

# 한국 근대 여성 구술 기록물을 통한 시소러스 개발에 관한 연구\*

## A Study on Thesaurus Development Based on Women's Oral History Records in Modern Korea

최 윤 경 (Yoon Kyung Choi)\*\*

정 연 경 (Yeon Kyoung Chung)\*\*\*

### 목 차

- |                |                         |
|----------------|-------------------------|
| 1. 서 론         | 4. 한국 근대 여성 구술사 시소러스 개발 |
| 2. 구술기록물과 시소러스 | 5. 결 론                  |
| 3. 시소러스 구축 과정  |                         |

### <초 록>

본 연구의 목적은 한국 근대 여성 구술기록물의 효율적인 주제 접근을 위한 시소러스를 개발하는 것이다. 이를 위해 문헌 연구와 국내외 시소러스 사례 분석을 수행하였고, 구술기록물에 입력된 색인어를 중심으로 시소러스를 구축하였다. 구축 과정은 총 5단계로 1단계는 한국 근대를 산 53명의 여성 구술자의 구술기록물에 입력된 한글 색인어 총 1,784개 추출하여 용어의 특성을 분석하였다. 2단계에서는 수차례에 걸쳐 해당 분야 전문가 회의를 거쳐 시소러스 대상 용어를 선정하였고, 3단계에서는 용어에 대한 등위·계층·연관 관계를 설정하였다. 4단계에서는 웹기반 시소러스 관리시스템을 개발한 후, 용어 및 관계 정보를 입력하였고, 5단계에서는 입력된 용어의 정보와 관계를 전문가 집단이 재검토하였다. 구축된 시소러스는 일반 주제어를 비롯한 지명, 장소명, 인명, 단체명 등 고유명사가 포함된 39개 대주제 영역의 1,076개의 용어로 구성되었으며 추후 추가적인 구술기록물을 바탕으로 확장될 것이다.

주제어: 시소러스, 통제어, 여성 구술사, 구술기록물, 한국 근대

### <ABSTRACT>

The purpose of this study is to develop a thesaurus for women's oral history in modern Korea. Literature review and case studies for four thesauri were performed for this study with which a thesaurus was built based upon the index terms in oral history records. The process of developing the thesaurus consisted of five steps. First, there are 1,784 index terms from the oral history records by 53 modern Korean women were extracted and analyzed. Second, possible terms for the thesaurus were selected through regular meetings with experts in the fields of information organization and women's oral history. Third, relationships between terms were defined by focusing on equivalence, hierarchy, and association. Fourth, after developing a Web-based thesaurus management system, terms and relationships were input to the system. Fifth, terms and relationships were again reviewed by experts from the relevant fields. As a result, the thesaurus comprise of 1,076 terms and those terms were classified to 39 broad subject areas, including proper nouns, such as geographic names, places, person's names, corporate names, and others, and it will be expanded with more oral history records from other people during the same period.

Keywords: thesaurus, controlled vocabulary, women's oral history, oral history records, modern Korea

\* 이 논문은 2010년도 정부재원(교육과학기술부 인문사회연구역량강화사업비)으로 한국연구재단의 지원을 받아 연구되었음(NRF-2010-322-B0022).

\*\* 이화여자대학교 사회과학대학 문헌정보학 박사과정(yunee20@gmail.com) (제1저자)

\*\*\* 이화여자대학교 사회과학대학 문헌정보학 전공 교수(ykchung@ewha.ac.kr) (교신저자)

■ 접수일: 2014년 1월 27일 ■ 최초심사일: 2014년 2월 4일 ■ 게재확정일: 2014년 2월 14일

■ 한국기록관리학회지 14(1), 7-24, 2014. <<http://dx.doi.org/10.14404/JKSARM.2014.14.1.007>>

## 1. 서론

### 1.1 연구 배경과 목적

한국의 근대는 단기간 내에 가장 빠르고 폭넓은 변화를 이룬 때로서 한국 사회를 형성하는데 큰 영향을 미친 시기이다. 조선왕조의 몰락부터 일제 강점과 식민지배, 해방, 6·25 전쟁, 그리고 분단국가로서 군사정권 하에서 산업화를 이루어냈다. 그러나 경제성장 중심의 사회변화가 정부 주도로 이루어지면서 개발독재라는 오명을 얻을 만큼 정치 및 사회 영역에서 비민주와 억압도 팽배하였고 가족과 성의 영역에서 비민주와 불평등을 정당화시켰다.

이로 이해 한국 근대사의 역사적 기록은 주로 국가나 민족, 사회가 주체가 되는 역사 서술이나 문서로 구성되어, 매우 이데올로기적이다. 그래서 지배층 또는 반정부적인 엘리트 집단의 시각을 반영하는 경향이 있으며 경제, 생산성, 노동 중심의 사회구조와 가치관의 지배로 인해 여성의 경험은 역사에서 나타나지 않았다(윤택림, 2012).

이러한 맥락에서 근대 여성에 대한 구술사는 여성의 일상적 경험에 대한 주관적 회상을 담은 기록을 통해 한국 근대사에서 무시되어온 여성의 경험에 가치를 부여한다. 그리고 기록되지 않았던 여성들에 대한 자료를 수집하고 보통 여성의 내러티브에도 관심을 가짐으로써 소외된 여성들의 풍부한 질적 자료를 얻을 수 있다. 또한 여성들의 경험을 탐색하고 정리하는 과정에서 여성들뿐만 아니라 남성들에게도 다수의 현실과 이해를 복원할 수 있어 궁극적으로 한국 근대사를 좀 더 풍요롭고 다채로운 관

점에서 재해석하고 새로운 역사 쓰기를 가능하게 한다.

구술기록물은 구술자와 연구자 간의 구술 행위를 통해 생성되는 기록물의 일종으로, 구술사를 비롯한 역사 연구의 중요한 정보 자료로 간주된다. 최근 우리나라에서도 구술기록물의 생산과 수집, 관리에 대한 필요성이 높아지면서, 구술 채록이나 구술 기록물의 수집·보존 기관이나 단체도 증가하고 있고, 이젠 단순한 축적 및 보존을 넘어 디지털 아카이브를 구축하여 웹 기반의 검색 및 접근, 이용까지 그 영역이 확대되고 있다.

그러나 구술기록물의 이용 및 접근에 대한 고민은 아직까지 미약하다. 특히 구술기록물에 대한 검색 및 접근은 대부분 전통적인 기록물 접근 방식인 출처 중심의 계층화된 접근 방식에 의존하는데, 출처 중심의 접근 방식은 기록에 담긴 내용에 대한 다양한 주제 접근을 제공하지 못하는 한계가 있다(설문원, 2010). 이는 주제에 대한 표준화된 접근을 가능하게 해주는 전문 용어집이나 시소러스와 같은 주제 접근 도구들이 구술기록물을 바탕으로 개발되지 않았기 때문이다. 아무리 귀중한 구술기록물이 수집·보존되더라도 이용과 접근을 용이하게 해주는 접근 도구가 없다면 구술 기록의 존재 가치는 떨어질 수밖에 없다.

이에 본 연구는 구술기록물의 주제 접근을 향상시키기 위한 방안으로 한국 근대 여성 구술사 시소러스 개발을 목적으로 하였다. 현재까지 국내에서 여성학이나 여성사, 구술기록물에 대한 시소러스는 거의 전무하기 때문에, 본 연구에서는 한국 근대 여성사라는 주제를 접근하는 도구로서 그 당시의 구술기록물을 통한 시소러스의

개발 과정 및 내용을 제시하고자 하였다. 특히, 실제 구술자들의 언어로 이루어진 용어들을 중심으로 시소러스를 어떻게 개발하고, 활용할 수 있을 것인가에 주안점을 두었다. 또한 한국 근대 여성 구술 기록물을 바탕으로 대상 용어와 우선어 선정 기준, 기술 방식 등을 설정하여 시소러스 개발 및 구축에 관해서도 논하였다.

