

## 조선시대 필사본 음식조리서의 음식 용어 의미 주석을 위한 연구\*

- 『음식디미방』을 중심으로 -

남길임 · 송현주  
(경북대학교)

<Abstract>

**Nam Kil-im, Song hyeon-ju. 2008. A sense annotation of culinary terminology in the Joseon dynasty culinary manuscripts. *Korean semantics* 26.** The purpose of this paper is to construct sense tagsets for tagging to food related terms used in Joseon dynasty culinary manuscripts. In addition to the morpho-syntactic annotation, the culinary terms in the corpus are annotated with domain-specific (in this case, culinary domain) sense information. For building sense tagsets are reviewed briefly Sejong Corpus, the study of lexical classes and food terms of the culinary scholars. After due consideration, we build the proper tagsets for this specialized corpus, the tagsets carries out the analyses of lexical classes. The major tagsets of this study is culinary terminology. Therefore, most words are hyponyms of 'FOOD', and they can be classified into the following: food name, ingredients, seasoning, garnish, kitchen utensils, tableware, measuring unit etc.

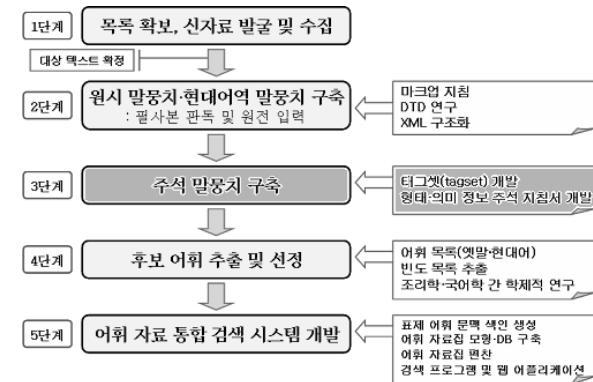
핵심어: 의미 주석 말뭉치(sense tagged corpus), 의미 주석(sense annotation), 조리 용어(culinary terminology), 조선시대 음식조리서(Joseon dynasty Culinary Manuscripts), 어휘 부류(lexical class)

\* 이 연구는 2007년 학술진흥재단(기초연구과제-인문사회분야)의 지원을 받았으며, 과제명은 '조선시대 한글 음식조리서 연구를 통한 조리 용어 통합 검색 시스템 개발'이다 (KRF-2007-322-A00050).

### 1. 들머리

본 연구는 조선시대 필사본 음식조리서의 검색 시스템 개발 과정의 한 단계로, 필사본 말뭉치에 나타난 음식 용어에 어휘 의미 주석을 함으로써 의미 주석 말뭉치를 구축하기 위한 기초 연구에 해당한다. 검색 시스템에서 검색 경로를 다양화하기 위해서는 어휘 형태 검색뿐만 아니라 의미 부류 검색이나 어휘 유형별 검색도 가능해야 한다. 이를 위해서 본 연구에서는 음식 이름, 음식 재료, 계량 단위, 조리 방법 등의 어휘 부류 체계를 구축하고, 이 부류 체계에 입각하여 음식조리서에 의미 주석을 수행하는 어휘 의미 주석의 단계에 대해 논의하고자 한다.

이 연구는 <그림 1>의 “조선시대 한글 음식조리서 연구를 통한 조리 용어 통합 검색 시스템 개발” 사업 중 3단계인 주석 말뭉치 구축 작업의 하위 작업에 해당된다. 즉, 의미 주석의 단계는 원시말뭉치 및 현대어역 말뭉치 구축, 마크업, 형태 주석 이후에 이루어지며, 이후 어휘 검색 시스템의 의미 부류별 검색을 위한 기초 자료로 활용된다.



<그림 1. 음식조리서 통합 검색 시스템 개발의 단계>

본 연구의 결과물인 검색 시스템은 17·18세기라는 특정 시대, 음식조리서라는 특정 전문 분야의 텍스트에 대한 일반인의 접근성을 높인다는 점뿐만

아니라, 형태별·의미 부류별·출전별 검색이 가능하도록 주석 말뭉치를 구축함으로써 국어학 및 조리학 등 전문가들에게도 도움이 될 것이다. 본 연구는 국어사적 연구인 동시에 어휘의미론적 연구이며, 조리학적 활용을 염두에 둔 연구이기도 하다. 여기에서는 이 중 어휘의미론적 관점에서 조리학적 정보를 주석하는 방법론에 대해 주로 논의할 것이다. 논의의 편의를 위하여 본 연구는 다음과 같이 구성된다.

2장에서는 의미 주석을 위한 기본적인 논의 사항으로 의미 주석 대상인 조선시대 필사본 음식조리서의 텍스트적 특성과 그 특성에 맞는 의미 주석을 위한 의미 부류 체계 구축에 대해 논의한다. 특히 여기서는 기존의 의미 부류 체계가 주로 범용적 목적을 위해 구축되었고, 사전 기반 의미 주석에 집중되어 왔음을 지적하고, 특정 전문 분야의 의미 부류 체계 구축을 위해 보완되어야 할 사항에 대해 논의할 것이다. 3장에서는 의미 주석 체계와 관련한 실질적인 작업들을 예비 작업 단계, 태그세트 개발 단계, 의미 주석 단계로 구분하여 제시할 것이며, 4장은 지금까지 진행되어 온 작업에서 향후 해결되어야 할 사항들에 대해 문제제기를 함으로써 결론을 대신하고자 한다.

본 연구의 주요 대상은 『음식디미방』<sup>1)</sup>의 어휘 약 5,000어절에 한하며, 이 연구의 결과는 향후 17·18세기 음식조리서 10만 어절 가량으로 확대 적용될 예정이다.

## 2. 의미 주석을 위한 기본적인 논의 사항

### 2.1. 조선시대 필사본 음식조리서의 특성

본 연구는 범용적인 용도가 아닌 특정 시기, 특정 분야의 텍스트를 특정

1) 『음식디미방』은 경북 북부 안동과 영양 일대에 살았던 정부인 안동 장씨(1598-1680)가 말년에 저술한 음식 조리서로 17세기 중엽 우리 조상들의 식생활 실상을 잘 알려 주는 문헌이다. 이 책은 면병류(18개), 어육류(74개), 주류 및 초류(54개)의 세 개 항목 아래 147개 항에 달하는 음식 조리법을 한글로 서술한 최초의 한글 조리서이다(백두현 2006: 9-23).

목적에 대해 검색하는 데 필요한 의미 주석 체계를 지향한다. 따라서 의미 주석의 대상이 되는 텍스트의 특성과 의미 정보의 특성을 분명히 해야 할 필요가 있다.

본 연구의 대상인 17·18세기 조선시대 필사본 음식조리서는 현대의 ‘요리책’에 해당하는 장르로, 부녀자들을 중심으로 전해 내려오는 당시의 일상적인 삶을 반영하는 생활 어휘가 녹아 있는 자료이다. 특히 국어학적으로 전통 음식과 관련된 각종 어휘 자원이 풍부하며 내용 특성상 재료명·조리 기구명·음식명·그릇명 등의 명사가 등재되어 있고, 각종 조리법과 관련된 동사가 풍부하다. 그리고 정도부사·상태부사·의성어·의태어 등의 고유어 부사가 자주 등장하는데, 특히 이러한 부사는 한글 자료에서만 제대로 확인할 수 있는 귀중한 한국어 어휘 자산이다.

다음의 <그림 2>는 『음식디미방』의 첫 페이지에 있는 ‘면병류’ 중 ‘면’의 원문과 현대어역이다<sup>2)</sup>. 여기서 면병류(麵餅類)란 ‘면으로 만든 국수류 음식과 떡 종류’를 뜻하며 요리 부류명에 해당한다.

[1] 원문  
음식디미방 면병류 면  
것모밀를 써어 하 면이 물너디 말고 알마초 물뵈어 쌀을 조히 아아 디홀 제 미리 물 품겨 축축이 햏야 듯다가 디홀 제 녹도 거피흔 쌀 조히 시어 건져 물 썩거든 모밀발 닳되에 물 부른 녹두 혼 복즈식 섯겨 지흔디 방하를 7만7만 지허 것글룰 처 버리고 키로 퍼 버리고 키 그테 흰 쌀이 나거든 그를 피화 다치 햏면 그 굴리 7장 희거든 면 물 제 더운물에 녹게 마라 누르면 비치 희고 조흔 면이 되느니라 교터는 식면 교터 7치 햏라.

