

〈고려대 한국어대사전〉의 형태 분석 정보

김양진 · 이현희*

1. 머리말
 2. 〈고려대 한국어대사전〉 형태 분석의 작업 지침
 3. 〈고려대 한국어대사전〉의 형태 분석 정보의 특징
 4. 맺음말
- 참고 문헌

1. 머리말

이 논문은 <고려대 한국어대사전>(이하 <고려대 한국어대사전>)에서 표제어부에 제시된 ‘형태 정보’ 가운데 ‘형태(소) 분석 정보’¹⁾

* 김양진: 고려대 민족문화연구원 사전편찬부 연구교수.

이현희: 고려대 민족문화연구원 사전편찬부 선임연구원.

- 1) 둘 이상의 언어 요소를 최소의 의미 단위로 쪼개어 보이는 과정을 ‘형태(소) 분석’이라고 한다. 원론적으로 볼 때 둘 이상의 언어 요소가 하나로 합쳐진 언어 단위(즉 복합어)를 낱말의 언어 요소로 나누는 과정은 ‘형태 분석’이라고 할 수 있을 것이다. 이렇게 분석된 결과를 바탕으로 동일한 의미와 기능을 가지면서 상보적 분포 관계에 있는 둘 이상의 형태를 하나의 추상적인 언어 요소로 대표화하는 과정을 ‘형태소 설정’이라고 할 때, 이러한 ‘형태소 설정’을 거친 뒤에 ‘형태 분석’의 결과를 조정한 내용을 ‘형태소 분석’이라고 할 수 있다. 이 둘을 아울러 표현하고자 할 때 ‘형태(소) 분석’으로 쓸 수 있을 것이다. 예를 들어, 본고의 논의에 따라 ‘찰밥, 차조, 찹쌀’에서의 ‘찰-, 차-, 찹-’을 동일한 기능을 가진 형태소 ‘찰’의 이형태라고 했을 때 ‘찰밥[=찰+밥]’에 대하여 ‘차조[=차(찰)+조]’, ‘찹쌀[=찰(찰)+쌀]’과 같은 분석을 보인 것을 ‘형태(소) 분석’이라고 할 수 있다는 것이다. 다만 모든 형태 분석이 이러한 이형태 관계를 보이는 것이 아니기 때문에, 논의의 편의를 위해서 앞으로 본고에서 ‘형태 분석’이라고 하면 ‘형태(소) 분석’의 개념을 전제한 것임을 밝힌다.

에 대해 소개하는 것을 목적으로 한다. ‘형태 정보’란 표제어의 바로 뒤에 ‘[]’ 안에 제시되는 정보로 해당 표제어의 원어 정보와 그에 대한 형태소 분석 정보를 아울러 이른다. 본고에서는 그 가운데 ‘형태(소) 분석 정보’의 구성 원리와 특징점을 보이고자 한다.

<고려대 한국어대사전>에서는 표제어로 선택된 전체 복합어(합성어, 파생어) 174,686개에 대해, 표제어의 구성 성분을 밝히고 각 구성 성분의 자격을 보이는 형태 표지를 부착한 형태 분석 정보를 제시하였다.²⁾ <고려대 한국어대사전>의 형태 분석 정보는 기존 국어사전들에서 표제어의 구조를 단순히 2분지로 제시하였던 것과는 전혀 다른 제시 방식이다. 이와 같이 대사전 편찬 사상 최초로 표제어 분석의 단계를 다분지로 확장하였다는 점, 표제어 구성 형식들의 형태론적 범주를 명시화하였다는 점에서 실용적 가치와 학술적 가치를 동시에 구현할 수 있게 되었다는 의의가 있다. 즉 <고려대 한국어대사전>의 다분지 분석을 통해 표제어의 단어 구조를 최대한 미시적인 차원까지 보일 수 있었고, 각 구성 요소들이 표제어의 구성에 어떤 자격으로 참여하는지를 명시적으로 제시하여 국어 조어법의 기초 자료를 포괄적으로 제시할 수 있게 되었다.

현재 <고려대 한국어대사전>의 총 386,889개의 표제어 중 형태 분석 정보를 포함한 표제어의 수는 174,686개에 이른다. 이 중 접사를 포함한 파생어 표제어는 접두 파생어 8,518개, 접미 파생어 77,413개³⁾, 접사를 포함하지 않은 합성어 표제어(단순 합성어)는

2) <고려대 한국어대사전>의 형태 정보는 김양진(1999)의 형태 사전 구성안을 초안으로 하고 고려대 민족문화연구원 국어사전편찬실의 논의를 거쳐 마련된 <고려대 사전 편찬지침>(2001)의 지침안을 기준으로 하여 2002년에는 대략적 열개를 갖추었던 것이다. 이에 대한 개략적인 기준과 실제에 대한 보고는 도원영(2003)에서 간략히 제시된 바 있다. <고려대 한국어대사전>의 형태 정보를 포함한 상세한 편찬 경위에 대해서는 <고려대 한국어대사전 편찬 백서>(2010, 예정)을 참고할 것.

86,197개이고, 그 외 어미를 포함한 표제어는 2,128개가 있다. 어중별로는 고유어 포함 복합어 111,446개, 한자어 포함 복합어 62,501개와 기타 736개⁴⁾의 표제어에 형태 분석 정보가 제시되어 있다. 물론 이 외에도 형태 분석 DB를 통해 다양한 항목의 통계치를 추출할 수 있다. 이렇게 분석된 형태 분석 DB의 구축은 다음과 같은 점에 의의가 있다.

첫째, <고려대 한국어대사전>의 형태 분석은 한국어의 조어 구성을 체계적이고 전체적으로 조망할 수 있는 토대를 마련했다는 점에서 그 의의가 있다. 둘째, 한국어의 공시적 언어 대상으로서의 형태소 목록을 추출하여 제시하고자 하였다. 셋째, 통계처리·말뭉치 구축, 형태 분석 말뭉치 혹은 신어, 번역 등 다양한 자연언어 처리의 토대 자료가 된다.

본고는 다음과 같이 구성된다. 우선 2장에서 사전 표제어에 형태 분석이 필요한 이유와 <고려대 한국어대사전>의 형태 분석의 작업 방향을 설명한다. 3장에서는 <고려대 한국어대사전>의 형태 분석이 가지는 장점에 대해 실제 예를 통해 보인다.

-
- 3) 이 수치는 ‘-하다’ 파생어 35,351개, ‘-되다’ 파생어 4,840개, ‘-거리다’ 파생어 2,424개, ‘-히다’ 파생어 2,387개, ‘-대다’ 파생어 2,376개, ‘-이다’ 파생어 1,541개, ‘-스럽다’ 파생어 765개, ‘-시키다’ 파생어 578개, ‘-지다’ 파생어 296개, ‘-뜨리다’ 파생어 139개, ‘-트리다’ 파생어 121개, ‘-롭다’ 파생어 108개, ‘-리다’ 파생어 108개, ‘-기다’ 파생어 99개, ‘-치다’ 파생어 91개 등을 포함한 것이다.
- 4) 기타 736개에는 <고려대 한국어대사전>의 접두사 표제어 204개, 접미사 표제어 351개 및 단일 어미를 포함한 기타 표제어 180여개가 포함된다.

2. <고려대 한국어대사전> 형태 분석의 작업 지침

한국어는 특히 조어법이 다양하고 복잡하게 발달한 언어이다. 기본적으로 고유어, 한자어, 외래어 3종의 어휘 체계를 가지고 있고 또 이들이 복잡다단하게 얹혀 있다. 특히 동일한 형태가 단어에 따라 다양한 의미로 실현되거나 각기 다른 자격으로 단어 구성에 참여하는 일도 빈번하다. 그런데 기존 사전의 표제어 구조 제시 방식만으로는 이에 대한 정확한 모습을 구현하는 데 한계가 있을 수밖에 없다. 표제어가 몇 조각으로 나누어질 수 있는지, 그때 각각의 조각들이 어떤 자격인지를 구체적으로 제시한 사전이 필요하다는 것이 우리들의 공통적인 문제 인식이었고 이를 시작으로 표제어 386,889개의 대사전에서 복합어 174,686단어의 형태 분석이라는 방대한 작업을 착수하기에 이르렀다.⁵⁾

2.1 형태 분석의 의의 및 분석 태도

<고려대 한국어대사전>의 형태 분석 정보 구축 작업을 통해 한국어 조어 연구와 교육에 있어 기존의 연구 성과를 아우르면서 대규모의 자료를 체계적으로 제공할 수 있는 새로운 인프라의 구축이 가능함을 확인하였다. 실제로 대사전 표제어에 대한 전면적인 형태 분석 작업은 그간 시도된 바조차 없었다. 이는 인력과 시간, 작업의

5) 국어사전에서의 형태 분석의 필요성에 대한 문제 제기는 이미 조재수(1989)에서 이루어진 바 있고 김양진(1999)에서는 국어 형태 사전의 구축에 반영할 정보를 중심으로 국어 고유어의 형태소 분석의 원리를 설정하고 그 분석의 결과를 목록화하여 제공할 바 있으며 송재영(2007/2008)에서 국어사전에서 형태 분석을 통해 역사적 정보까지를 포함한 국어 단어 구성의 정보를 구체화하는 방안이 꼼꼼히 논의된 바 있다. <고려대 한국어대사전>의 형태 분석에서는 이러한 기존의 문제 제기를 최대한 포함하고자 하였으나 종이 사전의 출간에는 많은 정보들이 함축되었다.

