

조선시대 필사본 음식조리서의 의미 주석 연구*

남길임 · 송현주 · 최준 · 현영희

(경북대학교)

<Abstract>

Nam Kilim, Song Hyunju, Choi Jun & Hyun Younghee. 2012. Semantic Annotations of Culinary Terms for Korean Culinary Manuscripts of Joseon Period. *Korean Semantics*, 39. This study presents a follow up study of the previously reported research project Compiling and Developing a Corpus of 17-19th Century Korean Culinary Manuscripts and a Customized Browser. The previous project, which was based on the morpheme based annotation, was not sufficient for extracting culinary terms since many terms are in the form of so-called 'extended lexical units(ELUs).' The main purpose of this study is to prove that the semantic unit of specific terms should be treated on phrase level. By extracting ELUs and morpheme units of culinary terms, firstly, richer information on the expressions could be obtained. Secondly, more accurate annotation of culinary terms has been achieved according as the context of each terms was taken into consideration. Lastly, the products attained from this study can be applied to compile domain-specific dictionaries and contribute to extend lemma status to multi-word items.

핵심어: 의미 주석(semantic annotation), 형태 단위 주석(morpheme units tagging), 확장된 어휘 단위 주석(extended lexical units tagging), 조리 용어(culinary terms), 조선시대 음식조리서 (Joseon dynasty culinary manuscripts)

* 이 논문은 2011-2012년 한국연구재단의 지원을 받아 수행된 연구임(NRF-2011-334-A00005).

1. 들머리

본 연구는 ‘조선시대 필사본 음식조리서 검색 시스템’의 의미 주석 결과를 대상으로 하여 의미 주석 방법과 의의를 논의하고, 형태 단위 의미 주석과 확장된 어휘 단위(extended lexical units:이하 ELU) 의미 주석 각각의 특성과 변별성을 제시하는 것을 목적으로 한다.

경북대학교 연구팀에서는 2007년~2009년과 2011년~2012년, 2회에 걸쳐 조선시대 필사본 음식조리서를 대상으로 하여 조리용어 통합 검색 시스템을 개발하고 용어 색인 DB를 구축하였다. 첫 번째 과제에서 구축한 ‘조선시대 음식조리서의 통합 검색 시스템’은 언어학자와 조리학자는 물론이고 일반인들에게 한국 전통 음식의 종류와 조리법에 대한 구체적인 정보를 제공하고, 역사 자료를 활용한 말뭉치 기반 사전편찬학의 새로운 방법론을 제시하는 데 기여하였다. 두 번째 과제에서는 조리 용어를 한국연구재단의 표준 지침에 맞는 색인 DB로 재구축하고, 사용자별로 이원화된 색인 DB를 구축함으로써, 사용자 집단별 서비스를 용이하게 하고 사용자 친화적 정보를 제공하는 데 초점을 맞추었다. 본 연구는 이러한 연구 결과를 기반으로 하여, 조선시대 음식조리서라는 특정 시기의 특정 장르에 적합한 의미 주석의 방법 개발의 의의와 주석 단위에 따른 의미 주석 결과를 비교하고, ELU를 고려한 주석이 필요함을 논의하고자 한다.

조선시대 필사본 음식조리서를 대상으로 한 말뭉치 구축 과정은 ‘원시 말뭉치 구축 → 원문-주해문 병렬 말뭉치 구축 → 형태 주석 말뭉치 구축 (→ 현대어 매핑 말뭉치 구축) → 의미 주석 말뭉치 구축’으로 단순화할 수 있다. 이 중 의미 주석은 의미 부류별 검색이 가능하도록 하여, 국어학자는 물론이고 국어학적 지식이 없는 조리학자나 일반인이 조선시대 음식 이름이나 조리 방법, 조리 용어 등에 쉽게 접근하는 데 기여한다. 그런데 지금까지의 의미 주석에 관한 연구는 대체로 범용적 목적으로 실시된 것이 대부분이었고, 무엇보다도 형태 단위를 넘어선 확장된 어휘 단위(ELU)에 관한 주석이 실제로 시도된 예는 거의 없는 형편이다.

이에 본 연구는 조선시대 필사본 음식조리서라는 전문 분야의 텍스트에 적합한 의미 주석 태그 세트를 개발하고, 이를 형태 단위뿐만 아니라 확장된

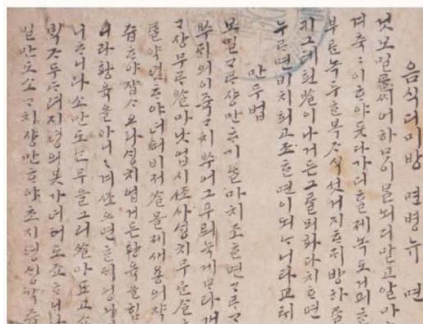
어휘 단위를 대상으로 의미 주석을 시도하였으며, 그 둘의 결과를 비교하여 살펴보았다는 점에서 기존 의미 주석에 대한 논의와 차별성을 가진다. 이러한 논의를 통해 의미 주석의 단위를 무엇으로 하느냐에 따라 의미 정보의 양과 질이 현저하게 달라진다는 것을 실증적으로 증명할 수 있을 것이다.

이 연구는 다음과 같이 구성된다. 2장에서는 조리 용어 통합 검색 시스템과 DB 구축, 실제 의미 주석 말뭉치 구축 과정에 대해 소개한다. 3장에서는 형태 단위에 대한 의미 방법과 주석 결과와 특징을 살펴보고, 4장에서는 확장된 어휘 단위에 대한 주석 방법과 주석 결과를 살펴, 형태를 넘어선 확장된 어휘 단위에 대한 주석이 갖는 의의를 찾아보고자 한다. 5장에서는 이상의 내용을 요약하고, 남은 문제를 제시한다.

2. 의미 주석을 위한 기초적 논의

2.1. 조선시대 필사본 음식조리서의 특성

본 연구에서 대상으로 삼은 텍스트는 17세기-20세기 초기에 작성된 조선시대 한글 필사본 음식조리서로서 총 28종이다. 근대국어의 모습을 보여주는 음식조리서인 이 자료들은 훈련된 연구자가 아니면 알아보기 힘든 초서체로 된 것이 대부분이며, 분량은 총 1,320면으로 110,693어절에 달한다.



<그림 1> 『음식디미방』 ‘면병뉴’의 일부

<그림 1>은 17세기에 나온 『음식디미방』의 ‘면병뉴’의 일부인데, ‘면’과 ‘만

두법'이 소개되어 있다.

조선시대 음식조리서는 현대의 조리서와는 정보를 전달하는 방식에서 차이가 있다. 현대 조리서의 경우 저자가 특정 독자를 염두에 두고 작성하는 것이 아니므로 조리와 관련한 정보가 비교적 객관적인데 반해, 조선시대 조리서는 대개 어머니나 시어머니가 딸이나 며느리에게 전하는 내용으로 되어 있으므로 서로 공유하고 있는 맥락 정보를 바탕으로 텍스트가 구성된다¹⁾. 따라서 이미 잘 알고 있는 조리 방법으로 다른 음식을 만드는 방법을 설명하기도 하고, 세상사의 지식을 나타내는 일상용어를 사용하여 전문적인 계량 단위나 조리 시기를 나타내기도 한다. 예를 들어, 음식 재료의 양이나 조리 시기 등을 표현하기 위해서 ‘되, 말, 정월’ 등 절대적인 양과 시기를 나타내는 말을 쓰기도 하지만, ‘주먹만큼, 상화 찌듯이, 배꽃 필 때’와 같은 표현을 사용하기도 하는 것이다. 이러한 특성을 고려할 때 조선시대 조리서를 대상으로 하는 의미 주석은 일상용어의 전문용어적 쓰임, ELU로서의 구 구성을 고려하여 수행될 필요가 있다.