[2] 현대어역  
음식디미방 면병류 면(麵) (메밀국수 만드는 법)  
겉메밀을 씻어 너무 많이 말리지 말고 알맞게 말려라. 메밀쌀의 잡것을 가려내고 깨끗이 해서 찜기 전에 미리 물을 뿜어 축축이 해 두어라. 찜을 즈음에 녹두 알갱이의

2) 『음식디미방』의 원문과 현대어역은 백두현(2006)을 따른 것이다. 또한 본 연구에서 대상으로 하는 조선시대 음식조리서는 대부분의 경우 원문에 띄어쓰기가 되어 있지 않다. 이에 형태 주석을 위하여 말뭉치 구축 시 입력자가 현대 정서법을 기준으로 하여 띄어쓰기를 하여 입력하였다. 다만 띄어 쓰는 것이 원칙이고 붙여 쓰는 것을 허용할 경우에는 붙여서 쓰도록 하였다.

껍질을 벗겨 깨끗이 씻은 다음 건져 내라. 물이 빠지거든 메밀쌀 다섯 되에 물에 불린 녹두를 한 복자씩 섞어 찜되, 방아를 살살 찜은 후 걸가루를 체로 치고 키로 까 붙어 키 끝에 흰 부스러기가 나오거든 그것을 모아 나누면 그 가루가 아주 희다. (그 가루로) 면을 반죽할 때에 더운 물에 눅게 반죽하여 누르면 빛이 희고 깨끗한 면이 되느니라. 고명은 세면(실국수)의 고명과 같게 하라.

<그림 2. 『음식디미방』 1a, ‘면병류’에서 ‘면’의 원문과 현대어역>

위 텍스트의 예에서 볼 수 있듯이, 17세기 필사본 음식조리서 『음식디미방』의 텍스트 구조는 ‘책제목 - 요리 부류명 - 요리 이름 - 요리 방법’으로 이루어져 있으며, 텍스트 내용은 주로 요리 방법, 요리 절차, 상차림법 등 ‘식생활’과 관련된 내용이다. 또한 요리 이름 하나당 분량은 다소 차이는 있으나 평균 50어절 정도이며<sup>3)</sup>, 이러한 텍스트 구조와 내용은 당시 집안에서 직접 조리를 담당했던 텍스트의 생산자와 수용자가 텍스트를 생산하고 이해하기 용이하도록 구조화되어 있다.

본 연구의 최종 결과물인 검색 프로그램의 주요 사용자는 조리학자 및 전통 조리학에 관심을 가지고 있는 일반인이 될 것으로 예상된다. 지금까지 원시 말뭉치, 현대어역 말뭉치가 구축되었으며, 형태 주석 말뭉치, 의미 주석 말뭉치 구축을 위한 연구가 진행 중이다. 모든 말뭉치는 TEI P5버전의 XML<sup>4)</sup>로 구축된다.

다음에서는 의미 주석 말뭉치 구축의 기초 작업에 해당하는 의미 부류 체계에 대해 논의하기로 한다.

## 2.2. 의미 주석 말뭉치 구축을 위한 의미 부류 체계의 구축

일반적으로 의미주석말뭉치(sense tagged corpus)는 텍스트의 개별 문맥

- 3) 각 요리마다 텍스트 길이에는 차이가 있으나 『음식디미방』은 총 8000여 어절로 구성되어 있으며 요리 항목은 147개이다. 이를 평균하여 얻은 값이 50이다.
- 4) XML(eXtensible Markup Language)이란 문서와 자료의 마크업(markup)을 위한 ‘일반 문법’을 정의하는데 이용되는 것으로 HTML과 같이 고정된 용도를 가진 제한된 어휘와 문법을 규정하는 것을 넘어서 사용자가 원하는 어떠한 형태의 마크업 문법이라도 규정할 수 있는 메타언어이다(이기황 2007: 149).

에 각 어휘의 의미를 특정 사전의 세부의미항목(sense)에 대응시켜 주석함으로써 구축한 말뭉치로 정의된다. 단, 이때 말뭉치의 구축 목적과 사용 용도에 따라 주석 대상 어휘 유형의 범위와 의미 주석의 수준은 각각 달리 주석될 수 있다. 따라서 의미주석말뭉치를 구축하기 위해서는 해당 텍스트에 어떤 범위의 의미 정보를, 어떤 수준까지 제공할 것인지에 대한 논의가 필요하다.

지금까지 국내에서 수행된 의미주석말뭉치 구축과 의미 부류 체계는 대부분 범용적인 목적 하에서 사전의 뜻풀이 체계를 기반으로 구축되었다는 점에서 전문 용어의 검색 기능을 주목적으로 하는 본 연구의 방향과는 다소 차이가 있다. 따라서 본 연구에서는 기존의 의미주석말뭉치와 의미 부류 체계 결과물을 적극적으로 수용하되, 본 연구의 특성에 맞게 수정·보완함으로써 본 연구의 연구 목적에 적합한, 합목적적인 의미 주석 체계를 구축하고자 한다.

우선, 참조할 수 있는 의미주석말뭉치의 대표적인 예로서, 21세기 세종계획 형태의미주석 말뭉치(약 1,100만 어절)와 서상규(2001)가 있다. 이들은 의미 주석 체계의 기반이 되는 대상 사전, 의미 주석의 수준, 주석 대상 범위의 세 가지 차원에서 달리 구축되었다. 즉 전자는 『표준국어대사전』(이하 『표준』)의 뜻풀이 정보를 활용하되 동형어 수준까지 구축되었으며 조사, 어미 등을 제외한 내용에 한하여 의미 정보를 부착한 반면, 후자는 『연세한국어사전』(이하 『연세』)의 뜻풀이 정보를 기반으로 다의어 수준까지 구축하였으며 기능어, 내용어의 모든 범주를 포괄하였다.<sup>5)</sup> 다음 <그림 3>은 21세기세종계획 기초자료구축 분과의 형태의미주석 말뭉치의 예이다.

4BS_B0110535700	대학정론	대학_01/NNG + 정론_01/NNG
4BS_B0110535710	쉬쉬하는	쉬쉬/MAG + 하/VV + 는/ETM
4BS_B0110535720	'대학평가	'/SS + 대학_01/NNG + 평가_03/NNG
4BS_B0110535730	결과'	결과_02/NNG + '/SS
4BS_B0110535740	대학교육협의회가	대학교육협의회/NNP + 가/JKS
4BS_B0110535750	최근	최근/MAG

- 5) 세종계획 전자 사전 분과에서는 구문 및 의미정보를 이용한 단어 의미 중의성 해소를 위해 체언뿐만 아니라 용언의 의미중의성 해소를 위해 용언의 여러 가지 격틀 정보들 중 문장의 구문에 가장 적합한 격틀 하나를 선택하는 시스템을 구축하여 테스트하였다. 이를 위해서 ‘세종’은 언어 정보, 선택 제약 정보를 이용하고 있다.

4BS_B0110535760	몇	몇/MM
4BS_B0110535770	년	년_02/NNB
4BS_B0110535780	동안	동안_01/NNG
4BS_B0110535790	실시해	실시_03/NNG + 하/XSV + 아/EC
4BS_B0110535800	오고	오/VX + 고/E

<그림 3. 21세기 세종계획의 형태미분석 말뭉치 중 일부>

이와 같은 사전 뜻풀이 의미 번호를 부착한 의미 주석 말뭉치는 전체 국어 어휘 자원의 체계적 의미 주석, 의미 빈도 목록 확보 등에는 용이하지만, 특정 전문 분야에 대한 사전의 뜻풀이는 일관성과 체계성을 가지고 있다는 보장이 없으므로 전문 분야의 의미 정보를 체계적으로 담기 어렵다는 한계가 있다. 특히 본 연구와 같이 실제 응용 시스템에 적용되는 전문 분야의 의미 정보를 구조적으로 제시하기 위해서는 사전 이상의 의미 부류 체계가 요구되는 경우가 있다. 다음의 예를 보자.