지난함 때문도 한 이유이겠으나 그보다는 이론적으로(국어학이나 언어학) 형태 분석 자체가 어렵거나 모든 표제어를 일관되게 처리할 수 있는 기준의 제시가 어려웠던 것도 중요한 원인이었다. 그럼에도 불구하고 한국어 어휘에 나타난 조어 방식의 다양함을 체계적으로 정리하고 설명하기 위해서는 한국어 어휘 전반에 걸친 형태 분석 작업이 반드시 이루어져야 함은 재론의 여지가 없다. 이러한 표제어의 형태 분석 작업의 필요성과 그 의의를 몇 가지로 정리하면 아래와 같다.

첫째, 한국어의 형태/형태소의 총목록을 제시할 수 있는 기반을 마련한다.

이는 <고려대 한국어대사전>의 형태 분석 정보를 통해 한국어 어휘의 전반적인 모습을 이해할 수 있는 토대가 마련됨을 의미한다. 한국어 어휘는 현재에도 계속 팽창하고 있다. 여기에 포함된 형태/형태소의 전체 목록이 어떠한지를 제시할 수 있는 사전은 그간 없었다. 그러나 이번 <고려대 한국어대사전>의 형태 분석 정보를 통해 한국어의 형태/형태소의 전체 목록은 물론 어떤 특정 형태/형태소가 조어에 어떤 방식으로 몇 회 사용되었는지 등의 세부적인 수치까지도 모두 추출해 낼 수 있게 되었다. 이는 이번 형태 분석 작업의 가장 큰 결과물이라 하겠다.

다음 (1)은 <고려대 한국어대사전>의 형태 분석 정보를 통해 추출할 수 있는 한국어 형태소의 목록 중 일부이다. 형태소 목록 중 일부는 표제어로 등재되지 않은 것들도 있다.

(1) 한국어 형태소 목록 일부⁶⁾

6) 동음이의어를 구별하기 위한 어깨번호는 <고려대 한국어대사전>(2009)의 표제어 및 뜻풀이 번호를 따른 것이다. 아래에 인용된 예들에서도 마찬가지이다.

고유어 자 립 형태소 목 록 (가 항)	<p>가1, 가2, 가계, 가까스로, 가꾸다, 가꾸로, 가꿈, 가납사니, 가누다, 가늘다, 가늠, 가다, 가다귀1, 가다리1, 가닥, 가달1, 가대기1, 가대기2, 가두다, 가둑2, 가뜩1, 가뜩2, 가라사대, 가라지1, 가라지2, 가락1, 가락2, 가락지, 가랑, , 가랏, 가래1, 가래2, 가래3, 가래4, 가래5, 가래6, 가로1, 가되, 가루1, 가르다, 가르랑, 가르릉, 가르마, 가르치다, 가리1, 가리2, 가리3, 가리4, 가리5, 가리다, 가리다, 가리단죽, 가리비, 가리사니, 가리산1, 가리새1, 가리새2, 가리, 3, 가마1, 가마2, 가마3, 가마4, 가마5, 가마6, 가마귀, 가마등, 가마우지, 가만, 가말다, 가무라기, 가문비, 가물, , 가물치, 가보1, 가쁘다, 가사리1, 가사리2, 가살, 가새1, 가슴1, 가슴츠레, 가시1, 가시2, 가시3, 가시다, 가시, , 가시랭이, 가오리, 가운데, 가위1, 가위2, 가위3, 가을1, 가을2, 가자미, 가장13, 가장귀, 가장이, 가재1, 가재기1, 가죽1, 가지1, 가지2, 가지3, 가지기, 가지다, 가축1, 가귀, 가탈1, 가탈2, 가팔막, 가폴막, 각다귀, 각단1, 각담1, 각시1, 간1, 간각1, 간나위, 간대로1, 간사위, 간살1, 간살이, 간수1, 간자미, 간정1, 간지라기, 간직, 갈1, 갈2, 갈3, 갈4, 갈5, 갈6, 갈7, 갈8, 갈가리2, 갈가위, 갈가지1, 갈갈, 갈개1, 갈겨니, 갈고리, 갈고자리, 갈고쟁이, 갈고지, 갈, 1, 갈기다, 갈기다, 갈다 , 갈래, 갈맞, 갈망1, 갈매1, 갈매2, 갈매기, 갈무리1, 갈부던, 갈비1, 갈비2, 갈비3, 갈새1, 갈쌍, 갈치, 갈귀1, 갈피1, 갈다, 감1, 감2, 감3, 감4, 감5, 감감, 감기1, 감다, 감다, 감자1, 감장1, 감장2, 감추다, 감탕1, 감태1, 감화보금, 갑자기, 갑절, 값, 갓1, 갓10, 갓2, 갓3, 갓4, 갓모, 강강술래, 강고도리, 강남조, 강냉이, 강다리1, 강다리2, 강대1, 강대2, 강동3, 강아지, 강장7, 강정1, 강중2, 강충이, 강치, 갓다, 갓다, 갈다, 갈다</p>
고유어 불완전 어 근 목 록 (가 항)	<p>가, 가꾸러, 가년, 가널, 가느스름, 가느스, 가늑, 가닐, 가닥, 가택, 가동, 가등, 가드락, 가든, 가든그, 가들, 가들막, 가뜬, 가란침, 가랍, 가랑, 가래, 가랑, 가리, 가리산, 가린, 가마노르케, 가마득, 가마말쑥, 가마무트름, 가마반드르, 가마반지르, 가마아득, 가막, 가막살, 가무, 가무끄름, 가무대대, 가무댕댕, 가무리, 가무래, 가무숙숙, 가무스레, 가무스름, 가무잡잡, 가무죽죽, 가무촉촉, 가무총총, 가무칙칙, 가무퇴퇴, 가무트름, 가물, 가못, 가분, 가불, 가못, 가뿐, 가뽏, 가새, 가새, 가수알, 가스랑, 가스리, 가슬, 가지, 가실, 가실, 가온, 가잡, 가장, 가죽, 가중, 가지런, 가직, 가짓, 가찰막, 가창, 가지작, 가칠, 가질, 가짓, 가팔, 가회, 각다분, 각죽, 각지, 각통, 간, 간간, 간닥, 간당, 간댕, 간둥, 간둥, 간질, 간드릿, 간드리, 간드작, 간들, 간막, 간실, 간잔지런, 간정, 간조통, 간중, 간질, 간짓, 갈, 갈(가라), 갈강, 갈개, 갈강, 갈고랑, 갈그랑, 갈근, 갈마, 갈매, 갈상, 갈신, 갈선, 갈쪽, 갈충, 갈평, 감, 감동, 감둥, 감뎌, 감성, 감승, 감실, 감작, 감정, 감제, 감쪽, 감창, 감당, 감파르족족, 갑, 갑갑, 갑작, 강, 강낭, 강당, 강동, 강쇠, 강중, 강팔, 갓</p>
고유어 접미사 형 태 목 록	<p>-口, -口직, -가마리, -갈-, -개, -거리8, -거리9, -거리10, -거리다, -게, -팡이, -구다, -굿다, -금-, -기29, -기30, -기다9, -기다10, -기다11, -까짓, -깔, -깟, -껏, -께, -꼴-, -꾸러기, -꾼, -꺼리, -나, -나다, -내, -내기, -네, -니, -넉, -다랄다, -다리, -다지, -답다, -달다, -대가리, -대기, -대다, -데기, -되다5, -되다6, -둥이, -드리다, -들6, -, -딱지, -판, -패기, -패기3, -뜨기, -뜨리다, -롭다, -리, -리다1, -리다2, -리다3, -막, -맞다, -매, -머리, -바가지, -바리, -바치, -바치, -반-, -발, -배기, -배기, -뱅이, -보, -보, -부리, -빼기, -뽕, -새, -새, -쇠, -쇠, -스럼다, -스레, -시키다, -실, -씨, -씩, -아치, -애다, -어치, -없다, -으키다, -음, -음직, -이26, -이27, -이28, -이29, -이30, -이31, -이다5, -이다6, -이다7, -이다8, -이키다, -작, -잡다, -장, -장이, -쟁이, -적, -적다, -정께, -죽, -지기11, -지기12, -지다, -직, -질, -짐, -짜, -짜리, -짜, -쩍다, -중, -쫘, -차, -잡다, -채, -추, -쥬다5, -추다6, -치, -치다, -치레, -코, -키다, -투성이, -통이, -트리다, -포, -하다, -히, -히다3, -히다4, -히다5</p>

