

# 한국어 어휘지도(UWordMap) 구축

2013. 07

울산대학교  
옥철영

## I. 한국어 어휘지도 개요

자연 언어의 어휘적 의미, 구문적 의미, 담화적 의미를 바탕으로 행위나 현상, 상태 등에 담긴 의미론적/개념론적 특성을 포함하고 있는 의미적 언어 자원 구축에 대한 연구는 다양하게 이루어지고 있다. 국외에서는 WordNet, Euro WordNet, Cyc, HowNet, Lexical FreeNet, EDR 등이 대표적이며, 국내에서는 카이스트의 CoreNet, ETRI의 어휘 개념망, 부산대의 KorLex 등이 대표적이라 할 수 있다.

KAIST CoreNet(코어넷)은 다국어 어휘 의미망으로서, 3,000여개의 개념이 한국어-중국어-일본어로 대응되어 동일한 개념체계를 공유한다. ETRI 명사 개념망은 한국어 명사 어휘들을 의미관계로 연결시켜 놓은 어휘 데이터베이스로서, 깊이 12레벨로 구성되어 있으며, 약 25여 만 고유명사들이 Instance\_Of 관계로 연결되어 있다.

UWordMap은 한국어의 공통적이고 개별적인 속성을 바탕으로 한국인의 보편적인 인지 체계와 개념 관계를 파악하여 이를 어휘의 의미적/개념적 네트워크를 형성한 온톨로지적 의미망이라 할 수 있다.

UWordsMap은 현재(2013년 6월 기준) 43만 여 어휘가 구축된 상태이다(용언 및 부사 포함). UWordMap의 구축 대상은 한국어 어휘 전체(모든 품사 및 언어 단위)로서, 핵심적 대상은 명사, 동사, 형용사이며, 부수적 대상은 부사, 관형사, 대명사, 감탄사, 조사, 수사, 의존명사 등이며, 북한어, 방언, 옛말, 전문용어, 고유명사, 어근, 어미 등 한국어 어휘 전체를 대상으로 연구 개발 중이다. <그림 1>는 구축 사례의 일부 모습이다.



## II. U-WIN(명사어휘의미망)

### 1. U-WIN의 기본 의미관계<sup>1)</sup>

U-WIN은 UWordMap의 명사어휘의미망으로, 한국어의 공통적이고 개별적인 속성을 바탕으로 한국인의 보편적인 인지 체계와 개념 관계를 파악하기 위하여, 2002년에 발표한 <시디용 표준>의 뜻풀이를 기반으로 다의적으로 구축된 어휘의 의미적/개념적 네트워크를 형성한 어휘의미망이다.

#### 1.1. U-WIN 최상위어 설정

U-WIN에서는 기초 작업 및 분석을 통하여 다음과 같은 기준에 따라 최상위어를 설정하였다.

<표 1> 최상위어 설정 기준

|                          |                                                                                                                                                                                                                                                                                                |
|--------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 최상위어<br>설<br>정<br>기<br>준 | <ul style="list-style-type: none"> <li>● 사전으로 기반으로 하는 만큼 최상위어는 사전에 등재되어 있는 어휘를 사용</li> <li>● 의미적으로 명확하게 인지되는 어휘를 사용</li> <li>● 형태적으로 사람들이 자주 사용하고 인식하는 어휘 사용</li> <li>● 다른 최상위어와의 개념적 중복성이 적은 어휘 사용</li> <li>● 하위어의 구성을 고려하여 선택</li> <li>● 기존 지식베이스에서의 최상위어 중 어휘 군집화를 통하여 공통된 어휘 사용</li> </ul> |
|--------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

<표 22>와 같은 최상위어 설정 기준을 바탕으로 U-WIN에서는 <표 23>과 같이 최상위어 23개를 설정하여 U-WIN의 상하 관계 기반 계층 구조의 틀을 마련하였다.

<표 2> U-WIN에서의 23개 최상위어

| Top-Level Word | Definition                                                             | Hyponymy(2 Depth)                                              |
|----------------|------------------------------------------------------------------------|----------------------------------------------------------------|
| { _0502}       | 물리적으로나 심리적으로 널리 퍼져 있는 범위. 어떤 물질이나 물체가 존재할 수 있거나 어떤 일이 일어날 수 있는 자리가 된다. | 곳_0101, {길_0101;도도_0700}, 자리_0101, 장소_0500, 지역_0302, 칸_0101... |
| {과정_0300}      | 일이 되어 가는 경로                                                            | 경과_0403, 단계_0300, 변화_0000,                                     |

1) 절은 최호섭(2007)에서 본 연구와 관련된 부분만 발췌하여 편집하였음.

|                    |                                                                                            |                                                                                              |
|--------------------|--------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------|
|                    |                                                                                            | _0101,<br>{역사_0401;사_0801;춘추_0202} ...                                                       |
| {관계_0501;계관_0101}  | 둘 이상의 사람, 사물, 현상 따위가 서로 관련을 맺거나 관련이 있음. 또는 그런 관련.                                          | 관련_0000, 비_0502,<br>{순서_0001;애차_0200;윤서_0200},<br>연관_0601, 유대_0200...                        |
| {기호_0100;심벌_0002}  | 어떠한 뜻을 나타내기 위하여 쓰이는 부호, 문자, 표지 따위를 통틀어 이르는 말.                                              | 논리기호_0000, 도형_0302,<br>부호_0301, 언어_0100,<br>{음성기호_0000;어음기호_0000;음성자<br>모_0000;표음기호_0000}... |
| {단위_0201;하나치_0000} | 길이, 무게, 수효, 시간 따위의 수량을 수치로 나타낼 때 기초가 되는 일정한 기준. 근, 되, 자, 그램, 리터, 미터, 초 따위가 있다.             | 국제단위_0000, 단위계_0000,<br>박자_0002, 언어단위_0000,<br>화폐단위_0000...                                  |
| {대상_1101}          | 어떤 일의 상대 또는 목표나 목적이 되는 것.                                                                  | {객체_0201;물격_0000},<br>{목표0001;표목_0201}, 상대_0402,<br>존재_0002....                              |
| {모양_0201}          | 겉으로 나타나는 생김새나 모습.                                                                          | {겉모양_0000;외양_0300},<br>모양새_0001, 상태_0101,<br>{양상_0801;양_2202}...                             |
| {물건_0001}          | 일정한 형체를 갖춘 모든 물질적 대상.                                                                      | 고물_0602, {구조물_0000;구조_0802},<br>기기_1300, 대상물_0001, 문선_0001...                                |
| {방법_0001}          | 어떤 일을 해 나가거나 목적을 이루기 위하여 취하는 수단이나 방식.                                                      | 가공법_0100, 공법_0100,<br>관리법_0001,<br>{농사법_0000;농법_0200},<br>{방식_0100;법식0002}, 수단_0101...       |
| {범위_0001}          | 테두리가 정하여진 구역.                                                                              | 범주_0501, 부문_0600, 부분_0100,<br>분야_0001, 영역_0302...                                            |
| {생물_0101;유생물_0000} | 생명을 가지고 스스로 생활 현상을 유지하여 나가는 물체, 영양·운동·생장·증식을 하며, 동물·식물·미생물로 나뉜다.                           | 동물_0001, 미생물_0000, 식물_0200,<br>{유기체_0002;생물체_0000}...                                        |
| {성질_0002;성분_0300}  | 사물이나 현상이 가지고 있는 고유의 특성.                                                                    | 가능성_0000, 물리적성질_0000,<br>병성_0000, 본질_0201, 성질_0001,<br>재질_0202...                            |
| {시간_0401}          | 어떤 시각에서 어떤 시각까지의 사이.                                                                       | {세월_0201;나날_0200;세화_0200;연광<br>_0103;연화_0200;오토_0202},<br>시점_0200, 대기시간_0000,<br>동안_0101...  |
| {요소_0401}          | 사물의 성립이나 효력 발생 따위에 꼭 필요한 성분. 또는 근본 조건.                                                     | 생산요소_0000, 성분_0104,<br>요소_0402, 인자_0701, 조건_0201,<br>토대_0003...                              |
| {인지_0803;인식_0002}  | 자극을 받아들이고, 저장하고, 인출하는 일련의 정신 과정. 지각, 기억, 상상, 개념, 판단, 추리를 포함하여 무엇을 안다는 것을 나타내는 포괄적인 용어로 쓴다. | 마음_0102,<br>{생각_0101;의려_0200;지려_0100},<br>앎_0000, 지식_0201...                                 |
| {작용_0101}          | 어떠한 현상을 일으키거나 영향을 미침.                                                                      | 감각_0201, 기능_0301, 영향_0400,<br>작용_0102, {힘_0101;파워_0002}...                                   |
| {재료_0101}          | 물건을 만드는 데 들어가는 감                                                                           | 감_0205, 금속재료_0000, 기재_0800,                                                                  |

|                              |                                       |                                                                                 |
|------------------------------|---------------------------------------|---------------------------------------------------------------------------------|
|                              |                                       | _0000,<br>{약재료_0200;약료_0001;약제_0200;약<br>중_0101}...                             |
| {정도_1101;정한_0300}            | 사물의 성질이나 가치를 양부, 우열<br>따위로 본 분량이나 수준. | 간격_0203, 강도_0501, 굳기_0000,<br>기준_0300, {길이_0101;장_1300},<br>{높이_0101;고_1400}... |
| {존재_0001}                    | 현실에 실제로 있음. 또는 그런 대상.                 | {공존_0001;공재_0200;구재_0600},<br>실재_0201, 존립_0002...                               |
| {종류_0201;종_0902;종속_04<br>00} | 사물의 부분을 나누는 갈래.                       | 갈래_0001, 인종_0102, 종목_0101,<br>직종_0000, 품종_0001...                               |
| {집단_0000}                    | 여럿이 모여 이룬 모임.                         | 가구_0303, 모임_0100, 무리_0100,<br>사회_0701, 조직_0002...                               |
| {행위_0001}                    | 사람이 의지를 가지고 하는 짓.                     | 개발_0001, 결정_0101,<br>경제행위_0000, 노력_0100,<br>부정행위_0100, 생활_0002...               |
| {힘_0103}                     | 어떤 일을 할 수 있는 능력이나 역량.                 | 경제력_0000, 권위_0001,<br>{능력_0201;역능_0000}, 인력_0101,<br>힘_0104...                  |

## 1.2. 의미 관계

의미 관계란 한 언어의 어휘 체계 속에 있는 어휘소 상호간에 가지고 있는 의미적 관계를 말한다. 다시 말하면, 한 어휘소와 다른 어휘가 의미상 가지고 있는 관계이다. 그러므로 의미 관계는 둘 이상의 어휘소 사이에서 성립하는 상대적 개념이다.

U-WIN에서의 의미 관계는 상하 관계, 동의 관계, 유의 관계, 반의 관계, 전체-부분 관계, 연관 관계 등 총 6가지 기본 의미 관계를 설정하였다. 이러한 U-WIN의 의미 관계를 설정하기 위하여, 『표준국어대사전』과 기타 사전을 바탕으로 하는 U-WIN 어휘 사전 데이터베이스(MRD)에서 추출될 수 있는 뜻풀이 및 용례 등 각종 어휘 정보를 활용하여 수작업 및 반자동 구축 작업을 병행하였으며, 기초 자원들을 활용하여 문장의 표면 구조(surface structure)를 중심으로 어휘의 계열적(paradigmatic) 관계와 통합적(syntagmatic) 관계를 분석하여 이러한 어휘의 의미 관계 설정에 활용하였다. U-WIN에서 사용하는 의미 관계를 정리하면 <표 24>와 같다.

<표 3> U-WIN의 기본 의미 관계

| Semantic Relation | Semantic Relation (detail) | direction |
|-------------------|----------------------------|-----------|
| 관계                | 상의관계                       | 단방향       |
|                   | 하의관계                       | 단방향       |
| 동의 관계             | 일반동의관계                     | 양방향       |
|                   | 오용어관계                      | 단방향       |
|                   | 방언관계                       | 단방향       |
|                   | 북한어관계                      | 단방향       |
|                   | 옛말관계                       | 단방향       |
|                   | 준말본말관계                     | 양방향       |
| 유의 관계             | 일반유의관계                     | 양방향       |
|                   | 순화어관계                      | 양방향       |
|                   | 거센말관계                      | 단방향       |
|                   | 높임말관계                      | 단방향       |
|                   | 낮춤말관계                      | 단방향       |
|                   | 은어관계                       | 단방향       |
|                   | 전문유의관계                     | 단방향       |
|                   | 전칭관계                       | 단방향       |
| 부분-전체 관계          | 부분관계                       | 단방향       |
|                   | 전체관계                       | 단방향       |
| 반의 관계             | 상보적반의관계                    | 양방향       |
|                   | 정도적반의관계                    | 양방향       |
|                   | 상대적반의관계                    | 양방향       |
| 연관 관계             | 관련어관계                      | 양방향       |

### 1.3. 상하 관계

#### 1.3.1. 상하 관계 처리 원칙

어휘의 대응(lexical correspondence)이라는 술어는 본래 하나의 개별 언어와 다른 하나의 언어에 관한 대조 연구에서 사용되던 술어이다. 어떤 단어와 다른 어떤 단어의 어휘 체계를 대조하면 한 언어의 단어 하나가 점유하고 있는 의미 영역(semantic range)이 다른 한 언어에서도 한 개의 단어에 의하여 동일하게 점유되어 있는 1:1 대응(one-to-one correspondence)이 나타나거나 하면, 한 언어에서는 한 개의 단어에 의해서 점유되어 있는 의미 영역이 다른 언어에서는 두 개 이상의 단어에 의하여 점유되는 1:n 대응(one-to-many correspondence)이 나타나는 경우도 있고, 그 반대로 n:1 대응(many-to-one correspondence)이 나타나는 경우가 있음을 알 수 있다.

U-WIN에서는 어휘 간의 상하 관계에서 어휘 간의 대응을 1:1 대응만을 인정하고 있다. 어휘 간의 상하 관계를 언어학적으로 정의하는 것은 그리 어렵지 않은 일이지만, 이론적 설명만으로 상하 관계를 처리한다는 것은 무리가 따른다고 볼 수 있다. 이는 몇 개의 어휘를 대상으로 하는 상하 관계 구축 작업은 쉬울 수도 있으나, 한국어 전체 어휘를 대상으로 상하 관계를 연결하고자 했을 때는 문제가 양산되리라고 본다. 그러나 U-WIN에서는 한국어 전체 어휘가 상하 관계로 모두 연결될 수 있을 것이라는 가능성을 전제로 상하 관계를 설정하였다.

U-WIN에서 사용한 상하 관계에 의한 계층 구조 형성의 기본 원칙은 <표 25>와 같다.



<표 4> U-WIN에서의 상하 관계 설정 원칙

|                                                                                                                                   |              |
|-----------------------------------------------------------------------------------------------------------------------------------|--------------|
| ① IS_A, KIND_OF                                                                                                                   | 상하 관계로 설정한다. |
| ② Top-Down/Bottom-Up 방식                                                                                                           | 모두를 이용한다.    |
| ③ 상하 관계는 어휘의 본질적 의미를 통해 연결된다.                                                                                                     |              |
| ④ 상하 관계 검증 방법 <sup>2)</sup>                                                                                                       | 을 적절히 이용한다.  |
| ⑤ 하위어의 정보를 상위어가 가지도록 한다.                                                                                                          |              |
| ⑥ 하위어는 상위어의 속성을 상속받는다. (제약 조건 포함)                                                                                                 |              |
| ⑦ 먼저 어휘의 형태적 측면(한자, 형태소 분석 정보)을 통해 상위어를 검토하고, 다음으로 각종 사전의 뜻풀이를 통해 검토하고, 마지막으로 말뭉치와 각종 문서를 통해 용례를 검토한다. 이들은 모두 단계적으로 분석되어야 할 대상이다. |              |
| ⑧ 전문용어의 경우, 전문용어의 특수성을 고려하여 분야별 전문용어 분류 체계를 최대한 반영하되, 단 여러 분류 체계가 있을 경우에는 일반적인 상하 관계 체계를 따른다.                                     |              |
| ⑨ 동의 관계와 유의 관계 중 양방향 어휘 관계를 제외한 나머지 관계는 상하 관계에 포함시킨다.                                                                             |              |
| ⑩ 계층성을 가지는 용언 어휘망과 부사 어휘망은 별도의 원칙을 설정한다.                                                                                          |              |

### 1.3.2. U-WIN에서의 상하 관계 설정

U-WIN에서의 상하 관계는 앞에서 기술한 상하 관계 기본 원칙을 바탕으로 언어학적인 분석을 통해 이루어졌다. 이러한 상하 관계 설정을 형태적인 측면에서의 상하 관계와 의미적인 측면에서의 상하 관계로 나누어 기술하면 다음과 같다. 물론 실질적으로 어휘의 상하 관계를 설정할 때에는 형태적인 측면과 의미적인 측면, 나아가 용례까지 모두 고려하여 상하 관계를 설정한다.

2) 관계 검증 방법은 Cruse(1986)의 검증 방법과 IS\_A 관계를 이용하였다.  
 {X는 Y의 한 종류이다/일종이다} - 계층적 구조  
 {X는 Y이다} - 계층적 구조  
 {X는 Y(으)로 사용된다} - 의미 분류적 정보

## ① 형태적인 측면에서의 상하 관계

### ■ 한자어

한자어로 구성된 단일명사의 경우, 한자가 갖는 핵심의미에 따라 상하 관계를 연결하였다. 예를 들어, ‘공립(公立), 사립(私立), 공설(公設), 관설(官設), 시립(市立), 창립(創立), 설비(設備)’ 등의 한자어들은 한자어 ‘立’과 ‘設’의 의미를 분석함과 동시에, 각 어휘의 뜻풀이에서 한자어 의미를 반영하는 요소를 분석하여 ‘설립(設立)’-‘설치(設置)’라는 상위어로 연결하였다. 이와 같은 방법으로 비슷한 유형의 한자어를 군집화하여 상하 관계를 설정하였다.

<표 5> 한자어의 상하 관계 설정 사례

|  |                                                                                                                                                                                                                                                                                                                                                                                           |
|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  | <p>_0100 : 지방 자치 단체가 <b>세워서 운영하는 일</b>. 또는 그 시설.</p> <p>사립_0400 : 개인이 자신의 자금으로 공익의 사업 <b>기관을 설립</b>하여 유지함.</p> <p>공설_0100 : 국가나 공공 단체에서 일반 사람들을 위하여 <b>만들어 세움</b>. 또는 그런 시설.</p> <p>관설_0000 : 국가 기관에서 <b>설립하거나 설치함</b>. 또는 그렇게 세운 시설.</p> <p>시립_0100 : 시의 경비로 <b>세워</b> 관리하고 유지하는 것.</p> <p>창립_0200 : 기관이나 단체 따위를 <b>새로 만들어 세움</b>.</p> <p>설비_0200 : 필요한 것을 <b>배풀어서 갖추</b>. 또는 그런 시설.</p> |
|--|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

### ■ 한자접미사

한자접미사가 붙은 명사는 한자어와 마찬가지로 한자접미사의 의미와 어휘 뜻풀이에서 한자접미사 의미를 반영하는 요소를 분석하여 상하 관계를 연결하였다. 예를 들어 ‘목공소, 발전소, 사무소, 거래처, 발행처, 공사장, 작업장’ 등과 같이 ‘~소(所), ~처(處), ~장(場)’ 등의 접미사가 붙은 명사는 ‘곳’이나 ‘장소’ 또는 더 큰 개념인 ‘공간’의 상위어와 연결되도록 하였다. 이러한 한자접미사의 상하 관계 설정은 대규모 어휘의 상하 관계 설정에서 간과할 수 있는 최상위어 구조까지의 개념화(conceptualization) 과정의 일관성을 가질 수 있다는 점에서 의미를 가진다고 할 수 있다.

다른 예로 한자접미사 ‘가(家)’의 경우는 다의어와 동형이의어인 접미사로서, 한자접미사를 면밀하게 분석함으로써 상하 관계의 일관적 구축과 세밀한 상하 관계 설정이 가능하다.

#### - 가(家)\_13 [접사]

- ① ‘그것을 전문으로 하는 사람’ 또는 ‘그것을 직업으로 하는 사람’의 뜻을 더하는 접미사.

- 예) 건축가, 교육가, 문학과, 비평가, 작곡가, 평론가 등
- ‘사람’이라는 공통된 상위어를 추출할 수 있으며, 뜻을풀이를 통해 ‘전문가’와 ‘직업인’이라는 중간 상위어를 설정할 수 있음.

② ‘그것에 능한 사람’의 뜻을 나타내는 접미사.

- 예) 병략가, 외교가, 이론가, 전략가, 전술가 등
- ‘사람’이라는 공통된 상위어를 추출할 수 있으며, 뜻을풀이를 통해 ‘전문가’ 또는 ‘재주꾼’ 등의 중간 상위어를 설정할 수 있음.

③ ‘그것을 많이 가진 사람’의 뜻을 나타내는 접미사.

- 예) 자본가, 장서가 등
- ‘사람’이라는 공통된 상위어를 추출할 수 있음.

