사 결과 현재 국내 일부 대학과 전문 기관에서는 로마자 번자를 하고 있는 것으로 밝혀졌는데, 대부분이 각기 다른 방식을 사용하고 있었고, 일서, 중국서, 러시아 자료에 대해서만 로마자 번자를 하고 있었다.

이는 KCR(한국목록규칙)이나 KORMARC(한국문헌자동화목록법)기술규칙내의 명확한 번자규정이 없고, 전거통제의 미비로 표목의 일관성이 지켜지지 않았기 때문일 것이다.

따라서 KCR이나 KORMARC의 번자 표기에 대한 표준을 제정하는 것은 반드시 필요한 일이라고 할 수 있겠다. 표준 번자 규정에는 구체적으로 어떤 번자 표기법을 따르라는 지침이 명시되어야 하며, 많은 예시들이 포함되어 있어야 한다. 또한 표목 전체에 로마자를 부기할 것인지 일부에만 부기할 것인지도 논의되어야 하며 로마자 표기 데이터를 기술하기 위해서는 사용할 필드, 한 필드 내에 로마자와 한글이 혼재 되어 있는 경우의 처리, 한자를 한글로 바꾸어 자동번자하는 경우에는 한자와 한글음이 다른 경우의 처리방법, 전거화일에 대한 작성 및 검색, 유지 보수 등의 처리 등에 대한 연구가 계속해서 이루어져야 한다.

또한 로마자 표기표준을 좀 더 잘 활용하기 위해서는 부정확한 로마자표기를 입력했을 때 빈도 등을 이용하여 정확한 한글로 대응시키는 시스템이 필요하다.<sup>8)</sup> 全文데이터를 색인하고 검색하는 경우에는 모든 색인어에 로마자를 부기하기가 어려울 것이기 때문이다. 이와 같은 경우에는 로마자-한글 역변환<sup>9)</sup>의 문제가 대두된다. 따라서 역변환을 할 경우 원래의 한글문자로 변환하기 위한 처리가 필요하게 된다. 한글글자 사이에 보이지 않는 코드를 생성할 수도 있을 것이며, 이의 처리는 방안만 결정된다면 그렇게 어려운 문제는 아닐 것이다.

젠 음역 또는 자역이란 용어가 사용되고 있다.  
8) 예를 들어 '권나문'이라는 저자를 검색할 때 이용자들은 각기 다른 표기로 검색할 수 있다. 'Kwon Na Moon', 'Kwon Nha Moon', 'Kwon Rna Mun', 'Gwon Na Mun', 'Kwun La Moon', 'Kwun Nah Mun' 등  
9) 로마자를 한글로 변환시키는 것을 말한다.

#### 4 결론 및 제언

본 연구에서는 정보검색 분야에서의 로마자 표기 표준의 활용에 관하여 살펴보기 위하여 일반적으로 활용되는 경우와 도서관에서 사용되는 MARC에서의 활용으로 나누어 살펴보았다.

일반적으로 활용되는 경우에는 로마자 표기 표준을 거의 따르고 있지 않은 인명, 회사명, 지명 등에 대한 문제를 주로 거론하였으며, 외국에서 우리 나라의 인명, 지명 등을 표기할 때 우리의 표준이 아닌 다른 방식을 사용하고 있어 국내뿐만 아니라 국외에서도 다른 표기가 이루어지고 있음을 살펴보았다.

MARC에서는 로마자 표기표준을 활용했을 때 나타나는 이점과 역변환의 문제, 표준 번자 표기 규정의 부재에 대한 문제를 다루면서 그 해결 방안을 모색해 보았다.

#### 참고문헌

- 국립국어연구원. 1996. 『국어의 로마자 표기 자료집』. 서울: 국립국어연구원.  
———. 2000. 『국어의 로마자 표기법』. 서울: 국립국어연구원. 김세중. 1997. 국어의 로마자 표기실태. 『새국어생활』, 7(2): 45-57.  
Lu., Suping. 1995. "A Study on the Chinese Romanization Standard in Libraries" *Cataloging & Classification Quarterly*, 21(1): 81-96.  
서기주. 1998. 『로마자 표기법과 영문관광안내문』. 석사학위논문, 계명대학교 국제학대학원, 미주학과.  
이익섭. 1997. 로마자 표기법의 성격. 『새국어생활』, 7(2): 5-25.  
이화여자대학교. 1998. 『동양서-CJK 입력기준에 관한 연구』. 서울: 첨단학술정보센터.  
정동우. 1998. 『규칙기반 이용 로마자표기 저자명 검색 시스템 설계』. 석사학위논문. 충북대학교 대학원, 전자계산학과.