## 1.2 연구 방법 및 내용

본 연구는 한국 근대 여성 구술사 시소러스 개발을 위한 기초 단계로 문헌 연구와 사례 조사를 실시하였다. 먼저 문헌 연구에서는 기본적인 구술 기록물의 특성과 유형에 대해 살펴보고, 시소러스 구축에 반영할 사항을 도출하였다. 사례 조사에서는 구술 기록을 비롯한 기록물 관련 국내외 시소러스 4개를 선정하고 각각의 구조와 특성을 파악하였다. 조사 대상 시소러스는 미국 USC(University of Southern California)의 Shoah Foundation Institute Thesaurus, 영국의 UK Archival Thesaurus(UKAT), 그리고 한국의 국가기록원 일반주제 시소러스와 국사편찬위원회 한국역사용어시소러스로 국내외에서 구술기록물 관련하여 접근 및 활용이 가장 활발하게 이루어지는 것으로 하였다.

시소러스의 구축은 용어 추출, 대상 용어 선정, 관계 설정 및 용어 정보 입력, 검수 및 수정 단계로 이루어졌다. 첫 번째, 용어 추출 단계에서는 ‘근대와 여성의 기억’ 연구 프로젝트를 통해 채록된 53명 구술자의 구술기록물 677건에 입력된 색인어 1,784개 중에 전처리를 통해 후보 용어 1,030개를 추출하였다. 이 연구 프로젝트는 이화여자대학교 한국여성연구원과 한국구

술사연구소가 한국연구재단의 지원을 받아 ‘여성구술생애사로 본 한국의 근대: 분단, 개발, 탈식민의 경험과 기억(이하 ‘근대와 여성의 기억’으로 약칭)’이라는 이름으로 2010년 9월부터 2015년 8월까지 5년간 진행되는 연구이다. 이 연구에서는 근대 여성의 삶에 대한 다각적인 이해를 위해 구축 채록과 더불어 수집한 구술기록물을 보존하고 활용하기 위한 아카이브 시스템을 개발하였으며, 수집 대상 자료는 여성 구술사 채록 자료뿐만 아니라 사진, 문헌, 동영상 등 물증자료들까지 광범위하다.

두 번째, 대상 용어 선정 단계에서는 여성학, 인류학, 사회학, 북한학, 문헌정보학 등 여성 구술사 및 정보 조직 분야의 학계 전문가들로 구성된 자문위원들이 정기 회의를 열어서 시소러스에 입력될 1,076개 용어를 일일이 선정하였다. 이 과정에서 시소러스 구축 범위와 용어 선정 기준, 기술 규칙까지 다루었다.

세 번째, 관계 설정 및 용어 정보 입력 단계에서는 전문가 자문회의를 통해 용어의 등위·계층·연관 관계를 정의하였으며, 네 번째, 웹 기반의 시소러스 관리 시스템을 개발하여 정의된 용어의 관계와 부가적인 정보를 입력하였다. 그리고 마지막 단계인 검증 및 수정 단계에서는 전문가 자문회의에서 입력된 용어의 관계와 부가 정보를 모두 검토하여 시소러스를 수정·보완하고, 향후 시소러스 관리에 고려할 사항을 논의하였다.

## 1.3 선행 연구

국내에서는 여성 구술사나 구술기록물에 관한 주제 접근이나 시소러스를 다룬 연구는 매우

드물다. 이는 기록관리학 분야에서 현재까지 기술기록에 관한 연구가 주로 수집이나 보존, 정리 등 관리에 관한 연구를 위주로 수행되었기 때문이다. 따라서 본 연구에서는 기록물과 관련된 시소러스 구축이나 전거 제어를 다룬 연구로 개념을 좀 더 확장하여 선행 연구를 고찰하였다.

국내의 경우, 기록관리 분야에서는 일반 주제어 중심의 시소러스 보다는 단체명이나 인명 중심의 전거 제어 연구들이 활발하게 이루어졌다(설문원, 2002; 이현정, 2006; 공귀영, 리상용, 2008). 이들은 기록물 전거제어 국제 표준인 ISAAR(CPF)을 중심으로 국내외 사례 분석을 통해 전거레코드의 기술 요소를 재구성하거나 전거 구축 시 고려 사항을 제안하였다.

기록물의 시소러스 구축에 관한 연구는 박부숙(2007)과 이해영 등(2008)의 연구가 대표적이다. 박부숙(2007)은 호주, 미국, UNESCO, 그리고 한국의 국가기록원의 시소러스 사례를 중심으로 기록물 검색 서비스를 위한 시소러스 개발 및 활용 현황을 간략하게 소개하였다. 이해영 등(2008)은 대학기록관에서 사용되는 용어를 대상으로 시소러스 구축 지침을 개발하고, 적용 과정에서 도출된 다양한 상황과 대응 방안을 제안하였다.

외국의 경우, 1990년대 이전부터 기록물에 대한 주제 접근의 필요성을 인식하여, 국가 기관을 중심으로 시소러스를 구축하고 이에 대한 연구가 지속적으로 수행되었다. McNitt(1989)은 기록물의 주제 색인어 시소러스인 대통령 도서관 정보 네트워크(Presidential Libraries Information Network, PRESNET) 시소러스를 중심으로 구축 과정과 적용 효과를 조사하

였다. 그리고 자체적인 기록물 데이터베이스를 관리하는 기록관리 기관들은 표준적인 대규모 시소러스를 활용하기 보다는 소규모의 전문화된 시소러스를 개발할 필요가 있다고 제안하였다.

Garrod(2000)는 영국 국가 디지털 데이터셋 아카이브(UK National Digital Archive of Datasets, NDAD)의 시소러스 개발 과정을 소개하였다. NDAD 시소러스는 인터넷을 이용한 검색 접근 도구라는 전제 하에 UNESCO 시소러스를 기반으로 구축되었고, 시소러스의 범위 선정, 디스플레이 방식 선정, 유지 관리 순으로 진행되었다. 그는 기존의 시소러스를 활용하는 것이 용어의 계층 관계를 참조할 수 있다는 점에서 상당히 유리하다고 주장하였다.

Giménez-Chorne와 Mercedes Escrig-Giménez(2011)은 스페인의 대표적인 역사 아카이브인 Archivo del Reino de Valencia에 소장된 기록물을 중심으로 시소러스 개발 과정을 소개하였다. 그들은 역사 기록물에 대한 시소러스가 정보의 가시화를 증진시키고 전문 이용자나 일반 대중에게 가장 유용한 접근 방식을 제공할 수 있다고 보았다.

이상과 같이 국내외 기록 관리 영역에서 전산화가 도입된 이후, 기록물의 주제 접근에 대한 고민이 계속적으로 이루어지면서 국가 차원에서 시소러스가 개발되었지만, 기술기록물에 대한 시소러스 구축이나 이와 관한 연구는 매우 미흡함을 알 수 있다.

## 2. 구술기록물과 시소러스

### 2.1 구술기록물의 특성

#### 2.1.1 개념과 유형

‘구술’은 구술자의 기억을 통해 재현된 내용을 기록화하는 작업으로, ‘구술사’는 구술 기록에 근거한 역사 기술이면서, 구술에 의해 생산된 자료라는 2가지 의미를 포함한다. 즉, 구술사는 단순한 구술 채록뿐만 아니라 채록 과정에서 산출된 구술 기록물을 분석하고, 그 경험을 재해석해서 새롭게 역사화하는 과정이 수반되어야 한다. 여기에서 구술기록물은 구술자의 기억과 경험 안에서 중요한 시대적 의미를 부여하기 위한 기초 자료가 된다. 『기록학용어사전(2008)』에서는 구술기록물을 ‘면담을 통해 개인의 기억 속에 남아있는 과거나 역사적 사건을 재구성하는 작업을 통해 만들어지는 기록’으로 정의하였다. 종합하면, 구술기록물은 당대의 사회상 및 집단기억을 형성하기 위한 수단, 영구적으로 보존·활용해야 할 중요한 정보자원이다(이정연, 2009).