이상적이고 궁극적인 사전(the ultimate dictionary)이 단어 단위가 아닌 ‘의미 단위(semantic unit)’에 대한 총체적 기술에 있다는(Sinclair et al. 2004: xxiv) 주장은 일반 언어 사전뿐만 아니라 전문용어 사전에도 매우 중요한 점을 시사한다.²⁾ 음식조리서는 수용자에게 정확한 정보를 전달하여 어떤 행위를 수행하도록 하는 호호테스트의 일종으로(Brinker 1992), 수용자는 음식조리서의 의미 단위의 정보를 얻기 위해 텍스트를 검색하게 된다. 그런데 이때 수용자가 원하는 전문 영역의 의미 단위는 단어 층위뿐만 아니라 구 층위 이상인 경우가 많다. 본 연구의 3, 4장에서는 형태 층위에서의 의미 주석과 구 단위 이상의 확장된 어휘 단위의 층위를 대상으로 한 의미 주석의 과정을 차례로 제시할 것이다.

2.2. 의미 주석 태그 세트 개발

- 1) 이러한 점은 『음식디미방』의 저자인 장씨 부인이 딸에게 남긴 다음 서문에서도 잘 드러난다. ‘이 칩을 이리 눈 어두운디 간신히 써시니 이 쓰줄 아라 이제로 시행하고 쫓조식들은 각각 벗겨 가오디 이 칩 가더 갈 생각올안 성심 말며 부디 상치 말게 간쇼호야 수이 찌러 버리다 말라.’
- 2) ELU에 대한 연구 동향과 말뭉치 주도 연구(corpus-driven approach)를 통한 ‘의미 단위’의 사전 기술에 대한 상세한 논의는 남길임(2011) 참조.

국어학이나 사전학계에서 개발된 의미 부류 체계는 주로 범용적·언어학적 목적을 위해 구축되었으므로, ‘조선시대 필사본 음식조리서’라는 특정 전문 분야의 어휘의 의미 주석에 활용하기에 적절하지 않다. 따라서 본 연구팀에서는 지난 음식조리서 통합 검색 시스템 개발 당시, 세종 전자사전과 국어 어휘 부류 사전을 참고하고, 조리학자의 자문을 얻어 전통음식 조리 용어를 주석하는 데 최적화된 의미 주석 태그 세트를 마련하였다.³⁾

(1)은 의미 주석을 위해 마련한 태그 세트의 층위별 대표 의미 부류를 일부 제시한 것이다.⁴⁾

형태 단위 의미 주석의 대상은 명사(의존명사 포함), 동사, 형용사, 부사 등의 내용어에 한하며, 의미 주석은 최대 깊이(depth)를 3 단계로 하며, 태그 세트는 제1층위 12개, 제2층위 49개, 제 3층위 87개로 구성된다. 이 가운데 ‘음식명, 조리방법’과 같이 하위 부류가 다양한 것은 제3층위까지, ‘음식재료, 양념고명, 부엌세간식기, 계량단위, 조리정도시기, 저장’과 같이 비교적 하위 부류가 단순한 것은 제2층위까지, ‘맛, 향미, 불’처럼 더 이상 구체화할 필요가 없는 경우는 제1층위에서 주석할 수 있도록 태그 세트를 마련하였다.⁵⁾

- (1) a. 제1층위(senseClass1): 음식명, 음식재료, 양념고명, 부엌세간식기, 계량단위, 조리방법, 조리정도시기, 저장, 맛, 향미, 불, 기타
- b. 제2층위(senseClass2): 주석, 친품, 떡, 한과, 음청류, 용량단위, 중량단위, 비가열 조리, 가열조리 등

3) 의미체계를 구축하는 방식으로는 크게 직접 구축과 참조 구축이 있다. 새롭게 독립적인 의미 체계를 갖추려는 직접 구축 방식은 하향식(top-down) 또는 상향식(bottom up)으로 이루어질 수 있다. 하향식은 고도의 배경지식을 가진 전문가가 특정 분야 온톨로지나 소규모의 개념망을 수동으로 개발할 때 사용된다. 상향식은 하위어에서 출발하여 더 포괄적인 의미를 가진 상위어를 찾아가는 방식이다. 주로 대규모 어휘의미망을 구축할 때 사전의 정의문을 분석하여 상위어를 반자동으로 추출하는 방법을 사용한다(윤애선 외 2009: 98 참조). 이 연구에서는 조리학이라는 특정 전문 분야의 텍스트라는 점과 근대국어로서 사전에 등재되지 않은 어휘가 적지 않은 점 때문에 하향식 직접 구축 방식을 사용하였다.

4) 본 연구팀에서 개발한 의미 주석 태그 세트 전체는 이 논문의 마지막에(부록 1) 실어 두었다.

5) 이은영(2009: 159-160)은 ACE(Automatic Content Extraction) 프로그램의 태그 세트를 적용한 후, 세계적으로 표준화되어 가는 추세인 의미 주석 도구의 활용도는 언어의 차이보다는 텍스트의 영역의 문제에 더욱 민감해 보인다고 한 바 있다. 따라서 의미 주석을 위해서는 범용적 목적의 의미 주석 태그 세트를 고려하되, 본 연구에서와 같이 대상 텍스트 장르를 충분히 고려하여 의미 주석 태그를 개발하여 사용하는 것이 더 유용함이 분명하다.

- c. 제3층위(senseClass3): 밥, 죽미음응이, 국수, 국, 탕, 찌개, 찻기, 까기, 굽어내기, 담그기, 삶기, 데치기, 끓이기, 띄우기 등

의미 주석을 위한 태그 세트 중에서 ‘음식명’에 관한 부분만을 제시하면 아래의 <표 1>과 같다.

<표 1> ‘의미 주석 태그 세트’의 일부

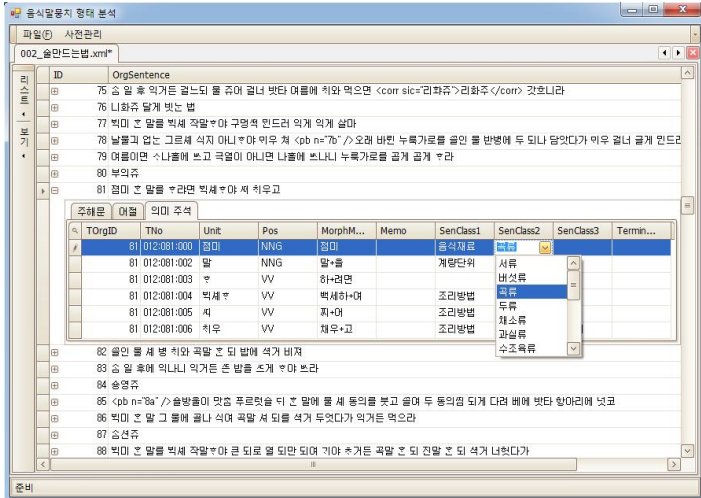
senseClass1	senseClass2	senseClass3
음식명	주식	밥, 죽미음응이, 국수, 만두, 떡국, 수제비, 주식 기타
	찬품	국, 탕, 찌개, 전골, 생채, 숙채, 찜선, 전구이, 조림, 볶음, 회, 김치, 장아찌, 젓갈, 튀각부각, 마른찬, 편육, 족편, 묵, 찬품 기타
	떡	찌는떡, 치는떡, 삶는떡, 지지는떡, 떡 기타
	한과	유밀과류, 유과류, 정과류, 다식류, 숙실과류, 과편류, 엿강정류, 한과 기타
	음청류	차, 화채, 음청 기타
	술	발효주, 증류주, 혼성주, 술 기타

본 연구 대상은 조리서라는 텍스트의 성격상 구체적인 음식의 이름을 나타내는 ‘음식명’과 음식을 만드는 방법을 나타내는 ‘조리방법’에 관한 태그가 가장 풍부하다.

3. 형태 단위 의미 주석

3.1. 형태 단위 주석의 실제

형태 단위에 대한 의미 주석은 아래 <그림 2>의 Food Corpus Tagger(이하 FCT)라는 도구를 자체 개발하여 수행하였다.