- (1) ㄱ. 만두소 장만키는 무을 ㄹ장 무릭 찰마(만두소 장만하기는 무를 아주 무르게 삶아, 『음식디미방』 1a)
  - ㄴ. 소를 기름지고 연흔 고기롤((만두)소로는 기름지고 연한 고기를, 『음식디미방』 8b-9a)
  - ㄷ. 소 아니면 싱치롤 ㄹ느리 ㅍ아 붓가 녀호라(채소를 쓰지 않는다면 꿩고기를 가늘게 다져 볶아 넣어라, 『음식디미방』 1b)
- (2) ㄱ. 개장 곳논 범(개장국 고는 범, 『음식디미방』 8b)
  - ㄴ. 찰과 옛 고아(쌀과 옛을 고아, 『병자기』 145)
- (3) ㄱ. 썩과 밥이 설면 사오납고 둌이 설거나(떡과 밥이 설익으면 좋지 않고 반죽덩어리가 설거나, 『음식디미방』 15a)
  - ㄴ. 무영 주머니에 닛 내듯 쳐 내여(무명 주머니에 넣어 잇꽃 물을 내듯 쳐내어, 『음식디미방』 1b)
  - ㄷ. 술독을 두터운 조희로나 유지로나 빠미라(술독을 두꺼운 종이나 기름종이로 싸매라, 『음식디미방』 15a)

6) 『음식디미방』은 다른 조리서와는 달리 ‘밥’과 ‘죽’의 요리법에 대해서는 언급하고 있지 않다(백두현 2006: 28).

위 (1)은 동형어 주석이 필요한 예로, (1ㄱ), (1ㄴ)과 (1ㄷ)은 각기 다른 어휘에 속한다. 즉 (1ㄱ), (1ㄴ)은 『표준』의 ‘소02’에 해당하는 것으로 ‘송편이나 만두 따위를 만들 때, 맛을 내기 위하여 익히기 전에 속에 넣는 여러 가지 재료’를 뜻하는 반면, (1ㄷ)은 ‘소(素)12’로 ‘고기나 생선 따위를 쓰지 아니하고 채소류만으로 만든 음식’ 즉 여기에서는 ‘채소’를 가리킨다. 또한 (2)는 다의어 주석이 필요한 예로서, (2ㄱ)에서 ‘긋는’은 현대어 ‘고다’의 뜻으로 『표준』의 ‘고다01’의 첫 번째 뜻 ‘고기 등을 오래 삶다’에 해당하며, (2ㄴ)의 ‘고아’는 ‘고다01’의 두 번째 뜻인 ‘옛 등의 액체를 오래 끓여 졸이다’의 뜻이다<sup>7)</sup>.

한편 (3)은 동형어, 다의어 주석으로도 해결되기 어려운 예를 보여 준다. 즉, ‘썩’과 ‘밥’은 보통은 음식 이름이 되겠지만<sup>8)</sup> 술을 빚을 때는 술 만드는 재료가 되기도 하는데, 검색 시스템의 사용자 관점에서, (3ㄱ)에서의 ‘썩’과 ‘밥’이 ‘요리 재료’로 쓰인다는 것은 매우 중요한 정보이다. (3ㄴ), (3ㄷ)의 경우도 동일한 대상이 조리학의 관점에서 다양한 의미 기능을 갖게 되는 경우를 보여 주는데, ‘주머니’나 ‘종이’가 ‘물품을 넣도록 만든 물건’이나 ‘글을 쓰거나 그림을 그리는 물건’이 아니라 조리 과정에서 사용하는 ‘조리 도구’로서의 의미를 획득한 것으로 볼 수 있다.<sup>9)</sup>

(1)~(3)과 같이 동형어, 다의어, 다의어 이상의 고려가 필요한 것, 텍스트 성격상 특별한 의미로 사용되는 것 등은 사전의 뜻풀이만을 기준으로 해서는 적절한 의미 주석을 할 수 없는 경우이다.

이러한 의미 정보의 다차원성을 해결하기 위해서는 주석 대상을 조리 용

7) 고다01

- ① 고기나 뼈 따위를 무르거나 진액이 빠지도록 끓는 물에 푹 삶다.
  - ② 졸아서 진하게 엉기도록 끓이다.
  - ③ 술 따위를 얻기 위하여 김을 내어 증류시키다.
- 8) 음식조리서 중의 하나인 『규합총서』를 보면 ‘오려쌀노 밥을 짓고(오려쌀로 밥을 짓고)’와 같은 용례를 찾을 수 있으며 이 때 ‘밥’은 일반 음식 이름인 ‘밥’으로서 ‘쌀, 보리 따위의 곡식을 씻어서 술 따위의 용기에 넣고 물을 알맞게 부어, 난알이 풀어지지 않고 물기가 잦아들게 끓여 익힌 음식『표준』’의 뜻이다.
- 9) 실제 문맥에서 각 어휘들이 조리 용어로 사용되었다고 판단하는 것은 주석자의 관점에서 다소 자의적일 수 있다. 이러한 자의성을 최대한 줄이기 위하여 본 연구팀에서는 조리학 연구자와 공동 연구를 수행할 것이며, 조리학 전공의 공동연구원에게 국어학적 분석 결과의 최종 검토를 의뢰할 계획이다.

어로 한정하고, 전문 영역 내에서 주석 체계의 일관성을 유지해야 할 필요가 있다. 즉, 주석 대상을 조리 관련 내용어에 한정하되, 주석의 수준은 동형어, 다의어를 포함하여, 조리학적 분류 체계에 기반한 단어 의미 중의성 해소(word sense disambiguation)까지를 목적으로 하는 것이다. 이를 위하여 본 연구에서는 사전의 뜻풀이 정보가 아닌, 별도의 조리 용어 의미 부류 체계를 마련함으로써 조리 용어 주석 작업을 수행하기로 한다.

본 연구에서 조리 용어 분류 체계를 구축하기 위해 참조한 의미 체계는 국내 어휘의미망<sup>10)</sup>, 어휘 분류 체계<sup>11)</sup>, 조리학 분야의 분류 체계 등이다. 이 중 어휘의미망과 어휘 분류 체계는 대량의 어휘 자원을 일관성 있게 분류하는데 기여할 수는 있으며, 조리학 분야에서 일반적으로 통용되는 분류 체계는 비록 국어학적으로 엄밀하지는 않으나 실제 조리학 관련 연구자들의 관점에

10) 어휘의미망 구축과 관련해서 국외 어휘망으로는 Wordnet, EuroWordnet, EDR, CYC, UMLS가 대표적이다. 국내의 어휘망은 CoreNet, U-WIN, ETRI 개념의미망, KorLex, 세종 의미 부류체계 등이 있다. CoreNet은 시맨틱웹 및 자연언어처리 시스템에서의 지식베이스로 활용하기 위해서 구축한 것으로 총 12단계의 하위 분류와 3,239개의 개념 노드가 한·중·일어로 대응되어 있다. ETRI의 개념의미망은 31개의 최상위 레벨의 어휘와 깊이 12레벨로 구성되어 있다. KorLex는 영한 번역을 통한 계층적 어휘의미망을 시작으로 현재 약 12만 개의 어휘로 구성되어 있다. 세종의 명사 분류는 의미 영역을 ‘구체물’, ‘집단’, ‘장소’, ‘추상적 대상’, ‘사태’ 등 다섯 개로 분류하고 해당 의미 영역을 공유하는 명사 어휘 목록과 함께 적정술어를 포함한다.

11) 국어학자들의 연구 결과에 따른 어휘 부분의 대표적인 예는 남영신(1987), 박용수(1989), 임지룡(1991), 국립국어원(1993), 신현숙 외(2000), 이순애(2006) 등이 있다. 다음은 이 중 조리용어가 포함된 부분만 제시한 것이다.