④ ‘그 특성을 지닌 사람’의 뜻을 나타내는 접미사.

- 예) 대식가, 명망가, 애연가 등
- ‘사람’이라는 공통된 상위어를 추출할 수 있음.

- 가(家)\_14 [접사]

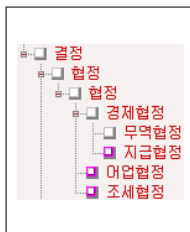
◀고유명사를 포함한 일부 명사 뒤에 붙어▶ ‘가문’의 뜻을 더하는 접미사.

- 명문가, 세도가, 재상가 등
- ‘가문(家門)’ 또는 ‘집안’이라는 공통된 상위어를 추출할 수 있음.

■ 복합명사의 경우

복합명사는 주로 오른쪽 성분이 핵심적인 의미정보 즉 유개념어를 가지는 경우가 많다. 따라서 형태적으로 오른쪽에 위치한 명사를 중심으로 상하 관계를 설정하였다. 예를 들어, ‘무역협정, 어업협정, 조세협정, 지급협정’ 등은 ‘협정(協定)’을 상위어로 갖는 어휘들이다.

<표 6> 복합명사의 상하 관계 설정 사례

|                                                                                     |                                                                                                                                                                                                                                                                                |
|-------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  | <p>_0000 : 나라와 나라 사이에 수출입 품목의 범위 따위의 무역에 관한 여러 조건을 규정하기 위하여 체결한 <b>협정</b>. (전문용어 분류 참조)</p> <p>지급<b>협정</b>_0000 : 두 나라 사이의 수출입 결제에 있어서 경화 사용을 절약하기 위하여 맺는 <b>협정</b>. (전문용어 분류 참조)</p> <p>어업<b>협정</b>_0000 : 두 나라 또는 여러 나라 사이의 협의에 의하여 국별, 어종별의 어획량을 따위를 결정(?)하는 <b>협정</b>.</p> |
|-------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

## ② 의미적인 측면에서의 상하 관계

사전을 기반으로 하고 있는 U-WIN은 의미가 그 중요한 위치를 차지한다. 사전의 뜻풀이를 이용하여 상하 관계를 설정하는 것은 형태적인 면을 고려하는 것과 함께 중요한 위치를 차지하고 있다.

### ■ 뜻풀이가 상위어 정보를 가진 경우

어휘의 뜻풀이가 상위어 정보를 갖는 경우, 즉 유개념어(類概念語)를 가지는 경우는 이를 상위어로 설정하였다. 예를 들어 '면세점(免稅點) : 외화 획득이나 외국인 여행자의 편의를 도모하기 위해 공항 대합실이나 시중에 설치된 비과세의 상점'의 경우 '상점(商店)'을 상위어로 설정하였다.

<표 7> 뜻풀이를 이용한 상위 관계 설정 사례

|  |                                                                                                                  |
|--|------------------------------------------------------------------------------------------------------------------|
|  | _0101 : 사람이나 물체가 차지하고 있는 .<br>길_0101 : 사람이나 동물 또는 자동차 따위가 지나갈 수 있게 땅 위에 낸 일정한 너비의 <b>공간</b> .                    |
|  | 사회_0701 : 같은 무리끼리 모여 이루는 <b>집단</b> .<br>조직_0002 : 특정한 목적을 달성하기 위하여 여러 개체나 요소를 모아서 체계 있는 집단을 이룸. 또는 그 <b>집단</b> . |

### ■ 같은 종류의 어휘의 뜻풀이가 다른 경우

유개념어를 통해 상하 관계를 확인한 후, 동일한 상하 관계로 연결하였다. 이와 같은 종류의 어휘들은 뜻풀이가 조금씩 다른 경우에 각각의 유개념어, 공통 개념, 의미 정보, 형태적 정보 등을 분석하여 동일한 상위어로 설정하였다.

<표 8> 같은 종류의 어휘의 뜻풀이가 다른 경우의 상위 관계 설정 사례

|  |                                                                                                                                                                                                                                                                                                                              |
|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  | _0000 : 아동들에게 기본적인 교육을 실시하기 위한 .<br>중학교_0001 : 초등학교와 고등학교 사이에 중등 보통 교육을 실시하기 위한 <b>학교</b> .<br>고등학교_0000 : 중학교를 졸업한 사람에게 고등 보통 교육과 실업 교육을 베푸는 <b>학교</b> .<br>대학_0101 : 고등 교육을 베푸는 <b>교육 기관</b> .<br>대학교_0000 : 예전에, <b>종합 대학</b> 을 단과 대학과 구별하여 이르는 말.<br>전문대학_0000 : 중견 직업인을 양성하기 위하여 전문적인 이론과 기술을 교수·연구하는 <b>고등 교육 기관</b> . |
|--|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

### ■ 상위어가 다의어일 경우

상위어가 다의어일 경우에는 어휘의 의미를 분명히 분석하여 상하 관계를 연결하였다.

<표 9> 다의어의 상위 관계 설정 사례

```

graph TD
    A[물건] --> B[기계]
    A --> C[집단]
    B --> D[장치]
    B --> E[기관]
    D --> F[기계장치]
    D --> G[교통기관]
    E --> H[교통기관]
    E --> I[내연기관]
    C --> J[조직]
    C --> K[기관]
    J --> L[기관]
    L --> M[교통기관]
    L --> N[국가기관]
    L --> O[사법기관]
    L --> P[합의제기관]
    L --> Q[의결기관]
    L --> R[의결기관]
    
```

■ 뜻풀이에 유개념어가 두 가지 이상 표현되어 있을 경우

국어사전에는 뜻풀이에 유개념어가 두 가지 이상 표현되는 경우가 다수 있다. 이러한 경우에는 유개념어가 각각 다른 상하 관계로 설정될 수 있다면, 다의어 설정 원칙인 어휘와 의미의 1:1로의 할당에 의거하여 새롭게 다의어를 생성시켜 상하 관계를 설정하였다<sup>4)</sup>.

<표 10> 유개념어가 두 가지 이상 표현되는 경우의 상하 관계 설정 사례

```

graph TD
    A[물건] --> B[대상물]
    B --> C[기록]
    C --> D[행위]
    D --> E[활동]
    E --> F[표현]
    F --> G[기록]
  
```

\_0201 : 주로 후일에 남길 목적으로 어떤 사실을 . 또는 그런 글.  
 기록\_0201 : 주로 후일에 남길 목적으로 어떤 사실을 적은 글.  
 기록\_0203 : 주로 후일에 남길 목적으로 어떤 사실을 적음.

■ 뜻풀이의 유개념어가 명확하지 못한 경우

뜻풀이의 유개념어가 있을 위치에 ‘것, 일, 곳, 말’ 등의 불명확한 어휘가 사용된 경우와 ‘나아감, 살핍, ~함’ 등 용언의 명사형으로 끝나는 경우들이 있다. 이와 같은 어휘들은 뜻풀이에서의 의미 정보, 말뭉치와 용례 분석 등 다양한 정보 분석을 하여 상하 관계를 연결한다.

3) [표준국어대사전]에는 통신 기관이 등재되어 있지 않다.

4) 여기서 중요한 점은 어휘의 사전 뜻풀이에 유개념어가 두 가지 이상 나타나지 않았는데 구축자가 임의적으로 새로운 뜻풀이를 추가할 수는 없다는 것이다.

## ■ 전문용어의 경우

전문용어들은 각 분야별로 용어 뜻풀이를 정리하고 있어, 전문용어의 상하 관계를 처리할 때에는 그 분야의 전문적인 분류 방식을 이용함과 동시에, 기본적인 상하 관계의 설정 원칙을 고려해야 한다. 의미의 변화가 없을 경우에는 기본적인 상하 관계 설정 원칙에 따라 연결하였다.

## 1.4. 동의 관계

동의 관계(absolute synonymy)는 둘 이상의 어휘소가 같은 의미를 가지고 있는 의미 관계를 말한다<sup>5)</sup>. U-WIN에서는 동의 관계를 6가지로 나누어 구축하였다.

### ① 일반 동의 관계

일반 동의 관계는 『표준국어대사전』 및 각종 사전에서 제공하는 동의어를 연결한 관계이다. 이러한 일반 동의 관계는 대부분은 자동으로 추출하여 연결하였다.

{편지\_02; 간독\_04; 간찰\_0002; 서간\_04; 서독\_03; 서소\_01; 서신\_02; 서장\_0301; 서찰\_02; 서척\_01; 서한\_02; 서함\_0001; 성문\_1001; 신\_0802; 신서\_03; 이소\_07; 찰한; 척간\_01; 척한; 편저\_01}

{중\_01; 결사\_0101; 범납\_01; 부도\_0703; 불승\_03; 상문\_0602; 석씨\_0003; 승\_13; 화합승}

{변소; 정방\_08; 청측; 측간\_01; 측실\_02; 측청\_01; 혼측; 회치장} ...

### ② 오용어 관계

오용어란 언어 사용자의 표기상의 오류나 잘못 사용하고 있는 어휘와 표준어를 연결시켜 놓은 관계로서, 의미적으로 동의 관계라 할 수 있다.

가늌쇠 - 말근쇠, 맑은쇠

5) Baldinger(1970) 지적처럼 전문어를 제외하고는 완전한 동의관계를 인정하지 않는 것이 언어학자들의 일반적인 견해이다. 그러나 본 연구에서는 동의 관계와 유의 관계를 분리하였다. 이는 [표준국어대사전]에서 제공하는 정보가 동의어와 유의어 구분하여 제공할 뿐만 아니라 PWN, EWN의 동의어 집합(synset)에 대한 의미 관계 설정이 아닌 어휘소 간 의미 관계 연결임 감안하여 이를 분리하였다.

가택질 - 가디약질, 갈괌\_0001, 갈괌질\_0001, 갈랙질, 개닥질

내리막길 - 내림길

쉬이여기다 - 수이여기다

기다랗다 - 길다랗다\_0001

데이터베이스 - 데이타베이스

### ③ 방언 관계

방언 관계는 방언과 표준어 사이의 관계를 말하는 것으로, 어휘 의미가 동일하기 때문에 동의 관계로 설정하였다.

개구리 - 가개비, 갈개비, 개구락지, 까구리, 깨고래기, 깨구락지, 깨구랭이, 깨구락데기, 머구리, 머우리,           떡장구, 멀구락지, 메구락지, 메구래기, 메구리, 맥자기, 맥장구, 맥재기, 떡자구, 떡장귀, 떡장기

벼\_01 - 나락, 나룩, 나룩, 노락

두름\_01 - 끄름지, 다랑구, 두럼, 두렁, 두레미, 두루미, 두리미, 드름, 못

### ④ 북한어 관계

북한어 관계는 북한어와 표준어 사이의 관계를 말하는 것으로, 어휘 의미가 동일하기 때문에 동의 관계로 설정하였다.

벼\_01 - 나락, 라룩, 화곡

가부장 - 부가장

가시밭길\_01 - 가시길\_01, 가시덤불길\_01

### ⑤ 옛말 관계

옛말 관계는 옛말과 현재 어휘 사이의 관계를 말하는 것으로, 동의 관계로 설정하였다.

가닥\_0001 - 가닭

살담배 - 지삼이, 툇담비

살림살이 - 사름사리

#### ⑥ 준말본말관계<sup>6)</sup>

준말은 단어의 일부분이 줄어든 말이며, 본말은 줄여지지 않은 본디 소리마디의 말로서, 의미에는 어떠한 영향을 주지 않는다. 그리하여 준말과 본말을 동의 관계로 설정하였다.

항공모함 - 공모\_03

대한제국 - 한국\_0501

복숭아 - 복사\_01

야간학교 - 야학\_0202

## 1.5 유의 관계

유의 관계(partial synonymy)는 둘 이상의 어휘소가 비슷한 의미를 가지고 있는 의미 관계를 말한다. U-WIN에서는 유의 관계를 8가지로 나누어 구축하였다.

#### ① 일반 유의 관계

일반 유의 관계는 『표준국어대사전』 및 각종 사전에서 제공하는 유의어를 연결함과 동시에, 각종 시소러스 및 유의어사전에서 수집한 유의어 집합을 이용하여 연결한 관계이다<sup>7)</sup>.

[학교; 학원\_0201]

[상소\_02; 배소\_03; 주서\_05; 주장\_10]

[서동\_01; 학동\_0001]

[선인\_0102; 도사\_1001; 도가자류]

#### ② 순화어 관계

---

6) ·도원영(2005)에서도 “본말과 준말은 음운, 형태적 차이에도 불구하고 의미적 차이가 없으므로 동의어로 처리한다”라는 원칙을 두고 있다.

7) 유의 관계와 상하 관계의 상관성에 대한 논의도 필요할 것으로 보인다. 예를 들어 김광해(2000)에서는 ‘학교’의 유의어로 ‘배움터, 학당, 학원, 학사, 문숙, 학창, 가숙’ 등을 설정하였는데, 여기에서 ‘배움터’는 ‘학교’의 집단적 의미와는 달리 공간적 의미를 가지는 어휘이다. 이러한 측면에서 상하 관계와 유의 관계의 상관성을 살피는 것도 연구적 의미가 있으리라 본다.



순화어란 불순한 요소를 없애고 깨끗하고 바르게 다듬은 말로서 지나치게 어려운 말이나 비규범적인 말, 외래어 따위를 알기 쉽고 규범적인 상태로 또는 고유어로 순화한 말을 이른다. 순화어 관계는 『표준국어대사전』에 기술되고 있는 순화어를 유의 관계로 설정한 것이다. 이러한 순화어는 특정한 말을 순화시키면서 그 의미가 변화되는 경우가 있으므로 유의 관계로 설정하였다.

이미지\_02 - 심상\_0402, 영상\_0102, 인상\_06

캐스터 - 진행자

커미션 - {구문\_02; 구전\_03; 두전\_02}, 수수료, 중개료

### ③ 거센말 관계

거센말은 어감을 거세게 하기 위하여 거센소리를 쓰는 말로서, 어감의 차이로 인해 어휘 의미를 미세한 차이를 가지고 있으므로 의미 관계로 설정하였다.

다사하다\_01 - 따사하다

방긋거리다 - 방끗거리다, 뽕긋거리다, 뽕끗거리다

가탈\_0101 - 까탈

### ④ 높임말 관계

높임말은 사람이나 사물을 높여서 이르는 말로서, 의미의 차이는 없으나 쓰이는 상황이 다르므로 유의 관계로 설정하였다.

{죽다0101; 운하다\_02} - 돌아가다\_0007

{사망\_0401; 사몰} - 연세\_03

{왕비; 군부\_0302; 궁비\_01; 왕후\_01} - {중궁전; 곤궁\_02; 곤전; 중궁\_02; 중전0202}

### ⑤ 낮춤말 관계

여기서의 낮춤말이란 사물이나 사물을 낮추어 이르는 말로서, 높임말과 마찬가지로 의미의 차이는 없으나 쓰이는 상황이 다르므로 유의 관계로 설정하였다.

{죽다0101; 운하다\_02} - 고꾸라지다\_0002, 웨지다

{게으름쟁이; 완낭} - {게으름뱅이\_01; 한산꾼\_02}

{어머니\_0101; 아미\_02} - 어미\_0101

#### ⑥ 은어 관계

은어란 어떤 계층이나 부류의 사람들이 다른 사람들이 알아듣지 못하도록 자기네 구성원들끼리만 빈번하게 사용하는 말로서, 상인·학생·군인·노름꾼·부랑배 따위의 각종 집단에 따라 다른데, 의태어·의성어·전도어·생략어·수식어 따위로 그 발생을 나눌 수 있다. 이러한 은어는 특정 어휘를 의미는 비슷하지는 쓰이는 환경이 다르므로 유의 관계로 설정하였다.

거짓말하다\_0001 - 뺑까다, 뺑놓다

고함\_01 - 곱\_02

널빤지 - 판때기

#### ⑦ 전문 유의 관계

특정한 분야에서 의미는 비슷하지만 형태적으로 다르게 사용하는 어휘와의 관계를 전문 유의 관계로 설정하였다.

사망\_0401 - 진\_01(무속)

수젓집 - 순가락집(불교)

황금\_0102 - 조금\_04(한의학)

#### ⑧ 전칭 관계

전칭 관계는 현재의 기관명, 벼슬명, 부서명 등을 중심으로 이전의 명칭과 연결시킨 것이다. 의미는 비슷하나, 역할이나 범위가 달라졌을 뿐만 아니라 형태적으로 다르기 때문에 유의 관계로 설정하였다.

낭중\_0201 - 제감\_0101 (벼슬)

병마절제사 - 병마단련사 (벼슬)

한국해양연구원 - 한국해양연구소 (기관명)

## 1.6 반의 관계

반의 관계는 두 어휘소가 서로 반대이거나 맞서는 의미를 가지고 있는 의미 관계를 말한다(윤평현 1995). U-WIN에서는 어휘의미론에서의 일반적인 반의 관계를 사용하여 3가지 반의 관계로 나누고 있다.

① 상보적 반의 관계(complementary antonymy)

상보적 반의 관계는 두 어휘소가 양분적 대립의 관계에 있는 것으로서, 한 쪽 어휘소의 의미를 부정하면 다른 쪽의 어휘소와 동의 관계를 유지하게 된다.

남자\_0201 ↔ 여자\_0201

추상\_01 ↔ 구상\_0501

기혼\_02 ↔ 미혼\_01

② 정도적 반의 관계(gradable antonymy)는 두 어휘소에 정도성(gradability)의 차이가 있는 관계가 있는 것으로서, 한 쪽을 부정하는 것이 바로 다른 쪽을 의미하지 않으며, 두 어휘소 사이에는 중간 상태가 있을 수 있다.

크다\_01 ↔ 작다\_01

뜨겁다\_0001 ↔ 차갑다\_0001

③ 상대적 반의 관계(relative antonymy)

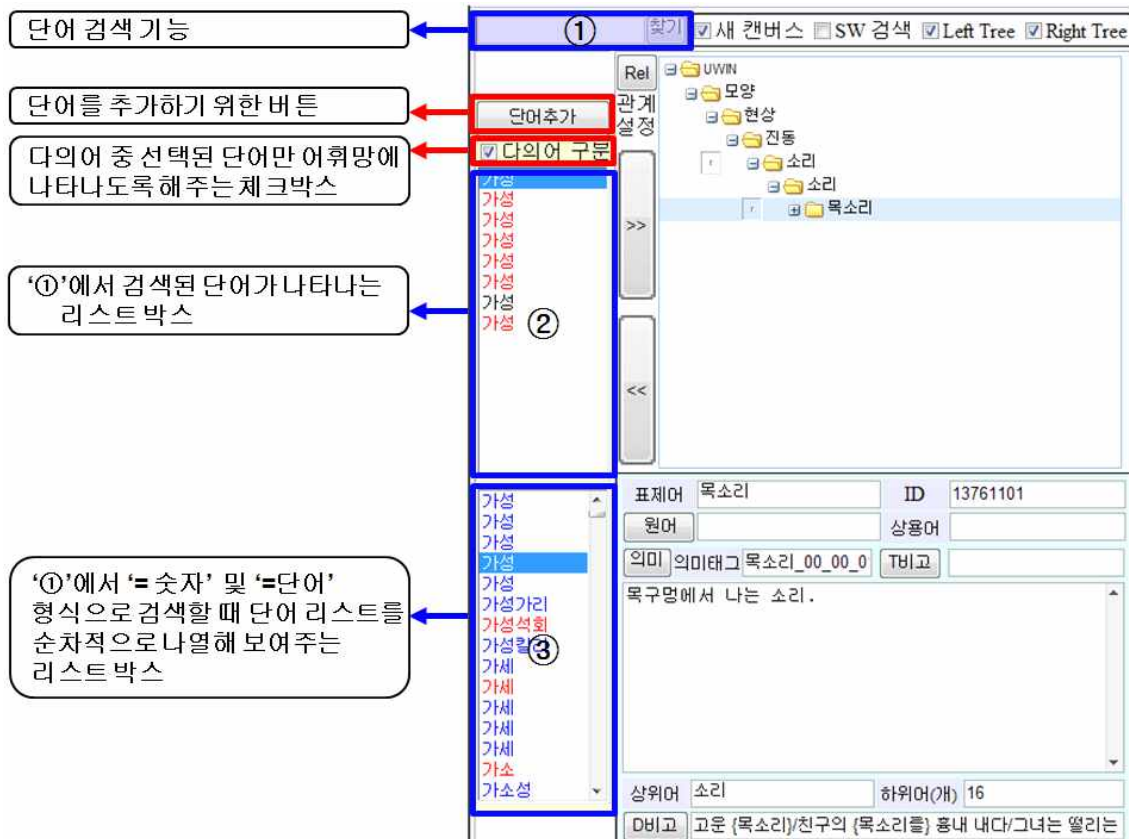
상대적 반의 관계는 두 어휘소가 대칭 관계를 이루는 경우를 말한다.