<그림 2> 의미 주석 도구(Food Corpus Tagger)

의미 주석 도구는 다음과 같은 특징을 갖는다. 첫째, 형태 정보뿐만 아니라 문맥상의 의미를 확인하는 데 용이하다. 즉, FCT는 선행 작업에서 구축된 원어-대역어 말뭉치와 원어 형태소 분석 및 대응 대역어 매핑 정보를 통합하여 확인할 수 있도록 하였다. 필요에 따라 <그림 2>에서 제시하는 바와 같이 원문을 기준으로 해당 주해문 정보의 확인과 의미 주석 작업이 동시에 가능한 창을 선택할 수도 있고, 별도로 작업의 효율성을 위해 의미 주석의 후보 형태인 내용어만을 제시하는 창을 선택할 수 있다.

둘째, 수동 주석 작업의 결과물을 기 구축 사전으로 활용할 수 있다. 대상 말뭉치가 조선시대 음식조리서이기 때문에 주석의 시작은 수동으로 진행할 수밖에 없다. 하지만 일정 규모의 의미 주석 작업이 수행된 이후에는 그 결과물을 기 구축 사전 형태로 FCT 의미 주석 사전에 입력하여 반자동으로 작업을 진행할 수 있게 하였다.⁶⁾

6) 이때 ‘디초/디초/디후/디조’ 등과 같이 동일한 유형에 해당하는 개별 형태의 음식적 변이형이 다양하기 때문에 원어 정보를 기준으로 기 구축 사전을 만들면 작업의 효율성이 낮아질 수 있다. 그에 따라 사전 입력의 기준 형태를 원어 형태로 삼지 않고 원어 형태에 1:1로 대응되는 현대어 형태 정보를 활용하여 사전 구축의 효율성을 높였다.

셋째, 의미 주석 작업의 결과는 선행 작업에서 구축한 데이터베이스에 자동으로 통합된다. 형태 단위 의미 주석 작업자들이 실제로 작업에 활용하는 FCT는 사용자에게 친숙한 인터페이스 형태로 제공된다. 이때 작업의 기반이 되는 선행 작업의 통합 결과물은 xml 구조로 설계되어 FCT에 연결되어 있기 때문에 의미 주석 정보는 <그림 3>과 같이 기존의 xml 구조에 자동으로 하위 의미 주석 세트 형태로 입력된다.

```
<sa>↓
<unit sid="80" tokseq="0" senseClass1="음식명" senseClass2="술" senseClass3="발효주">부의주</unit>↓
<unit sid="81" tokseq="0" senseClass1="음식재료" senseClass2="곡류">좁미</unit>↓
<unit sid="81" tokseq="2" senseClass1="계량단위" senseClass2="용량단위">말</unit>↓
<unit sid="81" tokseq="3">ㅎ</unit>↓
<unit sid="81" tokseq="4" senseClass1="조리방법" senseClass2="비가열조리" senseClass3="씻기">씻게</unit>↓
<unit sid="81" tokseq="5" senseClass1="조리방법" senseClass2="비가열조리" senseClass3="찜기">찌</unit>↓
<unit sid="81" tokseq="6" senseClass1="조리방법" senseClass2="비가열조리" senseClass3="비가열조리기타">취우</unit>↓
<unit sid="82" tokseq="0" senseClass1="조리방법" senseClass2="가열조리" senseClass3="끓이기">끓이</unit>↓
<unit sid="82" tokseq="1" senseClass1="음식재료" senseClass2="물">물</unit>↓
<unit sid="82" tokseq="3" senseClass1="계량단위" senseClass2="용량단위">컵</unit>↓
<unit sid="82" tokseq="4" senseClass1="조리방법" senseClass2="비가열조리" senseClass3="비가열조리기타">취우</unit>↓
<unit sid="82" tokseq="5" senseClass1="음식재료" senseClass2="발효제">폭말</unit>↓
<unit sid="82" tokseq="7" senseClass1="계량단위" senseClass2="용량단위">되</unit>↓
<unit sid="82" tokseq="8" senseClass1="음식명" senseClass2="주식" senseClass3="밥">밥</unit>↓
<unit sid="82" tokseq="9" senseClass1="조리방법" senseClass2="비가열조리" senseClass3="섞기">섞</unit>↓
<unit sid="82" tokseq="10" senseClass1="조리방법" senseClass2="조리방법기타" senseClass3="담그기">빚</unit>↓
<unit sid="83" tokseq="1" senseClass1="저장" senseClass2="저장기간">일</unit>↓
<unit sid="83" tokseq="2">후</unit>↓
<unit sid="83" tokseq="3" senseClass1="조리점도시기" senseClass2="조리점도">익</unit>↓
<unit sid="83" tokseq="4" senseClass1="조리점도시기" senseClass2="조리점도">익</unit>↓
<unit sid="83" tokseq="5" senseClass1="조리점도시기" senseClass2="조리점도">뜨</unit>↓
<unit sid="83" tokseq="6" senseClass1="음식명" senseClass2="주식" senseClass3="밥">밥</unit>↓
<unit sid="83" tokseq="7" senseClass1="조리방법" senseClass2="비가열조리" senseClass3="비가열조리기타">뜨</unit>↓
<unit sid="83" tokseq="9">뜨</unit>↓
</sa>↓
```

<그림 3> xml 구조에 포함된 의미 주석 결과

xml 구조로 입력된 원문, 대역문, 원어 형태, 현대어 형태 등의 정보와 함께 특정 식별자(예: 문장 및 어절 아이디)를 매개로 유기적으로 통합된 의미 주석 정보는 데이터베이스 사용자의 다양한 요구 사항에 따라 정보를 선택하여 접근할 수 있다. 의미 주석 정보는 국어학자는 물론이고, 조리 전문가나 요리에 관심이 있는 일반인들에 의해 활용될 수 있다. 이들 중 근대 국어의 형태에 익숙하지 않은 사용자들에게는 현대어 형태 단위를 기준으로 정보를 검색할 수 있게 할 필요가 있는데, 이때 의미 주석 결과는 음식과 관련하여 유용한 정보를 제공할 수 있다.

아래의 <표 2>은 ‘순향주’에 대한 의미 주석 결과를 바탕으로 ‘순향주’의

조리 정보를 구성한 것이다.

<표 2> ‘순향주’의 의미 주석 결과를 통해 추출한 조리 정보

senseClass1	senseClass2	내용
부엌세간식기	조리도구	술독, 관독, 독, 그릇, 체, 술뚜껑, 종이, 유지, 밥보자기, 단지, 청솔개비, 술가지, 쥘, 널빤지, 고자틀
음식재료	물	물, 탕수
	곡류	백미, 밀가루, 찹쌀, 멥쌀
	발효제	시루떡, 누룩, 누룩가루, 전술
	기타	청주
조리방법		씻기→찌기→식히기→우리기→찌기→섞기→씻기→가루내기→찌기→섞기→끓이기→체치기→섞기→씻기→찌기→갈기→끓이기...

<표 2>는 조선시대에 ‘순향주’를 만들기 위해 사용된 ‘부엌세간, 음식 재료, 조리 방법’에 대한 구체적인 항목을 제시해 준다. 이는 전통 조리학 연구자나 전통 요리에 관심을 갖고 있는 일반인이 조선시대의 요리의 이름이나 식재료, 요리법 등에 대한 정보를 검색하고자 할 때, 유용한 자료가 될 수 있다.

3.2. 형태 단위 의미 주석의 결과와 특징

여기에서는 형태 단위에 의미 주석된 결과를 품사별, 형태별 계량적 결과를 살펴보고, 그 특징을 논하기로 한다. 이 연구를 위해 구축한 조선시대 필사본 음식조리서 말뭉치는 110,693어절 규모이며, 주석 대상이 된 형태 수는 모두 183,758개이다. 이중 54,248개가 의미 주석의 대상이 되는데, 전체의 30%가량이다. 의미 주석의 대상이 되는 형태는 내용어이고 하나의 어절에 하나의 의미 주석이 달린다는 것을 고려하면, 전체 어절의 50%가량이 의미 주석 대상이 되는 것이다.