일반적으로 구술 자료의 종류는 구전, 구술 증언, 구술생애사로 구분한다. 구전은 여러 세대를 거쳐서 구전으로 전승되는 구술로 된 기록을, 구술 증언은 한 개인이 과거의 특정 사건이나 경험을 현 시점에서 서술하는 것을 뜻한다. 구술생애사는 한 개인이 태어나서부터 현재까지 살아온 경험을 현 시점에서 기술하는 것을 의미한다. 그러나 실제로 이 세 가지 유형은 연구 과정에서 명확하게 구별되기 보다는 혼재되어 나타나는 경우가 많다(윤택림, 함한희, 2006, pp. 57-59). 이외에도 기록물의 매체나 형태적

측면에서 구술 자료를 구술 채록 과정에서 산출된 다양한 유형의 자료, 녹취록, 동영상, 녹음 자료, 사진, 텍스트 자료 등으로 정의할 수 있다(이정연, 이정연, 유종덕, 이종윤, 2011).

#### 2.1.2 특성

구술기록물의 가장 큰 특성은 문어가 아닌 구어를 기반으로 한다는 것이다. 구술자 개인의 언어 행위의 특성을 반영하기 때문에, 구술기록물 상에서 구술자가 사용한 언어는 구술자가 속한 사회문화적 배경과 위상에 관한 정보인 계층적 차이, 직업, 세대, 지역성에 대한 많은 정보를 제공한다. 그리고 그 시대에 통용된 단어나 용어를 사용하기 때문에 구술자의 언어를 통해 시대성과 역사성을 보여줄 수 있다.

구술기록물은 사건 발생 자체보다는 구술자가 사건을 어떻게 기억하는지에 대한 과정과 해석에 초점을 두기 때문에 구술자가 사용한 언어와 그 언어의 전달 의미가 매우 중요하다. 또한 당시 시대를 설명하여 재해석을 가능케 하려면, 원본의 맥락이 잘 나타날 수 있도록 최대한 구술자의 언어로 표현되어야 한다. 이를 통해 생성된 구술기록물은 연구자 자신뿐 아니라 다른 연구자에게도 자료를 비롯한 내러티브의 맥락을 원본을 통해 활용될 수 있다(윤택림, 함한희, 2006, pp. 50-52; 이정연, 2009).

### 2.2 시소러스의 개념과 사례

#### 2.2.1 정의와 기능

시소러스는 특정 전문 분야에서 사용되는 개념어들을 바탕으로 용어의 정의보다는 용어 간 관계를 강조한다(Taylor & Jourdey, 2009).

등위, 계층 및 연관 관계를 중심으로 조직·통제된 색인어의 어휘집으로, 정보 검색을 위한 색인어 간의 동의어, 하위어, 관련어 등의 관계를 나타낸 사전이다. 여기서 색인어란 어떤 개념을 나타내는 단어로 주로 명사나 명사구 형태로 표현된다.

시소러스의 주요 기능은 같은 개념에 대해 다양한 표현을 사용하는 용어들을 통제하여 효율적인 탐색을 도모하는 것이다. 즉, 접근점으로 선정된 모든 어휘에 대해 표준 형식의 접근점을 선정하고, 이와 상이한 형식을 서로 연결하여, 특정 기록물과 관련된 모든 자료를 검색할 수 있다. 일반 주제어뿐만 아니라 변화 가능성이 높은 특정 사건이나 조직 명칭 등 고유명사들의 접근점을 통제함으로써 검색의 효율을 높이는 데 기여한다(이혜영 외, 2008). 또한 용어 간의 계층 및 연관 관계를 중심으로 색인자와 이용자가 용어의 의미적 맥락을 좀 더 명확하게 파악하는데 도움을 줄 수 있다. 그리고 동일한 용어가 다른 주제 분야에서 다양한 의미로 사용되는 경우 한정어를 이용해서 최대한 모호함을 배제시킬 수 있다.

궁극적으로 시소러스와 같은 통제어의 사용 목적은 색인어 선정 과정에서 효율성과 일관성을 제공하면서도 검색 시 이용자가 입력하는 검색어와 색인어 간의 관계를 연결시켜 검색의 재현율과 정확률을 증진시키는 것이다. 이러한 관점에서 구술기록물에서 시소러스를 활용한다면, 구술자와 연구자, 일반 이용자의 언어를 상호 연결시켜서 구술기록물의 주제 접근 효율성을 높일 수 있다.

특히 본 연구에서 수집된 구술기록물의 경우, 구술자가 사용한 용어 중에는 한국 근대에 통용

된 구어로써 연구자가 사용하는 전문 용어나 현대를 살아가는 일반 이용자의 구어와는 차이가 있다. 또한 구술자가 살아온 시대와 출신 배경 및 환경에 따라서 동일한 주제어가 다르게 표현되고 인식되고 있다. 예를 들어 's언니'는 '친한 언니'를 뜻하는 용어로 근대에 사용되었지만 현재는 통용어가 아니다. 또한 '경도'는 '월경'의 동의어로 분단 이전에는 통용된 용어이나 현재 남한에서는 거의 사용되지 않고 있다. 이러한 용어들의 관계를 설정하여 시소러스를 구축하고, 이용자 검색에서 활용할 수 있다면 찾고자 하는 정보를 모두 발견할 수 있다.

## 2.2.2 국내외 구술사 시소러스 사례

### 1) USC Shoah 재단 연구소 시소러스

USC Shoah 재단 연구소는 1994년 설립되어, 홀로코스트 생존자와 목격자의 증언 자료를 수집·관리하고 있는 기관으로 2005년 미국 사우스캐롤라이나 대학의 인문과학대학에 편입되었다. 1996년에는 디지털 시각 역사 아카이브(Visual History Archive)를 구축하였고, 현재 51,696명의 목격자 증언과 60,000개의 키워드와 120만 명의 인명, 70만점의 이미지를 보유하고 있으며, 수집한 구술기록물을 바탕으로 색인어를 추출하여 시소러스를 구축하고 있다(USC Shoah Foundation Institute Homepage, "About us").

시소러스는 ANSI / NISO Z39.19-2003: Guidelines for the Construction, Format, and Management of Monolingual Thesauri을 준용하여 개발되었다. 현재 5만개 이상의 용어를 수록하고 있으며, 동의어와 유사동의어 중심의 등위 관계와 계층 관계를 설정하였다. 연관 관계는 정의하지 않았으며, 범위주기를 이

용해서 용어 정의를 기술하였다. 17개의 최상 위어를 바탕으로 용어의 성격을 살펴본 결과, 이주, 문화, 일상생활, 차별, 감정·생각, 정치·경제 등과 같은 일반 주제성 용어와 함께 지명과 기관명, 인명, 사건명 등의 고유명사, 정치 영상 및 동영상과 같은 기록물 형태 사항도 포함되었다(USC Shoah Foundation Institute 2010: USC Shoah Foundation Homepage, "Cataloguing and Indexing").

## 2) UK 아카이브 시소러스

UK 아카이브 시소러스(UK Archival Thesaurus, UKAT)는 2003년부터 2004년까지 영국의 기록관리 영역에서 색인과 검색을 지원하기 위해 개발되었으며, 현재까지 용어가 지속적으로 갱신되고 있다. UKAT에서는 시소러스 개발 과정을 요구 사항 도출, 구축 범위 설정, 시소러스 용어 선정, 용어 관계 설정, 초기 데이터 반입 및 소프트웨어 테스트, 검수 및 수정으로 정의하였다(UKAT, 2003).

UKAT은 기존의 세계적인 아카이브 기관들이 색인에서 활용한 UNESCO 시소러스를 근간으로 초기에는 국가 디지털 데이터세트 아카이브(National Digital Archive Datasets, NDAD)에 수록된 National Archive의 PROCAT 주제 색인어를 중심으로 구축되었다. 시소러스의 구조와 형식은 영국 국가 표준 BS 5723-1987: British Standard Guide to Establishment and Development of Monolingual Thesauri와 ISO 2788-1986, ANSI/NISO Z39.19을 준용하였으며, 교육, 과학, 문화, 사회과학, 인문과학, 정보 및 커뮤니케이션, 정치학, 법학, 경제학 분야를 광범위하게 포괄하였다. 특히 이 시소러스는

아직 표출되지 않은(under-represented) 집단의 역사와 경험을 표현하는 용어에 주안점을 두었다.