한자어 접미사 형 태 목 록	-街, -家, -價, -歌, -哥, -家, -間(간18), -間(간19), -客, -鏡, -頃, -屈, -計, -糸, -係, -界, -高, -工, -館, -觀, -官, -狂, -鑛, -橋, -教, -具, -口, -國, -局, -圈, -權, -金, -氣, -記, -機, -器, -期, -難, -團, -談, -當, -代, -臺, -帶, -代, -隊, -宅, -島, -圖, -徒, -度, -欄, -亂, -量, -力, -曆, -令, -領, -嶺, -路, -爐, -錄, -論, -料(료1), -料(료2), -類, -流, -瘤, -律, -率, -裡/裏, -林, -魔, -末, -物, -民, -發, -方, -房, -輩, -白, -犯, -法, -別, -補, -服, -本, -夫, -婦, -部, -附, -分, -費, -事, -師, -社, -士, -詞, -史, -辭, -産, -上, -狀(상15), -狀(상16), -生(생6), -生(생8), -席, -線, -船, -選, -性, -所, -素, -囚, -手, -順, -術, -視, -室, -心, -氏, -兒, -愛, -額, -礦, -洋, -語, -業, -餘, -然, -慾, -用, -院, -員, -園, -元, -源, -願, -率, -律, -人, -者, -子, -作, -葬, -長, -狀, -丈, -帳, -的, -戰, -展, -店, -艇, -錠, -整, -劑, -祖, -朝, -族, -宗, -種, -座, -酒, -主, -紙, -地(지18), -地(지19), -池, -誌, -陣, -集, -次, -策, -責, -處, -廳, -體, -咱, -值, -癡, -通, -判/版, -品, -風, -畢, -下, -學, -漢, -行, -許, -形, -號, -畫/畫, -化, -會
--------------------------	--

<고려대 한국어대사전>에서는 이렇게 표제어 각각에 대한 형태 분석 정보들을 확정해 놓고 전체 복합어 174,686개에 대하여 형태 분석의 결과를 제시하였다. 아래 <표 1>은 각 품사에 따른 표제어 대비 형태 분석 정보를 포함하는 표제어 수의 비율을 제시한 것이다.

【표 1. 형태 분석의 통계치】

		전체 표제어수 ⁷⁾	표준어 표제어수	형태 분석 표제어수	표준어 대비 형태 분석 비율	기타
품 사	명사	254,622	233,282	101,805	43.6%	
	대명사	412	301	100	33.3%	
	수사	224	163	72	44.2%	
	동사	55,308	52,246	52,202(-2,077) ⁸⁾	99.9%	95.9%
	형용사	12,431	10,637	10,620(-432)	99.9%	95.8%
	관형사	745	691	612	88.6%	
	부사	14,012	12,028	8,857	73.6%	
	감탄사	693	508	155	30.5%	
	조사	340	210	84	40%	
	접사	555	545	545	100%	
	어미	1,502	936	681	72.8%	
총계(개)		340,844 ⁹⁾	311,547	175,733 ¹⁰⁾		
비율(%)		88%	80.5%	45.4%	56.4%	

7) <고려대 한국어대사전>의 표제어 중에는 3,192개의 다품사어가 포함되어 있다. 형태 분석 표제어 수를 산출하는 데에 이들을 일일이 고려하기 어려운 점

표 1에서 보인, <고려대 한국어대사전>의 전체 표제어 386,889개 중에는 46,045개의 구 표제어가 포함되어 있다. 구 표제어를 제외한 단어 340,844개 중에서 방언이나 비표준어를 제외한 표준어 어형은 311,547개로 <고려대 한국어대사전>의 전체 표제어를 대상으로 할 때, 80.5%의 표제어가 형태 분석 대상의 전체 기준이 된다. 따라서 <고려대 한국어대사전>에서 형태 분석의 결과를 반영한 175,733개의 표제어는 방언 및 비표준어를 포함한 전체 단어 대비 비율이 45.4%, 표준어 대비 비율이 56.4%로 제시된 것이다. 이는 국어 표준어에서 복합어(합성어 및 파생어)와 접사류(접두사, 접미사, 어미류 등)를 포함한 어휘의 비율이 단일어(고유어, 한자어, 외래어 포함)보다 약간 많은 정도라는 것을 의미한다.¹¹⁾

동사와 형용사의 경우, 표준어 전체 대비 99.9%의 형태 분석이 제시된 이유는 기본형을 이루는 어미 ‘-다’의 분석 때문인데 단일어간의 동사, 형용사가 얼마 되지 않기 때문에 이를 제외한 복합어간의 경우가 동사 표제어 전체 중에서 95.9%, 형용사 표제어 전체 중에서 95.8%로 형태 분석되었다는 점에서 동사와 형용사의 복

이 있어 다품사어는 우리 사전에서 첫 번째 품사로 제시한 품사에 귀속하여 수치를 산출하였다. 이 점에서 다품사어를 별도의 수치로 제시한 “<고려대 한국어대사전>의 종합적 고찰(도원영·차준경, 2009)”에서의 수치와 일부 차이가 있음을 밝혀둔다.

- 8) <고려대 한국어대사전>의 단일어 동사는 2,077개, 단일어 형용사는 432개로 추출되었다. 그러므로 이들을 제외한 나머지 예들은 모두 파생어이거나 합성어인 표제어라 할 수 있다.
- 9) 이 수치는 전체 표제어수 386,889개 중에서 구 표제어를 포함한 무품사어 45,163개, 준말 675개 등을 뺀 수치이다.
- 10) <고려대 한국어대사전>에서는 단독형 접사와 단독형 어미에도 형태 분석 표지를 하였다. 따라서 여기서 제공한 수치는 형태 분석이 이루어진 복합어의 수 174,686개에 단독형 접사와 단독형 어미의 수를 모두 포함한 것이다.
- 11) 단일어 동사는 2,077개, 단일어 형용사는 432개의 수를 단일어에 포함하더라도 이러한 상대치는 크게 달라지지 않는다.

합어 비율이 매우 높음을 알 수 있다. 그밖에 관형사의 형태 분석 비율(88.6%)이 높은 이유는 관형사 파생 접미사 ‘-的’ 파생어의 비율이 전체 관형사 중에서 매우 높은 비율을 차지하기 때문이며 부사는 첩어형 부사의 비율이 전체 부사 중에서 높은 비율을 차지하고 있기 때문에 형태 분석의 비율(73.6)이 높다. 어미 역시 통합형 어미가 많기 때문에 형태 분석의 비율이 전체 표제어의 72.3%에 달하는 반면 일반 명사나 수사 등은 전체의 40% 수준에서 형태 분석이 제시되었다. 감탄사의 형태 분석 비율이 다른 어휘 부류에 비해 낮은 이유는 감탄사는 전체가 하나의 형태로 재어휘화된 경우가 많다고 보고 형태 분석을 최소화했기 때문이다. 예를 들어, ‘아서라, 두어라’와 같은 감탄사에서 ‘-어라’ 혹은 ‘-라’는 의도적으로 분석하지 않았다.

둘째, 단어의 구조 및 의미에 대해 사용자의 정확한 이해를 보조할 수 있는 정보를 충분히 제공할 수 있다.

<고려대 한국어대사전>의 형태 분석은 단어의 구성을 사용자가 직접 확인함으로써 단어의 의미 파악에 도움을 주고 새로운 단어를 만드는 데 참조할 수 있도록 구성되었다. 예를 들어, 기존 사전에서 준말-본말의 관계로 다루어졌던 ‘다리미질’과 ‘다림질’, ‘다리미질하다’와 ‘다림질하다’는 형태 분석의 결과에 따라 서로 다른 단어 구성을 가진 것으로 평가되었으며 따라서 이들은 ‘준말-본말’의 관계가 아니라 단순한 유의 관계의 단어들이므로 처리되었다.

- (2) 가. 다리미질 [+다리미-질] [뽕] 다리미로 옷이나 천의 구김살을 문질러 펴는 일. ♣어머니는 와이셔츠 위로 물을 뿌리고 ~을 시작하였다. / ~이 필요한 빨래는 따로 세탁하는 게 좋겠다. <유의> 다림질

나. 다림질 [+다리-ㄹ-질] ㉢ 다리미로 옷이나 천의 구김살을 문질러 펴는 일. ㉣로 열을 가해야 구김살이 펴진다. / 석원은 ~을 금방 마쳐 줄이 끈게 선 바지를 입고 집을 나섰다. <유의> 다리미질

(2 가)의 ‘다리미질’과 (2 나)의 ‘다림질’은 실질적으로 동일하거나 거의 차이를 파악하기 어려운 정도로 유사한 의미로 사용된다. 하지만 단어 형성의 입장에서 이 두 단어는 서로 다른 구성을 가지고 있다. <고려대 한국어대사전>에서는 이러한 차이를 형태 분석의 결과로 제공함으로써 이들의 미세한 의미 차이에까지도 사용자들이 접근할 수 있도록 하였다. 즉 ‘다리미질’은 ‘다리미’라는 도구를 가지고 구김살을 펴는 동작을 가리킨다면 ‘다림질’은 물론 주로 ‘다리미’를 사용하겠지만 그와 상관없이 ‘다리는 일’ 자체를 가리키는 의미로 사용된다는 점을 명확히 할 수 있다는 것이다.¹²⁾ ‘다리미질’은