의미 주석이 된 형태의 품사별 통계와 각 품사에서 고빈도 형태로 나타나는 품사는 아래 <표 3>과 같다. 주로 음식재료, 음식명, 계량단위와 관련된 것으로 명사가 57.1%로 가장 높은 비율을 나타내며, 조리방법과 관련되는 동사가 35.4%, 주로 조리정도와 관련되는 부사와 형용사가 각각 4.1%, 3.3%로 그 뒤를 잇고 있다.

<표 3> 의미 주석 대상 형태의 품사별 통계 및 고빈도 형태⁷⁾

	명사(30,996)	동사(19,199)	부사(2,236)	형용사(1,817)
1	물(1,329)	섞다(849)	잠깐(290)	가늘다(199)
2	되(906)	치다(747)	많이(180)	얇다(172)
3	말(815)	찌다(723)	조금(171)	달다(105)
4	기름(507)	썰다(631)	정히(117)	잘다(88)
5	꿀(446)	끓이다(617)	죄(75)	곱다(68)

명사 중에서는 음식재료에 해당하는 ‘물, 기름, 꿀’이 각각 1, 4, 5위로 고빈도로 나타나며, 주로 곡류의 계량 단위로 사용되는 ‘되, 말’ 등의 단위성 의존명사도 높은 빈도로 나타난다. 동사의 경우 ‘섞다, 치다, 찌다, 썰다, 끓이다’ 등과 같이 구체적인 조리 행위를 나타내는 형태들이 고빈도를 보이고 있다. 이들 상위 5개의 동사들은 모두 ‘조리방법’ 하위의 제3층위 의미 주석의 명칭과 일치한다는 특성이 있다. ‘잠깐, 많이, 조금, 정히, 죄’ 등의 고빈도 부사들은 모두 ‘조리정도시기’의 하위 의미 주석 중 ‘조리정도’에 해당한다. 형용사 가운데 고빈도로 나타나는 ‘가늘다, 얇다, 잘다, 곱다’ 등은 부사와 마찬가지로 (2a)와 같이 부사형 어미 ‘-게’와 주로 결합해 부사어로서 조리 행위의 정도를 한정하며, ‘달다’는 (2b)에서처럼 음식이나 음식 재료의 맛을 나타내는 데 주로 사용된다.

7) 괄호 안의 숫자는 해당 품사 및 형태의 빈도를 나타낸다.

- (2) a. 입춘날 무우 가날게 치 쳐 싹고 무순 미느리 숯무우음 신검초 너허 빅비탕을
 끌혀(『시의전서』 ‘상갓침치’ 5b)
- b. 스오일 만의 너면 꿀 긋치 달거니와(『김충지맥주방문』 ‘감향주법’ 9b)

아래 <표 4>는 형태 단위를 기준으로 한 의미 주석의 계량적 결과이다.⁸⁾

<표 4> 형태 단위에 대한 의미 주석의 의미 부류별 분포

의미 주석	형태 빈도		유형수	빈도/유형수
조리방법	18,714	(34.5%)	764	24.5
음식재료	14,043	(25.9%)	1,773	7.9
조리정도시기	5,984	(11.0%)	963	6.2
음식명 ⁹⁾	4,791	(8.8%)	1,636	2.9
양념고명	3,294	(6.1%)	239	13.8
부엌세간식기	3,221	(5.9%)	407	6.5
계량단위	3,129	(5.8%)	138	22.7
맛	459	(0.8)	83	5.5
저장	413	(0.8)	107	3.9
불	144	(0.3)	14	10.3
향미	56	(0.1)	26	2.2
	계: 54,248	(100%)	계: 6,242	평균: 8.7

의미 주석 결과는 빈도를 기준으로 하면 ‘조리방법>음식재료>조리정도시기>음식명>양념고명>부엌세간식기>계량단위’의 순서로 나타난다. 이는 음식 조리 과정을 기술하는 음식조리서라는 텍스트의 특성이 반영된 결과로 보인다.

8) <표 4>에 제시한 의미 주석의 결과는 제1층위(senseClass1)만을 나타내며, 유형수에서는 모음 및 자음의 음소적 변이형, 연철 및 분철에 의한 변이형을 별개로 보지 않고 하나의 유형으로 통합한 것이다.

9) ‘음식명’의 빈도는 음식조리서에서 기술의 대상이 되는 음식항목(FoodItem)의 수를 나타내는 것이 아니라, 조리서 본문에 포함된 ‘음식명’의 빈도이다. 본 연구팀이 구축한 말뭉치에는 총 1,808개의 음식항목이 포함되어 있다.

빈도 대비 유형수의 비율은 ‘조리방법’과 ‘계량단위’가 20 이상의 수치를 나타내어 다른 의미 주석에 비해 상대적으로 높은 비율을 보임을 알 수 있다. 조리 과정에서의 동작을 나타내는 ‘조리방법’의 경우 한정된 형태를 반복적으로 사용할 개연성이 높다. ‘조리방법’은 전체 의미 주석 대상의 34%에 달하는 높은 형태 빈도를 보이지만, ‘조리방법’에 해당하는 형태는 (3)의 예와 같이 대체로 조리과정을 나타내는 한정된 동사 및 서술성 명사들이다.

- (3) a. 황^ㄴ날^ㄴ로 김체의 친 빅청을 다식 반^ㄴ독^ㄴ도^ㄴ른 즐게 혀야 젓갈^ㄴ니 계피 건강^ㄴㄴ
로 석겨(『규합총서 제일 하편』 ‘밥주악’ 29a)
b. 빅^ㄴ미 오 승 점미 일 두을 빅세^ㄴ혀^ㄴ여 들갓^ㄴ다가 각각 지^ㄴ에^ㄴ를 찌디(『보감록』
‘송절주’ 4b)
c. 츠쌀 서 되 빅세 작말 물 흔 눗동^ㄴ히^ㄴ로 죽 쑤어 디렁(『주방문』 ‘청명주’ 11b)

‘조리방법’을 나타내는 764개의 유형 가운데 6%에 해당하는 ‘섞다, 치다, 찌다, 썰다, 끓이다, 삶다, 담그다, 지지다, 볶다, 묻히다’ 등의 상위 48개 유형의 총 빈도는 13,135이다. 이는 ‘조리방법’ 형태의 전체 빈도의 70%에 해당하는 것으로, 특정한 동사 및 서술성 명사가 음식 조리의 특정 영역에 밀접한 관련이 있음을 알 수 있다.

또한 (4)에서 보듯이, 음식재료의 양을 제시할 때 사용되는 ‘계량단위’의 경우 대체로 음식재료의 상위 의미군과 밀접한 관련을 갖는 단위성 의존명사가 제한적으로 사용될 개연성이 높기 때문에 그 비율이 높게 나타난다.

- (4) a. 진말 한 되 가^ㄴ루^ㄴ누^ㄴ룩 한 되 범우^ㄴ려 무^ㄴ슨 항의 너허 찬 디 두되(『술 빛는 법』
‘백일주방문’ 1a)
b. 뜯맛 들거^ㄴ든 빅미 두 말 점미 두 말 빅세^ㄴ혀^ㄴ여(『술 빛는 법』 ‘백일주방문’ 1a)
c. 누룩^ㄴ가^ㄴ로 가^ㄴ늘^ㄴ게 혀여 닷 흠과 진말 닷 흠을 죽에 너허(『시의전서』 ‘두강주’
20a)

‘계량단위’에 해당하는 개별 형태 가운데 빈도 기준 상위 1~5위에 해당하는 것은 주로 곡류의 용량 단위로 사용되는 단위성 의존 명사 ‘되, 말, 흠, 두’ 그리고 길이의 단위로 사용되는 ‘치’ 등으로 이들의 총 빈도는 2,225이다. 이

는 ‘계량단위’ 전체의 71%를 망라하는 수치로 ‘조리방법’에서와 마찬가지로 몇몇 제한된 형태가 계량 단위를 나타내는 데 반복적으로 사용되고 있음을 보여 준다.