- ㄱ. 남영신(1987) - 밥, 떡, 음식, 반찬, 식료품·양념, 것갈, 술·담배, 요리, 식사·끼니
- ㄴ. 박용수 엮음(1989) - 식생활: 식품, 음식물, 기호식품, 조리·음식맛, 식사, 식생활 도구
- ㄷ. 임지룡(1991) - 의식주에 관한 어휘: 의생활, 식생활, 주생활, 생필품
- ㄹ. 국립국어원(1993) - 가정: 가정, 인칭, 친족, 집, 세간, 옷, 음식, 살림, 경조사, 생활상
- ㅁ. 신현숙 외(2000) - 식생활(음식 이름, 재료), 주방기구·주방용품을 가리키는 어휘, 음식의 맛이나 생태를 가리키는 어휘, 음식을 만드는 과정이나 먹는 행위와 관련지을 수 있는 어휘
- ㅂ. 이순애(2006) - 음식 이름에 관련된 어휘, 식재료에 해당하는 어휘, 미각이나 감각에 관련된 어휘, 조리법

서의 전문적인 분류 체계를 제시함으로써 실제 응용 시스템에서의 검색 기능을 강화하는 데 도움이 될 수 있다. 본 연구에서는 조리학자들의 조리어 분류 체계를 중심으로 하되, 어휘의미망, 어휘 분류 체계를 보완적으로 활용함으로써 의미 부류 체계를 구성하기로 한다.

이와 같이 조리어 분류 체계를 중심으로 의미 주석 체계를 설정하는 이유는 전문 영역의 분류 정보를 최대한 상세하고 체계적으로 반영하기 위한 것이다. 즉, 조리어 분류 체계를 중심으로 할 경우, 조리 용어에 한정된 주석 대상 어휘를 더욱 체계적으로 분류할 수 있으며, 분류를 상세하게 할 수 있다는 실질적인 효과가 있다.

예를 들어 ‘요리’ 분야에서 ‘백편’이라는 어휘는 단순히 ‘음식’ 또는 ‘떡’으로 분류되기보다는 좀 더 구체적으로 ‘찌는떡’으로 분류될 수 있을 것이며, 또한 주석 대상 어휘 역시 요리명이나 요리재료 등 명사뿐만 아니라 조리 동사나 조리 과정을 나타내는 정도 부사, 계량 단위를 나타내는 명사 등의 전문 영역의 다양한 어휘 분포 양상을 충분히 주석할 수 있을 것이다.

본 연구에서 주로 참조한 조리학의 부류 체계는 (4)와 같다.

- (4) ㄱ. 황혜성 외(1991) - 음식종류, 음식재료, 양념과 고명, 부엌세간과 식기, 장류와 저장가공식품, 조리의 기본 조작, 조리 기구와 썰기
- ㄴ. 윤서석(1991) - 밥류, 죽류, 미음·옹이죽류, 국수·수제비류, 만두류, 떡국류, 국류, 전골·찌개류, 나물·생채류, 구이류, 조림·지짐이류, 볶음·초(炒)류, 누르미·누름적·전류, 선·찜류, 강회·무침·수란·회류, 마른반찬류, 순대·죽편·편육류, 찜류, 김치류, 장아찌류, 것갈류, 떡류, 한과류, 엿류, 음청류, 술류, 장류, 식초류, 묵류
- ㄷ. 윤경숙 편저(1996)<sup>12)</sup> - 식생활과 관련된 말, 부록(동·식물의 다른 이름, 식품의 계량, 육류의 부위별 명칭과 조리법, 절식(節食)과 시식(時食), 통과와 레음식과 의례를 상징하는 의미, 외국의 조리용어)

12) 윤경숙 편저(1996)은 어휘 분류를 목적으로 하지 않으므로 사전의 대부분을 이루고 있는 ‘식생활에 관련된 말’에 해당하는 부분이 평면적으로(가나다순) 배열되어 있다. 다만 부록 부분의 일부는 세분화되어 있으므로 이를 참고하였다.

조리학에서는 조리 어휘 자체를 분류하기 위해서 어휘 분류 체계를 세운 바는 없으므로 이들은 대부분 조리학 분야에서 대체로 통용되는 분류 체계에 해당한다.

황혜성 외(1991)는 한국의 전통 음식을 개괄적으로 기술한 것으로 음식 종류 아래 ‘주식’ 그 아래에 ‘밥, 죽, 국수, 만두’가 하위분류되어 있다<sup>13)</sup>. 따라서 떡이름 ‘백편’은 ‘음식종류 - 병과류 - 떡 - 찌는떡’ 아래에 ‘꽃편, 숙편, 녹두편’ 등과 같은 층위에 함께 놓이게 된다. 또한 윤서석(1991)은 조선시대 이후 1980년대까지 출간된 문헌 자료를 통해 음식용어만 모아 각 용어에 대한 정의와 조리법을 해설한 사전으로 음식 종류에 따른 분류 어휘집이라 할 수 있다. 이 사전에서는 한국음식의 구분법을 기준으로 정리하여 어휘를 총 29가지로 분류하고 그 내용은 가나다순으로 제시하고 있다(윤서석 1991: 5~6).

이들 분류는 전문 영역의 특성을 적절하게 반영하고, 어휘 분류의 깊이(depth)가 깊으며 비교적 체계적으로 분류되어 있다는 장점이 있다. 본 연구에서는 황혜성 외(1991)와 윤서석(1991)의 체계를 상호 비교함으로써 의미 주석 태그셋트를 개발하고, 추후 『음식디미방』 5000여 어절에 대해 시험적인 의미 주석 과정을 수행함으로써 최종적인 주석 체계를 완성하고자 한다.

의미 주석을 위한 조리 용어 분류 체계와 의미 주석 태그셋트에 대해서는 3.2.에서 상세하게 논의될 것이다.

### 3. 의미 주석의 실제

이 장에서는 조선시대 음식조리서 말뭉치의 의미 주석 체계와 관련한 실질적인 작업들을 위한 3가지 단계 즉 예비 작업 단계, 태그셋트 개발 단계, 의미 주석 단계를 제시할 것이다.

13) 황혜성 외(1991)는 한국의 전통음식에 대한 개론서이므로 ‘한국의 풍토와 산물, 한국 음식 문화의 형성’과 같은 내용도 수록되어 있으나 본 연구에서 불필요하다고 생각되는 부분은 위의 어휘 부류 체계에서 제시하지 않았다.

### 3.1. 의미 주석을 위한 예비 작업

의미 주석을 위한 예비 작업은 형태주석말뭉치를 구축하는 데서부터 시작한다. 즉, 정확한 의미 주석을 위해서는, 형태 주석 단계에서부터 어휘 및 형태소의 경계를 분명하게 설정할 필요가 있으며 이는 유의미한 어휘 단위의 설정과 관련된다. 다음 (5)~(7)의 예를 살펴보자.

(5) ‘접사+명사/동사’류

- ㄱ. 것모밀를 써어 호 모이 몰너디 말고...(겉메밀을 씻어 너무 많이 말리지 말고..., 『음식디미방』 1a)
- ㄴ. 대추 빅즈를 짓두드려 무쳐 쓰라.(대추, 잣을 짓두드려 문혀 쓰라, 『음식디미방』 3b)

(6) ‘명사+명사’류

- ㄱ. 모밀밭 닛 되에 물 부른 녹두 혼 복자식...(메밀밭 다섯 되에 물에 불린 녹두를 한 복자씩..., 『음식디미방』 1a)
- ㄴ. 싱치 무른 술홀 츠혀 지렁기름의 붓가...(팽고기의 연한 살을 다져 기름간장에 볶아..., 『음식디미방』 1a)
- ㄷ. 밀로도 굴를 정히 상화 상화ㄹ러 ㄹ치 지허...(밀가루를 곱게 상화가루처럼 쥘어..., 『음식디미방』 1a)

(7) ‘명사/부사/어근+하/되/스럼/답’류

- ㄱ. 녹도 거피흔 밭 조히 시어 건져...(녹두 알갱이의 껍질을 벗겨 깨끗이 씻은 다음 건져..., 『음식디미방』 1a)
- ㄴ. 모밀ㄹ러 장만히기를 마치 조흔 면ㄹ러...(메밀가루 장만하기를 마치 깨끗한 면가루..., 『음식디미방』 1a)

(5)~(7)의 예는 ‘접사+명사/동사’, ‘명사+명사’, ‘명사/부사/어근+하/되/스럼/답’으로 구성된 복합어의 예로, 일반적인 형태주석말뭉치 구축에서 엄밀한 형태 분석 즉, 분석적 분석을 할 것이냐, 통합적인 어휘 분석을 할 것이냐에 대한 논쟁의 대상이 되어 온 것들이다. 즉, 이들은 <표1>과 같이 분석적 분

석과 통합적 분석이 가능하다.