12) 익명의 심사위원은 그럼에도 불구하고 ‘다리미질’과 ‘다림질’이 동일하게 뜻풀이한 이유를 제시할 필요를 제기하였다. 언어는 다양한 표현상의 이유로 동의어나 유의어를 만들어 낸다. 예를 들어, ‘귀걸이①’과 ‘귀고리’라는 단어는 전혀 별개의 형태 단위로 구성된 단어이지만 이들을 다르게 뜻풀이하기란 쉽지 않은 일이다. 물론 형태 관계가 다르면 해당 어휘에 대한 인식یدن 어감یدن 어떤 점에서든 차이가 있기 마련이지만 국어사전의 뜻풀이에서는 그 정도의 미세한 의미 차이는 구별하지 않는 것이 일반적이다. ‘다리미질’과 ‘다림질’은 전혀 다른 형태 관계에서 형성된 말이지만 둘다 “다리미를 가지고 옷이나 천 따위의 구김살을 문질러 펴는 일”이라는 동일한 과정을 지칭하는 말이다. 이들의 뜻풀이가 동일한 이유를 묻는 것은 ‘달걀~계란, 반대말~반의어~상대어’ 따위의 국어사전 속의 수많은 유의어들이 동일하게 뜻풀이된 이유를 묻는 것과 같다. 물론 ‘다리미질’에서는 ‘도구’의 의미가, ‘다림질’에서는 ‘행위’의 의미가 더 강조된 뜻풀이로 이들을 구별해 볼 수도 있을 것이다. 그러나 <고려대 한국어대사전>의 뜻풀이는 그 단계까지 진화하지는 못했다. 국어사전의 뜻풀이가 모든 유의어를 구별할 수 있을 때까지 진화할 수 있을지의 문제는 앞으로의 과제라고 하겠다. 다만 <고려대 한국어대사전>의 형태 분석이 뜻풀이만으로는 충분히 설명할 수 없는 이러한 미세한 차이에 대한 보충 설명의 역할을 할 수 있다는 점에 의의를 두고자 한다.

‘쟁기질, 바느질, 비질’ 등과 같은 계열의 단어라면 ‘다림질’은 ‘시침질, 달음질, 박음질’ 등과 같은 계열의 단어라는 점에서 차이가 있다. 이와 같이 형태 정보를 통하여 의미 관계와 음상이 유사한 두 단어의 어형상의 관계를 사용자들이 분명하게 식별할 수 있도록 한 데에 <고려대 한국어대사전> 형태 분석의 장점이 있다.

셋째, 동일 음상의 표제어들을 구분할 수 있는 언어 정보를 제시한다.

사전의 형태 분석 정보는 그 자체로서의 의의도 충분하지만 사전 내의 다른 정보들과의 유기적 상호 작용을 통해 그 필요성과 가치가 배가된다. 우선 동일 음상의 표제어들을 구분할 수 있는 언어 정보의 제시라는 점에서 형태 분석 정보의 필요성을 찾을 수 있다. 아래는 단어의 형태 정보의 차이가 표제어의 분리와 뜻풀이 기술에 반영된 예이다.

- (3) 가. 신짹¹ [+신+짹] ㉠ 신발의 날짹. ㉡수동이는 새 고무신이 닳을까 봐, ~을 벗어 들고 걸었다.
 나. 신짹² [+신-짹] ㉠ ‘신2’을 낫잡아 이르는 말. ㉡아무리 돈을 빌리러 왔다고 해도 사람을 무슨 헛 ~ 취급하듯 쫓아내는 건 잘한 일이 아니다.

‘신짹¹’과 ‘신짹²’는 기존 사전에 한 표제어의 다의로 처리된 예이다. 그러나 이러한 처리는 두 표제어의 구조가 각각 합성어와 파생어인 것을 보여주지 못한다. 앞에서 제시한 ‘다리미질’, ‘다림질’의 경우와 마찬가지로, <고려대 한국어대사전>과 같은 방식으로 형태 분석 정보를 제시해야 명사 ‘짹’과 접미사 ‘-짹’이 각기 다른 단어 구성에 참여하는 요소들이며 이에 따라 ‘신짹¹’과 ‘신짹²’이 동음이의어로 처리되어야 하는 단어임을 보여줄 수 있다.

같은 관점에서 형태 분석은 동음이의 표제어가 4개인 ‘달리다’ 계열 표제어들의 구분에도 유용한 정보를 제공할 수 있다.

- (4) 가. 달리다¹ [+달(단)-리_다] 1 ㉡자 ①‘단다’의 사동사.
 나. 달리다² [+달-리_다] ㉡자 … (무엇이 일정한 곳에) 걸리거나 붙은 채 떨어지지 않게 되다.
 다. 달리다³ [+달리_다] ㉡자 재물이나 기술 힘이 모자라다.
 라. 달리다⁴ [+달-리_다] ㉡타 (사람이 어떤 사람에게 다른 사람을) 오거나 가거나 할 때 옆에 있거나 따르게 하다. <유>팔리다.
 (5) 가. 달리다¹ : 내달리다[+내+달(단)-리_다], 내리달리다[+내리+달(단)-리_다], 말달리다[+말+달(단)-리_다], 치달리다[+치+달(단)-리_다]...
 나. 달리다² : 걸달리다[+걸+달-리_다], 덧달리다[+덧+달-리_다],

(4 가~라)에서처럼 <고려대 한국어대사전>에서는 ‘달리다¹’의 경우 ‘단다’와의 관련성을 보이는 정보를 함께 제시하고 ‘달리다²’나 ‘달리다⁴’는 피동과 사동의 접미사를 ‘달다’와 분리하여 제시할 수 있다. 마지막으로 ‘달리다³’은 단일어로 처리함으로써 이들이 단순히 동음이의어로 나누어지는 것을 넘어서 단어 구성상의 차이를 가지고 있는 단어들임을 한눈에 알아 볼 수 있게 하였다. 이러한 구별은 (5 가, 나)에서 보인 ‘달리다’ 부류의 복합어의 구분에도 확장된다. 사전 이용자들은 각각의 복합어들이 ‘달리다¹’~‘달리다⁴’ 중 어느 것에 해당하는지 형태 분석 정보를 보고 쉽게 알 수 있다.

이는 전체가 동일한 음상을 가진 표제어 뿐 아니라 다음과 같이 표제어의 일부 음절이 동일한 음상일 경우에도 마찬가지로 유용한 정보가 될 수 있다.

- (6) 가. 간고등어 [+간+고등어] ㉡
 나. 간곳없다 [+가_ㄴ+곳+없_다]

다. 간석기(-石器) [+가(갈)_ㄴ+石器] ㉞

라. 간암(肝癌) [+肝+癌] ㉞

(6 가~라)의 예에서 각 표제어에 공통적으로 포함되는 ‘간’은 그 의미나 기능이 모두 달라서 경우에 따라 고유어 명사이거나 동사의 활용형이거나 한자어 명사일 수 있다. 기존 사전의 원어 정보 제시 방식으로는 이들이 모두 다른 자격인 것을 명시적으로 드러내기 어렵다. 하지만 <고려대 한국어대사전>에서와 같은 방식의 형태 분석에서는 이들의 분석 정보를 통해 각 복합어를 구성하고 있는 구성 요소들의 자격이나 의미의 차이를 명확하게 보여줄 수 있다.

넷째, 동일 원어 정보의 표제어들을 구분할 수 있는 언어 정보를 제시한다.

일반적으로 한자어의 경우, 원어 정보가 한자로 제시되었다면 동형 혹은 동음의 표제어들의 구분이 용이할 것으로 생각하기 쉽다. 그러나 실제로 동형의 한자라 해도 단어의 구조에 따라 다른 의미로 실현되는 경우가 적지 않아 그 구분이 어려운 예들이 많다.

(7) 가. 타제2 [他製] ㉞ 다른 사람, 또는 다른 회사나 공장에서 만든 같은 종류의 제품. <원어> 타제품1(他製品).

가'. 타제품1 [+他製-品] ㉞ 다른 사람, 또는 다른 회사나 공장에서 만든 같은 종류의 제품. ㉞이 제품은 ~에 비해 성능과 품질 면에서 단연 앞섭니다. <약어> 타제2(他製).

나. 타제3 [他製] ㉞ 종류가 다른 제품. <원어> 타제품2(他製品).

나'. 타제품2 [+他+製品] ㉞ 종류가 다른 제품. <약어> 타제3(他製).

(7)의 예는 각각 형태 분석 정보를 통해 원어-약어 관계의 차이를 보여줄 수 있는 예이다. (7 가, 나)에 제시된 ‘타제2, 타제3’은

각각 <고려대 한국어대사전>에서 ‘타제품1, 타제품2’와 원어-약어 관계에 있는 것으로 제시된 것이다. 이들은 둘다 동일한 한자 ‘他製’와 ‘他製品’으로 구성되어 있어서 한자만으로는 그 의미 차이를 명확히 구별하기 어렵다. 하지만 이를 의미와 관련하여 형태 분석을 정밀하게 하려는 관점에 서게 되면 (7 가)의 ‘타제품1’은 [+他製-品]으로 (7 나)의 ‘타제품2’는 [+他+製品]으로 분석되게 된다. 따라서 ‘타제품1’과 ‘타제품2’가 동음이의어로 구별되어야 하는 것은 물론 ‘타제품1(他製品)’의 준말인 ‘타제2(他製)’와 ‘타제품2(他製品)’의 준말인 ‘타제3(他製)’이 동음이의어로 구별되어야 하고 ‘타제품2(他製品)’의 준말인 ‘타제3(他製)’과 ‘타제품1(他製品)’에 포함된 ‘他製’는 같은 형태일 수 있지만¹³⁾ ‘타제품1(他製品)’의 ‘他製’와 ‘타제2(他製)’의 ‘他製’는 동일한 형태 정보를 가지고 있다고 보기 어렵다는 점이 드러나게 되는 것이다.¹⁴⁾

(8)은 한자어 ‘客’이 <고려대 한국어대사전>에서 단일 한자어 명사, 한자어 접두사, 한자어 접미사로 쓰이고 있음을 보인 것이고, (9), (10)의 예는 ‘客’이 포함된 복합어에 대해 각각 형태 분석 정보를 통해 ‘客’에 대해서 명사-접두사-접미사 범주의 차이를 보여줄 수 있음을 보인 예이다.