4. 확장된 어휘 단위 주석

4.1. 확장된 어휘 단위 주석의 실제

앞에서 언급한 바와 같이 의미 주석의 대상을 형태소나 단어로 한정할 경우, 구 단위로만 추출될 수 있는 양질의 정보가 누락될 수밖에 없다. 따라서 전문 텍스트에서 의미 단위로 기능하는 확장된 어휘 단위(ELU)를 고려한 추가적인 의미 주석이 필요하다.

확장된 어휘 단위의 주석 결과 추가로 주석된 ELU는 5,154개로 54,248개의 형태 단위에 대한 의미 주석의 1/10에 해당한다. 아래의 <그림 3>은 확장된 어휘 단위에 대한 의미 주석의 결과를 일부 제시한 것이다.¹⁰⁾

morpheme_ID	ELU	senseClass1	senseClass2	senseClass3
100846-100848	<@>발인 맛</@>	맛		
065807-065809	<@>상흔 맛</@>	맛	맛	
000926-000928	<@>근는 채</@>	부엌세간식기	조리도구	
020348-020350	<@>굴근 채</@>	부엌세간식기	조리도구	
020441-020442	<@>복성화 나무가지</@>	부엌세간식기	조리도구	
065108-065112	<@>삼 년 묵은 거문참</@>	양념고명	양념	
014289-014291	<@>팍미 닷 되</@>	음식재료	곡류	
044755-044767	<@>훈쌀 혼 말</@>	음식재료	곡류	
019878-019880	<@>점미 혼 되</@>	음식재료	곡류	
073858-073859	<@>거란 흰주</@>	음식재료	알류	
076698-076700	<@>더든 받</@>	저장	저장장소	
135118-135122	<@>어지 안닛난 곳</@>	저장	저장장소	
069192-069194	<@>죽 뿌어</@>	조리방법	가열조리	끓이기
054869-054871	<@>졸 우어</@>	조리방법	가열조리	끓이기
147092-047094	<@>찌 바르고</@>	조리방법	비가열조리	끓여내기
161493-161496	<@>새줄 임시하고</@>	조리방법	비가열조리	비가열조리기타
141616-141618	<@>피 썬고</@>	조리방법	비가열조리	비가열조리기타
079325-079326	<@>스월 상수알</@>	조리절도시기	조리시기	
158030-158037	<@>스월의 작약찾치 성히 말 썬</@>	조리절도시기	조리시기	
019223-019224	<@>스월 초취알</@>	조리절도시기	조리시기	
148307-148309	<@>덕로 모양갓치</@>	조리절도시기	조리절도	
064875-064877	<@>씩 안치듯</@>	조리절도시기	조리절도	

<그림 3> 확장된 어휘 단위의 의미 주석 결과

10) <@>은 ELU의 시작점을 </@>는 ELU의 끝점을 나타내기 위해 주석자가 임의로 붙인 것이며, 형태 단위 주석과는 달리 ELU에 대한 주석을 위해서는 별도의 프로그램을 개발하지 않고, 엑셀(Excel)을 활용하여 의미 주석 작업을 수행하였다.

<그림3>에서 알 수 있듯이 형태 단위인 ‘맛, 체, 거문장, 되, 방’ 등은 조리 방법과 관련하여 정보를 제공할 수 있지만 그 구체성이 부족하다. 반면, ‘빌인 맛, ㄴ는 체, 삼 년 묵은 거문장, 푹미 닳 되, 더은 방, 죽 뿌어, 씨를 업시 하고, 스월 상스일’ 등은 의미 단위를 이루어 조리법과 관련한 정보를 담고 있으며, 의미 주석의 주요 대상이 된다.

ELU의 주석은 작업자가 문맥을 고려하여 수작업으로 일일이 주석해야 하므로 작업 효율성을 높이기 위해서 다음과 같이 두 단계로 나누어 진행하였다. 첫째 단계에서는 (5)에서와 같이 수식어의 수식을 받아 의미 단위를 형성하는 데 기여하는 ‘맛, 디, 째’ 등의 명사류, 상태를 나타내는 ‘ㄴ치, ㄹ고, 마곰’ 등의 조사류를 우선적으로 추출하여, 이들을 포함한 ELU를 주석하였다.

- (5) a. 맛: 맛
 b. 공간: 디, 곳
 c. 시간: 째, 제, 적, 날
 d. 단위: 되, 말, 종자, 홉
 e. 상태: ㄴ치, ㄹ고, 마곰
- (6) a. 맛: 빌인 맛, 상흔 맛
 b. 저장: 서늘흔 디, 희를 등지고 바람 당흔 곳
 c. 음식재료: 푹미 닳 되, 출발 흔 말, 접미 흔 되, 며도굴이 흔 말¹¹⁾
 d. 조리정도시기: 상화 씨기 ㄴ치, 푹곳 성히 필 째, 덩도 칩도 아닌 적

(5)의 ‘맛, 디, 째, 되, ㄴ치’ 등과 같은 형태는 맛, 장소, 시간 등을 나타내는 형태지만 단독으로는 유의미한 주석 단위가 되기 어렵다. 그러나 (6)에서와 같이 구를 이루어 쓰인 ‘빌인 맛, 서늘흔 디, 푹미 닳 되, 상화 씨기 ㄴ치’ 등의 예들은 조리법 텍스트 안에서 특별한 지위를 갖는 ‘의미 단위’ 즉 ELU가 된다. 이러한 유형은 ‘선행어(수식어) + 의존명사 상당 부류’의 패턴으로 나타나거나 ‘명사류+조사’의 형태로 나타나므로, 형태 단위가 아닌 구 단위 주석의 일차적인 대상이 된다.

11) ‘되, 말, 홉’ 등은 계량단위에 해당하는 것으로, ELU로는 ‘음식명/음식재료+양화사+계량단위’의 구성으로 나타나는 경향이 있다.

다음 단계에서는 이러한 특정 명사류로 한정할 수 없는 나머지 부류를 대상으로 유의미한 ELU 단위를 찾아 주석하였다.

- (7) a. 상화쇠 너름의 밧보면 거피흔 꽃홀 씨 얼경이로 처 청밀의 므라 녀코(『음식디미방』 ‘상화법’ 2a)
 b. 석이를 더운 물에 두어 되나 튀혀여 돌을 죄다 갈히여(『음식방문』 ‘석이버섯주약’ 11b)

‘팥’이나 ‘물’보다 좀 더 구체적인 정보를 제공할 수 있는 (7a)와 (7b)의 ‘거피흔 꽃, 더운 물’은 음식재료로서의 정보가 풍부한 ELU가 된다. 이러한 유형은 주석자가 문맥 정보를 파악하여 ELU로 적합한지 판단한 후 주석하였다.

4.2. 확장된 어휘 단위에 대한 의미 주석 결과와 특징

여기에서는 확장된 어휘 단위에 대한 의미 주석 결과와 형태 단위 주석의 결과를 계량적으로 비교하고, ELU에 대한 주석 결과에 대한 특성을 간략히 논의하기로 한다. 아래 <표 5>는 형태 단위에 대한 의미 주석 결과와 확장된 어휘 단위에 대한 의미 주석 결과를 비교한 것이다.¹²⁾

ELU에 대한 의미 주석은 ‘음식재료>조리정도시기>>저장>조리방법>부엌>세간식기>양념고명>>맛’의 순으로 나타난다. 형태 단위에 대한 의미 주석에서와 마찬가지로 ‘음식재료’와 ‘조리정도시기’ 빈도는 여전히 높은 빈도로 나타나는데 이는 조리법 텍스트라는 말뭉치의 특성에 기인한다. 반면, 형태 단위 의미 주석에서 0.8%에 불과하였던 ‘저장’의 의미를 갖는 ELU가 상당히 높은 비율로 증가하는 것을 주목할 필요가 있다. 이는 ELU를 고려하지 않는다면 ‘저장’에 관한 유의미한 정보가 상당 부분 누락될 수 있음을 의미한다.