주다\_0101 ↔ 받다\_0101

사다\_0001 ↔ 팔다\_0001

부모\_01 ↔ 자식\_0101





<그림 II-2> 명사 어휘망 검증 도구 기능

위의 <그림 II-2>는 명사 어휘망 검증 도구 중 좌측면의 기능을 설명한 것이다. 단어검색, 단어추가, 다의어 구분해서 보기, 검색된 단어보기, 단어리스트 중 검색된 단어부터 보기에 대한 컨트롤들이다.

이 중 '다의어 구분' 기능이 체크되어 있지 않다면, 검색한 단어가 다의어이고, 이 중 한 단어를 더블클릭한 경우 모든 다의어를 명사 어휘망에서 볼 수 있다. 그렇지만 다의어가 너무 많거나, 특정 다의어 하나만 보고 싶을 경우 '다의어 구분'을 체크하고 검색된 단어 중 하나의 단어를 더블클릭하면 선택된 단어만 어휘망에서 볼 수 있다. 아래의 <그림 II-3>에서 왼쪽 트리는 '다의어 구분'을 체크하지 않은 상태에서 '가성'이라는 단어를 더블클릭한 경우이며, 오른쪽 트리는 '다의어 구분'을 체크한 상태에서 '가성'을 더블클릭한 경우이다.



<그림 II-3> '다의어 구분' 기능의 차이

①검색창에서 '가성'과 같이 단순히 단어만 검색하면 형태가 동일한 단어들이 ②의 리스트에 나타나게 된다. ①검색창에서 '=가성'과 같이 검색하면 ③의 리스트에 '가성'부터 정렬된 순으로 단어들이 나타나게 된다. 그리고 '=100'과 같이 검색하면 100번째 단어부터 정렬된 순으로 단어들이 나타나게 된다.

③의 리스트는 선정된 어휘의 작업량을 확인하면서 어휘망을 검증하기 위해 만들었다.

## 2.2. 명사 어휘망에 어휘 추가 방법

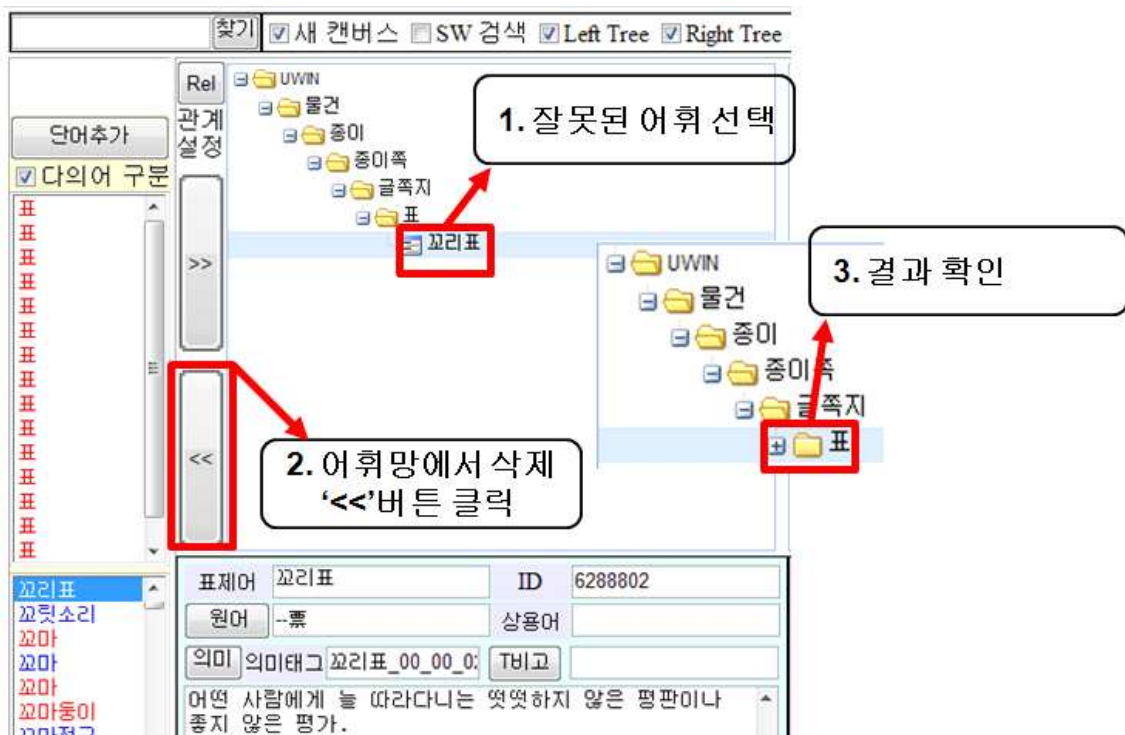
명사 어휘망 구축도구에서 명사를 추가하는 방법은 2가지가 있다.

그 중 첫 번째 방법은 <그림 II-2>에서 ②의 리스트의 단어를 선택한 뒤 상위어로 설정되어야 할 단어를 트리에서 선택한 후 '>>'버튼을 누르면 트리에 선택된 단어 아래에 리스트에서 선택한 단어가 들어가게 된다(<그림 II-4> 참조).



## 2.3. 명사 어휘망에 어휘 삭제 방법

구축된 명사 어휘망의 어휘 중 보다 세부적으로 분류되어야 하거나, 부적절한 위치에 추가되어있는 어휘는 삭제한 뒤, 다른 곳으로 옮겨야 한다. 삭제하기 위한 방법은 아래의 <그림 II-6>과 같이 어휘망 트리에서 어휘를 선택한 뒤, '<<'버튼을 눌러 삭제할 수 있다.



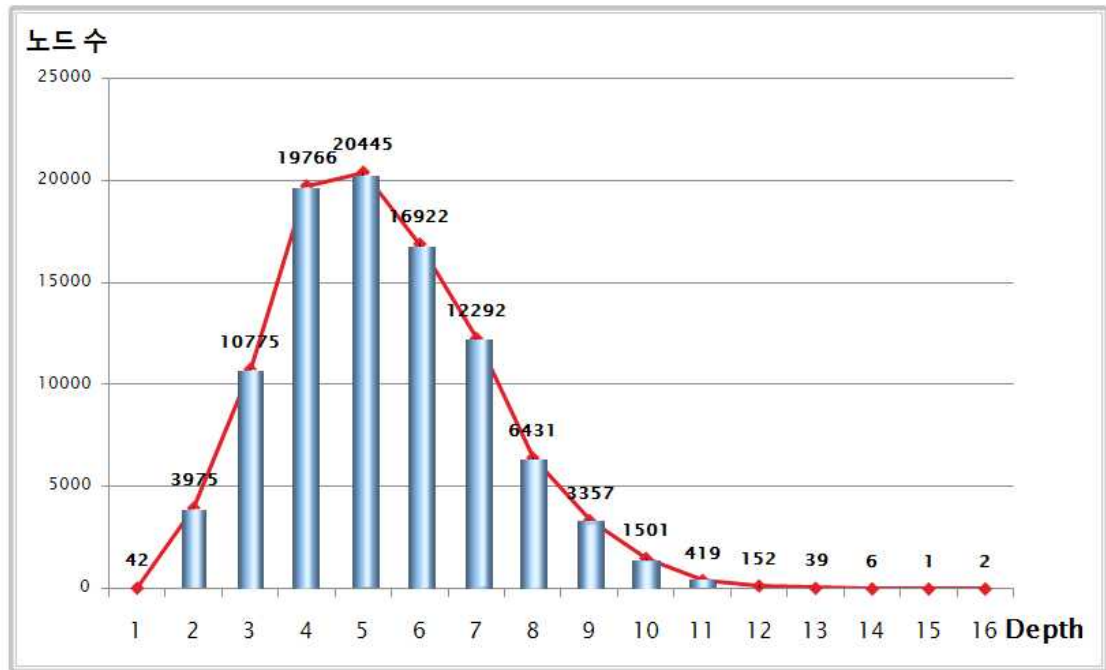
<그림 II-6> 상하 관계가 잘못 설정된 어휘 삭제 방법

## 3. 명사 어휘망 현황

현재까지 구축한 320,712개<sup>8)</sup>에 대한 명사 어휘망은 다음 <그림II-7>과 같은 분포를 이루고 있다.

8) 2013 6월까지 선정한 <표 I-4>의 의미관계 포함한 수치임.





<그림 II-7> 명사 어휘망의 노드별 어휘 분포

이 분포에서 알 수 있듯이 42개의 최상위 노드로 구성되어 있다. 이 중 현재의 최상위 노드에는 3.3절의 명사 상하 관계 설정이 어려워 인위적으로 추가한 10의 노드를 제외하면 32 최상위 노드로 구성되어 있다. 또한 최상위노드의 동의어(단말노드) 9개(동의어)를 제외하면 최종 23개의 노드로 구성되어 있다. 아래 표는 최상위 노드이다.

| 번호 | 최상위 노드      | U-WIN<br>비교 | 하위노드 수 | 비고 |
|----|-------------|-------------|--------|----|
| 1  | 공간_05_00_02 | ○           | 12     |    |
| 2  | 과정_03_00_00 | ○           | 47     |    |
| 3  | 관계_05_00_01 | ○           | 68     |    |
| 4  | 기호_01_00_00 | ○           | 26     |    |
| 5  | 단위_02_00_01 | ○           | 82     |    |
| 6  | 대상_11_00_01 | ○           | 60     |    |
| 7  | 모양_02_00_01 | ○           | 119    |    |
| 8  | 물건_00_00_01 | ○           | 420    |    |
| 9  | 방법_00_00_01 | ○           | 187    |    |
| 10 | 범위_00_00_01 | ○           | 34     |    |

|    |               |    |       |                    |
|----|---------------|----|-------|--------------------|
| 11 | 생물_01_00_01   | ○  | 30    |                    |
| 12 | 성질_00_00_02   | ○  | 392   |                    |
| 13 | 시간_04_00_01   | ○  | 46    |                    |
| 14 | 요소_04_00_01   | ○  | 34    |                    |
| 15 | 인지_08_00_03   | ○  | 11    |                    |
| 16 | 작용_01_00_01   | ○  | 79    |                    |
| 17 | 재료_01_00_01   | ○  | 64    |                    |
| 18 | 정도_11_00_01   | ○  | 173   |                    |
| 19 | 존재_00_00_01   | ○  | 29    |                    |
| 20 | 종류_02_00_01   | ○  | 39    |                    |
| 21 | 집단_00_00_00   | ○  | 43    |                    |
| 22 | 행위_00_00_01   | ○  | 293   |                    |
| 23 | 힘_01_00_03    | ○  | 91    |                    |
| 24 | 계관_01_00_01   | ○  | 0(단말) | ‘관계_05_00_01’의 동의어 |
| 25 | 성분_03_00_00   | ○  | 0(단말) | ‘성질_00_00_02’의 동의어 |
| 26 | 심벌_00_00_02   | ○  | 0(단말) | ‘기호_01_00_00’의 동의어 |
| 27 | 유생물_00_00_00  | ○  | 0(단말) | ‘생물_01_00_01’의 동의어 |
| 28 | 인식_00_00_02   | ○  | 0(단말) | ‘인지_08_00_03’의 동의어 |
| 29 | 정한_03_00_00   | ○  | 0(단말) | ‘정도_11_00_01’의 동의어 |
| 30 | 중_09_00_02    | ○  | 0(단말) | ‘종류_02_00_01’의 동의어 |
| 31 | 종속_04_00_00   | ○  | 0(단말) | ‘종류_02_00_01’의 동의어 |
| 32 | 하나치_00_00_00  | ○  | 0(단말) | ‘단위_02_00_01’의 동의어 |
| 33 | 용언/부사         | 추가 | 0(단말) | 용언하위범주화용 노드        |
| 34 | 지향작용_00_00_00 | 추가 | 0(단말) | ‘작용_01_00_01’의 동의어 |
| 35 | 시공_02_00_01   | 추가 | 0(단말) | “시간과 공간”           |
| 36 | 비유적으로_이르는_말   | 추가 | 1,133 |                    |
| 37 | 비유하여_이르는_말    | 추가 | 26    |                    |
| 38 | 상위어_선정불가      | 추가 | 194   |                    |
| 39 | 아울러_이르는_말     | 추가 | 230   |                    |
| 40 | 의미_순환_구조      | 추가 | 4     |                    |
| 41 | 통틀어_이르는_말     | 추가 | 1     |                    |
| 42 | 함을_이르는_말      | 추가 | 10    |                    |

또한, 전체 어휘망은 최대 깊이(depth) 16까지 구성되어 있으나 대체로 12까지 구성되어 있으며, 단말노드는 83,165개, 비단말노드(중간노드)는 96,232개이다. 이 중 비단말노드에는 용언의 의미군집을 하위범주화 정보 구축을 위해 용언 자체를 구분하기 위한 노드 및 용언-부사 간의 노드도 포함되어

있다.

위 표에서 추가한 노드는 용언의 최소 상계 노드와도 연결되어 있지 않으며 그의 하위에 속하는 어휘들은 배제 어휘에도 설정되지 않았다. 또한 임의로 추가한 노드들은 추후 연구자의 직관에 따라 적합한 상하 관계를 설정할 때 구분하기 위하여 추가된 노드들이다.

### III. 용언 하위범주 정보 구축

#### 1. 논항정보 구축 대상

표준국어대사전에 등재된 용언에 대해, 방언, 옛말, 북한말 등을 제외한 다음 표의 용언을 구축 대상으로 한다.

용언의 하위범주 정보 구축은 동사와 형용사의 논항정보를 구축하는 작업으로, 논항정보는 명사어휘망(U-WIN)을 기본으로 하여 연결한다. 논항(argument)은 문장의 의미를 구성하는 데에 참여하는 명사를 가리킨다. 따라서 부사는 논항정보 연결 작업에서는 제외되고 별도의 작업(IV장 참조)으로 구축하였다.

<표 IV-12> 구축 대상 용언

| 어휘  | 표준국어대사전 등재 어휘 | 어휘지도 구축 대상 |
|-----|---------------|------------|
| 동사  | 90,237        | 73,694     |
| 형용사 | 21,618        | 16,863     |
| 합계  | 111,855       | 90,557     |

#### 2. 용언의 하위범주 정보 구축 과정

##### 2.1. 용언의 하위범주 정보 추출

하위범주 정보는 <표준>에서 기술된 문형정보를 기준으로 하여 구축하였다. 예를 들어 '먹다'의 의미별 문형정보 및 뜻풀이, 용례는 다음 <표 IV-3>과 같다.

<표 IV-13> '먹다'의 뜻풀이 및 용례(표준국어대사전)

| 표제어  | 품사 | 문형     | 뜻풀이                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|------|----|--------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|      |    |        | 용례                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
| 먹다01 | 동사 | (...을) | <p>귀나 코가 막혀서 제 기능을 하지 못하게 되다.</p> <p>코 <b>먹은</b> 소리를 내다/이 사람은 귀가 <b>먹어서</b> 잘 못 들으니까 큰 소리로 말씀하셔야 돼요. ॥ 귀를 <b>먹었는지</b> 아무리 불러도 그냥 지나가더라.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |
| 먹다02 | 동사 | ...을   | <p>「1」 음식 따위를 입을 통하여 배 속에 들여보내다.</p> <p>밥을 <b>먹다</b>/술을 <b>먹다</b>/약을 <b>먹다</b>/물을 <b>먹다</b>/음식을 배불리 <b>먹다</b>/닭이 모이를 <b>먹다</b>/몸이 약해진 누나는 보약을 몇 차례나 <b>먹어도</b> 늘 골골거렸다.</p> <p>「2」 담배나 아편 따위를 피우다.</p> <p>담배를 <b>먹다</b>/아편을 <b>먹다</b>/정선이 야단 통에 두어 시간이나 담배를 끊었다가 <b>먹는</b> 담배라 머리가 아뜩한 것 같았다.《이광수, 흙》</p> <p>「3」 연기나 가스 따위를 들이마시다.</p> <p>연탄가스를 <b>먹다</b>/탄내를 <b>먹다</b>.</p> <p>「4」 어떤 마음이나 감정을 품다.</p> <p>양심을 <b>먹고</b> 투서를 하다/한번 <b>먹은</b> 마음이 변하지 않도록 하자./나는 마음을 독하게 <b>먹고</b> 그녀를 외면하였다.</p> <p>「5」 일정한 나이에 이르거나 나이를 더하다.</p> <p>네 살 <b>먹은</b> 아이/나이를 <b>먹다</b>/내년이면 삼십을 <b>먹는</b>구나.</p> <p>「6」 욕, 편견 따위를 듣거나 당하다.</p> <p>하루 종일 욕만 되게 <b>먹었네</b>./그래도 그는 속없는 소리를 하다가 가끔 편견을 <b>먹는</b> 것이었다.《이기영, 봄》</p> <p>「7」 (속되게) 뇌물을 받아 가지다.</p> <p>뇌물을 <b>먹다</b>/뇌물을 <b>먹고</b> 탈세를 눈감아 주다.</p> <p>「8」 수익이나 이문을 차지하여 가지다.</p> <p>남은 이익은 모두 내가 <b>먹어라</b>./시세가 마침 좋은 것 같아서 쌀을 붙였던 것인데 천 원을 <b>먹기</b>는 고사하고 본전 육백 원이 다 달아난 판이니...《채만식, 타류》</p> <p>「9」 물이나 습기 따위를 빨아들이다.</p> <p>기름 <b>먹은</b> 종이/김이 습기를 <b>먹어</b> 눅눅해졌다./숨이 물을 <b>먹어</b> 무겁다.</p> <p>「10」 어떤 등급을 차지하거나 점수를 따다.</p> <p>1등을 <b>먹다</b>/우승을 <b>먹다</b>/100점을 <b>먹다</b>/체육 대회에서 우리 반이 일 등을 <b>먹었다</b>.</p> <p>「11」 구기 경기에서, 점수를 잃다.</p> <p>상대편에게 먼저 한 골을 <b>먹었다</b>.</p> <p>「12」 (속되게) 여자의 정조를 유린하다.</p> <p>그는 벌써 여러 여자를 <b>먹었다</b>.</p> <p>「13」 매 따위를 맞다.</p> <p>상대의 센 주먹을 한 방 <b>먹고</b> 나가떨어졌다.</p> <p>「14」 남의 재물을 다루거나 맡은 사람이 그 재물을 부당하게 자기의 것으로 만들다.</p> <p>경리 직원이 회사의 공금을 <b>먹었다</b>.</p> |
|      |    | ...에   | <p>「1」 날이 있는 도구가 소재를 깎거나 자르거나 갈거나 하는 작용을 하다.</p> <p>이 고기에는 칼이 잘 <b>먹지</b> 않는다./대패가 잘 <b>먹는다</b>.</p> <p>「2」 바르는 물질이 배어들거나 고루 퍼지다.</p> <p>옷감에 풀이 잘 <b>먹어야</b> 다림질하기가 좋다./얼굴에 화장이 잘 <b>먹지</b> 않고 들뜬다.</p> <p>「3」 벌레, 균 따위가 파 들어가거나 퍼지다.</p> <p>사과에 벌레가 많이 <b>먹었다</b>./옷에 좀이 <b>먹어</b> 못 입게 되었다./얼굴에 버짐이 <b>먹다</b>.</p> <p>「4」 돈이나 물자 따위가 들거나 쓰이다.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |

|  |  |  |                                                                              |
|--|--|--|------------------------------------------------------------------------------|
|  |  |  | 공사에 철근이 생각보다 많이 <b>먹어</b> 걱정이다./낡은 집 수리에는 자칫 새로 짓는 것보다 비용이 더 <b>먹을</b> 수 있다. |
|--|--|--|------------------------------------------------------------------------------|

용언의 문형에 대한 전형적인 하위범주 정보는 기본적으로 각 뜻풀이의 용례에 기술되어 있으며, 본 과제에서는 II 장과 같이 다의어 수준까지 의미 태깅된 뜻풀이 및 용례에서 각 용언의 하위범주 정보 추출하였다.

뜻풀이/용례의 문장이 “.....(1)..... [용언\_1] .....(2)..... [용언\_2]”와 같을 경우,

- [용언\_1]의 문형의 하위범주 정보는 .....(1).....에서
- [용언\_2]의 문형의 하위범주 정보는 .....(2).....에서 추출하였다.

예를 들어, 뜻풀이 및 용례에서 사용된 ‘갈다’의 경우 다음 <표 IV-4>와 같은 하위범주 정보를 추출할 수 있다.