(8) 가. 객[客] ㉠ 남의 집에 찾아온 나그네. ㉡여관에서 ~을 마다하다니?

나. 객- ㉢두 일부 명사 앞에 붙어, ‘불필요하게 덧붙은’의 뜻을 더하는 말. ㉣물 / ~일 / ~소리 / ~식구 / ~순가락.

다. -객 ㉤미 일부 명사 뒤에 붙어, 구경을 오거나 거기에 참여하는 사람의 뜻을 더하여 명사를 만드는 말. ㉥관람~ / 등산~ / 방청~ / 풍류~.

13) 이렇게 볼 경우, ‘타제품1’은 ‘국화꽃’이나 ‘역전앞’에서와 같은 동어반복(tautology)의 유의어 형성과 유사한 과정을 거치는 것으로 볼 수 있다.

14) 이와 관련한 자세한 논의는 김양진(2009)를 참조할 것.

- (9) 가. 객방 [+客+房] ㉠ 손님이 머무르는 방. <유의> 빈실¹(賓室), 손님방¹(--房)
 나. 객소리 [=客+소리] ㉠ 쓸데없는 말. ㉡ 김 씨는 이제 겨우 다섯 살이 된 딸아이를 보고 미스 코리아감이네 어찌네 ~를 하다가 돌아갔다. <유의> 객담¹(客談), 객론¹(客論), 객설¹(客說)
 다. 관람객 [+觀覽-客] ㉠ 공연이나 전시회 따위를 구경하는 사람. ㉡ 박물관장은 ~으로 복적됐다.
- (10) 가. 객순갈¹ [+客+순갈] ㉠ 손님을 대접하기 위하여 마련한 순가락. ㉡ 내년에는 장사가 잘돼 ~이 좀 부족했으면 좋겠구먼. <본말> 객순가락¹(客---) <유의> 객술¹(客-)
 나. 객순갈² [=客+순갈] ㉠ 남의 밥을 빼앗아 먹으려고 들이미는 순가락. <본말> 객순가락²(客---) <유의> 객술²(客-)

이러한 형태 정보상의 차이는 기존의 이분지 분석(객-방, 객-소리, 관람-객, 객-순갈)만을 보이는 국어사전들의 경우나 형태 분석 대상에 대해 별도의 범주 표시를 하지 않은 경우라면 보여주기 어려운 것들이다.

다섯째, 형태 분석 정보와 뜻풀이 정보가 서로 연계되어 표제어의 뜻풀이에 일관성을 기할 수 있게 된다.

표제어의 형태 분석을 통한 구조적 정보와 사전 뜻풀이의 정보는 서로 연계될 필요가 있다. 이러한 연계는 뜻풀이의 일관성을 기하는 데에도 도움이 되고 또 형태 분석 표지를 통해 표제어가 어떤 의미 부류에 속하는 어휘에 해당하는지도 사용자가 쉽게 알 수 있다는 점에서 중요하다. 예를 들어 <고려대 한국어대사전>에서 한자 ‘초(草)’를 포함하는 표제어는 총 929건이다. 이 중 ‘사초(莎草), 연초(煙草), 개초(蓋草)’등 ‘초(草)’를 따로 분리해 내기 어려운 예들을 제외하고 형태 분석의 대상이 되는 단어는 총 231건이다. 물론 이들 단어들에 포함된 ‘草’는 그 의미가 다양하게 실현된다. 다음의 예들을 보자.

- (11) 가. 자립적 용법이 있는 ‘초(草)’ ⇒ 형태 분석에서 [+草]로 표시
 ① 풀 ② 건초(乾草)/마른풀 ③ 갈초 ④ 난초(蘭草) ⑤ 초안(草案) ⑥ 초서(草書)
- 나. 자립적 용법이 없는 ‘초(草)’ ⇒ 형태 분석에서 [±草]로 표시
 ① 건축용어로서의 ‘초’(단청에서 쓰임). ② 담배/연초(煙草).
- (12) 가. ① 가는기린초 [+가느(가늘)_ㄴ+麒麟+草] (식물)
 ② 글초 [+글+草] (초고)
- 나. ① 겹장구머리초 [+겹+장구+머리±草] (단청)
 ② 금연초 [+禁煙±草] (담배)

(11 가)는 ‘초(草)’가 <고려대 한국어대사전>에 명사 표제어로 등재된 예들이고 (11 나)는 단독 표제어로 처리되지 않지만 합성어나 파생어의 뜻풀이에는 반복적으로 나타나는 예들이다. (12 가)와 (12 나)는 <고려대 한국어대사전>에서 각각 (11 가)와 (11 나)의 구분을 통해 ‘초(草)’류 단어들을 실제 분석한 사례이다. 이러한 형태 분석을 통해 동일 형태라 할지라도 세밀하게 구별된 형태소가 확인될 수 있으며 이러한 형태소 단위의 구별이 뜻풀이에 반영되는 의미의 일관성을 좀더 명확하게 할 수 있다. 아울러 이러한 분석 결과는 추후에 ‘형태 분석 말뭉치’와 같은 실제 언어 자료의 형태 분석에도 유용한 기초 자료가 제공될 수 있다.¹⁵⁾

15) 홍종선(2001)에서는 자연언어의 전산 처리와 관련하여 국어 형태 단위인 단어와 접사(파생 접사와 굴절 접사 포함)의 형태론적 정보화를 전반적으로 다룬 바 있고, 민경모(2003)에서는 말뭉치를 대상으로 한 형태 분석의 기준을 꼼꼼히 제시한 바 있다. 말뭉치를 대상으로 한 형태 분석 결과물은 『한국어 형태소 및 어휘 사용 빈도의 분석1』(김홍규, 강범모, 2000), 『한국어 형태소 및 어휘 사용 빈도의 분석1』(김홍규, 강범모, 2004), 『한국어의 사용 빈도』(김홍규, 강범모, 2009)에서 종합적으로 제시된 바 있다. 김홍규·강범모(2000)은 150만 어절 규모의 세종 형태분석말뭉치에 기초한 것이고 김홍규·강범모(2004)는 550만 어절, 김홍규·강범모(2009)는 김홍규·강범모(2004)는 각각 550만 어절, 1500만 어절 규모의 세종 형태의미 분석 말뭉치를 대상으로 조사한, 한국어 단어와 형태소의 사용에 대한 명실상부한 종합적인 빈도 자료이

마지막으로 형태 분석을 통해 기존 사전에서 잘못 제공된 어휘 정보들을 수정할 수 있다. 표제어의 형태를 분석할 경우 형태와 의미의 연결을 확인해 보는 과정에서 기존의 사전들에서 잘못 기록되어 온 어휘 정보를 체계적으로 수정할 수 있다. 예를 들어, 기존 사전들에서 아래 ‘마름뇌’의 유의어로 제시된 ‘능뇌(能腦)’의 한자 원어 정보는 그 출처를 분명히 알기 어려운 예이다. <고려대 한국어대사전>에서는 이때의 ‘능뇌(能腦)’의 뜻풀이에 나타나는 ‘마름’이라는 고유어와 한자 ‘능(菱/菱)’과의 대비를 통해서 이 단어가 본래 ‘능뇌(菱腦)’로 표기되어야 하는 단어임을 밝히고 이를 바로 잡을 수 있었다.

- (13) 가. 마름 [𦉆] 𦉆 <식물>마름과에 속한 한해살이풀. 연못이나 늪에서 자란다. 뿌리는 진흙 속에 박고 줄기는 물위에까지 가늘고 길게 뻗는다. 잎은 줄기 꼭대기에 뭉쳐나고 삼각형이며, 잎자루에 공기가 들어 있는 부낭(浮囊)이 있어서 물위에 뜬다. 여름에 흰 꽃이 피며 마름모꼴의 열매가 열리는데 이것을 까서 먹는다.

나. 마름뇌 [+마름+腦] <유의>능뇌(能腦). ⇒ 능뇌(菱腦).

<고려대 한국어대사전>에서 (13)의 예를 통해 뜻풀이와 형태 정보(이 경우는 한자와 관련한 원어 정보)의 일치 여부와 관련하여 ‘능뇌(能腦)’를 ‘능뇌(菱腦)’로 바로 잡은 이후 이러한 정보는 ‘능형뇌(能形腦)’와 같은 단어의 어휘 정보를 수정하는 데에도 확장되었다.