12) 여기에서 ‘증가 비율’이라 함은 ELU로 추가 주석된 비율을 말한다. 예를 들어, ‘거피흔 꽃’은 형태 단위에 대한 의미 주석에서 ‘거피흔/조리방법, 꽃/음식재료’로 주석되었으며, 여기에 ‘거피흔 꽃’ 전체가 ELU로서 음식재료로 추가 주석된다. 따라서 ‘거피흔 꽃’은 이중으로 의미 주석이 되므로 정보량이 증가한다.

‘향미, 계량단위, 음식명, 불’ 등은 ELU로 의미 주석이 되는 경우가 거의 없는데, 이는 이들 단위가 ‘되, 흠, 만두’와 같이 하나의 어휘 단위로 존재하는 경우가 대부분이기 때문이다.

<표 5> 형태 단위 주석과 ELU 주석 결과 비교

의미 주석	형태 단위 기준		ELU 기준		증가 비율
조리방법	18,714	(34.5%)	320	(6.2%)	+1.7%
음식재료	14,043	(25.9%)	2,765	(53.6%)	+19.7%
조리정도시기	5,984	(11.0%)	1,264	(24.5%)	+21.1%
음식명	4,791	(8.8%)	0	(0)	-
양념고명	3,294	(6.1%)	205	(4.0%)	+6.2%
부엌세간식기	3,221	(5.9%)	255	(5.0%)	+7.9%
계량단위	3,129	(5.8%)	0	(0)	-
맛	459	(0.8)	6	(0.1%)	+1.3%
저장	413	(0.8)	338	(6.6%)	+81.8%
불	144	(0.3)	0	(0)	-
향미	56	(0.1)	1	(0)	+1.7%
	계: 54,248	(100%)	계: 5,154	(100%)	

<표 5>의 결과에서 확장된 의미 단위에 대한 의미 주석 결과에서 큰 증가를 보이는 ‘저장, 조리정도시기, 음식재료’에 주목할 필요가 있다. 먼저, ‘저장’은 형태 단위에 대한 주석 결과와 비교했을 때 83%나 새롭게 포착된다. 주석된 단위의 수는 338개로 전체의 6.6%에 불과하지만, ELU에 대한 고려가 특정 의미 부류를 포착하는 데 상당히 중요한 역할을 함을 잘 드러내 준다.

- (8) a. 메밥 버무린 것 흔 박아지 남겼다가 우를 덮혀 단단이 봉~~호~~야 벗 아니 드
 논 던 두엇다가(『술 만드는 법』 ‘두견주’ 3a)
- b. 그 밥 빨 제 물 두 되 쁘려 ㄹ장 더워서 미티 셋거 더운 더 빠 뒸다가 이틀
밭 디나거든 보야(『하생원택주방문』 ‘감주’ 5b)
- c. 치운 제어든 드슨 던 두고 더운 제어든 촌 던 뒸다가(『하생원택주방문』 ‘자
 주’ 4a)

(8a)의 ‘벗 아니 드는 디’, (8b)의 ‘더운 디’, (8c)의 ‘드손 디’, ‘춘 디’ 등은 ‘저장(저장장소)’에 해당하며, (8b)의 ‘이틀 밤’은 ‘저장(저장기간)’을 나타낸다. (8c)의 ‘치운 제’, ‘더운 제’는 ‘저장(저장시기)’에 해당한다. ‘저장장소’에 해당하는 ELU는 대체로 ‘관형어/절+장소 명사’의 패턴으로 나타나는데 이때 ‘장소 명사’는 ‘디, 곳’ 등 장소를 나타내는 의존명사가 대부분이다. 이들은 확장된 어휘 단위를 고려하는 관점에서만 ‘저장장소’로서 유의미한 단위로서 주석 대상이 된다.

또한 ‘저장시기, 저장기간’에 해당하는 ELU 역시 ‘관형어/절+시간 명사’의 패턴으로 나타나며 ‘저장장소’와 마찬가지로 형태 단위 주석에서는 포착되지 못하는 단위들이다. 이 연구에서 대상으로 삼은 텍스트에서 ‘저장’과 관련되는 표현의 절반가량은 형태 단위가 아닌 확장된 어휘 단위에 해당하였다.

두 번째로 높은 증가 비율을 보이고 있는 것은 ‘조리정도시기’이다. (9)는 조리시기에, (10)은 조리정도에 해당하는 것으로 ELU를 고려하였을 때, 상당히 유의미한 의미 주석 결과를 획득할 수 있다.

- (9) a. 정월 초희일에 빅미 화인 시승 혼 되를 가로 만들고(『술 만드는 법』 ‘산자 법’ 21a)
- b. 정월 첫 돛날 빅미 두 말 빅세 작말 탕슈 서 말 더렁 누룩 닷 홉 섯거 녀코(『하생원택주방문』 ‘삼해주’ 2a)
- c. 국화가 성이 필 씨 술이 혼 말이어든 쏫 두 되을 슈머이의 너어(『주식시의』 ‘꽃 향기를 넣은 술을 만드는 법’ 28a)

(9a)의 ‘정월 초희일’의 개별 형태 ‘정월’, ‘초희일’은 형태 단위 주석에서도 각각 ‘조리시기’로 주석되며, (9b)의 ‘정월 첫 돛날’의 경우 ‘정월’과 ‘돛날’은 형태 단위에서도 주석의 대상이 되지만, 관형사 ‘첫’은 주석 대상에서 제외된다. (9c)의 ‘국화가 성이 필 씨’의 경우 ‘국화, 성이, 필, 씨’ 각각은 의미 주석 대상이 되지 않지만, ‘국화가 성이 필 씨’는 ELU로서 ‘조리시기’로 의미 주석된다.

(10)은 ‘조리정도’를 나타내는 ELU이다. (10a)는 ‘명사+조사’, (12b)는 ‘관형어/절+의존명사’, (10c)는 ‘목적어+동사+부사형 어미’의 구성이다.

- (10) a. 안심슬노 풀닙갓치 골퓌죽쳐로 써셔(『시의전서』 ‘전골법’ 21b)
- b. 장국의 똥호고 표고 석이 더욱 너히 똥이 무를 만치 쓸히고(『주식방문』 ‘간막이탕’ 7b)
- c. 향의 저린 나모시와 캐캐 격지 노와 썩 안치듯 향의 골흠이 너코(『주식시의』 ‘즙장’ 30a)

(10)에서 ‘골퓌죽쳐로, 똥이 무를 만치, 썩 안치듯’은 형태 단위 주석에서는 ‘골퓌죽, 똥, 썩’은 ‘음식재료’로 ‘물, 안치’는 ‘조리방법’으로 주석된다. 그러나 실제 문맥을 고려하자면 ‘풀닙갓치, 골퓌죽쳐로, 똥이, 무를 만치, 썩 안치듯’ 전체가 ‘조리방법’에 해당하는 것으로 이는 ELU를 고려한 의미 주석에서만 정확히 포착해 낼 수 있는 새로운 정보이다. 이처럼 ELU는 형태 단위 주석에서 전혀 포착되지 않는 구성을 새롭게 포착한다는 의의도 있지만, 형태 단위에서 포착이 되는 형태라 하더라도, 더욱 상세한 정보를 제공할 수 있는 단위들을 연속 형태로 묶어서 보여 줄 수 있다는 의의도 함께 지닌다.

셋째로, 약 20%정도의 증가 비율을 보이는 ‘음식재료’ 또한 ELU 단위를 고려하였을 때 의미 있는 정보를 포착하게 해 준다. ‘음식재료’에 해당하는 ELU는 (11a) ‘음식재료/음식명+양화사+계량단위’, (11b) ‘관형어/절+음식재료명’ 구성으로 빈번하게 나타난다.