예	분석적 분석	통합적 분석
ㄱ. 것모밀롤	것/XPN+모밀/NNG+롤/JKO	것모밀/NNG+롤/JKO
ㄴ. 좃두드려	좃/XPV+두드리/VV+어/EC	좃두드리/VV+어/EC
ㄷ. 모밀찰	모밀/NNG+찰/NNG	메밀찰/NNG
ㄹ. 지령기름의	지령/NNG+기름/NNG+의/JKB	지령기름/NNG+의/JKB
ㅁ. 상화ㄹ라ㄹ치	상화/NNG+ㄹ라/NNG+ㄹ치/JKB	상화ㄹ라/NNG+ㄹ치/JKB
ㅂ. 다치ㅎ면	다치/MAG+ㅎ/XSV+면/EC	다치ㅎ/VV+면/EC
ㅅ. 장만ㅎ기롤	장만/NNG+ㅎ/XSV+기/ETN+롤/JKO	장만ㅎ/VV+기/ETN+롤/JKO
ㅇ. 거피ㅎ	거피/NNG+ㅎ/XSV+ㄴ/ETM	거피ㅎ/VV+ㄴ/ETM

<표 1. 형태소 분석의 예>

본 연구에서는 형태소 분석 시 통합적 분석의 방법을 취할 것인데, ‘것모밀/NNG, 좃두드리/VV, 메밀찰/NNG’ 등을 하나의 단위로 분석함으로써 어휘의미적으로 더 유의미한 어휘 단위를 추출할 수 있다

본 연구에서는 형태소 분석 시 통합적 분석의 방법을 취할 것인데, ‘것모밀/NNG, 메밀찰/NNG, 지령기름/NNG’의 체언류와 ‘좃두드리/VV, 다치ㅎ/VV, 장만ㅎ/VV’의 용언류 등의 복합어를 하나의 단위로 분석함으로써 검색 단위로 더 유의미한 어휘 단위를 추출할 수 있다.

예를 들어, 위 (ㄴ) ‘좃두드려’, (ㅇ) ‘거피ㅎ’의 경우, 통합적 분석을 함으로써 다양한 조리 방법과 관련된 동사류를 추출할 수 있으며, ‘좃-’, ‘두드리-’, ‘거피-’, ‘-ㅎ-’ 각각의 개별 단위로서는 확보할 수 없는 어휘 의미적 단위에 대한 빈도 정보를 쉽게 얻을 수 있다.

통합적 분석은 체언류에 더욱 적절하다. 예로, (ㄹ) ‘지령기름’의 경우 형태소 분석 단계에서 ‘지령’과 ‘기름’으로 분석하지 않는다. 만약 분석적 분석을 할 경우, 의미 주석도 필연적으로 ‘지령’은 장류로, ‘기름’은 양념으로 각각에

의미 주석이 되어, ‘기름을 친 양념장’이라는 ‘지령기름’의 본래 의미와는 무관한 의미 주석이 된다는 문제가 있기 때문이다. 실제로 검색 시스템 사용자의 관점에서 볼 때, ‘지령’ 또는 ‘기름’보다는 ‘지령기름’이라는 양념의 용법이 나 용례가 더욱 의미 있는 어휘가 될 것은 자명한 사실이다<sup>14)</sup>.

이와 같은 맥락에서 사전에 등재되지 않은 어휘의 경우에도 통합적 방법으로 분석한다. 예로 ‘상화ㄹ라’의 경우에도 ‘상화’와 ‘ㄹ라’로 분석하지 않고 ‘상화ㄹ라/NNG’로 형태 주석하도록 한다.

또한, 본 연구에서 의미 주석의 대상은 1차적으로 내용어 중에서 명사, 동사, 형용사, 부사에 한한다는 대원칙을 제시할 필요가 있다. 즉 형태 주석과는 달리 의미 주석에서는 모든 어휘를 대상으로 하지 않는데 그 이유는 대명사, 수사, 지정사, 감탄사, 조사, 어미 등은 음식조리서에서 유의미한 의미 부류에 속하지 않기 때문이다. <그림 2>에서 제시한 『음식디미방』 1a의 ‘면병뉴(76어절)’ 중에서 ‘면’을 형태 주석한 것의 일부인 8어절만을 보이면 <그림 4>와 같다. 의미 주석의 대상 어휘는 진하게 표시해 놓은 것으로 명사와 동사에 해당한다.

```

<pa>
  <s tid="1">
    <tok>
      <orth>것모밀롤</orth>
      <mor pos="NNG">것모밀</mor>
      <mor pos="JKO">롤</mor>
    </tok>
    <tok>
      <orth>씨어</orth>
      <mor pos="VV">씨</mor>
      <mor pos="EC">어</mor>
  </s>
  
```

14) 한 익명의 심사자는 복합어의 통합적 분석을 지향하는 본고의 관점에 대해, 형태 연구의 차원에서 분석적 분석의 필요성을 언급하였다. 필자 역시 복합어의 형태론적 특성을 고려하여 1차적인 형태소 분석의 결과를 그대로 반영하는 것에 의미가 있다는 심사자의 의견에 전적으로 동의한다. 단, 연구 초기에 이러한 형태 분석적 분석과 통합적 분석을 동시에 가능하게 하는 다차원적인 분석 체계를 제안되었음에도 불구하고, 현 상태의 주석 프로그램과 체계로는 다소 어려움이 있었다는 점을 여기서 밝혀 둔다. 형태 주석의 다차원적 분석에 대해서는 연구 팀 내에서 지속적으로 가능성을 모색 중이다.

```

</tok>
<tok>
  <orth>하</orth>
  <mor pos="MAG">하</mor>
</tok>
<tok>
  <orth>ㅁ이</orth>
  <mor pos="MAG">ㅁ이</mor>
</tok>
<tok>
  <orth>몰뇌디</orth>
  <mor pos="VV">몰뇌</mor>
  <mor pos="EC">디</mor>
</tok>
<tok>
  <orth>말고</orth>
  <mor pos="VX">말</mor>
  <mor pos="EC">고</mor>
</tok>
<tok>
  <orth>알마초</orth>
  <mor pos="MAG">알마초</mor>
</tok>
<tok>
  <orth>몰뢰여</orth>
  <mor pos="VV">몰뢰</mor>
  <mor pos="EC">여</mor>
</tok>
</s>
    
```

<그림 4. 『음식디미방』 1a ‘면병뉴’ 내 ‘면’의 형태 주석 결과 예시>

이와 같은 작업 과정에서 조리 용어로서 무의미한 어휘들은 음식디미방 5,000여 어절을 분석한 결과에 따라 별도로 불용어(stop word)<sup>15)</sup> 목록을 선정하여 목록화하고 10만 어절을 대상으로 의미 주석 작업을 하는 중에 계속

15) 실제로 어떤 문헌 컬렉션에서 80% 이상의 문헌에서 공통으로 출현하는 단어는 검색에 쓸모가 없는데 이러한 단어를 불용어(stop word)라고 하며, 보통 미리 걸러내어 배제시킨다. 따라서 관사, 전치사, 접속사 등이 불용어 리스트의 자연스러운 후보가 된다. 불용어의 제거는 다른 중요한 이점도 있는데 이는 색인 구조의 크기를 상당히 줄여준다는 것이다(임수연 2004: 32).

해서 보충해 나가기로 한다. 예를 들어 <그림 4>에서 추출할 수 있는 불용어는 조사 ‘를’, 연결어미 ‘-어, -디, -고’, 보조용언 ‘-말’, 부사 ‘하, ㅁ이, 알마초’ 등으로, 이들은 음식조리만을 위한 특별한 어휘라고 보기 어렵다.

### 3.2. 의미 주석을 위한 태그세트(tagsets) 개발

2장에서 살펴본 바와 같이 의미 주석을 위한 태그세트 개발을 위해서 본 연구에서는 국어학적 어휘 분류 체계, 조리학적 분류 체계 참조하였다. 이는 기존의 상향식(bottom-up) 분류 체계의 단점을 보완하기 위해 하향식(top-down) 전문어 분류 체계를 도입함으로써, 분류 체계의 전문성과 실용성을 보장하기 위한 방법이다.