다. 다만 본고에서 논의된 <고려대 한국어대사전>에서의 형태 분석에서와 같은 정밀한 형태 분석 사전이 기반되었다면 김홍규·강범모(2009)에서 분석된 대규모의 언어 자료는 단순히 한국어의 어휘 빈도와 대표 형태소 빈도를 넘어서 기계 번역이나 자동 번역과 같은 고차원적 수준의 언어 자료로까지도 활용될 수도 있게 될 것이다.

(14) 가. 능뇌 [能腦] ⇒ [菱腦]

나. 능형뇌 [能形腦] ⇒ [±菱形+腦]

‘능뇌’, ‘능형뇌’는 백과사전이나 한의학 전문 분야에서는 흔히 ‘菱腦’, ‘菱形腦’로 사용되어 오던 것이나 어떤 이유에서인지 <표준국어대사전>(1999/2009)에서는 ‘能腦’, ‘能形腦’로 기록되어 온 것이다.¹⁶⁾ <고려대 한국어대사전>에서는 이와 같이 형태소를 일치시키는 과정에서 기존의 국어사전들에서 답습되어 온 다양한 유형의 오류들이 일관성 있게 수정되었다.

2.2 형태 분석 정보의 범주별 표지

<고려대 한국어대사전> 형태 분석 작업의 가장 큰 원칙은 ‘표제어의 형태에 충실한 분석’이다. 표제어에서 해석된 언어 형태를 최대한 분석하고 최대한 일관성이 있게 공시적인 관점에서 최대한 ‘해석’한다는 것이 이 형태 분석 과정의 가장 큰 원칙이다.¹⁷⁾

-
- 16) <금성판 국어대사전>에서는 ‘능형뇌(能形腦)’를 표제어로 등재해 놓고 그 뜻풀이 끝에 유의어로 ‘능뇌(菱腦)’를 두었으나 ‘능뇌(菱腦)’나 ‘능뇌(能腦)’는 표제어로 올리지 않았다. <고려대 한국어대사전>에서 ‘능형(菱形)’을 불완전 어근으로 처리하여 ‘ㅅ’로 표시한 것은 이 단어가 ‘마름모꼴’의 이전 용어로 평가되어 공시적 쓰임을 인정하지 않게 된 데 따른 것이다. ‘능뇌(菱腦)’를 [±菱+腦]와 같이 분석하지 않은 이유는 ‘능뇌(菱腦)’를 ‘능형뇌(菱形腦)’의 준말로 보고 준말은 형태 분석하지 않는 원칙에 따른 것이다.
- 17) 이러한 형태 분석의 원칙은 김양진(1999)의 형태 해석 원리와 형태 보존 원리에 따른 것이다. 김양진(1999)에서는 형태 해석 원리와 형태 보존 원리를 다음과 같이 형상화하였다. (1)형태 해석 원리1 : 형태는 해석되어야 한다. (2)형태 해석 원리2 : 형태 해석은 일관되어야 한다. (3)형태 보존 원리 : 해석된 형태는 보존되어야 한다.
- (1)형태 해석 원리1은 다음과 같이 구체화할 수 있다. (ㄱ)어떤 형태가 모든 환경에서 동일한 의미로 실현될 경우, 그 형태는 하나의 단일한 형태소로 해석된다. (ㄴ)어떤 형태가 어떤 환경에서 상이한 형태로 실현될 경우, 분포가 넓은 형태가 그렇지 않은 형태에 대하여 우선적으로 해석된다. (ㄷ)어떤 형

<고려대 한국어대사전>의 형태 분석 작업에서 다음의 형태 표지들은 단순히 복합어를 둘 혹은 셋으로 분리하는 분리 기호가 아니라 분석된 형태 단위의 형태론적 범주를 보여 주는 형태 범주 표지임에 유의할 필요가 있다. 모든 형태 범주 표시는 분석된 형태의 왼쪽에 일관되게 표시하여 분석된 각각의 형태들이 각각의 형태 부류에 포함되도록 하였다. 국어 단어 분석의 결과에 따른 형태 범주는 아래와 같이 5가지로 유형화하여 보였다.

【표 2. 형태 분석 표지와 그 예】

	해당 범주	예시	특이사항
+	완전 어근/ 자립어	+사람, +그, +둘, +잡, +예쁘, +에그머니, +는	조사 포함
±	불완전 어근/ 비자립어	±궁금, ±깨끗, ±대단, ±따뜻, ±마땅, ±쓰러, ±아름, ±어떠, ±엉뚱, ±조용, ±중얼, ±홀륭, ±簡單, ±深刻, ±燦爛, ±華麗	
=	접두사	=개, =첫, =假, =生, =超	
-	접미사	-개, -기, -이, -하, -되, -生	
-	어미	-다, -어, -고, -시,	선어말어미 포함

‘+’는 <고려대 한국어대사전>에 표제어로 등재된 단일어(완전 어

태가 어떤 환경에서 상이한 형태로 실현될 경우, 어형을 가장 잘 보존할 것 같은 환경에서의 형태가 그렇지 않은 형태에 대하여 우선적으로 해석된다.

(2)형태 해석 원리2는 어떤 언어 결합형에서 분석된 형태가 다른 결합형에서도 일관되게 해석되어야 한다는 것이다. (3)형태 보존 원리는 형태와 형태의 결합에서의 형태 변화는 최소로 이루어져야 하며 이때의 형태 변화는 본래 형태를 최대한 유지하는 방향으로 일어난다는 원리이다. 이상의 모든 원리는 “모든 언어 형태는 해석되어야 한다”는 김양진(1999)의 완전 해석 원리에 따른 것이다.

근(자립어)에 표시한다. 여기에는 명사, 대명사, 수사, 관형사, 부사, 동사 어간, 형용사 어간, 감탄사 등이 포함되며 학교 문법에 따라 단어의 구성에 참여하는 조사도 ‘+’로 분석하였다.¹⁸⁾ 불완전 어근 범주의 표지인 ‘ㅁ’은 <고려대 한국어대사전>에 표제어로 등재된 단어, 접두사, 접미사, 어미로 분석된 요소가 아닌 나머지 어휘 요소에 표시하였고, 접두사 표지 ‘=’와 접미사 표지는 ‘-’는 각각 <고려대 한국어대사전>에 등재된 접두사와 접미사로 선정된 형태의 왼쪽에 표시하였다. 특히 단어 구성에 참여하는 어미에 이르기까지 별도의 구별 표지()를 두어 형태 분석의 대상으로 하였다. ‘_’ 표지는 단어 형성에 참여하는 용언 활용형에 나타나는 어미(부사형 어미, 관형사형 어미, 선어말어미 등)를 용언 어간과 구별하여 보일 필요 때문에 설정한 표지이다.¹⁹⁾

18) 조사는 실질형태소라기보다는 형식형태소이기 때문에 단어로 취급하기보다는 접사로 처리하는 시도가 좀더 국어의 특성에 맞을 수도 있다. 하지만 다른 형식형태소들과 달리 조사의 경우는 둘 이상의 조사가 통합하여 합성적 의미, 기능을 나타내는 측면이 두드러진다고 보아 이들 ‘조사+조사’의 구성을 합성 조사(에서, 에도, 로서, 만도 등)로 보고 실사에서의 합성어와 같은 차원의 처리가 되도록 하였다. 이는 둘 이상의 어미가 통합하여 단일한 기능을 나타내는 경우를 통합형 어미(-는다, -습니다, -도다, -어라, -것다 등)로 보아 합성어로 처리하지 않은 것과 대비된다. 이는 단순히 조사는 단어로 처리하고 어미는 단어가 아닌 것으로 처리하는 문법적 태도와 관련한 것이 아니라 ‘합성’이나 ‘파생’과 같은 단어 형성 기제에 대한 형태론적 개념을 염두에 둔 것으로 어미류의 통합이 문법화와 같은 특정 형태소의 기능 약화와 관련되어 있는 데 비하여 조사의 합성은 각 조사의 어휘 혹은 기능 의미가 합성의 과정에서 일정한 원리를 가지고 참여하고 있다고 보았기 때문이다. 이에 대한 자세한 논의는 별도의 자리를 통하여 상세히 논하고자 한다.

19) 명사형 어미 ‘-ㅁ/-음’과 ‘-기’는 <고려대 한국어대사전>에서 파생 접미사로 설정한 ‘-ㅁ/-음’, ‘-기’와의 혼동을 고려하여 모두 접미사로 처리하였다. 이들이 단어 구성 내부에서 여전히 명사형 어미로서의 기능을 가지고 있는지 아니면 이미 이러한 기능을 상실하고 접미사화한 ‘-ㅁ/-음’, ‘-기’로 기능이 통합되었는지는 후후의 연구 과제이다.

2.3 형태 분석의 실제

2.2절의 [표2]에 제시된 형태 표지를 이용하여 <고려대 한국어대사전> 형태 분석의 전형적인 실제 사례를 몇 가지 제시하면 다음과 같다.

형태 위주의 분석

<고려대 한국어대사전>에서는 형태 단위의 분석을 기본으로 보이고 형태소는 보일 필요가 있는 경우에 한하여 추가 정보로 괄호 안에 부연하였다.