UKAT에는 2014년 1월 다운로드 파일 기준으로 22,607개 용어가 수록되었으며 교육, 과학, 문화, 사회과학 및 인문학, 정보와 커뮤니케이션, 정치 및 법, 경제학, 국가 구분, 사건 등 8개 주제 영역으로 구성되었다. 용어 관계는 등가 관계, 계층 관계, 관련 관계로 정의되었으며, 범위주기를 사용하여 용어 정의를 제공하였다(UKAT Homepage, "About UKAT").

## 3) 국가기록원 일반주제 시소러스

2007년부터 국가기록원은 기록물의 생산·정리·검색 과정에서 출현하는 용어 통제를 시도하였으며, 2007년 2월부터 '기록물 관리·검색용 시소러스(검색어 사전)' 구축 계획을 수립하고 2007년 10월 3만개 수준의 일반주제 시소러스를 개발하였다. 전문용어가 아닌 일반어를 주축으로 한 주제어 사전으로 일반 명사를 중심으로 구축되었고 전년도 20개 중앙행정기관 생산기록물에 대한 키워드 입력 사업을 통해 선정된 12,000개 용어를 주축으로 하였다. 이후 업무 기능과 더불어 지명, 외국인명, 기관·단체명, 법률명, 행사·회의명, 사건·사고명 등 고유명사가 추가되었다(박부숙, 2007).

시소러스는 국제 표준인 ISO 2788과 국가 표준인 KS X 15489: 2007 문헌정보 - 기록관리를 준용하였다. 관계는 등위·계층·연관 관계를 중심으로, 등위 관계는 우선어(USE), 비우선어(UF) 이외에 UP(Upward Posting)를 두어 우선어보다 범위가 작지만 문서작성이나 검색 시 우선어와 구분 없이 혼용하여 사용하는

용어를 정의하였다. 또한 PT(이전 용어)와 LT(이후 용어) 관계를 두어 용어의 시대적 변화와 이력을 관리할 수 있도록 하였다. 계층 관계는 BT(상위어), BTI(상위어: 사례관계), NT(하위어), NTI(하위어: 사례관계), 관련 관계로 RT를 정의하였다. 이 밖에 대응 외국어로서 중국어, 영어, 스페인어 등 16개 외국어 코드가 있었다. 범위주기에 필요한 경우 용어의 정의를 기술하였으며, 한정어에는 업무 분야명 등 짧은 용어를 기술하고, 참조 정보로 한자정보를 기재하였다(국가기록원, 2009).

#### 4) 국사편찬위원회 한국역사용어시소러스

한국역사용어시소러스는 2000~2005년까지 총 6만여 단어(인명, 지명, 서명, 문화재 1만 건 포함)의 한국역사용어 시소러스 사전으로 구축되었으며, 한국사 DB와 한국 역사정보통합 시스템과 연계되었다. 이 시소러스는 국제 표준인 ISO-2788을 준용하였으며, 역사 용어가 시간과 공간적 속성을 동시에 갖는다는 특성을 용어의 선정과 배열에 반영하였다(류주희, 2004).

구축 대상 용어는 중등학교 국사교과서 및 역사 사전류, 한국 역사정보통합시스템에 구축된 원전자료에서 용어를 추출해서, 용어의 범주를 13개의 주제와 인명, 지명, 서명, 문화재로 구분하였다.

표목은 한글명(한자)(한정어) 형식으로 띄어쓰기 없이 기술하였다. 용어 관계는 상위어(BT), 하위어(NT), 관련어(RT), 유의어(USE-UF)를 중심으로 입력되었으며, 이외에도 영문명, 범위주기, 분류(해당 용어의 주제 카테고리), 연도(해당 용어와 관련된 연도), 시대(해당 용어가 일어난 시대명)를 기술하였다.

## 3. 시소러스 구축 과정

### 3.1 구축 절차

ANSI/NISO Z39.18에서는 시소러스 구축 절차를 구축, 테스트 및 평가, 유지, 관리 시스템 운용으로 나누고, 세부적인 구축 방법으로는 하위 용어부터 수집하여 계층 용어를 수집하는 상향식 방식과, 상위 용어부터 하향식으로 계층 용어를 구축하는 하향식 방식을 제안하였다. 그리고 ISO 25964에서는 시소러스 구축을 계획, 구현 초기, 구현, 시소러스 공개, 업데이트로 구분하고, 각 단계별 내용과 고려 사항을 제시하였다. 먼저 계획 단계에서는 목적과 대상, 업무 및 책임분담, 사용하고자 하는 소프트웨어를 제시해야 한다고 하였다. 구현 단계에서는 이용자를 비롯한 이해당사자와 시소러스 전문가를 참여시키도록 권고하였다. 공개 단계에서는 주제 분야, 언어, 참고한 표준, 약어 및 특수문자에 대한 설명, 총 용어 수, 용어 선택 및 복합어, 우선어 선택 등 시소러스에 적용된 규칙 등에 대해 설명해야 한다고 권고하였다. 그리고 배포를 위해 전자 시스템에서 시소러스를 관리할 것을 제안하였다(박옥남, 2011).

본 연구에서는 국내외 시소러스 개발 사례와 ISO 25964와 ANSI/NISO Z39.18을 참고하여 한국 근대 여성 구술사 시소러스의 구축 과정을 '용어 추출 및 후보군 생성 ⇒ 대상 용어 선정 ⇒ 용어 관계 설정 ⇒ 관리 시스템 개발 및 용어 입력 ⇒ 검증 및 수정'으로 이루어진 5단계를 <표 1>과 같이 정의하였다.

〈표 1〉 시소러스 구축 단계

| 단계                  | 주요 활동  |
|---------------------|--|
| 1. 용어 추출, 후보군 생성    | 색인어 추출 및 중복 용어 제거, 시소러스 입력 후보 용어군 생성                                   |
| 2. 대상 용어 선정         | 후보 용어군에서 구축 대상 용어 선정 및 기준 정립   |
| 3. 용어 관계 설정         | 시소러스에서 사용될 관계 정의, 대상 용어에 대한 등위·계층·연관 관계 설정, 기준 정립                      |
| 4. 관리시스템 개발 및 용어 입력 | 시소러스 관리시스템 개발: 주요 기능, 입력 요소, 인터페이스 구현<br>대상 용어 일괄 반입, 용어 관계 및 용어 정보 입력 |
| 5. 검증 및 수정          | 구축된 시소러스에 대한 관계 검증 및 수정  |

### 3.2 구축 단계별 내용

#### 3.2.1 용어 추출 및 후보군 생성

시소러스 구축 1단계로 ‘근대와 여성의 기억’ 연구 프로젝트를 통해 채록된 53명 구술자의 구술기록물 677건에 입력된 색인어 1,784개를 전처리를 통해 중복 용어를 제거하고, 후보 용어군 1,030개를 추출하였다. 구술기록물에 입력된 색인어는 채록을 담당한 박사급 이상의 연구자가 구술자의 생애에서 중요한 경험이나 사건을 나타내는 명사형 키워드로 부여하였다. 색인어는 기본적으로 구술자가 사용한 언어로 기재하되, 동시에 이에 해당하는 표준말이나 전문 용어가 있다면 함께 부여하였다. ‘합정동 연립주택 앞 가족사진’과 같이 구의 형태로 표현된 용어는 ‘합정동’, ‘연립주택’, ‘가족사진’으로 용어를 분절시켜 후보군을 생성시켰다.