<표 IV-14> 다의어 태깅된 뜻풀이 및 용례에서 추출한 ‘갈다’의 하위범주 정보

| 용언 | 용언keyID | 명사   | 명사keyID  | 빈도 | 격조사 | 비고 | 용언 | 용언keyID | 명사  | 명사keyID  | 빈도 | 격조사 | 비고 |
|----|---------|------|----------|----|-----|----|----|---------|-----|----------|----|-----|----|
| 갈다 | 702001  | 건전지  | 1632800  | 1  | 을   | 0  | 갈다 | 702001  | 부속품 | 17596100 | 2  | 을   | 0  |
| 갈다 | 702001  | 공기   | 3213401  | 5  | 을   | 1  | 갈다 | 702001  | 부품  | 17751200 | 1  | 을   | 0  |
| 갈다 | 702001  | 관    | 3582800  | 1  | 을   | 0  | 갈다 | 702001  | 불   | 17976001 | 1  | 으로  | 0  |
| 갈다 | 702001  | 구두창  | 4120300  | 1  | 을   | 0  | 갈다 | 702001  | 불씨  | 18105001 | 1  | 을   | 0  |
| 갈다 | 702001  | 굽    | 4705804  | 1  | 을   | 0  | 갈다 | 702001  | 성   | 21318200 | 2  | 을   | 0  |
| 갈다 | 702001  | 굽창   | 4717200  | 1  | 을   | 0  | 갈다 | 702001  | 술   | 23024200 | 1  | 을   | 0  |
| 갈다 | 702001  | 금상   | 5352700  | 1  | 을   | 0  | 갈다 | 702001  | 스토리 | 23211100 | 1  | 을   | 0  |
| 갈다 | 702001  | 기저귀  | 5826100  | 2  | 으로  | 1  | 갈다 | 702001  | 시트  | 23620101 | 1  | 을   | 0  |
| 갈다 | 702001  | 기저귀  | 5826100  | 1  | 을   | 0  | 갈다 | 702001  | 신발창 | 23803101 | 1  | 을   | 0  |
| 갈다 | 702001  | 깃    | 6062201  | 1  | 을   | 0  | 갈다 | 702001  | 심지  | 24173701 | 1  | 을   | 0  |
| 갈다 | 702001  | 깃털   | 6072100  | 1  | 을   | 0  | 갈다 | 702001  | 앞니  | 25421300 | 2  | 을   | 0  |
| 갈다 | 702001  | 깎창   | 6151300  | 1  | 을   | 0  | 갈다 | 702001  | 애자  | 25530700 | 1  | 을   | 0  |
| 갈다 | 702001  | 깍데기  | 6247102  | 1  | 을   | 0  | 갈다 | 702001  |     | 26304300 | 1  | 을   | 0  |
| 갈다 | 702001  | 나무못  | 6632000  | 1  | 을   | 0  | 갈다 | 702001  | 연탄  | 27005100 | 2  | 을   | 1  |
| 갈다 | 702001  | 널빤지  | 7387800  | 1  | 을   | 0  | 갈다 | 702001  | 연탄불 | 27005700 | 2  | 을   | 1  |
| 갈다 | 702001  | 돌    | 10059304 | 1  | 을   | 0  | 갈다 | 702001  | 유리  | 29535300 | 2  | 을   | 0  |
| 갈다 | 702001  | 뒤창   | 10734600 | 1  | 을   | 0  | 갈다 | 702001  | 유리창 | 29545900 | 1  | 을   | 0  |
| 갈다 | 702001  | 마른자리 | 12459200 | 3  | 으로  | 0  | 갈다 | 702001  | 이   | 30317701 | 1  | 을   | 0  |
| 갈다 | 702001  | 말굽   | 12757601 | 1  | 을   | 0  | 갈다 | 702001  | 이빨  | 30494800 | 1  | 을   | 0  |
| 갈다 | 702001  | 물    | 14451801 | 9  | 을   | 1  | 갈다 | 702001  | 이영  | 30594900 | 1  | 을   | 0  |
| 갈다 | 702001  | 물건   | 14463101 | 1  | 을   | 0  | 갈다 | 702001  | 자리  | 31627201 | 1  | 을   | 0  |
| 갈다 | 702001  | 물고기  | 14471100 | 1  | 을   | 0  | 갈다 | 702001  | 전구  | 32874600 | 3  | 을   | 1  |
| 갈다 | 702001  | 미끼   | 14699501 | 1  | 을   | 0  | 갈다 | 702001  | 전등  | 32953500 | 1  | 을   | 0  |
| 갈다 | 702001  | 밀창   | 15022801 | 1  | 을   | 0  | 갈다 | 702001  | 전지  | 33205800 | 1  | 을   | 0  |
| 갈다 | 702001  | 바퀴   | 15227501 | 1  | 을   | 1  | 갈다 | 702001  | 전지약 | 33206800 | 1  | 을   | 0  |
| 갈다 | 702001  | 배넛니  | 15955100 | 1  | 을   | 0  | 갈다 | 702001  | 젓니  | 33858500 | 2  | 을   | 0  |
| 갈다 | 702001  | 번    | 16335101 | 1  | 을   | 1  | 갈다 | 702001  | 진자리 | 36097602 | 3  | 을   | 0  |
| 갈다 | 702001  | 병    | 16804101 | 2  | 을   | 1  | 갈다 | 702001  | 참쌀  | 36788902 | 1  | 을   | 0  |

|    |        |      |          |   |    |   |    |        |     |          |    |   |   |
|----|--------|------|----------|---|----|---|----|--------|-----|----------|----|---|---|
| 갈다 | 702001 | 찾길   | 36791402 | 1 | 을  | 0 | 갈다 | 702102 | 손가락 | 22144900 | 1  | 을 | 1 |
| 갈다 | 702001 | 창    | 36793202 | 1 | 을  | 0 | 갈다 | 702102 | 쇠   | 22331301 | 1  | 을 | 0 |
| 갈다 | 702001 | 창유리  | 36835900 | 1 | 을  | 0 | 갈다 | 702102 | 쇠불이 | 22359401 | 2  | 을 | 0 |
| 갈다 | 702001 | 책상보  | 36927900 | 1 | 을  | 0 | 갈다 | 702102 | 쇠불이 | 22359402 | 4  | 을 | 0 |
| 갈다 | 702001 |      | 38749800 | 1 | 으로 | 0 | 갈다 | 702102 | 숫돌  | 23069900 | 1  | 을 | 0 |
| 갈다 | 702001 | 타이어  | 39078401 | 1 | 을  | 1 | 갈다 | 702102 | 안쪽  | 25130201 | 1  | 을 | 0 |
| 갈다 | 702001 | 탄    | 39125902 | 1 | 을  | 0 | 갈다 | 702102 | 암석  | 25305200 | 1  | 을 | 0 |
| 갈다 | 702001 | 탄불   | 39140300 | 1 | 을  | 0 | 갈다 | 702102 | 양면  | 25823701 | 1  | 을 | 0 |
| 갈다 | 702001 | 탄창   | 39171900 | 1 | 을  | 0 | 갈다 | 702102 | 양편  | 25943000 | 1  | 을 | 0 |
| 갈다 | 702001 | 탄환   | 39179202 | 1 | 을  | 1 | 갈다 | 702102 | 연마재 | 26867700 | 1  | 을 | 0 |
| 갈다 | 702001 | 털    | 39393701 | 3 | 을  | 0 | 갈다 | 702102 | 연장  | 26966801 | 1  | 을 | 0 |
| 갈다 | 702001 | 패킹   | 40276801 | 1 | 을  | 0 | 갈다 | 702102 | 옥   | 27856500 | 2  | 을 | 0 |
| 갈다 | 702001 | 퓨즈   | 40985700 | 1 | 을  | 0 | 갈다 | 702102 | 옥돌  | 27870100 | 2  | 을 | 1 |
| 갈다 | 702001 | 피    | 41074800 | 1 | 을  | 0 | 갈다 | 702102 | 옥석  | 27885901 | 1  | 을 | 0 |
| 갈다 | 702001 | 활자   | 43239700 | 1 | 을  | 0 | 갈다 | 702102 | 유리  | 29535300 | 1  | 을 | 0 |
| 갈다 | 702002 | 관원   | 36771100 | 1 | 을  | 0 | 갈다 | 702102 | 인쇄판 | 30933800 | 1  | 을 | 0 |
| 갈다 | 702002 | 면도날  | 13358302 | 1 | 을  | 0 | 갈다 | 702102 | 인조석 | 30981301 | 1  | 을 | 0 |
| 갈다 | 702002 | 벼슬   | 16645201 | 4 | 을  | 0 | 갈다 | 702102 | 철봉  | 37221501 | 1  | 을 | 0 |
| 갈다 | 702002 | 벼슬아치 | 16645800 | 3 | 을  | 0 | 갈다 | 702102 | 칼   | 38567200 | 14 | 을 | 1 |
| 갈다 | 702002 | 상대   | 20166601 | 1 | 을  | 0 | 갈다 | 702102 | 판재  | 40156801 | 1  | 을 | 0 |
| 갈다 | 702002 | 인물   | 30889501 | 1 | 으로 | 1 | 갈다 | 702102 | 평판  | 40511801 | 1  | 을 | 0 |
| 갈다 | 702002 | 임금   | 31277900 | 1 | 을  | 0 | 갈다 | 702102 | 표면  | 40755601 | 4  | 을 | 0 |
| 갈다 | 702002 | 임원   | 31307600 | 1 | 을  | 1 | 갈다 | 702102 | 행실  | 42038000 | 1  | 을 | 0 |
| 갈다 | 702002 | 전문가  | 32998300 | 1 | 으로 | 1 | 갈다 | 702103 | 가루  | 145200   | 1  | 을 | 0 |
| 갈다 | 702002 | 직무   | 35917400 | 1 | 을  | 0 | 갈다 | 702103 | 감자  | 867202   | 2  | 을 | 0 |
| 갈다 | 702002 | 책임자  | 36932500 | 2 | 을  | 1 | 갈다 | 702103 | 개암  | 1237901  | 1  | 을 | 0 |
| 갈다 | 702002 | 처    | 36942100 | 1 | 을  | 1 | 갈다 | 702103 | 겉면  | 1813400  | 1  | 을 | 0 |
| 갈다 | 702002 | 친구   | 38346601 | 1 | 을  | 1 | 갈다 | 702103 | 결석  | 2006400  | 1  | 을 | 0 |
| 갈다 | 702102 | 가공품  | 29200    | 1 | 을  | 0 | 갈다 | 702103 | 고기  | 2554201  | 4  | 을 | 1 |
| 갈다 | 702102 | 거울   | 1495301  | 1 | 을  | 0 | 갈다 | 702103 | 곡식  | 3006100  | 5  | 을 | 0 |
| 갈다 | 702102 | 겉면   | 1813400  | 2 | 을  | 0 | 갈다 | 702103 | 과일  | 3543300  | 2  | 을 | 0 |
| 갈다 | 702102 | 고체   | 2921300  | 1 | 을  | 0 | 갈다 | 702103 | 국물  | 4425201  | 1  | 을 | 0 |
| 갈다 | 702102 | 공예품  | 3365100  | 1 | 을  | 0 | 갈다 | 702103 | 귀비탕 | 4896600  | 1  | 을 | 0 |
| 갈다 | 702102 | 공작물  | 3386901  | 2 | 을  | 0 | 갈다 | 702103 | 깨   | 6172100  | 1  | 을 | 0 |
| 갈다 | 702102 | 굽    | 4705803  | 1 | 을  | 0 | 갈다 | 702103 | 껍데기 | 6247101  | 2  | 을 | 0 |
| 갈다 | 702102 | 기록자낫 | 5810000  | 1 | 을  | 0 | 갈다 | 702103 | 껍질  | 6250001  | 1  | 을 | 0 |
| 갈다 | 702102 | 날    | 6869300  | 5 | 을  | 1 | 갈다 | 702103 | 나무  | 6621301  | 3  | 을 | 0 |
| 갈다 | 702102 | 날붙이  | 6898800  | 1 | 을  | 0 | 갈다 | 702103 | 나무  | 6621302  | 1  | 을 | 0 |
| 갈다 | 702102 | 낫    | 7055400  | 4 | 을  | 1 | 갈다 | 702103 | 난알  | 6867101  | 4  | 을 | 0 |
| 갈다 | 702102 | 대검   | 9018400  | 1 | 을  | 0 | 갈다 | 702103 | 난알  | 6867102  | 1  | 을 | 0 |
| 갈다 | 702102 | 대팻날  | 9374900  | 1 | 을  | 0 | 갈다 | 702103 | 녹두  | 7717900  | 8  | 을 | 1 |
| 갈다 | 702102 | 덕    | 9487801  | 1 | 을  | 0 | 갈다 | 702103 | 논밭  | 7776500  | 2  | 을 | 0 |
| 갈다 | 702102 | 도    | 9647002  | 1 | 을  | 0 | 갈다 | 702103 | 닭고기 | 8805100  | 1  | 을 | 0 |
| 갈다 | 702102 | 도끼날  | 9681001  | 1 | 을  | 0 | 갈다 | 702103 | 돌   | 10059301 | 2  | 을 | 0 |
| 갈다 | 702102 | 돌    | 10059301 | 6 | 을  | 0 | 갈다 | 702103 | 동부  | 10279801 | 1  | 을 | 0 |
| 갈다 | 702102 | 돌레   | 10645401 | 1 | 을  | 0 | 갈다 | 702103 | 두부콩 | 10580200 | 3  | 을 | 0 |
| 갈다 | 702102 | 메스   | 13279301 | 1 | 을  | 0 | 갈다 | 702103 | 땅바닥 | 11167401 | 1  | 을 | 0 |
| 갈다 | 702102 | 면    | 13350602 | 3 | 을  | 0 | 갈다 | 702103 | 땅콩  | 11176101 | 1  | 을 | 0 |
| 갈다 | 702102 | 면도칼  | 13358600 | 2 | 을  | 0 | 갈다 | 702103 | 메밀  | 13270602 | 3  | 을 | 0 |
| 갈다 | 702102 | 무예   | 14186200 | 1 | 을  | 0 | 갈다 | 702103 | 무   | 13960101 | 1  | 을 | 0 |
| 갈다 | 702102 | 반대쪽  | 15357700 | 1 | 을  | 0 | 갈다 | 702103 | 무   | 13960102 | 2  | 을 | 1 |
| 갈다 | 702102 | 배달말  | 15960100 | 1 | 을  | 0 | 갈다 | 702103 | 무리풀 | 14075600 | 1  | 을 | 0 |
| 갈다 | 702102 | 부엌칼  | 17633300 | 1 | 을  | 1 | 갈다 | 702103 | 물건  | 14463101 | 2  | 을 | 0 |
| 갈다 | 702102 | 북부   | 17800400 | 1 | 을  | 0 | 갈다 | 702103 | 물고기 | 14471100 | 2  | 을 | 0 |

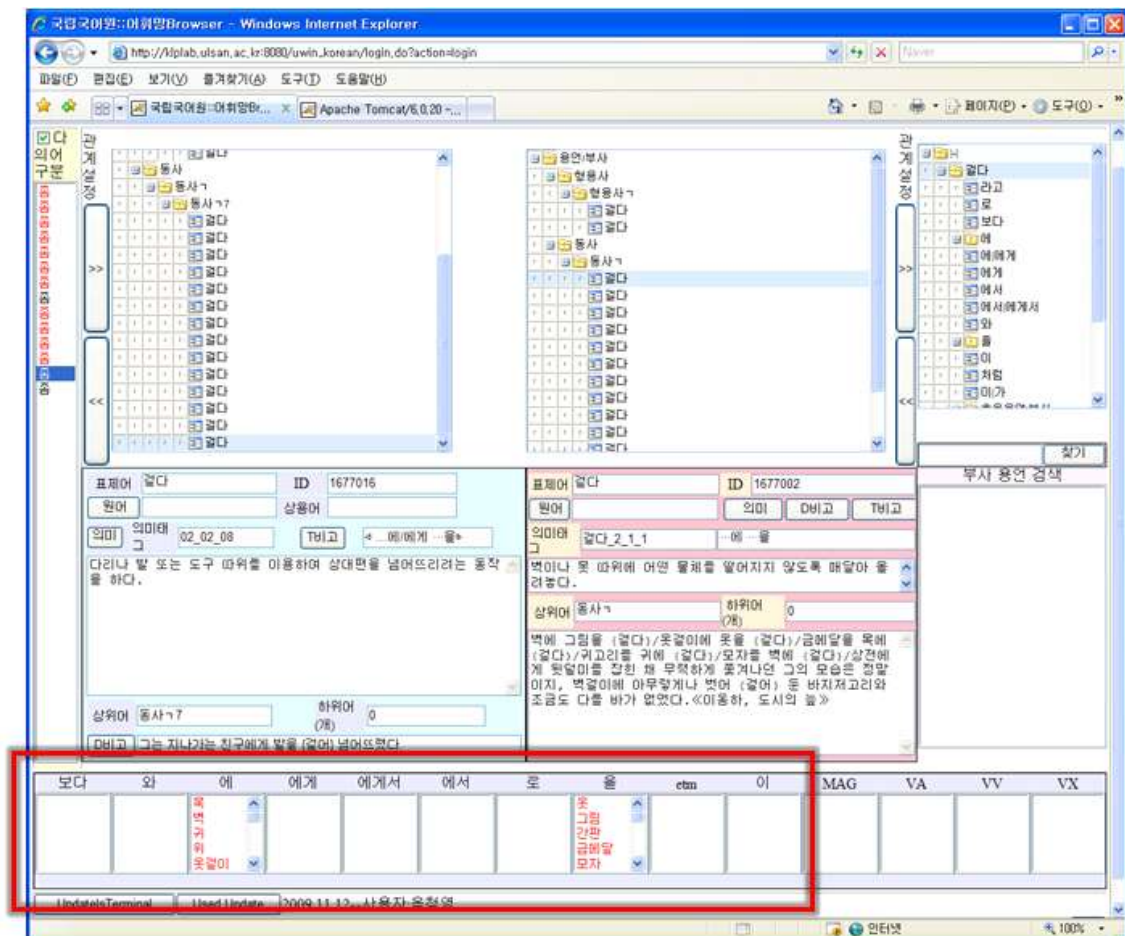
|    |        |     |          |   |   |   |    |        |      |          |   |   |   |
|----|--------|-----|----------|---|---|---|----|--------|------|----------|---|---|---|
| 갈다 | 702103 | 물질  | 14592703 | 1 | 을 | 0 | 갈다 | 702103 | 야채   | 25678302 | 2 | 을 | 0 |
| 갈다 | 702103 | 밀   | 14941100 | 1 | 을 | 0 | 갈다 | 702103 | 약    | 25690601 | 3 | 을 | 0 |
| 갈다 | 702103 | 밤   | 15672900 | 2 | 을 | 0 | 갈다 | 702103 | 약물   | 25706900 | 1 | 을 | 0 |
| 갈다 | 702103 | 밭   | 15914501 | 1 | 을 | 0 | 갈다 | 702103 | 약재   | 25737600 | 4 | 을 | 0 |
| 갈다 | 702103 | 백반  | 16138400 | 1 | 을 | 0 | 갈다 | 702103 | 얼음   | 26299101 | 2 | 을 | 0 |
| 갈다 | 702103 | 벼   | 16627601 | 5 | 을 | 0 | 갈다 | 702103 | 연근   | 26827900 | 1 | 을 | 0 |
| 갈다 | 702103 | 보리쌀 | 16952400 | 2 | 을 | 0 | 갈다 | 702103 | 연자매  | 26963100 | 1 | 을 | 0 |
| 갈다 | 702103 | 빼   | 18851001 | 1 | 을 | 0 | 갈다 | 702103 | 열매   | 27073200 | 3 | 을 | 0 |
| 갈다 | 702103 | 뿌리  | 18899801 | 2 | 을 | 0 | 갈다 | 702103 | 옛기름  | 27242200 | 2 | 을 | 0 |
| 갈다 | 702103 | 사과  | 19036800 | 1 | 을 | 1 | 갈다 | 702103 | 오렌지  | 27619400 | 1 | 을 | 0 |
| 갈다 | 702103 | 사항  | 19523300 | 1 | 을 | 0 | 갈다 | 702103 | 옥수수  | 27890802 | 3 | 을 | 0 |
| 갈다 | 702103 | 살   | 19821705 | 3 | 을 | 0 | 갈다 | 702103 | 은    | 29985600 | 1 | 을 | 0 |
| 갈다 | 702103 | 생강  | 20560802 | 1 | 을 | 0 | 갈다 | 702103 | 잡곡   | 32108700 | 1 | 을 | 0 |
| 갈다 | 702103 | 섬유질 | 21293401 | 1 | 을 | 0 | 갈다 | 702103 | 장목수수 | 32252400 | 1 | 을 | 0 |
| 갈다 | 702103 | 수수쌀 | 22638300 | 1 | 을 | 0 | 갈다 | 702103 | 조각   | 34116001 | 1 | 을 | 0 |
| 갈다 | 702103 | 쌀   | 24279101 | 2 | 을 | 0 | 갈다 | 702103 | 주사   | 34874500 | 1 | 을 | 0 |
| 갈다 | 702103 | 쌀   | 24279102 | 4 | 을 | 0 | 갈다 | 702103 | 차    | 36541902 | 1 | 을 | 0 |
| 갈다 | 702103 | 씨   | 24454001 | 1 | 을 | 0 | 갈다 | 702103 | 참쌀   | 36788901 | 1 | 을 | 0 |
| 갈다 | 702103 | 알약  | 25237301 | 1 | 을 | 0 | 갈다 | 702103 | 채소   | 36893400 | 1 | 을 | 0 |

<표 IV-4>에서 용언keyID 및 명사keyID는 U-WIN 데이터베이스 관리번호이며, 각각 구축 도구에서 해당 어휘로 직접 매핑한다. 빈도는 용언의 논항으로 나타난 명사의 빈도수이며, 격조사는 하위범주의 격이다. 비고는 용언의 용례에서 명사의 논항이 출현한 경우는 '1'로, 뜻풀이나 용언이 아닌 다른 품사의 용례에서 출현한 경우는 '0'이다. 이는 이번 하위범주 정보 구축 작업을 용례에서 나타난 논항만을 우선 처리하기 위해 구분하였다.