- (15) 가. 돌아보다 [+돌_아+보_다]
나. 아드님 [+아드(아들)-님]

(15 가)와 같은 분석은 ‘형태 분석’이고 (15 나)와 같은 분석은 ‘형태(소) 분석’인데 이와 관련한 논의는 각주1)의 내용을 참조하기 바란다. 다만 이를 [+아들+님]과 같이 ‘형태소 분석’의 방식으로 보이는 데에는 표제어와의 불일치와 같은 여러 가지 문제가 있음을 지적하고자 한다.

가장 미시적인 수준까지 분석

모든 형태 분석의 결과는 다분지로 제시하되 자소 단위까지 구체적으로 보였다.(다분지의 장점에 대해서는 3.1의 논의를 참조할 것)

- (16) 떠나다 [+ㅌ(ㅌ)_어+나_다], 썩먹다 [+ㅍ(ㅍ)_어+ㅊ_다]

자소 단위의 분석은 형태 분석을 공시적 관점에서 최대한으로 분

석하는 태도로부터 온 것이다. 공시적으로 분리 가능한 형태를 분리하고 남겨진 요소는 자소 이상의 단위라면 별개의 형태로 인정된다. ‘떠나다’의 ‘떠’는 ‘깨어나다, 일어나다, 벗어나다, 태어나다, ...’의 ‘깨어, 일어, 벗어, 태어, ...’ 등의 ‘-어’와의 동형성에 의해 ‘ㄷ- + -어’로 분리되며 이때의 ‘ㄷ-’는 ‘뜨-’의 이형태로 파악된다. 다만 <고려대 한국어대사전>에서는 ‘ㄷ-’이 ‘뜨-’의 이형태임을 보이기 위해 ‘ㄷ-’의 뒤 괄호를 열어 이것이 ‘뜨-’의 이형태임을 밝혔다.

불규칙한 활용형의 분석

국어의 다양한 불규칙 활용형들의 분석 방안을 고려하였다.

- (17) 걸음 [+걸(걷)-음], 지은이 [+지(짓)_은+이], 가까이 [+가까(가깝)-이], 노래지다 [+노래(+노랑_아)+지_다], 빨리 [+빨리(빠르)-이], 피내다 [+피(푸)_어+내_다], 그리하여[+그리+하_여], 해내다 [+해(+하_어)+내_다]

불규칙 용언에서 나타나는 불규칙한 변화형들은 형태소를 보일 필요가 있는 경우에 준하여 괄호 안에 대표 형태를 밝힘으로써 의미를 분명히 하였으나 불규칙형 가운데 일부는 자소 단위의 분석으로도 분리할 수 없는 경우가 있으므로 이런 경우에는 부득이 ‘노래(+노랑_아)’와 같은 방식으로 처리하였다. 이들은 공시적인 활용의 예이므로 굳이 ‘노래(<+노랑_아)’처럼 역사적 처리를 하지 않았다.

고유어 불완전 어근에 대한 처리

고유어 복합어에서 자립어(완전 어근)와 접사 부분을 제외하고 남겨진 부분은 불완전 어근으로 처리하였다.

- (19) 가. 오솔길 [±오솔+길], 오두방정 [±오두+방정], 박쥐 [±박+쥐], ...
 나. 궁금하다 [±궁금-하_다], 깨끗이 [±깨끗-이], 따뜻하다 [±따뜻-하_다], 마땅히 [±마땅-히], 아름답다 [±아름-답_다], 어퍼하다 [±어퍼-하_다], 조용하다[±조용-하_다], 중얼거리다[±중얼-거리_다], 홀륭하다 [±홀륭-하_다]

한자어 불완전 어근에 대한 처리

한자어 복합어의 경우도 한자어 자립어(완전 어근)와 접사 부분을 제외하고 남겨진 한자 부분은 불완전 어근으로 처리하였다.

- (18) 가. 타제품 [±他+製品] ㉠ 종류가 다른 제품. <약어> 타제3(他製)
 나. 화려체 [±華麗-體] ㉠ 『문학』 수사법과 수식어를 동원하여 화려하게 꾸민, 정감이 풍부한 문체.

중첩형 음성 상징어의 분석

중첩형 음성 상징어의 분석에서, 단일어로 쓰이지 않는 일음절 중첩의 음성 상징어는 분석하지 않았으나 2음절 중첩의 음성 상징어는 어근 부분이 완전 어근인지, 불완전 어근인지 즉 자립어인지 비자립어인지의 여부와 상관없이 모두 분석하는 태도를 취하였다.

- (20) 가. 박¹³ ㉠ ① 얇고 질긴 종이나 천 따위를 한 번에 야무지게 찢는 소리를 나타내는 말. 또는 그 모양을 나타내는 말. ㉠찰민이는 종이를 ~ 찢었다. <큰말> 벅^{3②} <센말> 짹⁴
 나. 박박¹ [±박+박] ③ 얇고 질긴 종이나 천 따위를 자꾸 야무지게 찢는 소리를 나타내는 말. 또는 그 모양을 나타내는 말. ㉠수형은 화난 표정으로 편지를 ~ 찢어 버렸다. / 입지 못할 웃이라고 그렇게 ~ 찢는 것이냐? <큰말> 벅벅^{1③}, 북북^② <센말> 짹짹^{1④} <참고> 북북^{1②}
 다. 박박² ㉠ 얼굴 따위가 마맛자국이 많아서 얇어 있는 모양을 나타내는 말. ㉠동네 아이들은 내 마맛자국을 보고 ~ 얇은 곰보라고 놀려 댔다. <큰말> 벅벅² <센말> 짹짹³

(21) 가. 까딱 [ㄱ ㄹ] ㄱ 고개나 손발 따위를 앞뒤로 또는 아래위로 세게 살짝 움직이는 모양을 나타내는 말. 『말없이 일어난 효민은 고개만 ~ 흔들며 건성으로 인사를 했다. <큰말> 끄떡¹ <여린> 까닥¹

나. 까딱까딱¹ [+까딱+까딱] [ㄱ ㄹ] ㄱ 고개나 손발 따위를 앞뒤로 또는 아래위로 세게 자꾸 조금씩 움직이는 모양을 나타내는 말. <큰말> 끄떡끄떡¹ <여린> 까닥까닥¹

다. 까딱까딱² [±까딱±까딱] [ㄱ ㄹ] ㄱ 잘난 체하며 버릇없이 교만하게 구는 모양을 나타내는 말. <큰말> 꺼떡꺼떡¹ <여린> 까닥까닥²

(20 나)의 ‘박박¹’은 (20 가)에서 보인 ‘박¹³’의 예에 따라 중첩 구성의 합성어로 분석하였지만 (20 다)의 ‘박박²’는 1음절 중첩어이기는 하지만 1음절의 단일어가 표제어로 인정되지 않았기 때문에 단일어처럼 처리한 것이다. ‘하하, 흑흑, 철철, 빌빌’ 등의 예가 이에 해당한다. (21 나)처럼 2음절 중첩어의 경우도 단일어 표제어가 존재하는 경우는 당연히 합성어로 분석하고 (21 다)의 2음절 중첩어의 경우 1음절 중첩어와 달리 비록 단어로 인정되지 않는 불완전어근일지라 하더라도 분석을 하였다. 이는 우리말의 음절 구조가 2음절을 기본으로 하고 있다는 판단에 근거한 것이다.²⁰)

20) 예를 들어, 우리말 두음법칙과 같은 음운 규칙의 경우, 단순한 2음절 단어에서는 어두에 오지 않는 ‘ㄹ’은 아무 문제없이 실현되지만 4음절 이상의 구성에서는 두음법칙이 단어성 여부와 상관없이 2음절 단위로 끊어서 적용되는 양상을 보인다.

가. 남녀(男女), 녹리(鹿梨), 노력(努力), 강림(降臨), 보류(保留), 권리(權利)

나. 남부여대(男負女戴), 오비이락(烏飛梨落), 무실역행(務實力行), 배산임수(背山臨水), 호사유피(虎死留皮), 원형이정(元亨利貞)

나)의 예에서 ‘여대(女戴), 이락(梨落), 역행(力行), 임수(臨水), 유피(留皮), 이정(利貞)’ 등은 국어에서 단어 혹은 형태소의 자격을 인정하기 어려운 예들이지만 각각 두음법칙과 같은 음운 규칙이 적용되고 있다. 이는 두음법칙과 같은 음운 규칙들이 국어의 형태소 여부와는 다른 형태론적 단위를 이룰 때 2음절 이상의 음절을 단위로 하여 적용되고 있기 때문인데 중첩의 경우도 이러한 특성을 고려하여 2음절 이상의 중첩 부분은 별개의 형태 단위로 분석하

문법 형태소의 처리

<고려대 한국어대사전>에서는 국어의 모든 복합형 문법 형태소를 분석하였다. 국어에서 단어로 인정되는 조사뿐 아니라 단독형의 접사와 어미들도 모두 형태 분석의 결과 추출된 것이므로 동일한 형태 범주 표시를 각 형태의 왼쪽에 표시하였다. 다만 각주 14)의 언급처럼 조사의 복합은 합성어로 평가하고 어미류의 통합은 단순한 통합형 어미로 처리한 점에 차이가 있다.²¹⁾

(22) 가. -ㅁ에도 [ㅁ+에+도] [문]연 모음이나 ‘ㄹ’로 끝나는 용언, ‘이다’의 어간 또는 선어말 어미 ‘-으시-’의 뒤에 붙어, 주로 ‘-ㅁ에도 불구하고’의 구성으로 쓰여, ……을 나타내는 말.