#### 3.2.2 대상 용어 선정

시소러스 구축 2단계로 후보 용어군에 속한 전체 1,030개 용어를 대상으로 전문가 회의를 거쳐 시소러스에 수록될 용어를 선정하였다. 이를 위해 여성 구술사와 정보 조직 분야 전문가들이 수차례 정기 회의를 실시하여, 선정 기준과 시소러스의 구축 범위, 기술 방식에 대해 논의하고 결정하였다. 후보 용어군으로 ‘s언니’, ‘s

동생’, ‘스웨터페매기’, ‘샷바느질’, ‘카바레’, ‘기지촌여성’, ‘땅콩장사’, ‘이불가게’ 등과 같이 학술적인 용어가 아닌 여성의 일상사와 관련된 용어가 많았으며 용어 자체에 방언이나 북한말, 비속어, 인명, 국가 및 지역명, 기관 및 단체명, 사건명 등 고유명사도 다수 포함되었다. 그래서 지속적인 정기 회의를 통해 최종 시소러스 대상 용어로 총 995개가 결정되었으며 이 중 지명 87개가 포함되었다.

#### 3.2.3 용어 관계 설정

시소러스의 핵심은 바로 용어 간의 관계를 규명하고 정의하는 것인데, 본 연구에서는 단일 언어를 위한 국제 표준 시소러스 개발 지침인 ISO 2788-1986(E) 『Documentation - Guidelines for the Establishment and Development of Monolingual Thesauri』와 국내외에서 적용한 사례를 참고하여 등가관계(USE, UF), 계층관계(BT, NT), 관련관계(RT)를 용어의 기본 관계로 설정하였다.

등가관계는 동의어와 유사동의어를 비롯하여, 이전 용어(PT), 이후 용어(LT), 방언(DT), 북한어(NK) 등 동일 개념을 표현하는 다양한 유형의 이형을 수용할 수 있도록 관계를 확장했다.

계층관계는 상위어(BT)와 하위어(NT)를 중심으로 설정 과정에서 필요한 경우 새로운 용

어를 추가하였다. 예를 들어, '아버지'와 '부모'는 모두 '가족'의 하위 개념이나, 두 용어 간의 성격이 다르기 때문에 '가족 지위'와 '가족 관계'라는 용어를 추가하여, '아버지'는 '가족 지위', '부모'는 '가족 관계'로 상위어를 설정하였다.

또한 지역의 경우, 국가명은 '지명' 아래 '아시아', '아프리카', '유럽', '오세아니아', '북미/남미' 등 대륙명을 추가하고, 대륙명 아래 국가명, 국가명 아래 지역명을 하위어로 설정하였다. 예를 들어, '길림성'의 상위어는 '중국'을, '중국'은 '아시아'를 상위어로 설정하였다. 한국 지역명은 시·군·구·동까지 지역별로 수준이 다양하므로, 한국십진분류법 제5판의 지역구분표를 참고하여 특별시와 광역시, 도 수준의 지명은 생성한 다음, 하위 군, 구, 동은 현 행정구역 기준에 준하여 계층 관계를 정의하였다. 예를 들어, '합정동'은 '서울특별시' 아래 '마포구'를 추가하고, '마포구' 아래 '합정동'을 하위어로 설정하였다.

연관관계는 계층 관계로는 적절하지 않지만, 주제명 간에 강한 연관성이 있기 때문에 그 관계성을 명시하는 것이 적당한 경우에 설정하였다. 예를 들어, '강제 결혼'과 '매매혼', '고부'와 '고부갈등'은 계층 관계는 아니나 이 두 개념 간에 관련성이 높다고 판단되어 연관관계를 설정하였다. 이외에도 해당 주제명이 포함된 용어나 정책, 유명 인물, 유관 기관 및 단체, 직업, 사건 등이 포함되었다.

### 3.2.4 관리 시스템 개발 및 용어 입력

본 연구에서는 효과적인 시소러스의 구축과 용이한 용어 관리를 지원하고자 웹 기반 시소러스 관리시스템을 개발하여 대상 용어를 일괄 반입한 후, 시스템을 이용해서 약 2개월 간 용어와

용어 정보 및 관계를 입력했다.

이를 위해 관리시스템의 주 이용자를 과업(task)에 따라 3가지 유형으로 구분하고, 각 이용자별 핵심 과업을 다음과 같이 정의하였다. 첫째, '입력자'는 시소러스에 등록된 기존 용어를 메타데이터에 입력하거나, 메타데이터 입력 단계에서 새로운 용어를 시소러스 후보어에 추가할 수 있다. 둘째, '관리자'는 전문가 회의를 통해 선정된 용어를 시소러스에 입력하거나, 기존 용어의 정의나 관계 등을 수정·삭제할 수 있다. 셋째, '검수자'는 시소러스에 새롭게 등록된 용어 메타데이터를 검토하여 이를 수정하거나 확정할 수 있다.

다음으로 관리시스템의 주요 기능을 용어 보기 기능과 시소러스 구축 기능, 용어 반출 기능, 카테고리 관리 기능으로 구성하였다. 첫째, 용어 보기 기능은 구축된 용어의 정보를 확인하고, 수정할 수 있으며, 부가적으로 용어 검색과 용어 구조 브라우징 기능을 구현하였다. 용어 검색에서는 구축된 용어를 대상으로 검색할 수 있고, 용어 구조 브라우징에서는 전체 용어에 대한 계층 구조를 볼 수 있게 하였다. 또한 용어 상태에 따라 고립어와 미검증어 리스트를 별도로 확인할 수 있는데, 고립어는 새롭게 입력된 용어나 아직 관계가 맺어지지 않은 용어를, 미검증어는 검증자가 용어 정의와 관계를 아직 확인하지 않은 용어로 정의하였다.

둘째, 시소러스 구축 기능은 용어의 정의와 관계, 기타 부가적인 정보를 입력할 수 있도록 구현하였다. 기본적인 용어 정보는 용어명(한글), 용어명(영문), 용어명(한자), 해당분야/주제, 카테고리, 출처, 일반적 정의, 맥락적 정의로 정의하였다(〈표 2〉 참조).

<표 2> 용어의 주요 입력 정보

| 구분      | 요소명        | 설명  |
|---------|------------|---|
| 용어정의    | 용어명(한글)    | • 용어의 한글명을 기재하고, 개념 단위로 띄어 씀  |
|         | 용어명(영문)    | • 용어의 영문명<br>• 해당 용어가 외래어인 경우 기재함   |
|         | 용어명(한자)    | • 용어의 한자를 기재함<br>• 특정 글자의 한자가 없는 경우는 한글로 기재함  |
|         | 해당분야/주제    | • 동음이의어를 식별하기 위한 항목으로 명사 또는 명사구로 간략하게 한글로 기재함   |
|         | 카테고리       | • 고유명사와 일반주제어를 구분하여 해당 유형을 선택함<br>• 카테고리명의 추가, 수정, 삭제는 '카테고리 관리' 메뉴에서 설정함                                     |
|         | 일반적 정의     | • 해당 용어의 사전적 정의를 문장 형태로 기술하고, 반드시 출처를 기재함<br>• 여러 정의가 존재하는 경우 1), 2) ... 순으로 기재하며, 출처가 다를 경우 각각 기재함           |
|         | 맥락적 정의     | • 여성주의적 측면, 또는 구술사와 연구자의 시각에서 일반적 정의와는 다른 정의나 해석이 존재하는 경우 기재함<br>• 출처가 있으면, 반드시 기재하고 출처가 없으면 연구자의 이름을 반드시 기재함 |
|         | 출처         | • 일반적 정의와 맥락적 정의에서 외부 자료를 인용한 경우 반드시 기재함  |
|         | 용례         | • 필요한 경우, 용어 사용의 예시를 기재함<br>• 출처가 있는 경우, 반드시 기재함  |
|         | 참조파일       | • 해당 용어에 관한 용례나 정의를 보충하기 위한 파일이 있는 경우 업로드함  |
|         | 검증         | • 검수자가 해당 용어의 정의와 관계가 적절하게 기재되었는지 확인한 다음, 체크박스에 표시하고 저장함  |
|         | 입력자/입력일    | • 시스템 자동생성  |
|         | 수정자/수정일    | • 시스템 자동생성  |
| 검증자/검증일 | • 시스템 자동생성 |   |

대부분의 국내 시소러스에서는 용어의 한글 명 다음에 한자를 각괄호 안에 부기하고, 한정어를 원괄호 안에 기재하고 있으나, 본 연구에서는 특수 문자 입력에 오류가 날 수 있으므로 개별 요소로 분리하여 입력하도록 설계하였다. 또한 요소명을 일반인들이 쉽게 이해할 수 있도록 범위주기(SN) 대신 '일반적 정의'로, 한정어 대신 '해당분야/주제'로 변경하였다. 한편, 용어의 여성주의적 측면이나 구술사와 연구자의 시각에서 일반적 정의와는 다른 정의 또는 해석이 존재할 수 있다는 점을 감안하여 '맥락적 정의'를 추가하였다.