실제로 사전 뜻을풀이를 이용한 상향식(bottom-up) 의미 체계는 ‘조리학’이라는 전문 용어 검색을 목적으로 하는 본 연구의 특성과 관련하여 다음과 같은 문제가 발견되었다. 다음 (8)에서 밑줄 친 부분은 『음식디미방』의 명사 중 일부를 『표준』의 뜻을풀이를 이용하여 상향식(bottom-up)으로 상위어를 추출할 때 예상되는 상위어 주석 유형들이다.

- (8) 모밀ㄹᄇᆞᆫ: 메밀의 열매를 찢어서 낸 가루.  
 풀: 쌀이나 밀가루 따위의 진분질에서 빼낸 끈끈한 물질.  
 의이죽: 울무쌀을 불려 갈아서 쑨 죽.  
 만도쏘: 만두 속에 넣는 재료.  
 무: 십자화과의 한해살이풀 또는 두해살이풀의 뿌리.  
 지렁기름: 기름을 친 양념장.  
 심치: 썩은 살코기.  
 술: 사람이나 동물의 뼈를 써서 몸을 이루는 부드러운 부분.  
 빅즈: 잣나무의 열매.  
 호초: 후추나무의 열매.

즉 『음식디미방』의 ‘만두법’ 중에서 일부 명사인 ‘모밀ㄹᄇᆞᆫ, 풀, 의이죽, 만도쏘, 무, 지렁기름, 심치, 술, 빅즈, 호초’만을 『표준』의 뜻을풀이를 따를 경우,



상위어로 ‘가루, 물질, 죽, 재료, 뿌리, 양념장, 살코기, 부분, 열매’가 추출된다. 이를 그대로 주석 체계에 적용할 경우, ‘가루, 물질, 뿌리, 부분, 열매’ 등은 전문 영역 어휘를 주 대상으로 하는 본 연구의 말뭉치 주석에서는 무의미한 의미 정보가 될 수 있다<sup>16)</sup>. 즉 사용자들은 ‘잣’이나 ‘후추’를 열매로 보기보다는 ‘고명’ 또는 ‘음식 재료’, ‘양념’으로 분류할 것이며, 따라서 사용자의 입장에서 친숙한 의미 부류는 사전의 뜻풀이와는 달리 음식 재료, 음식 이름, 조리 방법 등이 될 가능성이 높다.

이러한 문제를 해결하기 위해 본 연구에서는 황혜성 외(1991)와 윤서석(1991)의 조리서와 조리용어 사전을 종합적으로 고려하여 조리학이라는 전문 영역의 입장을 중심으로 태그세트를 설정하였다. 이러한 하향식 의미 부류 체계는 추후 실제 의미 주석 과정에서 더욱 체계적으로 수정될 가능성도 있는데, 이때 말뭉치 어휘 용례를 중심으로 한 상향식 방법이 적용될 것이다. 또한 향후 연구의 경과에 따라 기구축된 태그세트와 의미 주석 결과는 조리학 전문가에게 의뢰하여 감수를 받도록 할 예정이다.

1차로 설정한 태그세트는 음식 이름, 음식 재료, 양념과 고명, 부엌 세간과 조리 도구, 저장 식품, 계량 단위, 조리 방법, 조리 정도, 보관 방법의 9가지로 구성되며 깊이(depth)는 3으로 한다. 태그세트는 각 단계별로 2자리 숫자로 구성하여 최대 6자리 숫자가 된다. 그 일부를 보이면 (9)와 같다.

- (9) 음식이름 01
  - 주식 0101
    - 밥류 010101, 죽류 010102, 응이·미음류 010103, 국수류 010104, 수제비류 010105, 만두류 010106, 떡국류 010107 ...
  - 찬품 0102
    - 국류 010201, 탕류 010202, 전골류 010203, 찌개류 010204, 나물·생채류 010205, 구이류 010206, 조림·지짐이류 010207 ...

16) 이에 대해 익명의 심사자는 ‘가루, 물질, 뿌리, 열매’ 등과 같은 뜻풀이 중차 개념 또한 의미 주석에서 중요하게 다루어야 하는 문제라고 지적하였다. 필자 역시 문맥적 의미뿐만 아니라 단어가 가진 본래 고유의 의미를 동시에 고려하여 의미 태그세트를 마련하는 것이 학술적으로뿐만 아니라 활용에 있어서도 도움이 된다는 심사자의 의견에 동의한다. 이에 후속 연구와 작업에서는 이러한 점을 보완하여 태그세트를 정제할 예정이다.

떡 0103  
 찌는떡 010301, 치는떡 010302, 빛는떡010303, 지지는떡010304 ...

### 3.3. 의미 주석 작업의 실제

여기에서 논의할 의미 주석 작업은 3.2.의 태그세트 설정 이후의 작업으로 의미 주석의 대상과 태그세트는 3.1.과 3.2.의 논의에 따른다. 여기에서는 실제로 의미 주석 시 사용될 대원칙 및 실제 의미 주석의 실패를 제시하고, 주석 작업 시 발생한 몇 가지 문제에 대해서 언급하고자 한다.

우선, 의미 주석 시 정확한 의미 주석을 위해 마련한 의미 주석 지침의 대원칙은 다음과 같다.

- (10) 가. 형태분석 말뭉치의 원문(원어절과 분석 결과 모두)은 가급적 수정하지 않는다. 단, 형태분석 말뭉치에서 발견되는 명백한 오류에 대해서는 이를 별도로 기록한 후 형태분석 말뭉치의 개선에 반영될 수 있도록 한다.
  - 나. 분석 대상 품사는 원칙적 명사, 동사, 형용사, 부사만으로 한다. 다만 이들 4개의 품사에 해당하는 어휘라도 의미 주석이 필요하지 않은 어휘는 불용어 목록을 참고하여 주석하지 않는다.
  - 다. 의미 주석은 조리 용어 의미 부류 체계의 의미 부류 코드를 사용한다.
  - 르. 『표준』에 등재되지 않은 어휘가 나타날 경우에는 <조리용어사전>과 ‘현대어역’을 고려하여 주석한다<sup>17)</sup>.

17) 예를 들어 ‘이의죽’은 『표준』에 등재된 단어가 아니며 이 어휘의 의미 주석을 위해서는 <우리말조리어사전>을 참고하였다. 즉 ‘응이’라는 표제어 아래 ① ‘이의’의 발음이 순한 대로 불려서 된 것. 옛날에는 울무를 갈아 가루를 내어 두었다가 꿀물을 타서 먹었다. ② 갈분이나 울무, 수수 등을 곱게 갈아 고운체에 쳐서 냉수에 풀어 끓여 걸쭉하게 만들어 꿀을 타서 먹는 음식’라는 뜻풀이가 있어 ‘이의죽’은 울무죽으로 현대어역하고, 그 의미 주석은 아래와 같이 된다.

```
<tok>
                                <orth>이의죽</orth>
                                <w sem="010103">이의죽</w>
</tok>
```

이에 근거하여 의미 주석한 결과를 일부 보이면 <그림 5>와 같다<sup>18)</sup>. 의미 주석의 대상이 아닌 어휘에는 기본(default) 값으로 '000000'을 할당한다.

```

<sa>
  <s tid="1">
    <tok>
      <orth>갯모밀롤</orth>
      <w sem="020107">갯모밀</w>
      <w sem="000000">롤</wr>
    </tok>
    <tok>
      <orth>쌌어</orth>
      <w sem="0702">쌌</w>
      <w sem="000000">어</w>
    </tok>
    <tok>
      <orth>하</orth>
      <w sem="000000">하</w>
    </tok>
    <tok>
      <orth>막이</orth>
      <w sem="000000">막이</wr>
    </tok>
    <tok>
      <orth>몰너디</orth>
      <w sem="0703">몰너</w>
      <w sem="000000">디</w>
    </tok>
    <tok>
      <orth>말고</orth>
      <w sem="000000">말</w>

```

18) 논의의 편의를 위해 <그림 4>와 <그림 5>에서는 형태 주석과 의미 주석의 예시를 따로 제시하고 있으나, 실제로 말뭉치는 원문·현대어역 병렬 말뭉치와, 주석 말뭉치 모두가 한 개의 파일로 구성되며, 필요에 따라 개별 말뭉치를 별도로 추출될 수 있도록 설계되어 있다. 또한, 본 연구의 주제인 형태·의미 주석은 원문에 한해서만 주석되는데, 현대어역 말뭉치와 병행하게 주석하지 않은 이유는 기술적인 문제보다는 시간과 비용상의 문제에 기인함을 여기서 밝혀 둔다.

```

      <w sem="000000">고</w>
    </tok>
    <tok>
      <orth>알마초</orth>
      <w sem="000000">알마초</wr>
    </tok>
    <tok>
      <orth>몰뢰여</orth>
      <w sem="0703">몰뢰</w>
      <w sem="000000">여</wr>
    </tok>
  </s>

```

<그림 5. 『음식디미방』 1a ‘면병뉴’의 의미 주석 결과 예시>

<그림 5>는 형태 주석의 예를 보인 <그림 4>와 똑같은 부분을 의미 주석한 것이다. 본 연구에서 의미 주석은 명사, 형용사, 동사, 부사만을 대상으로 하고, 이들 내용어 중에서도 불용어를 제외한 어휘만을 의미 주석하므로 <그림 5>와 같은 의미 주석 결과를 얻게 된다.