## 22 정보 구축 도구

뜻풀이 및 용례에서 추출한 각 용언 문형의 하위범주 정보를 명사 어휘망의 최소상계노드와 연결하기 위한 구축 도구는 다음 <그림 IV-1>과 같다.



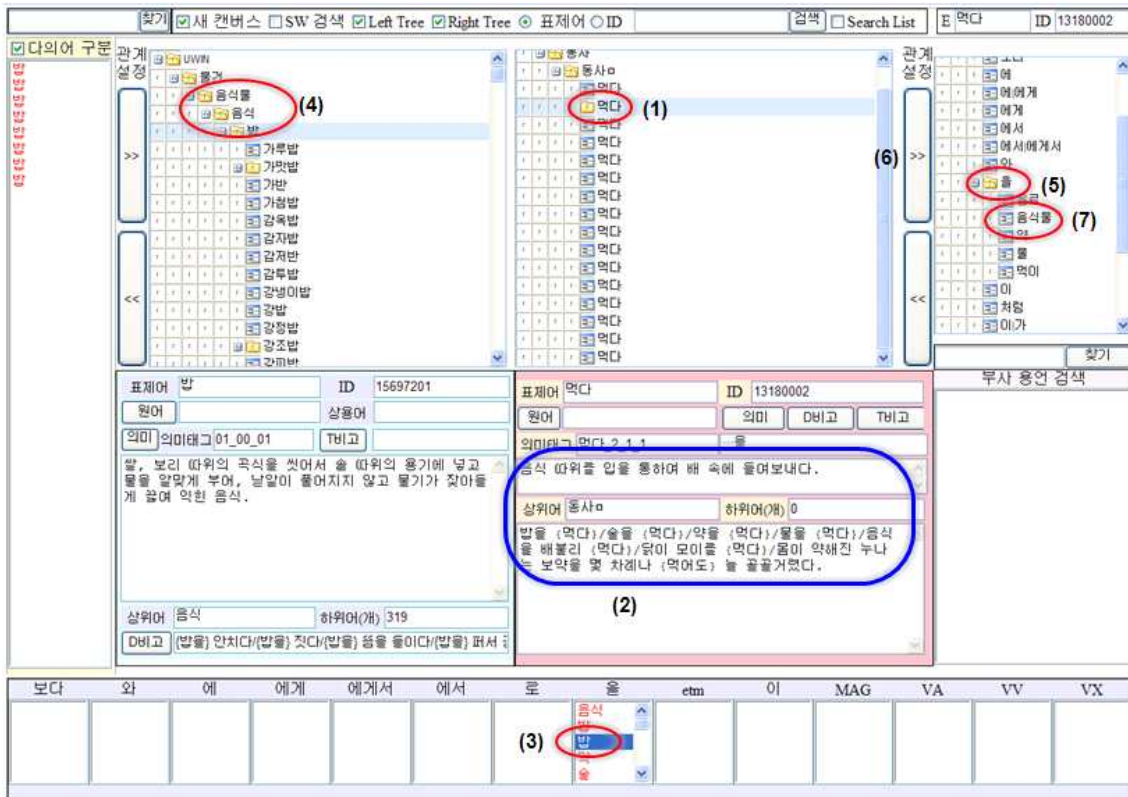


<그림 IV-8> 하위범주 구축 도구

<그림 IV-1>의 아래 쪽(붉은색 박스)에 추출된 하위범주 정보가 나타나며, 개별 하위범주는 직접 명사 어휘망과 연결되어 있다.

## 2.3. 정보 연결 과정

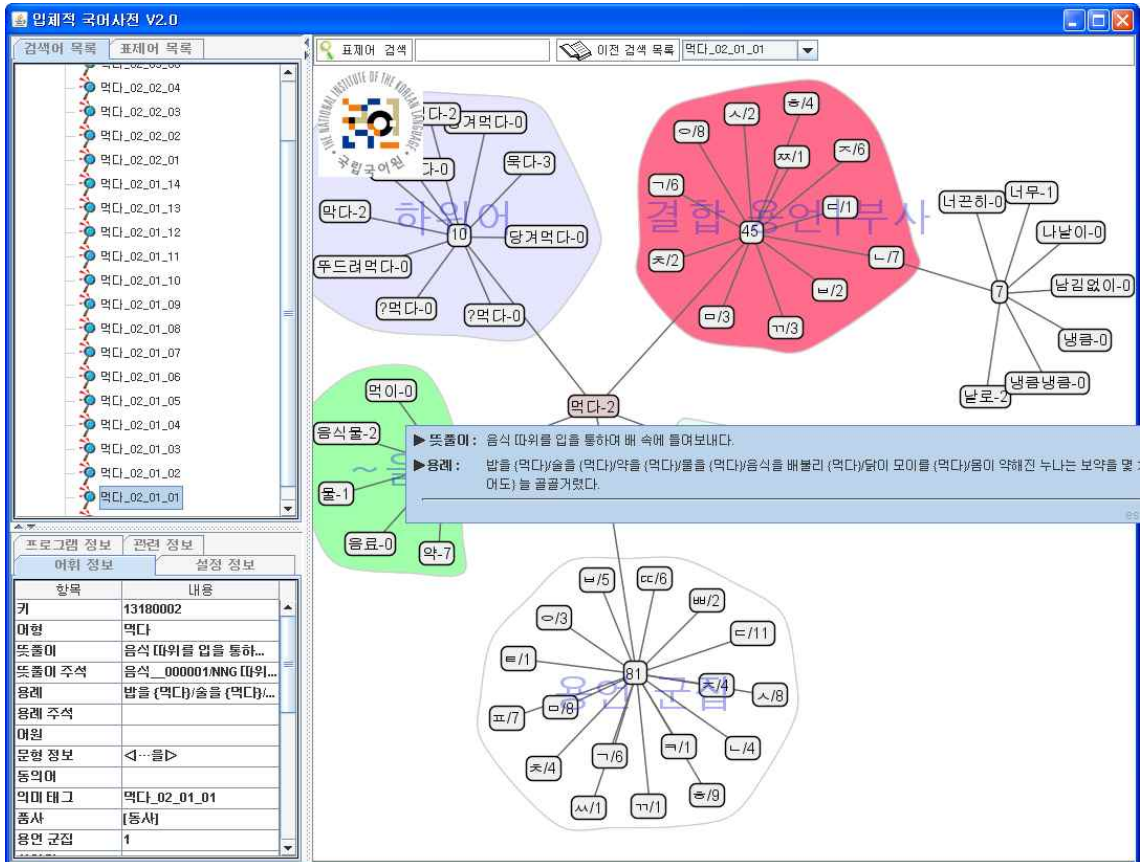
용언의 하위범주 정보 구축은 어휘망 편집 브라우저를 통해 이루어진다. 이 어휘망 편집 브라우저에서 용언의 하위범주 정보 연결 과정은, 용언의 용례와 의미에서 논항 후보를 추출하고 명사 어휘망에서 구축 지침에 따라 최소상계노드의 논항을 선택하여 관계 트리에다 넣는 것이다. 논항 후보의 다의어 수가 많을 경우에는 어휘망 편집 브라우저에서 다의어 구분을 사용한다.



<그림 IV-9> 하위범주 정보 연결 과정

- ① 검색 창에서 '먹다'를 입력하면 선정된 다의어 '먹다'가 출력되고, 이 다의어 중에서 특정 용언(예, '먹다\_02\_01\_01')을 선택한다.
- ② 선택된 용언의 뜻풀이와 용례를 확인한다.
- ③ 용언의 하위범주 정보(예, '밥')를 더블클릭하면, 명사 어휘망에서 {밥}의 계층구조 및 상위어/하위어를 확인한다. 여기서 뜻풀이와 용례에서 추출한 논항정보 중에 용언의 용례에서 추출한 논항정보는 붉은색으로 표기되며, 뜻풀이나 다른 품사의 용례에서 추출한 논항정보는 검은색으로 표기된다. 이번 연구과제에서는 붉은색으로 표기된 논항정보를 우선 처리하였다
- ④ '먹다\_02\_01\_01'과 의미적으로 목적격("을")으로 연결될 수 있는 최소상계 노드를 결정한다.
- ⑤ 구축하고자 하는 하위범주(결합관계, 예, "을")를 선택한다.
- ⑥ 관계설정 버튼을 누르면 ⑦과 같이 선택된 명사의 최소상계노드가 용언의 하위범주 관계로 설정된다.

이렇게 설정된 하위범주 정보는 <그림 IV-3>과 같이 입체적 국어사전에서도 확인할 수 있다.



<그림 IV-10> 입체적 국어사전에서 '먹다\_02\_01\_01'에 연결된 논항정보

위 입체적 국어사전 브라우저에서 동사 '먹다\_02\_01\_01'을 중심으로 '먹다\_02\_01\_01'의 목적어("을")로 올 수 있는 명사들을 명사 어휘망에서의 최소 상계노드인 {먹이, 음식물, 물, 음료, 약}이 연결되어 있으며, '먹다\_02\_01\_01'와 결합하는 부사 {너끈히, 너무, 나날이, 남김없이, 냉큼, 냉큼냉큼, 날로} 등이 구축되어 있음을 확인할 수 있다.

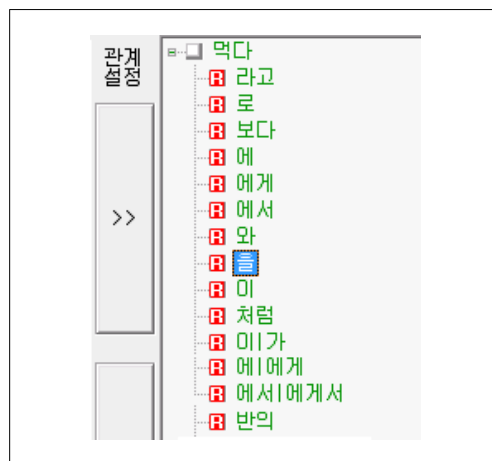
### 3. 하위범주 정보구축 지침

일관성 있는 하위범주 정보 구축을 위해서 다음과 같은 지침을 마련하였다.

㉠ 용언의 논항정보 구축 지침은 2008년 지침을 기본으로 삼는다.

(1-1) 용언 어휘망의 문형정보는 <표준>의 뜻풀이에 따른다.

2008년의 문형정보가 <표준>의 뜻풀이였으므로, 2009년 문형정보도 역시 그 뜻풀이에 따른다. 다만, 2009년은 편집기에서 문형정보 수를 <그림 IV-4>와 같이 간소화하였다. “-라고, -로, -보다, -에, -에게, -에서, -와, -을, -처럼, -이, -처럼, -이/가, -에/에게, -에서/에게서”로 간소화하였다. 이것은 구축 정보의 명확성을 보다 높이기 위한 것이다.

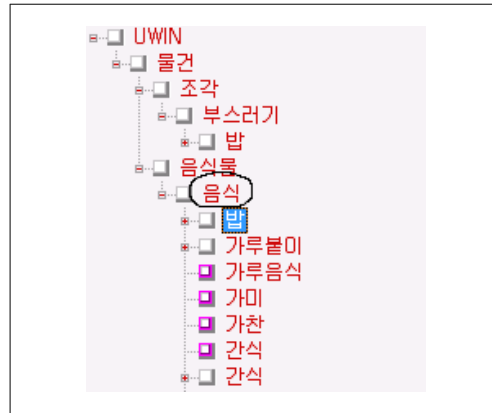


<그림 IV-11> 간소화된 문형정보

문형정보를 간소화하면 복잡 난해한 논항정보 연결 과정을 해결할 수 있다. 예를 들어, 문형정보를 간소화하기 전에는 {-로}가 지금처럼 하나가 아니라 “-로/대로, -로/라고, -로/에, -로/에/에게, -로/에/에게/에게로, -로/에게, -로/에게로, -로/에게로/에, -로/을” 등과 같이 여러 형태로 구성되어 있어서 논항정보 연결의 혼선을 빚을 수 있었다.

(1-2) 문형정보 해당 명사는 명사 어휘망의 최소상계노드와 연결한다.

기본적으로 2009년 용언의 논항정보 구축 작업에서도 최소상계노드를 기준으로 삼는다. 이것은 2008년 작업에서 기준이던 최소상계노드를 따르는 것이다. 이 기준은 국립국어원 2008년 정책연구용역제안서와 최종보고서에 근거한다. “용언 어휘망의 ‘먹다\_02\_01\_01’에는 문형정보 {-을}에 해당하는 목적어들을 명사 어휘망 최소상계노드와 연결해야 한다.”



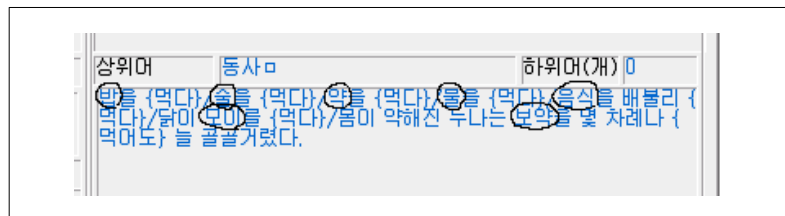
<그림 IV-5> {밥}의 최소상계노드

위의 사례처럼, '먹다\_02\_01\_01'은 "{밥}을 먹다"에서 {밥}을 바로 논항정보로 처리하지 않고 {밥}의 최소상계노드 명사인 {음식}을 논항정보로 처리한다. '먹다\_02\_01\_01'은 "{밥, 국수, 반찬, 빵}을 먹다"로 쓰인다. 작업에서는 이들 명사를 하나씩 논항으로 연결하기보다 모두를 아우르는 최소 범위의 상위노드 명사로서의 {음식}을 논항정보로 처리한다.

## ② 용언의 논항정보는 각 용언의 용례와 의미를 참조하여 구축한다.

### (2-1) 용언의 논항정보는 용례에 있는 논항을 기본정보로 삼는다.

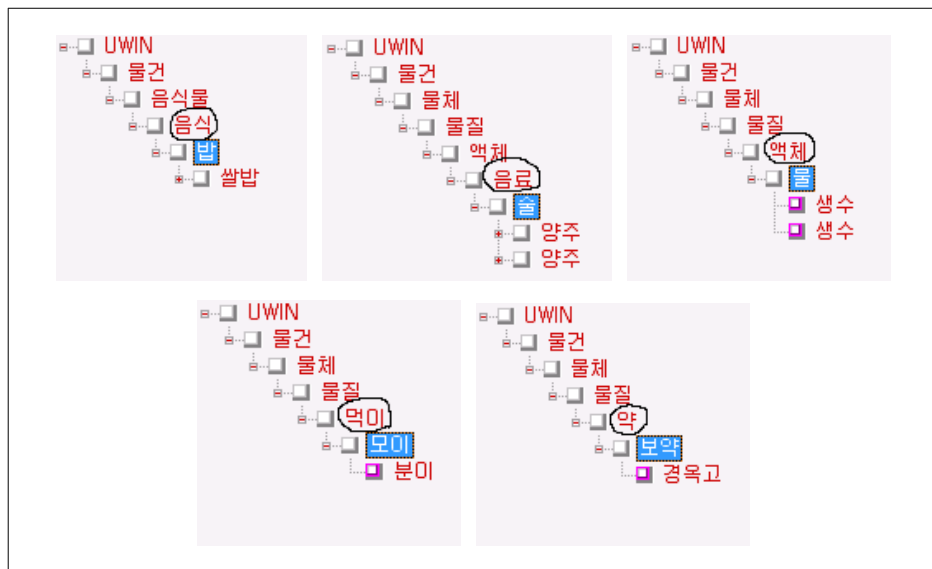
이 지침은 어휘망 편집 브라우저에 용언의 용례 정보가 들어있으므로 이것을 충분히 활용하고자 하는 입장에 준거한다. 용례에서 논항을 추출하는 것은 논항정보 구축 작업 자체의 경제적 효율성에도 관계가 있다. <그림 IV-6>의 '먹다\_02\_01\_01' 용례에서 '먹다'의 문형정보 {~을}과 결합된 논항 명사 {밥, 술, 약, 물, 음식, 모이, 보약}을 추출할 수 있다.



<그림 IV-6> '먹다\_02\_01\_01'의 용례

(2-2) 각 논항의 정보는 명사 어휘망에서 상위노드 명사로 연결한다.

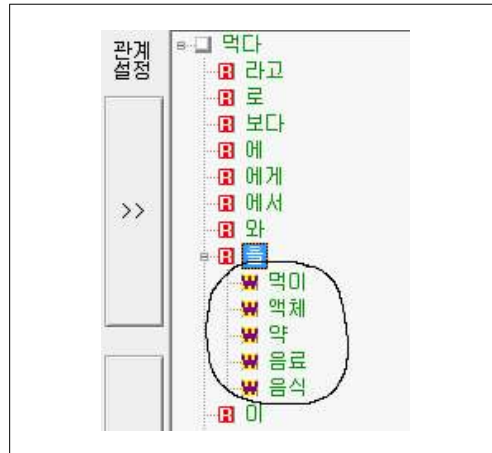
용례에서 추출한 명사들은 어휘망에서 바로 가져와서 용언의 논항정보로 넣는 것이 아니라, 지침 (1-2)의 최소상계노드 연결 기준에 따라 각각의 상위노드 명사를 선택하여 논항정보로 연결한다. 명사 {밥, 술, 약, 물, 음식, 모이, 보약}은 최소상계노드 {음식, 음료, 액체, 먹이, 약}으로 논항 설정되어 동사 ‘먹다\_02\_01\_01’의 논항정보로 처리된다.



<그림 IV-7> 논항 명사들의 최소상계노드 명사

지침 (2-2)까지의 작업 결과, ‘먹다\_02\_01\_01’의 논항정보 연결 상태를 살펴보면 <그림 IV-8>과 같다. 용례 정보가 풍부하더라도, 논항정보가 풍부해지는 것은 아니다. 논항이 지침 (2-2)에 따라 연결될 때는 논항 명사들의 상위노드 명사가 이들을 아우르면 논항정보 개수가 축소된다. 따라서 관계 트리에 보이는 논항정보 양의 다소와는 관련성이 적다.



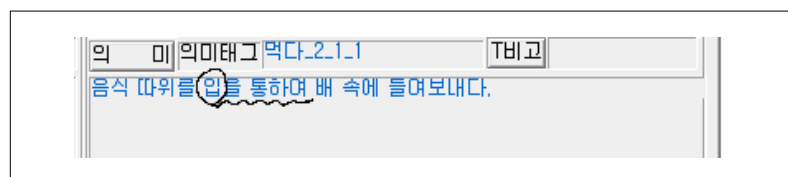


<그림 IV-8> 논항정보 연결 상태

### ③ 의미정보상 최소상계노드 연결이 부적합하면 노드를 조정한다.

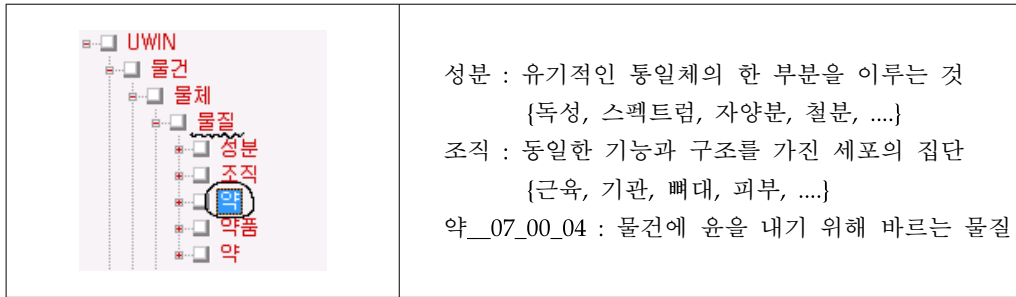
(3-1) 상위노드가 논항의 동위노드를 아우르지 못하면 자기노드로 연결한다.

지침 (2-1)의 <그림 IV-6>에 나오는 논항 명사들 중에서 “{약}을 먹다”의 ‘약\_07\_00\_01’은 최소상계노드로 바로 연결할 수 없다. 만일 그대로 연결하면 ‘먹다\_02\_01\_01’의 의미에 위배되는 논항정보가 된다. ‘약\_07\_00\_01’의 상위노드 명사는 “입을 통하여” 먹을 수 없는 약까지 포함하기 때문이다. 그래서 이 지침에 따라 ‘약\_07\_00\_01’은 자기노드로 연결한다.