관계 구축은 이미 설정된 용어 관계인 USE, UF, RT, BT, NT, PT, LT, SK, NK, ZT의 관계를 정의하고 하나의 용어가 여러 종류의 관계를 설정할 수 있도록 하였다. 또한 대응 관계는 자동으로 입력될 수 있도록 했는데 예를 들어 '가족'에서 '가족관계'를 NT로 설정한 경우, '가족관계'에서 '가족'에 대해 BT 관계가 자동 생성되도록 하였다.

셋째, 용어 반출 기능은 구축된 시소러스를 파일 형태로 반출하는 것으로, 용어리스트, 용어관계리스트, 자모순시소러스 형식으로 엑셀 파일을 다운로드할 수 있도록 구현하였다.

넷째, 카테고리 관리 기능은 일반 주제어와 고유명사의 성격을 가진 용어를 구분하기 위해 용어에 범주를 정의하는 것으로 주제어(일반), 법률명, 한국인명, 외국인명, 국내기관, 외국기관, 국명, 회의명으로 구분하였다.

### 3.2.5 검증 및 수정

입력된 용어 메타데이터를 대상으로 전문가 회의에서 입력된 용어 정보와 관계를 재검토하였으며 필요한 경우 용어 및 관계를 수정·삭제·추가하였다. 또한 시소러스에 수록된 용어나 우선어 선정 기준, 기술 방식 등에 대한 추가적인 논의와 합의도 이루어졌다. 특히 우선어 선정 기준을 로마자 약어/두문자어/숫자, 단체명, 국가명/지명, 표준어/방언/북한어, 속명/학술명 등으로 구분하여 기준을 좀 더 구체적으로 기술하였다.

## 4 한국 근대 여성 구술사 시소러스 개발

### 4.1 대상 용어

일반적으로 시소러스는 전문 분야의 개념을

나타내는 전문용어를 구축 대상으로 삼는다. 그러나 '한국 근대 여성 구술사 시소러스'는 한국 근대 여성의 생애와 관련한 여성학, 사학, 사회학, 북한학 등 사회과학 분야의 전문용어를 비롯하여 원 구술기록물에서 중요하게 다루어져야 하는 구술자의 중요한 경험이나 사건 등 여성의 일상적 삶과 관련된 일반 용어, 구술자가 사용한 일반 용어까지 구축 대상 범위를 확장하여 개발하였다. 또한 구술자와 연구자명을 제외한 개인명, 단체명, 지리명, 사건명 등 고유 명사도 시소러스 대상 용어로 포함시켰다.

이를 위해 1차적으로 '근대와 여성의 기억' 연구 프로젝트를 통해 채록된 구술기록물의 색인어를 사용하였으며, 2차적으로는 용어 관계, 특히 등위 관계와 계층 관계 설정에 필요한 새로운 용어가 추가되었다. 이렇게 1차 용어 선정 이후, 실질적인 시소러스 구축과 최종 검증 과정을 거쳐 전체 1,076개의 용어가 시소러스에 입력되었으며, 구조적으로 최상위어에 해당하는 용어가 총 39개로 설정되었다(〈표 3〉 참조).

### 4.2 우선어 선정 기준

모든 용어는 원칙적으로 '한글 완전명'을 우

〈표 3〉 근대 여성 구술사 시소러스의 최상위어

|    |      |     |      |        |
|----|------|-----|------|--------|
| 가족 | 기관   | 생애  | 의생활  | 정치사회운동 |
| 개발 | 매체   | 성   | 이주   | 종교     |
| 경제 | 문학   | 시기  | 인간관계 | 주생활    |
| 관습 | 민족   | 식생활 | 인물   | 직업     |
| 교육 | 분단   | 여가  | 임신   | 통일     |
| 교통 | 법    | 예술  | 장소   | 행사     |
| 국가 | 사회가치 | 육아  | 전쟁   | 행정     |
| 국방 | 사회생활 | 의료  | 정치   |        |

선어로 설정하였으며, 나머지 이형은 모두 비우선어로 처리하였다. 그리고 구술자가 표현한 구어가 표준어가 아닐 경우, 표준어를 우선어로 채택하고, 구술자가 사용한 용어는 비우선어로 설정하였다. 사건명의 경우 이형이 많았는데, 이런 경우 참고 정보원에서 가장 많이 채택하고 있는 형태를 우선어로 설정하였다. 예를 들어 ‘한국 전쟁’은 ‘6.25’, ‘육이오’, ‘6.25 전쟁’ 등 다양한 명칭으로 불리므로 『한국민족문화대백과』와 같은 참고정보원을 참고하여 ‘한국 전쟁’을 우선어로 채택하였다.

숫자의 경우, 한글로의 표기가 친숙하지 않은 경우를 제외하고는 숫자를 우선어로 처리하고, 한글표기는 비우선어로 처리하였다. 예를 들어, ‘4.19혁명’은 우선어로 처리하고, ‘사일구혁명’은 비우선어로 처리하였다. 단체명의 경우 기관명이나 행정기구명에 대해서는 정식명칭을, 명칭이 바뀐 경우는 새로운 기관명을 우선어로 채택하였다.

지명의 경우, 국명은 공식명을 우선어로 채택하고, 비공식명이 더 일반적인 경우에는 이를 우선어로 삼았다. 판단이 어려운 경우, 한국십진분류법 제 5판의 지역구분표에 명시된 표목명을 참고하였다. 예를 들어, ‘북한’은 우선어로 선정하고, ‘조선민주주의인민공화국’은 비우선어로 설정하였다. 그리고 지명과 행정구역명이 동시에 존재하는 경우에는 현재 사용하는 행정구역명을 우선어로 채택하였다.

한편, 시소러스 상에 방언과 북한어가 많았는데, 기본적으로 표준어 형식을 우선어로 채택하고, 구술자가 사용한 방언과 북한어는 동의어인 비우선어로 취급하되, 북한어는 별도의 관계인 NK(North Korea)로 설정하였다. 그러나 대응

되는 적절한 표준어가 없는 경우, 방언과 북한어를 우선어로 채택하였다.

### 4.3 관계 설정

기본 시소러스 관계를 중심으로 수집된 구술 기록물의 특성을 반영하여 남한어(SK)와 북한어(NK)의 관계를 추가하였다. 이외에도 향후 상세한 관계 표현을 위해 이전용어(PT), 이후용어(LT), 대응 외국어 등을 추가적으로 마련하였다. 대응 외국어 코드는 ISO-639-3:2007을 적용하였으며, 세계 10대 언어에 동남아시아 국가에 해당하는 기호를 추가하였다(〈표 4〉 참조).