- (11) a. 찰쌀 혼 말 밥 지어 누룩가로 닷 흠을 물은 말고 전술 병 반에 고로고로 비 즈면 맛시 쑈 갓트니라(『시의전서』 ‘청감주’ 22a)
- b. 굴희 싸 운독 너코 쓸혀 시근 물의 소곰 프러 브어 똥다가(『하생원택주방문』 ‘청대콩침 만드는 법’ 26a)

(11a)의 ‘찰쌀, 누룩가로, 전술’ 등은 형태 단위 주석에서도 ‘음식재료’로 주석되지만, ELU 단위에 의미 주석을 할 경우 음식재료명과 더불어 음식재료별 계량(용량) 정보를 보다 더 상세하게 얻을 수 있다. 특히 (11b)에서 ‘물’이라는 형태 단위에 대한 주석은 ‘쓸혀 시근 물’이라는 ELU에 대한 주석에 비해 조리영역에서는 정보의 양이 현저히 부족하다.

지금까지 살펴본 ELU는 <표 6>에서와 같이 개별 구성 요소에 대한 형태

단위 주석과의 처리 방식의 차이에 따라 크게 세 가지 유형으로 나뉜다.

<표 6> 형태 단위 주석과 ELU 주석의 차이에 따른 유형

	형태 단위 주석	ELU 주석
1 유형	[복성곳]_∅ [펼]_∅ [째]_∅ [서늘흔]_∅ [디]_∅ [콩죽]_∅ [ㄹ치]_∅	[복성곳 펼 째]_<조리시 기> [서늘흔 디]_<저장장소> [콩죽 ㄹ치]_<조리정도>
2 유형	[굴근]_∅ [체]_<조리도구> [더운]_∅ [방]_<저장장소> [치]_∅ [쳐]_<썰기>	[굴근 체]_<조리도구> [더운 방]_<저장장소> [치 쳐]_<썰기>
3 유형	[정월]_<조리시기> [상허일]_<조리시기> [고기]_<수조육류> [복근]_<볶기> [즙]_<음 식재료기타> [빅미]_<곡류> [두]_∅ [되]_<용량단위>	[정월 상허일]_<조리시기> [고기 복근 즙]_<음식재료 기타> [빅미 두 되]_<곡류>

제1유형은 형태 단위에 의미 주석을 할 경우 주석 대상이 되지 않던 것이 ELU를 고려할 경우 새롭게 주석 대상이 되는 경우이다. ‘복성곳’, ‘펼’, ‘째’ 등과 같은 개별 형태소는 음식 조리 과정에서 유의미한 단위가 되지 않는다. 그러나 ‘복성곳 펼 째’를 하나의 단위로 보면 ‘조리정도시기(조리시기)’를 나타내는 유의미한 단위로 포착이 가능하다. ‘콩죽 ㄹ치’의 ‘콩죽’은 ‘음식명’에 해당하지만 문맥을 고려한 ‘콩죽 ㄹ치’는 ‘조리정도시기(조리정도)’로 새로운 의미 정보를 나타내 주는 ELU가 된다. 따라서 이 유형은 의미 주석의 기본 단위를 개별 형태소로 삼는 형태 단위 주석의 한계가 가장 잘 드러내 주며, ELU에 대한 의미 주석의 필요성을 분명하게 해 준다.

제2유형은 형태 단위에서보다 ELU를 고려할 경우 좀 더 구체적이고 상세한 의미 정보를 제공하는 경우이다. 형태 단위 주석에서는 ‘체, 방, 치-’는 선행요소인 ‘굴근’, ‘더운’, ‘쳐’ 등과 한 단위로 주석되지 않으므로, 정보의 양이 빈약하다. 반면 ELU를 고려하면 ‘굴근 체, 더운 방, 치 쳐’로 더 많은 정보를 제공할 뿐만 아니라, ‘ㄹ는 체, 촌 방, 회 치고’ 등의 ELU와 변별되는 정보를 제공할 수 있다는 점에서 의의가 있다.

제3유형은 형태 단위에 대한 의미 주석 결과와 일부 중복되는 경우이다.

형태 단위 주석에서는 ‘정월’과 ‘상희일’에 각각 ‘조리정도시기(조리시기)’의 동일한 의미 주석이 할당되어 ‘정월’과 ‘상희일’이 서로 어떤 관련을 갖는지 알기 어려운 반면, ELU에 대한 의미 주석에서는 이들을 하나의 단위로 ‘조리정도시기(조리시기)’로 주석된다. ‘고기 복근 줍’의 경우에도 형태 단위 주석에서는 ‘음식재료, 조리방법, 음식재료’ 각각으로 주석되지만, ELU를 고려하면 ‘음식재료’라는 하나의 단위로 주석된다.

아래의 (12)는 『음식디미방』의 ‘이화주법’의 원문이며, <표 7>은 형태 단위에 대한 의미 주석 결과와 ELU와 형태 단위 의미 주석을 통합하여 추출한 결과이다.

```
(12) <org class="술" name="이화주법 닷 말 비지" page="17a">
  <s id="538">이화주법 닷 말 비지</s>
  <s id="539">비긋 성히 필 때에 백미 닷 말 백세 세말혀</s>
  <s id="540">김체로 처 구무씩 민드라 빨마</s>
  <s id="541">드스흔 체 두 손으로 극히 밋이 처 덩이 업계 하고</s>
  <s id="542">누룩ㄹ러 혼 말 셋거 녀헛다가 닉거든 브라</s>
</org>
```

<표 7> 형태 단위 주석과 ELU 및 형태 단위 통합 주석 추출의 예

senseClass1	senseClass2	형태 단위 추출	[형태 단위 + ELU] 통합 추출
조리정도시기	조리시기	-	배꽃 성히 필 때
음식재료	곡류	백미	백미 다섯 말
	발효제	누룩가루	누룩가루 한 말
부엌세간식기	조리도구	김체	김체
계량단위	용량단위	말	말
조리방법		씻기→가루내기→ 체 치기→∅→삶기 →반죽하기→섞기	씻기→가루내기→체치기→ 반죽 하기 →삶기→반죽하기→섞기

<표 7>에서도 보듯이, 의미 주석 결과를 통해 ‘이화주’에 관한 조리 정보를 추출할 경우 형태 단위에서 주석한 결과만으로 추출한 결과에 비해, 형

태 단위 주석 결과에 ELU에 대한 의미 주석 결과를 통합한 자료가 더 상세한 정보를 제공한다. 즉 <표 7>을 통해 이화주의 ‘조리시기’에 대해 새로운 정보를 제공해 줄 뿐 아니라, 이화주를 만들 때 사용하는 ‘음식 재료’에 대해 구체적으로 알려 주고, ‘조리방법’에 관해 추가 정보를 제시해 줌을 확인할 수 있다.

5. 마무리

지금까지 조선시대 음식조리서를 대상으로 한 의미 주석의 방법과 실제, 형태 단위에 대한 의미 주석과 확장된 어휘 단위에 대한 의미 주석의 결과에 대해 논의하였다. 여기에서는 이상의 내용을 요약하고, 남은 문제를 제시하기로 한다.

첫째, 조선시대 필사본 음식조리서의 특성을 살펴보고, 본 연구 팀에서 개발한 의미 주석 태그 세트의 일부를 제시하였다. 연구에서 대상으로 삼은 말뭉치는 17~19세기 조선시대 필사본 음식조리서로 총 110,693어절, 183,758개 형태로 구성되며, 형태 단위에 대한 의미 주석의 대상이 되는 것은 54,248개 이고, 확장된 어휘 단위로서 주석되는 항목은 모두 5,154개였다. 이는 확장된 어휘 단위를 고려함으로써 양적으로 10% 가량의 유의미한 의미 단위를 더 확보할 수 있을 뿐만 아니라, 질적으로도 정보의 정확성과 정밀성을 높일 수 있다는 것을 의미한다.