단, 다음과 같은 문제는 향후 논의가 더 필요한 부분으로 판단된다.

우선, 대부분 반자동 작업에 의해 수행되는 의미 주석의 핵심은 주석자가 문맥 의미를 판단하는 것인데, 문맥에 따라 어휘 의미가 정확하지 않은 경우가 있다. 다음 예의 경우, 해당 어휘가 사전에 등재되어 있지 않아 주해자가 국어사적 지식을 활용하여 어휘 의미를 파악한 것으로 어휘 의미가 정확하지 않은 경우이다.

(11) ㄱ. ㄱ장 늑게 흥여 관질그르식 안치와(아주 늑늑하게 하여 두레박 모양의 질그릇에 가라 앉혀라, 『음식디미방』 1b)

ㄴ. ㄱ장 여문 밀흘 보리 느무드시 지허(가장 잘 여문 밀을 보리 대끼듯이 쪼여, 『음식디미방』 2a)

위의 (11)의 예들은 백두현(2006)에서 (11ㄱ)의 ‘관질그르식’에서 ‘관’이 향아리뿐 아니라 두레박의 의미도 갖고 있으므로 두레박 모양의 질그릇으로 불

수 있을 것 같다고 하였다. 또 (11ㄴ) ‘느무드시’에 대해서는 ‘느물-(느무-)’이라는 어간이 존재하였던 듯하다고 하였으며 이를 근거로 ‘애벌 찜다’라고 해석하고 있다. 이와 같은 경우에는 주해자가 국어사적 지식을 가진 전문가이기 때문에 이를 존중하여 주해자의 뜻풀이에 기대어 의미 주석을 하기로 한다. 즉, 주해자의 현대어역을 따를 경우, (11ㄱ)은 조리 도구로, (11ㄴ)은 조리 방법으로 주석될 것이다.

다음으로, 문맥에 따라 보통 때와는 다른 의미로 사용되는 어휘 또는 한 개 어휘가 여러 가지로 분류되는 경우에 대한 문제가 있다.

- (12) ㄱ. 쌀 흰 말을 빅세혀여 ㄱ장 닉게 써(쌀 한 말을 깨끗이 씻어 잘 익게 써서, 『음식디미방』 21b)
- ㄴ. 녹도 거피흔 쌀 조히 시어 건져 물 썩거든 메밀쌀(녹두 알갱이의 껍질을 벗겨 깨끗이 씻은 다음 건져 내라. 물이 빠지거든 메밀 알갱이, 『음식디미방』 1a)
- ㄷ. 식면 디어 쓸 체면 밀굴룰 ㄱ장 조히 혀여(실국수를 만들려고 쓸 때는 밀 가루를 가장 깨끗이 하여, 『음식디미방』 1b)
- ㄹ. 교터는 식면 교터 ㄱ치 하라(고명은 녹말국수의 고명과 같게 하라, 『음식디미방』 1a)

(12)에서 본 바와 같이 (12ㄱ)의 ‘쌀’은 ‘쌀(米)’의 뜻이지만 (12ㄴ)에서의 ‘쌀’은 문맥상 ‘쌀(米)’이 아니라 ‘알갱이’라는 뜻이다. 또한 (12ㄷ)의 ‘식면’은 ‘실국수’의 뜻이나, (12ㄹ)의 경우 문맥을 고려하여 조리법의 내용을 살펴보면 ‘녹말국수’라는 뜻이다. 이와 같은 문제는 다의어와 동형어 처리와 마찬가지로 반자동적 의미 주석 이후에 수작업으로 말뭉치를 수정하는 후처리 단계에서 문맥을 확인하여 정확하게 주석하기로 한다.

마지막으로 일부 어절 또는 구 구성의 경우 때에 따라 무의미한 주석이 될 여지가 있다.

- (13) ㄱ. 주머니에 닛 내듯 쳐 내여(주머니에 잇꽃 물을 내듯 쳐 내어, 『음식디미방』 1b)
- ㄴ. 촌물에 쎄여들가 희거든 효근 약과낫ㄱ치 사흐라 쓰는니라(찬물에 담가

- 희어지면 작은 약과낫만하게 썰어 쓴다, 『음식디미방』 1b)
- ㄷ. 개곰낫마곰 쎄에 비즈라(개암알 크기만큼씩 떼어 빻어라, 『음식디미방』 1a)
- ㄹ. 비곗 성히 필 때에 빅미 닛 말(배꽃이 한창 필 때 백미 다섯 말, 『음식디미방』 17a)

(13ㄱ)에서 ‘주머니에 닛 내듯 쳐 내여’에서 ‘주머니, 닛, 내다, 치다’는 음식을 조리하기 위해서 ‘쳐내’는 행위를 ‘주머니에 잇꽃 물을 내듯 치’하는 것으로 어절을 하나하나 분석하는 것보다는 구 전체가 ‘조리 방법’으로 의미 주석되는 것이 더 유의미하다고 볼 수 있다. 또한 (13ㄴ)은 ‘명사+부사’의 구성 (13ㄷ)은 ‘명사+의존명사’ 구성인데, 이들 또한 각각 분석하지 않고 하나의 단위로 보는 것이 좋다. (13ㄷ)의 ‘개곰낫만큼’에서 ‘개곰’은 현대어의 ‘개암’에 해당하는데 문맥에서의 의미를 무시하고 ‘개곰’만을 의미 주석한다면 ‘개곰’은 ‘요리 재료’로 주석된다. 하지만 문맥에서의 의미를 고려하여 ‘개곰낫마곰’을 한 개의 단위로 주석하게 되면 이는 ‘계량 단위’로 주석될 것이다. 마찬가지로 (13ㄹ)에서도 밀줄 친 부분 전체가 한 개의 단위로서 요리를 하는 적당한 때를 의미하는 것이므로 의미 주석의 한 단위가 되는 것이 더 좋다<sup>19)</sup>.

이러한 구성은 ‘-듯, ㄱ치, 마곰, 쎄’ 등의 형태를 가진 구성을 한꺼번에 추출하여 이들을 하나의 단위로 의미 주석할 수 있기 때문에 추후에 수정 작업 시 일괄적으로 처리할 수 있을 것이다. 이를 위해 전체 말뭉치를 대상으로 주석 대상 패턴을 목록화할 필요가 있다.

#### 4. 마무리

지금까지 조선시대 필사본 음식조리서의 검색 시스템 개발 과정의 한 하위 작업으로 조리용어 의미주석말뭉치 구축 과정에 대해 논의하였다. 2장에

19) 구어 말뭉치 구축 시에도 비슷한 문제가 있었듯이, 이와 같은 구 단위는 형태 주석 단위와 의미 주석 단위가 달라지는 것, 어디까지를 유의미한 구 단위로 볼 것인가 하는 구 단위 설정의 문제 등 여러 가지 문제점이 있어 실제 주석에서 구를 하나의 단위로 보고 주석하는 데는 어려움이 있다. 이에 추후 연구를 통해 구 단위 이상의 의미 단위에 대한 의미 주석의 방법을 모색하고자 한다.