### 4.4 기술 방식

#### 4.4.1 용어명

한글 용어명을 위해 한글 및 외래어 표기는 ‘한글 맞춤법’ 문교부 고시 제88-1호, ‘표준어 규정’ 문교부 고시 제88-2호를 따르고, ‘외래어 표기법’ 문교부 고시 제85-11호(1986년 1월 7일) ‘국어의 로마자 표기법’ 문화관광부 고시 제2000-8호를 따랐다. 시소러스의 용어는 대부분 명사로 구성되나, 필요한 경우 형용사 또는 형용사적으로 쓰인 다른 명사나 동사의 수식을 받는 명사구로 기재하였다. 일반 명사는 단수형으로 기술하였으며, ‘6.25’나 ‘4.19’와 같이 숫자가 포함된 경우, 중간점(·)이 아닌 온점(.)을 사용하였다. 복합 명사, 외래어, 로마자, 고유명사의 처리는 우선적으로 국립국어원의 표준어 규정을, 사건명은 『한국민족문화대백과』를 참고하였다. 띄어쓰기의 경우, 모든 용어는 개념단

〈표 4〉 한국 근대 여성 구술사 시소러스의 주요 관계

| 약어  | 완전명                     | 설명   |
|-----|-------------------------|--|
| SN  | Scope note<br>(범위주기)    | 색인어 내에서 색인어의 의미를 나타내기 위하여 부속된 설명<br>본 시소러스 관리시스템에서는 '정의(일반)'으로 표현함 |
| USE | Use<br>(~대신 사용하라)       | 동어어와 유사동어어 간의 선택이 필요한 경우 지시되는 우선어, 즉 우선어에<br>대한 접두기호로 사용           |
| UF  | Used for<br>(~로 사용하라)   | 뒤에 수반되는 용어가 비우선 동어어 혹은 유사동어어를 나타내는 것으로 비우선어<br>에 대한 접두기호로 사용       |
| TT  | Top term<br>(최상위어)      | 특정 용어의 가장 상위어 개념으로 시소러스의 자모순 배열에서 사용                               |
| BT  | Broader term<br>(상위어)   | 좀 더 확장된 의미를 가진 개념을 표현하는 상위어를 지시하는 기호                               |
| NT  | Narrower term<br>(하위어)  | 더욱 세분화된 의미를 가진 개념을 표현하는 하위어를 지시하는 기호                               |
| RT  | Related term<br>(관련어)   | 관련 있으나 동어어, 유사동어어, 상위어, 하위어를 표현하지 않는 관련어                           |
| PT  | Prior Term<br>(이전용어)    | 용어의 이전, 이후 관계를 표현할 때 이전 용어를 지시하는 기호                                |
| LT  | Later Term<br>(이후용어)    | 용어의 이전, 이후 관계를 표현할 때 이후 용어를 지시하는 기호                                |
| ST  | Standard Term<br>(표준어)  | 표준어와 방언 간의 관계를 표현할 때 표준어를 지시하는 기호                                  |
| DT  | Dialect Term<br>(방언)    | 표준어와 방언 간의 관계를 표현할 때 방언을 지시하는 기호                                   |
| SK  | South Korean<br>(남한어)   | 북한어에 대응되는 국어를 지시하는 기호  |
| NK  | North Korean<br>(북한어)   | 북한어를 지시하는 기호   |
| FT  | Foreign Term<br>(대응외국어) | VIE, LAO, THA, MAY, IND, FIL, JPN, CHI 등                           |
| ZT  | 기타                      | 앞서 제시한 관계에 부합하지 않으나, 차후 관계성을 고려할 필요가 있는 경우<br>설정함                  |

위로 띄어 쓰는 것을 원칙으로 하되, '효부상'과 같이 붙여 쓴 형태가 더 일반적인 경우는 띄어 쓰지 않았다. 인명, 행정부서명, 학과명, 질병명, 단체명, 지명 등 고유명사는 붙여서 기재하였다.

영문 용어명은 한자로 표기할 수 없는 외래어인 경우, 해당하는 영문명의 정식명칭을 기재하고 고유명사를 제외한 일반명사는 소문자로 기입하였다. 띄어쓰기의 경우, 사전에 기재된 형

식을 준용하여 용어의 영문명을 띄어 썼다. 단, 인명은 '성, 이름'의 순으로 기술하고 성과 이름 사이를 띄어 썼다.

한자명은 한자 요소에 기입하고, 한자 병기가 부분적으로 불가능한 경우, 불가능한 자리에 한글 표기를 그대로 사용하였다. 완전 한글 또는 외래어인 경우에는 한자 병기를 생략하였으며, 띄어쓰기 없이 기재하였다.

#### 4.4.2 해당분야/주제

동음이의어를 식별하기 위한 요소로 식별구를 명사 또는 명사구 형태로 간략하게 기입하였다. 일반적으로 시소러스에서 '한정어'란 용어를 사용하지만, 본 연구에서는 일반 이용자가 이해하기 쉬운 명칭으로 바꾸었다. 식별구는 시소러스에 입력된 우선어를 사용해서 식별하도록 하고, 만약 우선어가 없다면 임의로 기재하도록 하였다. 특히 동일명칭의 지명이 두 가지 이상 있는 경우에는 각 지명이 포함되는 광역 지명을 기재하였다. 예를 들어 '오음리'는 전라도와 강원도에 모두 존재하는 지명이므로, 해당분야/주제 요소에 '전라도'와 '강원도'로 입력하여 총 2개 용어로 생성하였다.

#### 4.4.3 용어 정의(일반) 및 출처

원래 시소러스의 범위주기(Scope Note, SN)를 '용어 정의(일반)'로 변경해서 사용했으며, 용어의 정의를 외부의 주제별 전문용어사전, 백과사전, 일반 사전 등을 참조하여 기술하도록 하였다. 정의를 여러 개로 기술하는 경우, 1), 2), 3)을 차례대로 사용하고 참고한 외부 전문 정보의 출처는 '출처' 요소에 기재하도록 하였다.

#### 4.4.4 맥락적 정의

여성주의적 관점에서 '일반적 정의'에 기재한 내용과 다른 해석 또는 정의가 있는 경우에 해당 용어에 대한 맥락적 정의를 기술하도록 하였다. 여성주의적 관점을 반영하여 개념의 재정립이 필요한 경우나, 구술자가 해당 용어를 사용한 의도가 일반적 정의와는 다르다고 연구자가 판단할 경우에 기재할 수 있도록 하였다. 맥락적 정의에 참고한 논문이나 자료가 있는 경

우 반드시 정의 다음에 출처를 명시하고, 없는 경우 정의를 입력한 연구자의 이름과 직위를 입력하도록 하였다.

## 5. 결론

2000년대 이후부터 역사에 기록되지 않은 민중의 삶을 조명하기 위한 구술채록 작업이 정부 및 연구 기관을 중심으로 이루어졌다. 그러나 구술채록이 주로 사건 중심적이고, 남성중심적이며, 사회 사건의 증거자료 수집을 목적으로 이루어졌다. 더욱이 여성 구술자는 거의 없었으며, 주로 특정 역사적 사건의 '피해자'로서 증언으로 채록되는 경향이 있었다. 그 결과 지금까지의 구술기록물은 한국의 근대적 사회 변화 속에서 여성 경험을 제대로 보여주지 못하였다. 그러므로 근대 경험에 관한 여성 구술채록은 여성 연구뿐만 근대사를 재조명하기 위해서 매우 중요하며, 이 과정에서 생성된 다양한 구술기록물을 수집하고 관리하는 것뿐만 아니라 이제는 효과적인 접근 및 이용도 제공해야 한다.

축적된 구술기록물은 일회성 자료가 아니라 한국 근대사 연구에 지속적으로 활용될 수 있기 때문에, 해당 분야 연구자를 비롯한 이용자가 다양한 방식으로 접근할 수 있도록 해야 한다. 특히 구술기록물은 내용 중심의 주제 접근 방식이 필요하지만, 지금까지 이에 대한 연구나 시도가 거의 이루어지지 않았다. 이에 본 연구는 한국 근대 여성의 생애와 관련된 주제어를 중심으로 시소러스를 구축하였다.