<그림 IV-9> ‘먹다\_02\_01\_01’의 의미

‘약\_07\_00\_01’의 동위노드 명사들은 의미연관성 면에서 부적합하다. 그것의 동위노드에는 ‘먹다\_02\_01\_01’이 “입을 통하여” 들여보낼 수 없는 것을 포함하고 있는 {성분, 조직}뿐만 아니라 전혀 먹을 수 없는 물질인 ‘약\_07\_00\_04’도 있다. 상위노드로 연결하면 맞지 않다. 이런 경우가 더러 있으니, 논항정보를 연결할 때 지침 (3-1)로도 동시에 살펴야 한다.



<그림 IV-10> '약\_07\_00\_01'의 동위노드 명사들

(3-2) 한 용언의 두 논항이 상하 관계에 있으면 상계 상위노드로 연결한다.

한 용언의 두 논항이 지침 (2-2)에 따라 최소상계노드로 처리되어도, 그 결과가 상하 관계에 놓여있으면 지침 (3-2)로 조정해야 한다. 앞서 '먹다\_02\_01\_01'의 논항 {음식, 밥}은 상위노드 {음식물}로 처리될 수 있겠으나, 이것은 일차적인 것이다. <그림 IV-11>처럼 {음식, 밥}은 상하 관계에 있어서 {밥}은 {음식}보다 더 상위 단계의 {음식물}로 최종 연결한다.



<그림 IV-11> 상하 관계의 {음식, 밥}

그러면 앞선 <그림 IV-8>의 논항정보 연결 상태는 지침 (3-2)의 적용으로 <그림 IV-12>와 같은 최종 연결 상태로 수정된다. 용언의 논항정보를 연결할 때는 하나의 지침만으로 결정되는 것이 아니다. 한 용언의 논항들 내에서도 다양한 형태의 상하 관계가 나타나기도 하므로, 다수의 지침으로 동시에 살펴봐야 한다. 그때 이런 수정 과정이 발생된다.

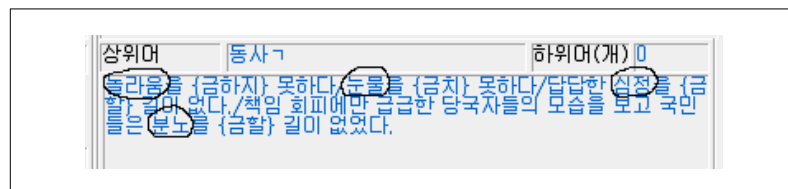




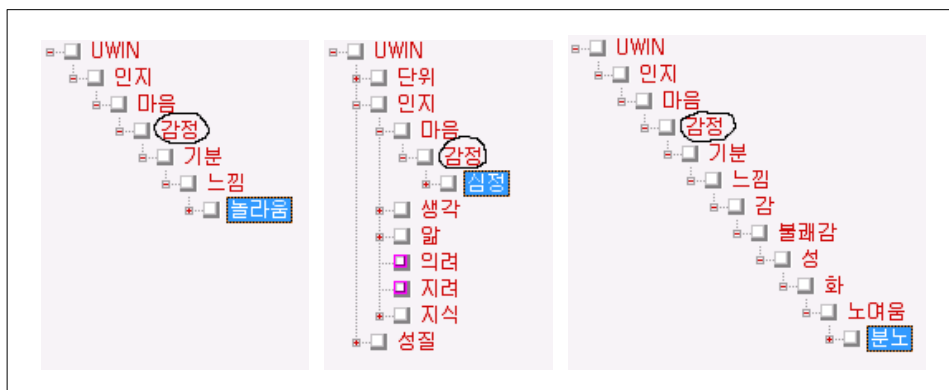
<그림 IV-12> 최종 연결 상태

(3-3) 논항들이 유사 의미장에 같은 줄기이면 공통 상위노드로 연결한다.

한 용언의 논항들이 유사 의미장에 속하고 최소상계노드를 설정할 때도 결국 같은 줄기의 하위어이면 공통 상위노드로 연결한다. ‘금하다\_\_00\_00\_02’에서 {~을}과 결합하는 논항 {놀라움, 심정, 분노}는 최소상계에 있는 노드가 각각 <그림 IV-14>와 같이 {느낌, 감정, 노여움}이지만 상위노드 {감정}의 한 줄기에 있으므로 {감정}을 공통 상위어로 처리한다.



<그림 IV-13> ‘금하다\_\_00\_00\_02’의 용례



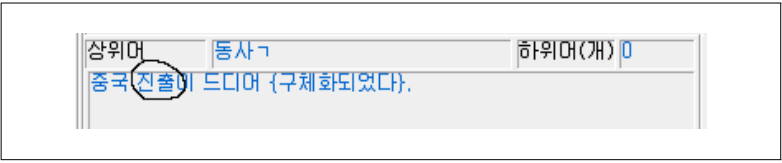
<그림 IV-14> 같은 줄기의 상위노드

지침 (3-3)은 (3-2)에서 확대되는 논항들을 취급한다. 대개는 지침 (3-2)로

처리되는 편이지만, 유사 의미장의 논항들 수가 둘 이상으로 늘어날 때에 처리하기 유용한 지침이다. 만일 추출했을 때는 유사 의미장의 논항들이지만 같은 줄기에 속해 있지 않으면, 각각의 상위노드를 선택해야 하고 제일 높은 공통 상위노드의 명사를 선택할 수 없다.

(3-4) 두 논항 형태가 같고 상하 관계이면 의미가 넓은 상위노드로 연결한다.

‘구체화되다\_\_00\_01\_02’의 용례에서 문형정보 {-이/가}와 결합하는 명사 {진출}을 추출할 수 있다. 이것을 명사 어휘망에서 살펴보면 상하 관계의 두 논항 형태로 나타난다. 이들이 의미상으로 확연히 구분되는 것이면 논항정보를 연결하기가 쉽지만, 그렇지 않을 경우 의미가 넓은 상위노드 쪽을 선택해야 한다. 전체 용언에서 이 같은 경우가 많다.

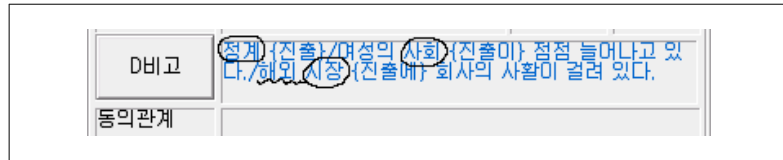


<그림 IV-15> ‘구체화되다\_\_00\_01\_02’의 의미

|  |                                                                                                                                                  |
|--|--------------------------------------------------------------------------------------------------------------------------------------------------|
|  | <p>진행 : 앞으로 향하여 나아감.<br/>         진출__02_00_01 : 어떤 방향으로 활동 범위를 넓혀 나아감.<br/>         전진 : 앞으로 나아감. ≒ 진행<br/>         진출__02_00_02 : 앞으로 나아감.</p> |
|--|--------------------------------------------------------------------------------------------------------------------------------------------------|

<그림 IV-16> 상하 관계에 있는 두 {진출}

두 논항 중에 상위의 것을 선택하면 의미가 넓은 상위노드이다. <그림 IV-16>과 같이 {진출\_\_02\_00\_02}보다 상위의 {진출\_\_02\_00\_01}을 선택하여 그것의 최소상계노드를 논항정보로 연결한다. <그림 IV-17>처럼 {진출\_\_02\_00\_01}의 의미정보를 보면 적합한 정도를 알 수 있다. “활동 범위를 넓혀 나가는” 진출의 의미가 ‘구체화되다\_\_00\_01\_02’에 더 적합해 보인다.

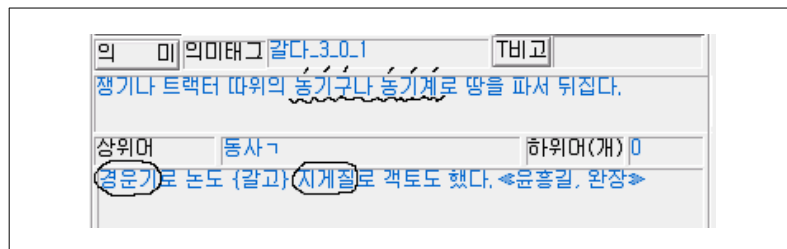


<그림 IV-17> '진출\_02\_00\_01'의 의미

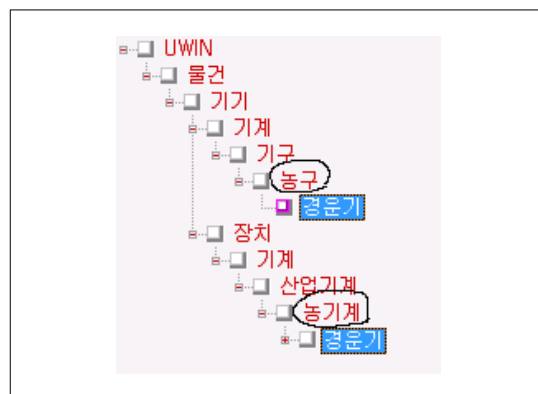
#### 4] 논항이 여러 노드에 걸쳐 나타나면 의미정보로 노드를 설정한다.

(4-1) 논항의 두 동형이 의미정보에 포함되면 복수 상위노드로 설정한다.

'갈다\_03\_00\_01'의 용례에서 {~로}와 결합된 {경운기}가 나오는데, 명사 어휘망에서 {경운기}를 찾으려면 두 개의 {경운기}가 나타난다. 지침 (3-3)이나 (3-4)에 예처럼 한 줄기에 소속된 것이 아니므로 <그림 IV-18>처럼 의미정보를 살펴볼 필요가 있다. '갈다\_03\_00\_01'의 의미정보에는 "농기구나 농기계"로 나오므로 각각의 상위노드를 결정할 수 있다.



<그림 IV-18> '갈다\_03\_00\_01'의 의미

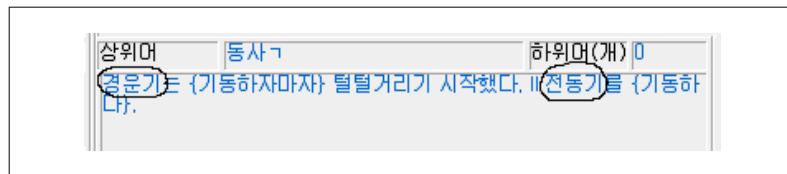


<그림 IV-19> 다른 줄기의 {경운기}

지침 (4-1)은 ‘갈다\_03\_00\_01’의 경우처럼 두 {경운기} 중에 어느 한쪽을 논항정보로 골라 관례적으로 처리하는 문제를 해결한다. ‘갈다\_03\_00\_01’의 의미정보는 “손으로 가는 {농구}로서의 {경운기}”와 “동력을 사용하는 {농기계}로서의 {경운기}”는 각각 포함하므로, 다른 줄기인 것을 감안하여 복수 상위노드를 설정하는 것이다. 의미적 요소가 중요하다.

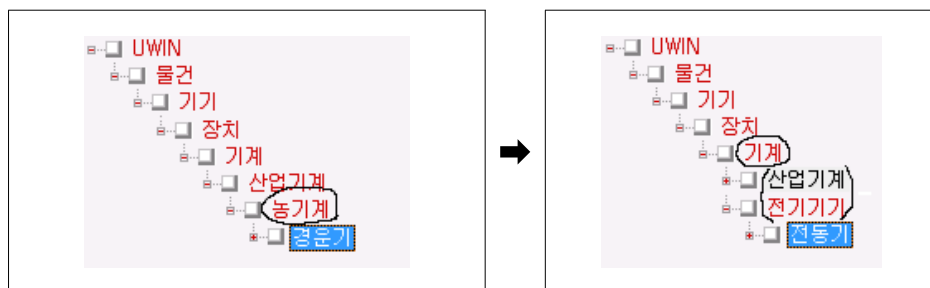
(4-2) 논항이 동형이라도 의미정보가 다르면 단수 상위노드로 설정한다.

‘기동하다’는 논항에 {경운기}가 들어있지만, 지침 (4-1)이 적용되는 ‘갈다\_03\_00\_01’과 {경운기}와 달리 의미정보의 제약을 받는다. 이때의 경운기는 <그림 IV-20>와 같이 “{농구}로서의 {경운기}”가 아니라 “동력을 사용하는 {경운기}”를 의미하기 때문이다. 그러면 복수 상위노드의 것을 설정할 수 없고 의미에 맞는 단수 상위노드의 것을 선택해야 한다.



<그림 IV-20> ‘기동하다’의 의미

여기에서 단수 상위노드로 처리된 ‘기동하다’의 논항 {경운기}는 한 용언 내의 다른 논항 {전동기}와의 관계에 의해, 지침 (4-2)에 이어 (3-2)의 적용을 받는다. {경운기}는 {전동기}와 한 줄기에 소속되어 있으므로, {농기계<산업기계}로서의 {경운기}와 {전기기기}로서의 {전동기}는 {기계}를 공통 상위노드로 받고, {기계}가 최종 논항정보이다.

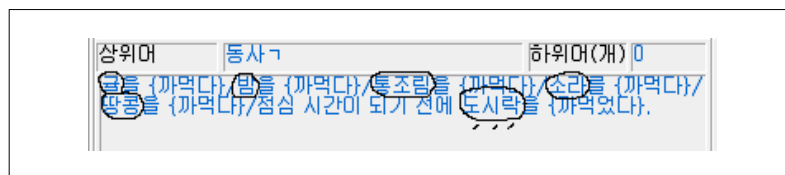


<그림 IV-21> {경운기, 전동기}의 상위노드

⑤ 논향이 상위노드를 설정하지 못하는 상태이면 제자리를 유지한다.

(5-1) 논향의 상위노드를 설정하면 그 의미가 부적합할 때 자기노드에 둔다.

최소상계노드를 준수하는 지침 (2-2)에 의해 대부분의 논향은 상위노드를 설정하지만, 그렇게 하면 용언 본래의 의미에 부적합한 경우도 있다. ‘까먹다\_\_00\_00\_01’에서 {도시락}은 상위노드의 {그릇}을 설정하면 의미상 {까먹다}에 적합하지 않다. “{도시락}을 까먹다”는 통용되어도 “{그릇}을 까먹다”는 어색하다. 그래서 {도시락}을 그대로 연결한다.



<그림 IV-22> ‘까먹다\_\_00\_00\_01’의 용례



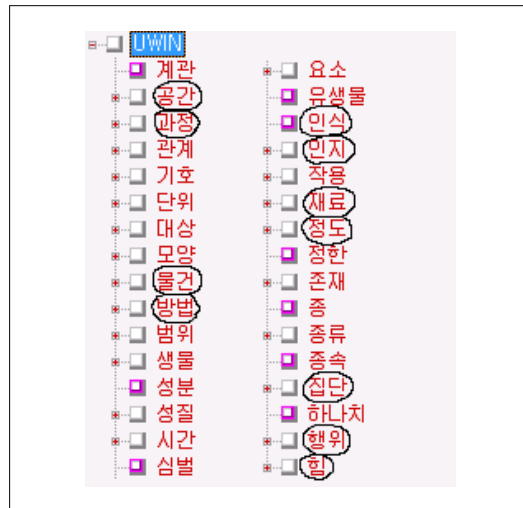
<그림 IV-23> {도시락}의 위치

(5-2) 논향이 최상위노드일 때 그 의미가 포괄적이더라도 자기노드에 둔다.

상위노드를 설정하다 보면 최상위노드까지 올라가는 경우가 있다. 최소상계노드로 결정하여 최상위노드 논향이 되거나, 최상위노드 논향 자체가 자기노드여서 더 이상 올리지 못하는 경우도 있다. 논향이 최상위노드이면 최소상계노드의 역할 면에서 적절하지는 않다. 최상위노드의 논향 자체가 너무 큰 덩어리여서 의미상 포괄적인 면도 크다.

하지만 최상위노드 논향들에 어떤 제약을 가하는 것은 일단 주의할 필요가 있다. 앞선 지침들로 논향정보를 구축하는 작업의 일관성과 관계있기 때문이다. 지침을 따로 부가해도 다른 지침의 적용 형평성에 영향을 끼칠 수 있다.

의미론적 논의가 필요하다. 그래서 최상위노드 논항은 의미가 포괄적이더라도 자기노드에 두는 지침을 적용한다.



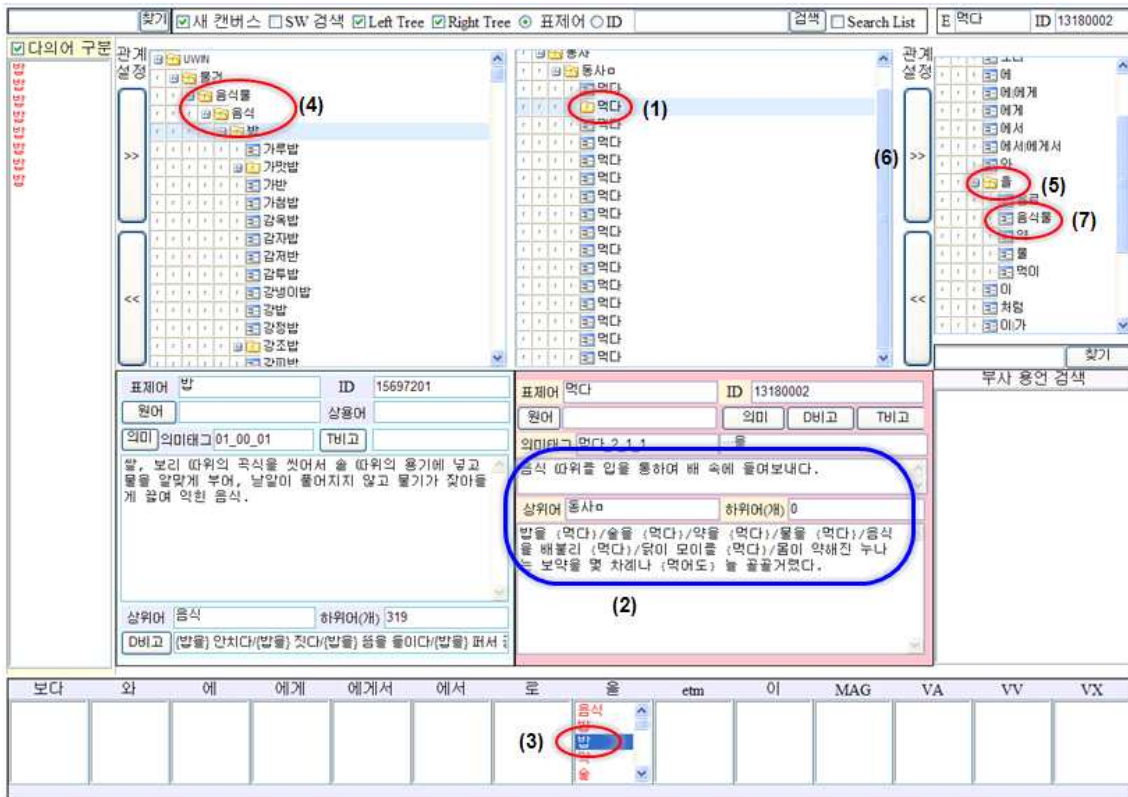
<그림 IV-24> 최상위노드 목록

<그림 IV-24>는 명사 어휘망의 최상위노드에 있는 명사 리스트이다. 동그라미로 표시한 것은 논항정보를 연결할 때 최상위노드 중에서도 출현 빈도가 높은 항목들이다. 위의 {공간, 과정, 물건, 방법, 인식, 인지, 재료, 정도, 집단, 행위, 힘} 중에서도 특히 {물건, 방법, 행위, 힘} 등은 가장 빈번하게 상위노드의 논항정보로 연결되고 있는 것들이다.

## 4. 배제 어휘(N항) 처리

### 4.1. 하위범주 정보 구축

용언의 하위범주 정보 구축 작업은 어휘망 편집 브라우저를 통해 이루어진다. 어휘망 편집 브라우저에서 용언의 하위범주 정보 연결 과정은, 용언의 용례와 의미에서 논항 후보를 추출하고 명사 어휘망의 최소상계노드의 논항을 선택하여 관계 트리에 연결하는 작업이다. 논항 후보의 다의어 수가 많을 경우에는 어휘망 편집 브라우저에서 다의어 구분을 사용한다.