둘째, 형태 단위에 대한 의미 주석에 대해 살펴보았다. 형태 단위에 대한 주석 작업을 위해서는 자체 개발한 Food Corpus Tagger를 활용하였으며, 품사별·의미 주석별 결과를 살펴보았다. 형태 단위에 대한 의미 주석 결과는 ‘조리방법>음식재료>>조리정도시기>음식명>양념고명>부엌세간식기>계량단위의 빈도순으로 나타났다. 이러한 형태 단위 주석 결과는 향후 전통 조리학 분야에서 유용한 자료로 활용될 수 있는데, 대표적으로 조선시대 조리서의 식재료와 양념의 변천사 전통 조리 방법의 경향성 등을 연구하는 데 기초자료로 활용될 것이다.

셋째, 확장된 어휘 단위에 대한 의미 주석의 방법과 유형, 결과를 살펴보았

다. ELU에 대한 의미 주석 결과, ‘음식재료>조리정도시기>>저장>조리방법>부엌세간식기>양념고명>>맛’의 빈도순으로 나타나는데, 이중 ‘저장, 조리정도시기, 음식재료’에서 정보량이 현격히 늘어남을 확인할 수 있었다. 특히, 형태 단위에 대한 의미 주석과 확장된 어휘 단위에 대한 의미 주석 결과를 계량적으로 비교한 결과, ELU를 고려하는 것이 형태 단위 의미 주석이 포착하지 못한 부분을 보완하고, 해당 단위에 대한 의미 정보를 보다 풍부하게 제공해 준다는 점을 확인할 수 있었다.

본 연구의 결과는 향후 의미 주석의 단위를 형태에서 확장된 어휘 단위까지 고려할 수 있는 데 기여하였다. 또한 음식조리학 용어 사전을 만드는 데 활용될 수 있으며, 단어 중심의 표제어 선정을 보완하는 데 도움을 줄 수 있을 것이다. 무엇보다도 본 연구는 전문용어 사전학에 대한 새로운 대안을 제시하였다는 점에서 의의가 있다. 다만, 형태 단위에 대한 의미 주석과 ELU에 대한 의미 주석을 위해 별개의 프로그램을 활용하였으므로, 그 결과가 자동적으로 통합되지 않기 때문에 이 점은 추후 기술적으로 보완되어야 할 부분이다.

참고문헌

- 강범모(1998). “한국학 문헌의 전산화를 위한 TEI 부호화 방안의 응용과 확장”. 학술진흥재단 연구결과보고서.
- 남길임(2011), “확장된 어휘 단위에 대한 연구 동향과 한국어의 기술”, 한국사전학 제 18호, 한국사전학회, 73-98.
- 남길임, 송현주(2008), “조선시대 필사본 음식조리서의 음식 용어 의미 주석을 위한 연구”, 한국어 의미학 26, 한국어의미학회 47-72.
- 백두현(2009), “옛 편지에 나타난 식생활과 음식문화”, 선비문화 16권, 남명학연구원, 56-67.
- 안의정, 박진양, 남길임(2008), “한글 필사본 음식조리서 말뭉치 구축을 위한 마크업 방안 연구”, 언어와정보 12권 2호, 한국언어정보학회, 95-114
- 안의정, 차경희, 남길임(2011), “한글 필사본 음식조리서의 텍스트언어학적 연구”, 어문론총 54권, 한국문학언어학회, 5-38.
- 윤애선, 황순희, 이은령, 권혁철(2009), “한국어 어휘의미망 KorLex 1.5의 구축”, 소프

- 트웨어 및 응용 36권 1호. 한국정보과학회, 92-108.
- 이은령(2009), “19세기 문헌 국역본의 개체명 인식 및 관계 추출을 위한 기초 연구”, 언어학 제53호, 한국언어학회, 141-162.
- 이흥식(2008), “음료 명사의 사전 기술에 대하여”, 한국어의미학 27, 한국어의미학회, 199-224.
- Biber, D. et al.(1999). *The Longman Grammar of Spoken and Written English*. London: Longman.
- Biber, D., Kim, Y. & Nicole, T.(2010). A Corpus-driven Approach to Comparative phraseology: Lexical Bundles in English, Spanish, and Korean. *Japanese/Korean Linguistics* 17, 75-94.
- Brinker, K.(1992). *Linguistic text analysis*. 3rd Edition. Berlin.
- Nam, K.I., Song, H.J., Choi, J. and Hyun, Y.H.(2012). “Extracting and Annotating Extended Lexical Units of Culinary Terms for Korean Culinary Manuscripts of Joseon Period”. *Euralex 2012 15th European Association for Lexicography International Congress*.
- Paek, D.H., Nam, K.I., Lee, M.H., Ahn, E.J., Song, H.J.(2008). Compiling and Developing a Corpus of 17-19th Century Korean Culinary Manuscripts and a Customized Corpus Browser, *Euralex 2008 13th European Association for Lexicography International Congress*. 741-746.
- Sinclair, J. McH.(2004). *Trust the Text; Language, Corpus and Discourse*. Routledge.
- Stubbs, M.(2009). The Search for Units of Meaning; Sinclair on Empirical Semantics, *Applied Linguistics*, 30(1), 115-137.

남길임

702-701 대구광역시 북구 대학로 80
 경북대학교 인문대학 국어국문학과
 전화번호: 053-950-5111
 전자우편: nki@knu.ac.kr

송현주

780-714 경상북도 경주시 동대로 123
 동국대학교 교양교육원
 전화번호: 010-3827-3767
 전자우편: songhj@knu.ac.kr

최준

172 남길임 · 송현주 · 최준 · 현영희

702-701 대구광역시 북구 대학로 80
경북대학교 인문대학 국어국문학과
전화번호: 010-9493-2618
전자우편: c-juni@daum.net

현영희

702-701 대구광역시 북구 대학로 80
경북대학교 인문대학 국어국문학과
전화번호: 010-9376-0203
전자우편: dew840713@hanmail.net

원고 접수일: 2012년 10월 30일

원고 수정일: 2012년 12월 14일

계재 확정일: 2012년 12월 21일

[부록 1: 의미 주석 태그 세트]

senseClass1	senseClass2	senseClass3
음식명	주식	밥, 죽미음옹이, 국수, 만두, 떡국, 수제비, 주식기타
	찬품	국, 탕, 찌개, 전골, 생채, 숙채, 찜선, 전구이, 조림, 볶음, 회, 김치, 장아찌, 젓갈, 튀각부각, 마른찬, 편육, 족편, 묵, 찬품기타
	떡	찌는떡, 치는떡, 삶는떡, 지지는떡, 떡기타
	한과	유밀과류, 유과류, 정과류, 다식류, 숙실과류, 과편류, 옛강정류, 한과기타
	음청류	차, 화채, 음청기타
	술	발효주, 증류주, 혼성주, 술기타
조리방법	비가열조리	씻기, 까기, 굵어내기, 담그기1, 거르기, 말리기, 절이기, 가루내기, 채치기, 찌기, 누르기, 무치기, 썰기, 갈기, 찜기, 다지기, 찢기, 두드리기, 섞기, 식히기, 빻기, 묻히기, 반죽하기, 양념하기, 비가열조리기타
	가열조리	삶기, 데치기, 끓이기, 졸이기, 찌기, 부치기, 볶기, 튀기기, 굽기, 훈연하기, 달이기, 익히기, 가열하기기타
	조리방법기타	떡우기, 담그기2
음식재료	물, 서류, 버섯류, 곡류, 두류, 채소류, 과실류, 수조육류, 어패류, 해초류, 알류, 약이성향신채류, 당류, 발효제, 꽃나뭇잎, 당류, 음식재료기타	
양념고명	양념, 고명, 양념고명기타	
부엌세간식	부엌공간, 부엌세간,	

기	조리도구, 식기수저, 부엌세간식기기타
계량단위	용량단위, 중량단위, 수량단위, 길이두께단위, 시간단위, 계량단위기타
조리정도· 시기	조리정도, 조리기간, 조리시기, 조리정도시기기타
저장	저장방법, 저장장소, 저장기간, 저장시기, 저장기타
맛	
향미	
불	
기타	