구축 과정은 제 1단계로 53명 구술자의 녹취록, 텍스트, 이미지 메타데이터에 수록된 한글

색인어 총 1,784개 추출하였으며, 중복 용어 처리 후 국내 지역명과 국가명을 포함하여 후보용어 1,030개를 생성하고 용어의 특성을 분석하였다. 제 2단계에서는 여성 구술사 및 정보 조직 분야 전문가들이 정기 회의를 거쳐 시소러스 대상 용어를 선정하였다. 제 3단계에서는 용어에 대한 등위·계층·연관 관계를 설정하고, 우선어 선정 기준이나 표기 방식 등을 결정하였다. 제 4단계에서는 효율적인 시소러스 구축을 위한 웹기반 관리시스템을 개발하고, 용어 및 관계 정보를 입력하였다. 관리시스템의 입력 요소와 관련하여 용어와 관련된 핵심적인 정보와 관계를 메타데이터로 정의하였다. 제 5단계에서는 입력된 용어를 대상으로 용어 정보와 관계를 재검토하였으며 필요한 경우 용어 관계를 수정·삭제·추가하였다. 최종 구축된 시소러스는 39

개의 대주제 영역의 1,076개의 용어로 구성되었으며 추후 지속적인 한국 근대를 산 여성의 구술기록물을 바탕으로 추가되고 확장될 것이다.

본 연구에서 개발한 시소러스는 '여성'이라는 젠더의 특성과 '한국 근대사'라는 시대적 특성, 그리고 '구술기록물'이라는 자료의 특성이 복합적으로 반영되었다. 따라서 본 시소러스는 구술기록물을 이용하는 연구자와 일반 이용자에게 다양한 접근점을 제공할 수 있는 효과적인 검색 도구로 활용될 수 있으며 여성을 중심으로 근대를 새롭게 조명할 여성들의 구술생애기록을 지속적으로 수집하고 축적하여 서비스하는데 기반이 될 것이다. 더 나아가 본 연구에서 제시한 시소러스 개발을 위한 실제 용어 수집 과정이나 설계 및 개발 과정은 향후 다양한 주제 분야의 시소러스 개발의 참고자료가 될 것이다.

## 참 고 문 헌

- 공귀영, 리상용 (2008). 대한민국입시정부기록물의 단체명 전거레코드 설계 방안에 대한 연구. 서지학 연구, 40, 181-212.
- 국가기록원 (2009). 국가기록원 시소러스 지침(1.1). 대전: 국가기록원.
- 류주희 (2004). 문화콘텐츠 기반으로서의 한국역사용어 시소러스. 인문콘텐츠, 4, 199-211.
- 박부숙 (2007). 기록물 검색 서비스를 위한 시소러스 개발과 활용. 기록인, 1, 56-57.
- 박옥남 (2011). ISO25964를 통해 살펴본 시소러스 구축 표준에 관한 연구. 정보관리연구, 42(2), 43-65.
- 설문원 (2002). 기록물을 위한 단체 전거레코드 연구. 한국기록관리학회지, 2(2), 39-68.
- 설문원 (2010). 기록 검색도구의 발전과 전망. 기록학연구, 23, 3-43.
- 윤택림 (2012). 왜 구술사인가: 한국 근현대사에서 구술 기록의 중요성. 기록인, 20, 18-23.
- 윤택림, 함한희 (2006). 새로운 역사쓰기를 위한 구술사연구방법론. 서울: 아르케.
- 이정연 (2009). 구술사 기록물 아카이브 구축을 위한 메타데이터 모델링 및 표준 요소 개발에 관한 연구. 정보관리학회지, 26(1), 163-184.
- 이정연, 이정연, 유종덕, 이종윤 (2011). 여성구술생애기록물 맥락 표현을 위한 메타데이터 구조화에

- 관한 연구. 기록학연구, 30, 57-88.
- 이해영, 이미영, 이은영, 이혁준, 이현정, 최영실, 박미자 (2008). 대학기록관 시소러스 구축 지침의 개발 연구: 명지대학교 대학사료실의 사례를 중심으로. 한국기록관리학회지, 8(1), 189-210.
- 이현정 (2006). 수집형 기록관의 전거제어시스템 분석: 민주화운동기념사업회 사료관리시스템의 사례를 중심으로. 기록학 연구, 13, 91-134.
- Garrod, Peter (2000). Use of the UNESCO Thesaurus for archival subject indexing at UK NDAD. *Journal of the Society of Archivists*, 21(1), 37-54.
- Giménez-Chornet, V. & Escrig-Giménez, M. (2011). Designing a thesaurus to give visibility to the historical archives in the Archivo del Reino in Valencia. *Knowledge Organization*, 38(2), 154-166.
- McNitt, William H. (1989). Development of the Presnet subject descriptor thesaurus. *The American Archivist*, 52(3), 358-364.
- Taylor, Arlene G. & Jourdre, Daniel N. (2009). *The organization of information*, Westport, CT: Libraries Unlimited.
- UKAT (2003). UK Archival Thesaurus(UKAT): Construction and editing methodology. Retrieved January 10, 2014, from <http://www.ukat.org.uk/downloads/methodologyV6.pdf>
- USC Shoah Foundation Institute (2010). USC Shoah Foundation Institute Thesaurus (Abridged). Retrieved January 10, 2014, from [http://sfi.usc.edu/download/SFI\\_Thesaurus.pdf](http://sfi.usc.edu/download/SFI_Thesaurus.pdf)

[홈페이지]

- 국가기록원 관련용어 검색. 검색일: 2014. 1. 26.  
<http://www.archives.go.kr/next/search/thesaurusSearch.do>
- 국사편찬위원회 한국역사용어 시소러스. 검색일: 2014. 1. 26. <http://thesaurus.history.go.kr/>
- USC Shoah Foundation Institute Homepage. Retrieved January 26, 2014, <http://sfi.usc.edu/>
- UKAT Homepage. Retrieved January 26, 2014, <http://www.ukat.org.uk/>

• 국문 참고자료의 영어 표기

(English translation / romanization of references originally written in Korean)

- Gong, Gui-Young & Lee, Sang-Yong (2008). A study on the modeling of archival authority records for corporate bodies of the provisional government of the Republic of Korea: 1919-1945. *Journal of the Institute of Bibliography*, 40, 181-212.
- Lee, Hyun-Jeong (2006). Analysis of authority control system in collecting repository from

- the case of archival management system in Korea Democracy Foundation. *Korean Journal of Archival Studies*, 13, 91-134.
- Lee, Jung Yeon, Lee, Jung Yeoun, Ryoo, Jong Duk, & Lee, Jong Yoon (2011). A study on metadata structuralization for context representation of women's oral life history. *Korean Journal of Archival Studies*, 30, 57-88.
- Lee, Jung Yeoun (2009). A study on modeling metadata and developing standard elements to establish oral history archives. *Journal of the Korean Society for Information Management*, 26(1), 163-184.
- National Archives of Korea (2009). *Guide for thesauri of National Archives of Korea (1.1)*. Daejeon: National Archives of Korea.
- Park, Boo Sook (2007). Development and usage of thesauri for archive searching services. *Girok-in*, 1, 56-57.
- Rieh, Hae-young, Lee, Mi Young, Lee, Eun Young, Lee, Hyuk Joon, Lee, Hyun Jeoung, Choi, Young Sil, & Park, Mi Ja (2008). Study on the development of guidelines for thesaurus construction at university archives: Case study of Myongji University Archives Center. *Journal of Records Management & Archives Society of Korea*, 8(1), 189-210.
- Ryu, Ju-hee (2004). Korean history thesaurus as a basis of cultural content. *Humanities Content*, 4, 199-211.
- Seol, Moon-Won (2002). A study of archival authority records for corporate bodies. *Journal of Records Management & Archives Society of Korea*, 2(2), 39-68.
- Seol, Moon-Won (2010). A study on development and prospects of archival finding aids. *Korean Journal of Archival Studies*, 23, 3-43.
- Yoon, Taek-Lim & Hahm, Han-Hee (2006). *Oral history research methods*. Seoul: Arche.
- Yoon, Taek-Lim (2012). Why oral history?: Importance of oral history records in history of modern Korea. *Girok-in*, 20, 18-23.