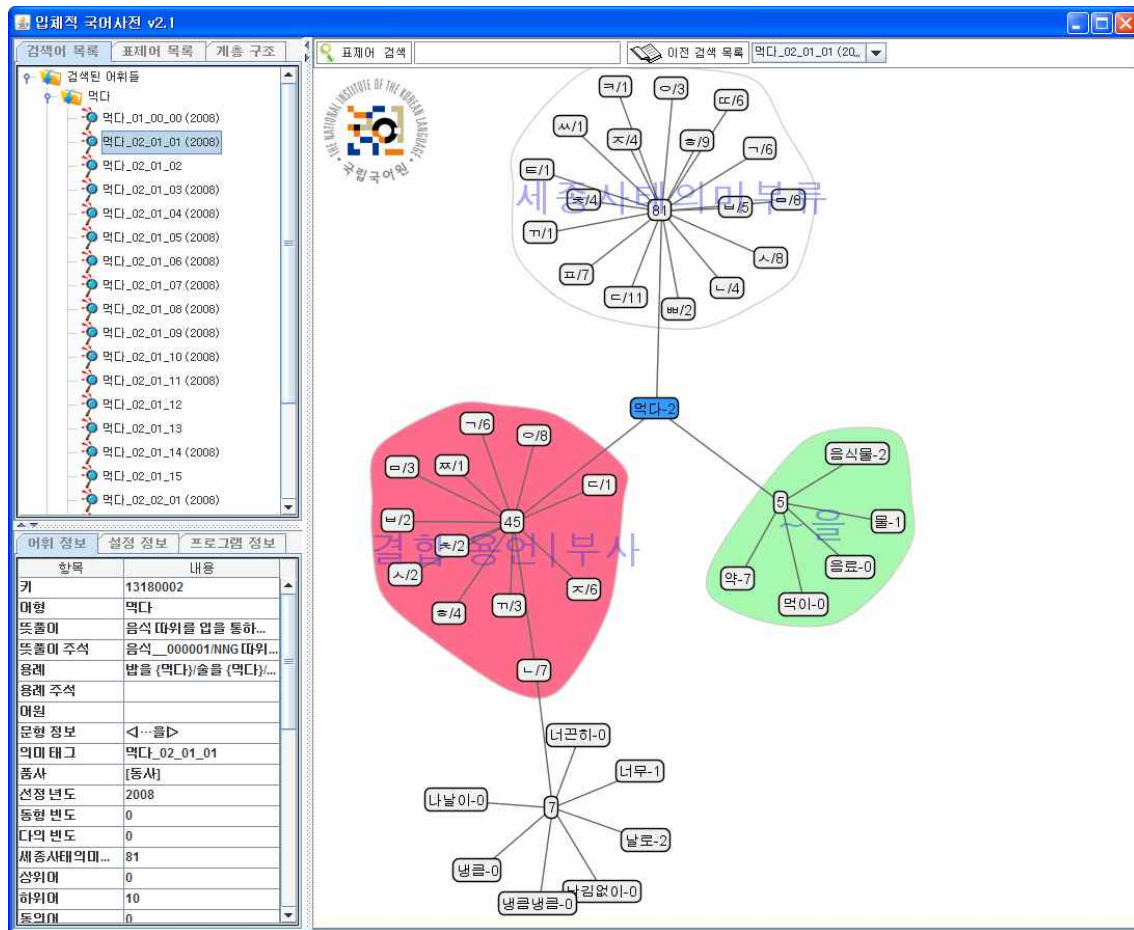


<그림 III-43> 하위범주 정보 구축 과정

- ① 검색 창에서 '먹다'를 입력하면 선정된 다의어 '먹다'가 출력되고, 이 다의어 중에서 특정 용언(예, '먹다\_02\_01\_01')을 선택한다.
- ② 선택된 용언의 뜻풀이와 용례를 확인한다.
- ③ 용언의 하위범주 정보(예, '밥')를 더블클릭하면, 명사 어휘망에서 {밥}의 계층구조 및 상위어/하위어를 확인한다. 여기서 뜻풀이와 용례에서 추출한 논항정보 중에 용언의 용례에서 추출한 논항정보는 붉은색으로 표기되며, 뜻풀이나 다른 품사의 용례에서 추출한 논항정보는 검은색으로 표기된다.
- ④ '먹다\_02\_01\_01'의 목적격("을")으로 의미적으로 연결될 수 있는 최소상계노드(여기서는 '음식물')를 결정한다.
- ⑤ 구축하고자 하는 하위범주(결합관계, 예, "을")를 선택한다.
- ⑥ 관계설정 버튼을 누르면 ⑦과 같이 선택된 명사의 최소상계노드가 용언의 하위범주 관계로 설정된다.

이렇게 설정된 하위범주 정보는 <그림 III-2>와 같이 입체적 국어사전에서

도 확인할 수 있다.



<그림 III-44> 입체적 국어사전에서 '먹다\_02\_01\_01'에 연결된 논항정보

위 입체적 국어사전 브라우저에서 동사 '먹다\_02\_01\_01'을 중심으로 '먹다\_02\_01\_01'의 목적어("을")로 올 수 있는 명사들을 명사 어휘망에서의 최소 상계노드인 {먹이, 음식물, 물, 음료, 약}이 연결되어 있으며, '먹다\_02\_01\_01'과 결합하는 부사 {너끈히, 너무, 나날이, 남김없이, 냉큼, 냉큼냉큼, 날로} 등이 구축되어 있음을 확인할 수 있다.

## 4.2. 배제 어휘(N항)

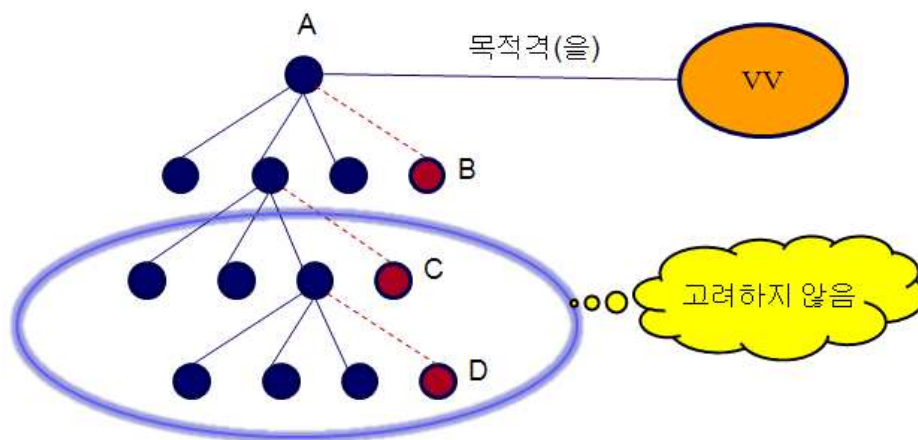
4.1절의 과정을 통해 구축된 용언의 하위범주 정보는 용언의 해당 의미와 결합할 수 있는 명사 어휘망의 최소상계노드이다. 명사 어휘망은 명사의 뜻



풀이의 본질적 의미를 통해 상하위 관계가 설정된 것으로 의미의 개념적 포괄 관계를 나타낸다. 그렇지만 개념적인 상하 관계의 어휘가 특정 용언과 함께 사용될 때는 의미적으로 결합되지 못하는 경우도 있다. 이런 점에서 U-WIN은 세종의미부류와 다르다.

예를 들어, <그림 III-2>의 동사 '먹다\_02\_01\_01(음식 따위를 입을 통하여 배 속에 들여보낸다.)'의 목적어 논항에 연결된 최소상계노드 {먹이, 음식물, 물, 음료, 약}에서 {약(병이나 상처 따위를 고치거나 예방하기 위하여 먹거나 바르거나 주사하는 물질.)}은 U-WIN 상에서 368개의 하위어를 가지고 있으며, 이 하위어 중 {각질용해제, 고도리, 금약, 근이완제, 금황산, 붕대액, 비듬약, 비누약, ...} 등의 70개의 하위어는 동사 '먹다\_02\_01\_01'와 의미적으로 결합할 수 없는 어휘{바르는약, 붙이는약, 주사약, 해충약 류}들이다.

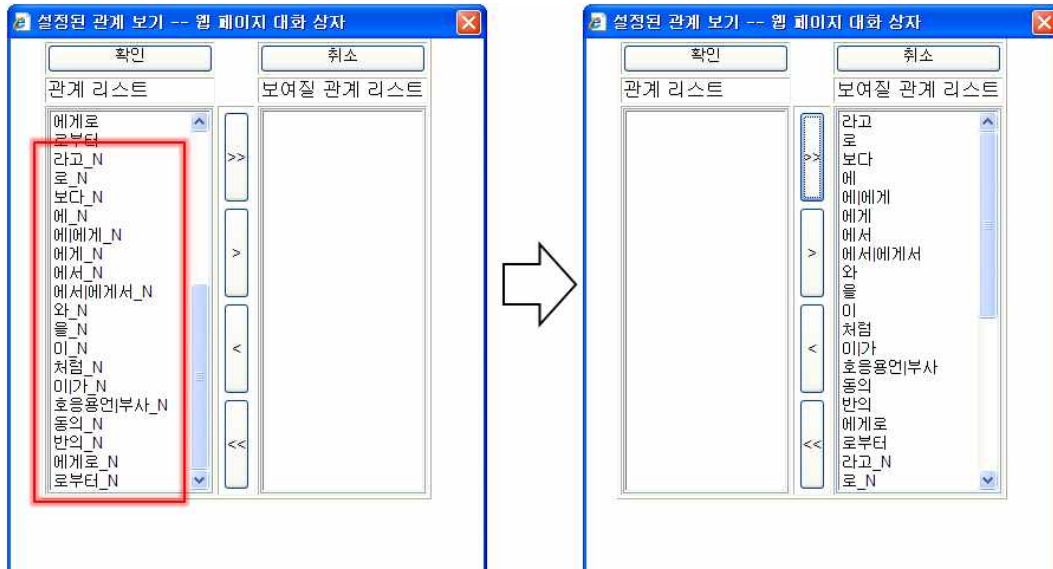
또한 최소상계노드의 하위어가 다음 그림과 같이 또 다른 하위어를 가진 중간 노드일 수 있다.



위 그림에서 B, C, D 노드가 용언 VV와 의미적으로 결합할 수 없는 배제 어휘라고 할 때, B뿐만 아니라 C와 D로 배제 어휘로 설정되어야 한다. 그러나 C나 D와 같이 최소상계노드의 하위어의 하위어 등 상하 관계의 모든 하위어를 살펴보는 것은 현실적으로 많은 시간이 소요된다. 따라서 이번 연구에서는 최소상계노드 A의 바로 하위어만을 대상으로 배제 어휘를 설정하였다.

#### 4.3. 배제 어휘 처리를 위한 어휘망 편집 도구

배제 어휘 설정을 위해 어휘망 편집 도구를 다음과 같이 기능 개선하였다.



<그림 III-3> 의미 관계 설정 창

위 그림의 왼쪽 부분에 배제 어휘를 설정하기 위한 의미 관계 {이\_N, 을\_N 등}을 볼 수 있다. <그림 III-1>의 어휘망 편집 브라우저에서 용언을 선택한 후 마우스의 오른쪽 버튼을 누르면 위 <그림 III-3>의 의미 관계 설정 창의 왼쪽 리스트가 화면에 나타나고 설정할 의미 관계를 선택한 후 화살표 버튼을 눌러 선택된 의미 관계를 포함시킬 수 있다. 이후 최소상계노드의 하위어를 보면서 배제시킬 어휘를 선택하면 해당 어휘가 배제 관계의 배제 어휘로 설정된다<sup>9)</sup>.

배제 어휘는 의미별 용언에 연결된 최소상계노드의 모든 하위어를 살펴보고 결정해야 한다.

9) 편집 브라우저의 사용법은 『어휘의미 체계 기반 입체적 국어사전 확장』 (옥철영2009) 참조.

## V. 결합관계 정보 구축

### 1. 부사의 결합관계 구축 과정

#### 1.1. 부사의 결합관계 추출

부사는 용언을 비롯한 문장의 여러 성분(체언, 관형사, 부사 등)을 수식하는 성분부사와, 문장 전체를 수식하는 문장 부사로 나뉜다. 또한 부사는 형태 구성의 관점에서도 품사가 (1) 원래 부사인 것, (2) 파생부사, (3) 합성부사, (4) 조사가 결합한 형태, (5) 어미가 결합한 형태 등 다양한 형태가 있다.

- (1) 품사가 부사 : 매우, 좀, 아주, 썩, ...
- (2) 파생부사 : 자연히, 없이, 같이, 많이, 달리, 높이, 빨리, ...
- (3) 합성부사 : 밤낮, 요즈음, 잘못, 좀더, ...
- (4) 조사가 결합한 형태 : 각중에, 순식간에, 정말로, 별도로, 되는데로, 그나마, 그만큼, ...
- (5) 어미가 결합한 형태 : 가볍게, 격렬하게, 되도록, 어쩌면, 끝없이, 각설하고, 못해도, ...

또한, 특정 부사(예, '결코')는 의미적으로 부정형태 혹은 보조용언과도 결합하는 등 다양한 결합 형태를 보인다.

본 과제에서는 2008년에 선정된 2,640개의 부사에 대해, 부사-부사, 부사-용언과 의미적으로 결합하는 성분부사에 한해서 결합관계를 구축하였다.

용언의 하위범주 정보 추출과 마찬가지로 다의어 수준까지 의미 태깅된 뜻풀이 및 용례에서 각 부사의 결합관계를 추출하였다.

다의어 태깅된 뜻풀이/용례의 문장이 “.....(1)..... [부사\_1][부사\_2] .....(2)..... [용언\_1][용언\_2]”와 같을 경우,

- [부사\_1]의 결합관계는 [부사\_2], [용언\_1], [용언\_2]을
- [부사\_2]의 결합관계는 [용언\_1], [용언\_2]을 추출할 수 있다.

예를 들어, 다음 표는 ‘가급적’과 ‘가까스로’에 대해 뜻풀이와 용례에서 추

출된 결합관계의 부서, 용언이다.

<표 V-1> 뜻풀이와 용례에서 추출한 {가급적, 가까스로}와 결합하는 부서 및 용언

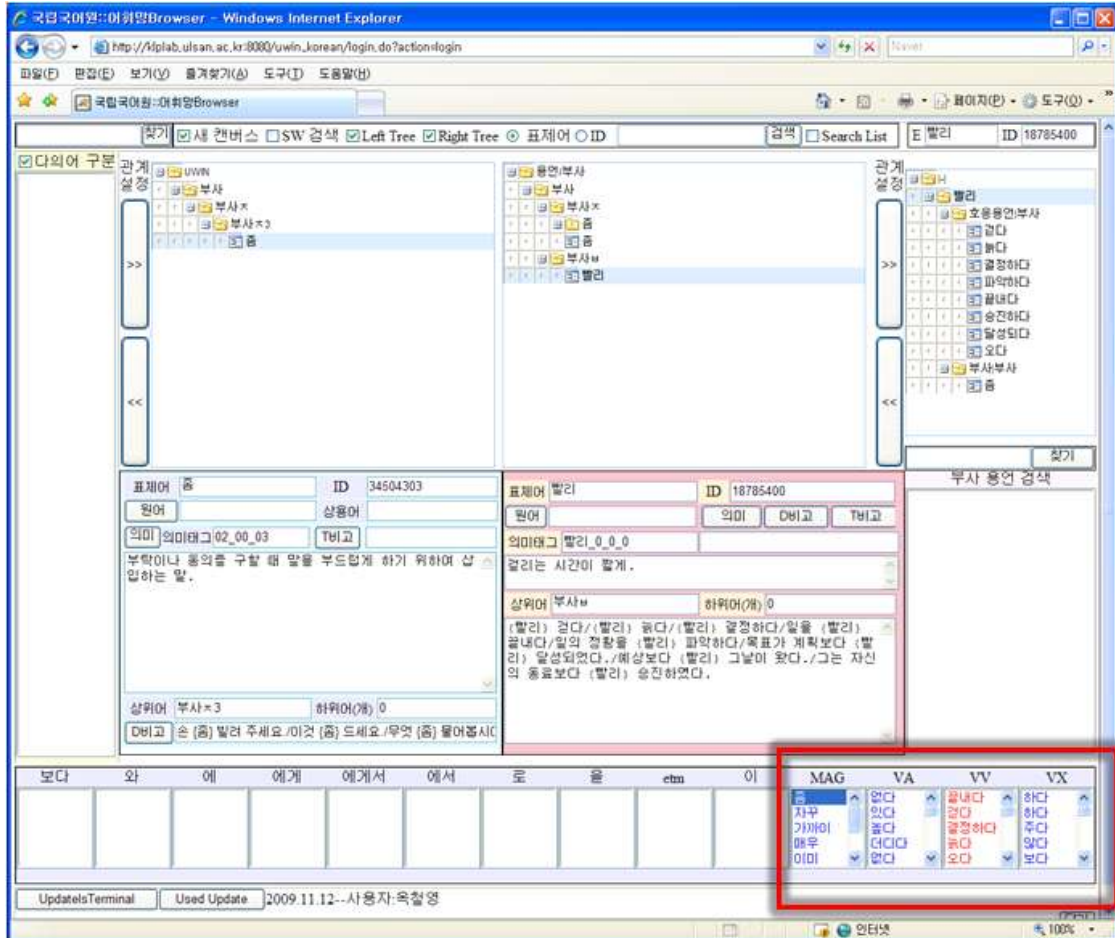
| 부서   | 부서keyID | 호응부사/용언 | 결합부사/용언 keyID | 품사  | 빈도 | 비고 |
|------|---------|---------|---------------|-----|----|----|
| 가급적  | 40802   | 많이      | 12743400      | MAG | 1  | 1  |
| 가급적  | 40802   | 안       | 25001300      | MAG | 1  | 0  |
| 가급적  | 40802   | 노력하다    | 7590002       | VV  | 2  | 0  |
| 가급적  | 40802   | 도와주다    | 9845200       | VV  | 1  | 1  |
| 가급적  | 40802   | 되다      | 10449612      | VV  | 1  | 0  |
| 가급적  | 40802   | 먹다      | 13180005      | VV  | 1  | 0  |
| 가급적  | 40802   | 바라다     | 15123901      | VV  | 1  | 1  |
| 가급적  | 40802   | 복원하다    | 17212404      | VV  | 2  | 0  |
| 가급적  | 40802   | 빠르다     | 18762403      | VA  | 1  | 1  |
| 가급적  | 40802   | 상대되다    | 20166607      | VV  | 1  | 0  |
| 가급적  | 40802   | 않다      | 25168402      | VX  | 1  | 0  |
| 가급적  | 40802   | 적다      | 32732400      | VA  | 2  | 0  |
| 가급적  | 40802   | 크다      | 38859502      | VA  | 1  | 0  |
| 가급적  | 40802   | 해치다     | 41962301      | VV  | 1  | 0  |
| 가급적  | 40802   | 획득하다    | 43477804      | VV  | 1  | 0  |
| 가까스로 | 42701   | 꿀꺽꿀꺽    | 6439802       | MAG | 1  | 1  |
| 가까스로 | 42701   | 가다      | 68941         | VX  | 1  | 0  |
| 가까스로 | 42701   | 가혹하다    | 473402        | VA  | 1  | 0  |
| 가까스로 | 42701   | 건지다     | 1645403       | VV  | 1  | 0  |
| 가까스로 | 42701   | 건다      | 1665902       | VV  | 1  | 0  |
| 가까스로 | 42701   | 걸치다     | 1702006       | VV  | 1  | 0  |
| 가까스로 | 42701   | 계속하다    | 2437313       | VV  | 1  | 0  |
| 가까스로 | 42701   | 구하다     | 4053503       | VV  | 1  | 1  |
| 가까스로 | 42701   | 내다      | 7130823       | VV  | 1  | 0  |
| 가까스로 | 42701   | 내다      | 7130833       | VX  | 1  | 0  |
| 가까스로 | 42701   | 내뿜다     | 7278800       | VV  | 1  | 0  |
| 가까스로 | 42701   | 넓히다     | 7405601       | VV  | 1  | 0  |
| 가까스로 | 42701   | 넘기다     | 7408003       | VV  | 2  | 0  |
| 가까스로 | 42701   | 넘다      | 7409803       | VV  | 1  | 0  |
| 가까스로 | 42701   | 놓다      | 7903233       | VX  | 1  | 0  |
| 가까스로 | 42701   | 늪다      | 8107100       | VV  | 1  | 0  |
| 가까스로 | 42701   | 대하다     | 9000305       | VV  | 1  | 0  |
| 가까스로 | 42701   | 더부러지다   | 9462800       | VV  | 1  | 0  |
| 가까스로 | 42701   | 데리다     | 9598801       | VV  | 1  | 0  |
| 가까스로 | 42701   | 들다      | 10852829      | VV  | 1  | 1  |
| 가까스로 | 42701   | 들어서다    | 10877602      | VV  | 2  | 0  |
| 가까스로 | 42701   | 뚫다      | 10915804      | VX  | 1  | 0  |
| 가까스로 | 42701   | 떠받들다    | 11214502      | VV  | 1  | 0  |
| 가까스로 | 42701   | 뚫다      | 11316904      | VV  | 1  | 0  |
| 가까스로 | 42701   | 뜨다      | 11338701      | VV  | 1  | 0  |
| 가까스로 | 42701   | 만나다     | 12646603      | VV  | 1  | 0  |
| 가까스로 | 42701   | 만들다     | 12663402      | VV  | 1  | 0  |
| 가까스로 | 42701   | 말리다     | 12785701      | VV  | 1  | 0  |
| 가까스로 | 42701   | 말하다     | 12748116      | VV  | 2  | 0  |
| 가까스로 | 42701   | 매달리다    | 12978703      | VV  | 1  | 0  |
| 가까스로 | 42701   | 멀다      | 13218901      | VA  | 1  | 0  |
| 가까스로 | 42701   | 면하다     | 13350406      | VV  | 1  | 0  |
| 가까스로 | 42701   | 묻다      | 14449601      | VV  | 1  | 0  |
| 가까스로 | 42701   | 밀치다     | 14991400      | VV  | 1  | 0  |
| 가까스로 | 42701   | 벗어나다    | 16530402      | VV  | 1  | 0  |
| 가까스로 | 42701   | 벗어나다    | 16530404      | VV  | 1  | 0  |
| 가까스로 | 42701   | 빌다      | 18678803      | VV  | 1  | 0  |
| 가까스로 | 42701   | 삼키다     | 20082703      | VV  | 1  | 1  |
| 가까스로 | 42701   | 성공하다    | 21329204      | VV  | 1  | 0  |