서는 의미주석말뭉치를 구축하기 위한 기초적인 논의로, 필사본 음식조리서의 텍스트 구조와 특성에 대해 논의하였고, 이러한 텍스트의 특성을 고려한 의미 부류 체계 설정의 필요성을 논하였다. 3장에서는 의미주석말뭉치의 구축 과정을 소개하였는데, 예비 작업 단계, 태그셋 개발 단계, 의미 주석 단계의 단계로 작업 내용을 제시하였다.

본 연구는 조선시대 필사본 조리서를 말뭉치로 구축하고 형태·의미 주석을 가하여 검색 시스템을 개발함으로써 이를 필요로 하는 조리학자 및 일반인의 국어사 자료에 대한 접근성을 높이기 위한 연구이다. 본 연구 결과물은 검색시스템의 의미 부류 검색 기능을 추가하는 데 유용하게 활용될 것이다. 또한 이를 활용하여 조리 용어에 대한 어휘분류집을 만든다면, 다음과 같이 표제 어휘에 대한 의미 부류 정보를 제시하는 데 유용한 자료로 활용될 것이다.

word	POS	Sem_class	Freq	Index	KWIC <sup>20)</sup>
표제어	품사	의미 부류	빈도 <sup>21)</sup>	위치	용례(괄호 안은 현대어역)
모밀ㄹ	NNG	음식 재료_곡류	1	음식디미방 1a	...모밀ㄹ 장만ㅎ기를 마치... (메밀가루 장만하기를 마치)
의이죽	NNG	음식명_주식_죽	1	음식디미방 1a	...풀 뿌디 의이죽ㄹ치 뿌어... (풀을 쑤디 울무죽같이 쑤어서)
뿌	VV	조리 방법_가열하기_쑤기	1	음식디미방 1a	...의이죽ㄹ치 뿌어 그 푸리... (울무죽같이 쑤어서 그 풀을)
쑤	VV	조리 방법_갈기·다지기	6	음식디미방 1a	...낫 업시 쑤사 싱치 무른... (덩어리 없이 다지고 썰고기의 연한)

<표 2. 조선시대 음식조리서의 어휘 검색 결과 예시>

20) 본 연구에서는 지면 관계상 용례를 표제어를 포함하여 3~4어절 정도만 보여주고 있지만, 실제로는 사용자가 용례를 확인할 때는 확인하고자 하는 표제어를 포함한 앞뒤 어절 수를 직접 결정할 수 있도록 설계된다.

21) 여기에서 보인 빈도는 『음식디미방』 5000어절을 대상으로 했을 경우이며, 최종적으로 표시될 빈도는 본 연구에서 대상으로 하는 조선시대 음식조리서 10만 어절의 분석이 끝난 후 자동적으로 추출될 것이다.

이러한 정보를 활용할 경우, 사용자는 표제어나 품사를 검색하는 것은 물론이고 의미 부류를 통해 필요한 내용을 검색할 수도 있다. 즉 사용자가 조선시대 음식재료로 사용된 것 중에서 어떤 곡류가 사용되었는지, 그러한 곡류는 어떤 이름으로 불리었는지가 궁금하다면 본 연구에서 주석한 의미 부류에 따라 필요한 내용만을 불러와 볼 수 있게 된다.

불용어의 유형 및 범위 설정, 조리학적으로 유의미한 다중 어휘 단위(multi lexical unit)의 의미 주석 문제 등은 향후 본 연구에서 해결되어야 할 문제이다. 또한 의미 주석을 위한 의미 부류 체계의 깊이 및 범위 조정 역시 작업 과정에서 구체화되어야 할 것이며, 사용자 친화적(user-friendliness)인 프로그램 개발을 위해서 조리학 전공자와 국어학 전공자의 협업이 필요하다.

### 참고문헌

강범모(2003), 언어, 컴퓨터, 코퍼스 언어학, 고려대학교 출판부.  
 김광혜(1993), 국어 어휘론 개설, 집문당.  
 김한샘(1999), “현대국어 관용구의 계량언어학적 연구”, 연세대학교 국어정보학 협동과정 석사학위 논문.  
 문유진(1996), “의미론적 어휘개념에 기반한 한국어 명사 WordNet의 설계와 구축”, 서울대학교 컴퓨터공학과 박사학위 논문.  
 백두현(2006), 음식디미방 주해, 글누림.  
 서상규(2001), “말뭉치의 주석과 한국어 기본 어휘 의미 빈도 사전”, 계량언어학 1, 박이정.  
 서상규(2001), “어휘사 연구와 국어정보화”, 한국어학 14권, 한국어학회, 49-69.  
 이기황(2007), “XML을 이용한 주석 말뭉치의 구조화와 활용”, 제1회 한국 언어·문학·문화 국제 학술 대회 발표 논문집, 연세대학교 국어국문학과 BK21 한국 언어·문학·문화 국제인력양성 사업단, 148-156.  
 이숙의(2006), “한국어 동사 의미망 구축 연구”, 충남대학교 국어국문학과 박사학위 논문.  
 이순애(2006), “한국어 교재의 식생활 어휘 연구”, 가톨릭 대학교 국어국문학 석사학위 논문.  
 임수연(2004), “전문용어의 의미관계 정보를 이용한 도메인 온톨로지의 구축”, 경북대학교 컴퓨터공학과 박사학위 논문.

임지룡(1991), “국어의 기초 어휘에 대한 연구”, 국어교육연구 23, 국어교육학회, 87-131.  
 임지룡(1992), 국어의미론, 탑출판사.  
 차재은 · 강범모(2002), “다의 설정의 방법에 대하여”, 한국어학 15권, 한국어학회, 259-284.  
 채상용(2006), “온톨로지 기반 지능형 규칙 구성 요소 추출에 관한 연구”, 연세대학교 컴퓨터과학 · 산업시스템공학과 석사학위 논문.  
 최호섭 외(2006), “대규모 우리말 어휘지능망 구축 방법”, 한글 273, 한글학회, 125-151.  
 최호섭(2007), “대규모 사용자 어휘지능망 구축과 활용”, 울산대학교 정보통신공학 전공 박사학위 논문.  
 한복려 외(1999), 다시 보고 배우는 음식디미방, 궁중음식연구원.  
 한유석 · 설근수(2004), 한국어 시소리스 연구, 한국문화사.  
 황혜성 외(1991), 한국의 전통 음식, 교문사.  
 Cruse, D.A. (1986), Lexical Semantics, Cambridge Univ. press.(임지룡 · 윤희수 역 (1989), 어휘의미론, 경북대학교 출판부.)  
 Doo-hyun Paek, Kil-im Nam, Mi-hyang Lee, Eui-jeong Ahn, Hyeon-ju Song. (2008), Development of the Integrated Concordancer for the Corpus of the 17th to 19th Century Culinary Manuscripts, Proceedings of the 13th EURALEX 2008.  
 Fellbaum, C. (1998), WordNet, The MIT Press.  
 Jackendoff, R. (1995), Semantic Structures, The MIT Press.  
 Murphy M. Lynne (2003), Semantic Relations and the Lexicon, Cambridge Univ. press.(임지룡 · 윤희수 역(2007), 의미관계와 어휘사전, 박이정.)  
 Nirenburg, S. & Raskin, V. (2004), Ontological Semantics, The MIT Press.  
 Piet van Sterkenburg et al.(2003), A Practical Guide to Lexicography, John Benjamins Publishing Company.  
 Saint-Dizier, P. & Viegas, E.(1995), Computational Lexical Semantics, Cambridge University Press.

남길임  
 대구시 북구 산격동 1370  
 경북대학교 인문대학 국어국문과  
 702-701  
 전화번호: 053-950-5111  
 전자우편: nki@mail.knu.ac.kr

송현주  
 대구시 북구 산격동 1370  
 경북대학교 사범대학 국어교육과  
 702-701  
 전화번호: 053-950-5827  
 전자우편: songhj@mail.knu.ac.kr

원고 접수일 : 2008년 7월 10일  
 게재 확정일 : 2008년 8월 13일

<사전류>

국립국어원 편(2001), 표준국어대사전, 두산동아.  
 남영신(1987), 우리말 분류사전, 한강문화사.  
 박용수 엮음(1988), 우리말 갈래 사전, 한길사.  
 연세대학교 언어정보개발연구원 편(1998), 연세 한국어 사전, 두산동아.  
 윤서석(1991), 한국의 음식 용어, 민음사.  
 윤숙경 편저(1996), 우리말 조리어 사전, 신광출판사.