|      |       |        |          |    |   |   |
|------|-------|--------|----------|----|---|---|
| 가까스로 | 42701 | 세우다    | 21597912 | VV | 1 | 0 |
| 가까스로 | 42701 | 쉬다     | 23104901 | VV | 1 | 0 |
| 가까스로 | 42701 | 안돈하다   | 25038210 | VV | 1 | 0 |
| 가까스로 | 42701 | 어렵다    | 26019302 | VA | 1 | 0 |
| 가까스로 | 42701 | 열다     | 27055309 | VV | 1 | 0 |
| 가까스로 | 42701 | 열리다    | 27068902 | VV | 6 | 0 |
| 가까스로 | 42701 | 오다     | 27580113 | VV | 1 | 0 |
| 가까스로 | 42701 | 오다     | 27580131 | VX | 1 | 0 |
| 가까스로 | 42701 | 이렇다    | 30411900 | VA | 2 | 0 |
| 가까스로 | 42701 | 이루다    | 30420403 | VV | 1 | 0 |
| 가까스로 | 42701 | 일부러스럽다 | 31133800 | VA | 1 | 0 |
| 가까스로 | 42701 | 일어나다   | 31175712 | VV | 1 | 0 |
| 가까스로 | 42701 | 일어서다   | 31175902 | VV | 1 | 0 |
| 가까스로 | 42701 | 잇따르다   | 31422303 | VV | 2 | 0 |
| 가까스로 | 42701 | 있다     | 31425326 | VX | 1 | 0 |
| 가까스로 | 42701 | 있다     | 31425327 | VX | 1 | 0 |
| 가까스로 | 42701 | 귀어짜다   | 35398605 | VV | 1 | 0 |
| 가까스로 | 42701 | 짚다     | 36344301 | VV | 1 | 0 |
| 가까스로 | 42701 | 참다     | 36731001 | VV | 6 | 1 |
| 가까스로 | 42701 | 참다     | 36731002 | VV | 2 | 0 |
| 가까스로 | 42701 | 참다     | 36731003 | VV | 1 | 0 |
| 가까스로 | 42701 | 찾아가다   | 36857501 | VV | 1 | 0 |
| 가까스로 | 42701 | 피하다    | 41075407 | VV | 1 | 0 |
| 가까스로 | 42701 | 하다     | 41273902 | VV | 2 | 0 |
| 가까스로 | 42701 | 하다     | 41273940 | VX | 1 | 0 |
| 가까스로 | 42702 | 가라앉다   | 99106    | VV | 1 | 0 |
| 가까스로 | 42702 | 건지다    | 1645403  | VV | 1 | 0 |
| 가까스로 | 42702 | 구조되다   | 4317502  | VV | 1 | 0 |
| 가까스로 | 42702 | 나다     | 6579740  | VX | 1 | 0 |
| 가까스로 | 42702 | 나타나다   | 6727302  | VV | 1 | 0 |
| 가까스로 | 42702 | 내전하다   | 7256906  | VV | 1 | 0 |
| 가까스로 | 42702 | 놓다     | 7903233  | VX | 1 | 0 |
| 가까스로 | 42702 | 돌아오다   | 10120802 | VV | 1 | 0 |
| 가까스로 | 42702 | 되다     | 10449620 | VV | 2 | 0 |
| 가까스로 | 42702 | 드나들다   | 10781104 | VV | 1 | 0 |
| 가까스로 | 42702 | 맞추다    | 12954412 | VV | 1 | 1 |
| 가까스로 | 42702 | 면궁하다   | 13355602 | VV | 1 | 0 |
| 가까스로 | 42702 | 면하다    | 13350405 | VV | 1 | 0 |
| 가까스로 | 42702 | 모면하다   | 13595204 | VV | 1 | 0 |
| 가까스로 | 42702 | 물리치다   | 14516901 | VV | 1 | 0 |
| 가까스로 | 42702 | 불잡다    | 18245703 | VV | 1 | 0 |
| 가까스로 | 42702 | 비다     | 18379702 | VV | 1 | 0 |
| 가까스로 | 42702 | 비집다    | 18574801 | VV | 1 | 0 |

<표 V-1>에서 부사keyID 및 결합부사/용언keyID는 U-WIN 데이터베이스 관리번호이며, 각각 구축 도구에서 해당 어휘로 직접 매핑한다. 품사는 결합되는 어휘의 품사(부사(MAG), 동사(VV), 형용사(VA), 보조용언(VX))이며, 빈도는 부사와 함께 나타난 부사 및 용언의 빈도수이다. 비고에서는 부사의 용례에서 출현한 경우는 '1'로, 뜻풀이나 부사가 아닌 다른 품사의 용례에서 출현한 경우는 '0'이다. 이는 이번 부사 결합관계 정보 구축 작업을 용례에서 나타난 경우를 우선 처리하기 위해 구분하였다.

## 1.2. 부사의 결합관계 구축 도구

뜻풀이 및 용례에서 추출한 각 부사의 결합관계를 구축하기 위한 구축 도구는 다음 그림과 같다.



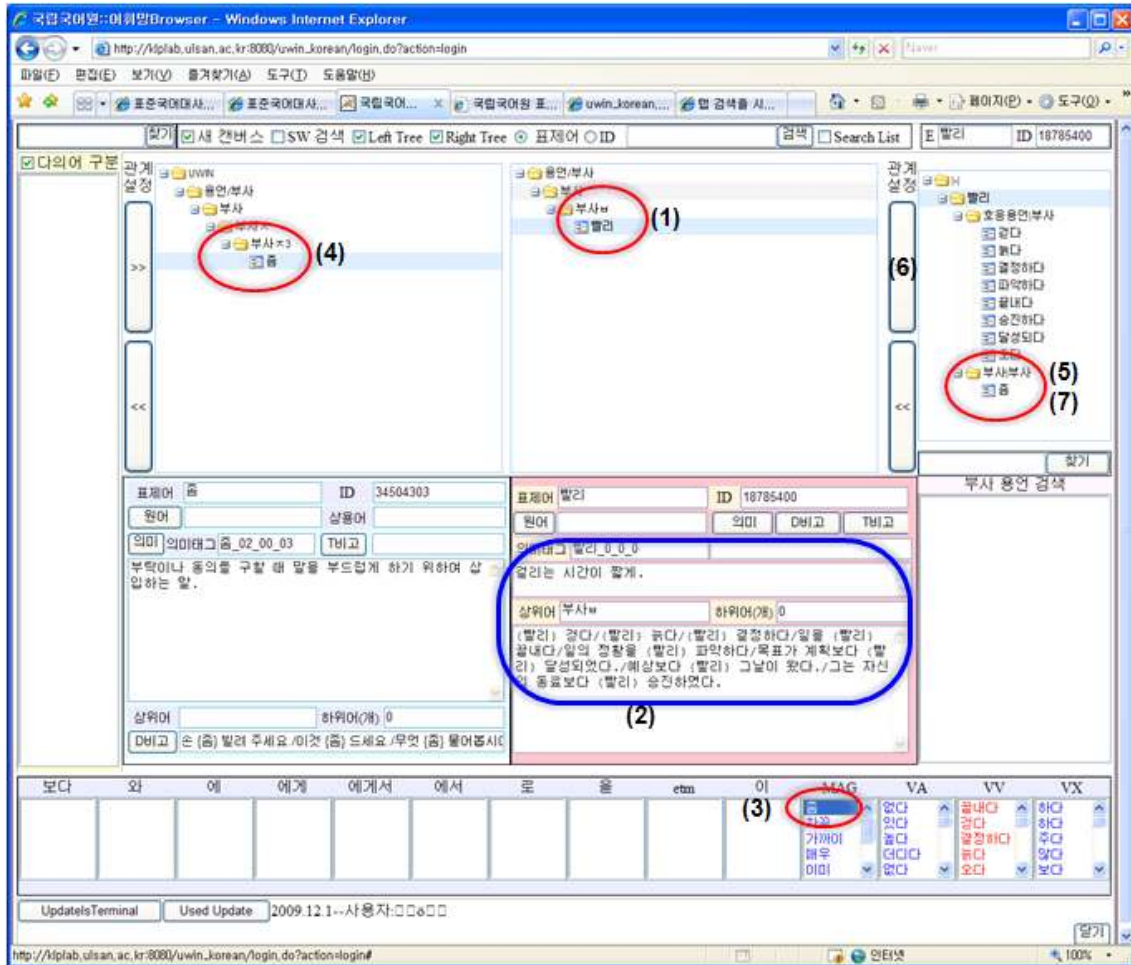
<그림 V-1> 부사의 결합관계 구축 도구

위 그림의 아래 쪽(붉은색 박스)에서 뜻풀이/용례에서 추출된, 특정 부사와 결합할 수 있는 부사, 용언(형용사, 동사, 보조용언)을 나타내며, U-WIN과 연결되어 있다.

### 1.3. 부사의 결합관계 연결 과정

부사의 부사와 용언과의 결합관계 구축은 어휘망 편집 브라우저를 통해 이루어진다. 이 어휘망 편집 브라우저에서 용언과 부사의 용례에서 추출한 결합관계 정보를 부사의 문법적인 수식 지침(V.2 장)에 따라 호응용언|부사 혹

은 부서|부사 결합관계 트리에다 넣는 것이다.



<그림 V-2> 부사의 결합관계 연결 과정

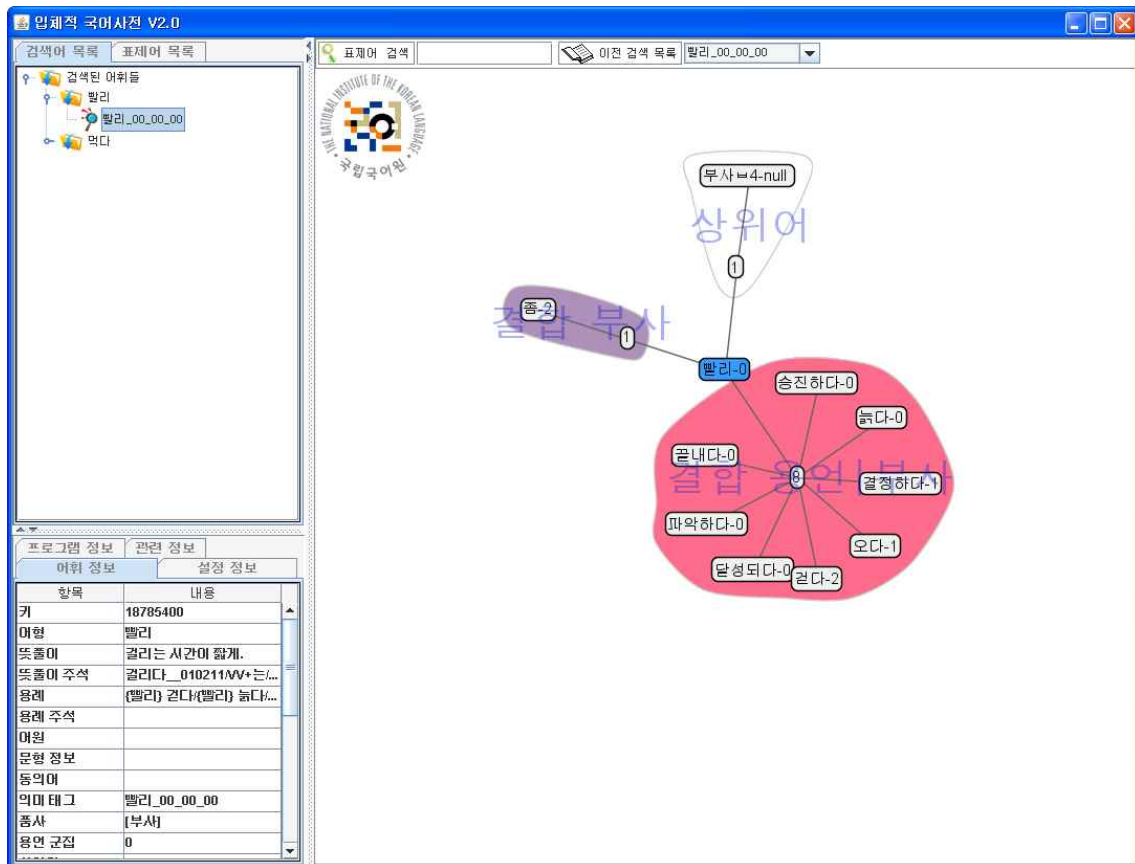
- ① 검색 창에서 선정부사(예, '빨리')를 입력하면 선정된 다의어 '빨리'가 출력되고, 이 다의어 중에서 하나를 선택한다.
- ② 선택된 부사의 뜻풀이와 용례를 통해 여러 다양한 용례를 확인한다.
- ③ 용례에서 추출된 부서\_부사의 결합관계인 '좁'을 더블클릭한다. 여기서 뜻풀이와 용례에서 추출한 부서 결합관계 정보에서 부사와 용언의 용례에서 추출한 결합관계 정보는 붉은색으로 표기되며, 뜻풀이와 다른 품사의 용례에서 추출한 결합관계 정보는 검은색으로 표기된다. 이번 과제에서는 붉은색으로 표기된 어휘를 우선 처리하였다
- ④ U-WIN의 '용언/부사'에서 '좁'의 뜻풀이와 용례(D비고)를 확인할 수 있다.
- ⑤ '빨리'와 '좁'이 의미적으로 부서\_부사 결합관계가 성립되면 부서\_부사



결합관계를 선택한다.

- ⑥ 관계설정 버튼을 누르면 ⑦과 같이 선택된 부서 '좁'이 '빨리'와 부서\_부사 결합관계로 설정된다.

이렇게 설정된 부서\_부사 혹은 부서\_용언의 결합관계 정보는 다음 입체적 국어사전에서 {결합 부서}, {결합 용언|부사}로 확인할 수 있다.



<그림 V-3> 입체적 국어사전 브라우저에서 부사의 결합관계

위 입체적 국어사전 브라우저에서 부서 '빨리\_00\_00\_00'를 중심으로 부서\_부사 결합관계인 {좁}과 부서\_용언 결합관계인 {걷다, 늑다, 달성되다, 파악하다, 끝내다, 결정하다, 승진하다, 오다} 등이 구축되어 있음을 확인할 수 있다.



## 2. 부사의 결합관계 구축 지침

<표준>에 등재된 용례를 기반으로 다음의 지침에 따라 {부사-용언, 부사-부사} 결합관계 정보를 구축하였다. <표준>에서 용례를 확인할 수 없는 경우나 문장부사 등 결합관계를 밝힐 수 없는 경우들은 처리하지 않도록 한다. 이는 임의로 용례를 만들거나 다른 말뭉치를 통해서 확인을 할 수도 있겠지만, 정확성을 기하기 위함이며 다음의 미처리 기준을 설정하였다.

### ① 사전에서 용례를 확인할 수 없는 경우 따로 처리하지 않는다.

- 가득가득\_\_01\_00\_03 : 냄새나 빛 따위가 공간에 매우 널리 퍼져 있는 상태. ≡가득가득히 「3」.
- 가득가득\_\_01\_00\_04 : 감정이나 정서, 생각 따위가 매우 많거나 강한 모양. ≡가득가득히 「4」.
- 살랑살랑\_\_01\_00\_02 : 물이 끓어오르며 이리저리 자꾸 움직이는 모양.
- 하늘하늘\_\_00\_00\_03 : 어디에 매인 데 없이 멋대로 한가하게 놀고 지내는 모양.

### ② 문장 부사는 따로 처리하지 않는다.

- 그러나\_\_00\_00\_00 : 앞의 내용과 뒤의 내용이 상반될 때 쓰는 접속 부사.  
¶ 우리는 열심히 손을 흔들었다. 그러나 선수 중 아무도 돌아보는 사람이 없었다./아내는 조용히 그러나 단호하게 말했다./성기는 잠자코 밥숟가락을 들었다. 그러나 밥은 반도 먹지 않고 상을 물려 버렸다.《김동리, 역마》/근심스러운 소식, 듣기 싫은 소식, 그러나 또한 심중 팔구는 반드시 나올 소식을 그들은 겹먹은 마음으로 기다리고 있는 것이었다.《김동인, 운현궁의 봄》

### ③ 'N+이다' 서술어는 처리하지 않는다.

- 사사건건\_\_00\_00\_00 : 해당되는 모든 일마다. 또는 매사에. ≡건건사사 [II]·사사건건이.

¶ 동생은 사사건건 말썽이다./엄마는 내가 하는 일에 사사건건 간섭이다./사사건건 걸고넘어지니 나보고 어찌란 말이야?/내버려진 춘복이를 거두어 기른 정이 애뜻한 공배로서는, 춘복이한테 하찮은 일도 대수롭게 넘기지 못하고 사사건건 토를 달고 나서게 되는 것이었다.《최명희, 혼불》

- 어느새\_\_00\_00\_00 : 어느 틈에 벌써.

¶ 입학한 지가 어제 같은데 어느새 졸업이다./코흘리게 어린애가 어느새 자라서 청년이 되었구먼./한라산은 비구름에 허리까지 잠뎀 잠겨 있는데 감질나게 찢끔거리던 가랑비는 어느새 그쳐 있었다.《현기영, 변방에 우짖는 새》

#### ④ ‘용언+-이’의 형태 중 부사로 처리하기 힘든 경우는 처리하지 않는다.

- 상관없이\_\_00\_00\_01 : ⇒ 상관없다 「1」.

¶ 그는 몹시 피곤한 것 같았고 옷차림하곤 상관없이 초라하고 헐벗은 것처럼 보였다.《박완서, 도시의 흉년》

- 아랑곳없이000 : 어떤 일에 참견을 하거나 관심을 둘 필요가 없이.

¶ 들몰댁이 역전에 이르렀을 때는 서늘한 날씨는 아랑곳없이 전신은 땀으로 젖어 있었다.《조정래, 태백산맥》/갈증으로 입이 탄 사병들은 죽고 사는 것은 아랑곳없이 서로 미친 듯 물을 찾았다.《홍성원, 육이오》

#### ⑤ ‘구 구성, 절’을 수식하는 경우는 처리하지 않는다.

- 설령\_\_01\_00\_00 : ((뒤에 오는 ‘-다 하더라도’ 따위와 함께 쓰여)) 가정해서 말하여. 주로 부정적인 뜻을 가진 문장에 쓴다. ≡설사02(設使)·설약(設若)·설혹(設或)·억혹(抑或)·유혹01(猶或).

¶ 저들이 설령 우리를 이곳에서 내보내 준다 해도 아주 좋아주지는 않을 것이다./그런 일은 없겠지만 설령 이번 일이 안된다고 하더라도 너무 실망은 하지 마라./설령 그가 그것을 흠뻑다 할지라도 너는 그를 믿어 주어야 한다.

⑥ 사전에서의 용법을 살핀 후 처리할 수 없는 것은 처리하지 않는다.

- 마치\_\_03\_00\_00 : ((흔히 ‘처럼’, ‘듯’, ‘듯이’ 따위가 붙은 단어나 ‘같다’, ‘양하다’ 따위와 함께 쓰여)) 거의 비슷하게.

¶ 마치 선녀처럼 고운 얼굴/반장은 마치 자기가 담임 선생님인 듯이 아이들에게 이래라저래라 했다./그녀의 목소리는 마치 천상에서 울리는 음악 소리 같다./전방에는 마치 사원 비슷한 높은 누각이 당당하게 막아서 있었다.《홍성원, 육이오》/창선은 마치 사막에 고개만 파묻은 타조 모양으로 혼자 눈을 감고 아무것도 보지 않았다.《나도향, 뉘우치려 할 때》

- 어찌\_\_00\_00\_04 : ((감탄 표현의 ‘-ㄴ지’, ‘-는지’ 따위와 함께 쓰여))동작의 강도나 상태의 정도가 대단함을 나타낸다.

¶ 그 꽃을 보는 순간 어찌 예쁘던지 녀를 잃고 바라보았다.

⑦ 용례가 잘못되어 부사로 보기 어려운 것들은 처리하지 않는다.

- 까르륵\_\_00\_00\_02 : 여자나 아이들이 자지러지게 웃는 소리. 또는 그 모양.

¶ 누가 무슨 말을 했는지 옆방에 한 패가 와글와글 떠들더니, 까르륵, 여자도 섞인 웃음소리가 터진다.《최명희, 혼불》

- 으하하\_\_00\_00\_00 : 입을 크게 벌리며 거리낌 없이 크게 웃는 소리. 또는 그 모양.

¶ 으하하! 저 녀석 보게.