나. -ㅁ에라 [ㅁ+에+라] [문]중 모음이나 ‘ㄹ’로 끝나는 용언, ‘이다’의 어간 또는 선어말 어미 ‘-으시-’의 뒤에 붙어, ……을 반문하는 형식으로 강조하여 나타내는 말. 해라체로 쓰인다.

(23) 가. -ㄹ는지1 [ㄹ(려)_느_ㄴ지] [문]연 ① 모음이나 ‘ㄹ’로 끝나는 용언, ‘이다’의 어간 또는 선어말 어미 ‘-으시-’의 뒤에 붙어, 주로 ‘알다’나 ‘모르다’와 함께 쓰여, …… 스스로 의문을 나타내는 말.

나. -ㄹ는지2 [ㄹ(려)_느_ㄴ지] [문]중 모음이나 ‘ㄹ’로 끝나는 용언, ‘이다’의 어간 또는 선어말 어미 ‘-으시-’의 뒤에 붙어, …… 스스로 의문을 나타내는 말.

(24) 가. -ㄹ러니 [ㄹ(리)_러(더)_니] [문]연 모음이나 ‘ㄹ’로 끝나는 용언, ‘이다’의 어간 또는 선어말 어미 ‘-으시-’의 뒤에 붙어,

지만 단어가 아닌 1음절의 중첩은 따로 분석하지 않는 근거로 삼았다. ‘국란사양상(國亂思良相)’이나 ‘인중용(人中龍), 등용문(登龍門)’ 등과 같이 한문 구성의 독법에서도 이러한 형태론적 인식 단위들이 엿보이는데 이에 대해서는 좀더 깊이 있는 연구가 선행될 필요가 있다.

- 21) 조사의 통합은 합성어로 처리하고 어미의 통합은 통합형으로 처리하는 태도는 여러 가지 문제를 야기하지만 이에 대해서는 별도의 다른 논문에서 구체적으로 논의하기로 하고 본고에서는 <고려대 한국어대사전>에서 이들 문법 형태소들을 최대한 공식적인 분석을 전제로 하여 분석하였음을 지적하는 것으로 그친다.

……을 나타내는 말. 주로 옛 말투에 쓰이며, ‘-겠더니’의 뜻을 나타낸다.

나. -리러라 [리(리)_러(더)_라(다)] ㉞종 모음이나 ‘ㄹ’로 끝나는 용언, ‘이다’의 어간 또는 선어말 어미 ‘-으시-’의 뒤에 붙어, ……을 나타내는 말. 해라체로, 주로 옛 말투에 쓰이며, ‘-겠더라’의 뜻을 나타낸다.

(25) 가. -리만치 [리(리)+만치] ㉞연 모음이나 ‘ㄹ’로 끝나는 용언, ‘이다’의 어간 또는 선어말 어미 ‘-으시-’의 뒤에 붙어, ……을 나타내는 말. 앞 절의 사실을 인정하면서 그것이 뒤 말에 대한 근거가 됨을 나타낸다.

나. -리만큼 [리(리)+만큼] ㉞연 모음이나 ‘ㄹ’로 끝나는 용언, ‘이다’의 어간 또는 선어말 어미 ‘-으시-’의 뒤에 붙어, ……을 나타내는 말. 앞 절의 사실을 인정하면서 그것이 뒤 절에 대한 근거가 됨을 나타낸다.

형태 분석의 제외 대상

다만 <고려대 한국어대사전>에서는 다음과 같은 경우는 형태 분석의 대상으로 삼지 않았다.

(26) 가. 구 표제어 : 사회 제도 [社會制度], 결사의 자유 [結社의自由]

나. 고유 명사 : 춘향전 [春香傳], 대한민국 [大韓民國]

다. 제식 훈련의 구령이나 일부 서술어의 활용형이 명사나 감탄사로 규정된 말 : 앞으로가, 뒤로 돌아 가, 아서라

라. 한자 부수의 이름 : 코비 [코鼻], 돌석 [돌石], 큰대[큰大], 갈지 [갈之]²²⁾

마. 둘 이상의 단어가 뒤섞여 하나의 단어로 된 경우 : 직간접 [直間接], 동식물 [動植物]

22) ‘큰대자’, ‘갈지자’의 경우는 한자 부수로 관습화된 경우가 아니라 주로 조사 ‘로’와 함께 쓰여 ‘큰 대(大) 글자와 같은 모양’이나 ‘갈 지(之) 글자와 같은 모양’이라는 의미로 합성적 의미를 만드는 단어이기 때문에 각각 ‘[+크_ㄴ+大+字+로]’, ‘[+가_ㄹ+之+字+로]’와 같이 형태 분석을 하였다는 점에 차이가 있다.



- 바. 방언 및 비표준어 : 머리말 [+머리+말], 머릿말 [] ㄹ머리말.
 바. 단, 방언 및 비표준어 중에서 대응 표준어와 동일하지 않은 한자어나 외래어가 포함된 경우 해당 한자나 외래어 원어를 따로 밝힘. : 토방[土房] 《방언》 ‘뜰’의 방언(전라), 탁배기(濁배기)[濁배기] 《방언》, 그로테스크하다[프.grotesque하다] ㄹ기괴하다(奇怪--).
 사. 현재 공시적인 쓰임이 없는 「역사」 전문어 : 국자감 [國子監] 「역사」, 집현전 [集賢殿] 「역사」
 아. 「민속」, 「음악」, 「무용」 분야의 표제어가 방언형이나 비표준어 형이 분명한 경우 : 가마싸움 「민속」 경북 의성에서… , 띄뱃소리 「음악」 거문도의 뱃노래…
 가세진 [가세陣] 「민속」 호남 지방 농악(農樂)의…

<고려대 한국어대사전>의 형태 분석에서는 원칙적으로 역사적 변화 과정을 거치면서 공시적으로 분석하기 어려운 어형들은 비록 그 어원에 의한 단어 분석이 널리 알려진 경우라 하더라도 분석하지 않았는데(개구리, 고름, 무덤, 기둥 등) 그 외에도 (26 가~아)까지에서 보인 유형들은 비록 단일어는 아니지만 공시적인 형태 분석에서 배제한 예들이다.

(26 가)는 단어가 아닌 구 표제어들로 <고려대 한국어대사전>의 구 표제어는 46,045개(무품사어 45,163개, 준말 675개 포함)이다. 구 표제어들은 이미 표제어에서 띄어쓰기에 의해 형태가 구별되고 복합어만을 형태 분석한다는 <고려대 한국어대사전>의 입장에 따라 형태 분석 대상에서 제외된 것이다.

(26 나)는 고유명사인데 고유명사는 비록 복합어로 구성되었다 하더라도 일반 명사의 단어 구성과는 차이가 있다. 대사전을 대상으로 할 때, 일반적인 복합어의 형태 분석 대상은 최대 목록에 해당한다고 할 수 있기 때문에 그 분석상의 통계치에 대한 확인 등에 의의가 있을 수 있겠으나 고유명사는 그렇지 못하다. 따라서 고유

명사에 대한 분석은 국어사전이 아닌 별도의 연구를 통해 이루어져야 하며 국어사전에서 고유명사에 대한 분석은 불필요한 노력으로 판단하여 <고려대 한국어대사전>의 형태 분석에서는 배제하였다. 다만 고유명사를 이루는 데 주로 참여하는 접사는 별도의 표제어로 등재하여 고유명사를 이루는 양상을 간략히 밝혔다. <고려대 한국어대사전>에 등재된 고유명사의 수는 총 10,016개이다. 고유명사와 일반명사가 다의어로 묶어서 처리된 경우는 일반명사에 준하여 처리하였다.

(26 다)는 ‘앞으로 가, 뒤로 돌아’ 따위 제식 훈련의 명령이나 ‘아서라, 두어라’ 따위 발어사화된 감탄사 등이다. 이들은 공시적으로 비록 하나의 단어로 처리되어 이나  등의 어휘 범주를 부여받았다 하더라도 이들은 일반적인 복합어와는 다른 차원의 단어 형성 과정을 거친 것으로 판단된다. 이를 인용적 단어라고 할 것인지 통사적 단어라고 할 것인지, 구적 단어라고 할 것인지 등에는 여러 가지 논란의 요소가 있겠으나 적어도 일반적인 복합어와는 다른 원리에 의해 단어가 된 점이라는 데 근거하여 이들은 공시적인 형태 분석의 대상에서 배제하였다.

(26 라) 한자 부수의 이름 역시 특정한 관습에 따라 둘 이상의 언어 요소를 하나로 묶어서 이르는 말에 지나지 않으므로 일반적인 단어 형성과 궤를 달리한다고 보아 형태 분석의 대상으로 삼지 않았다. 이와 같이 특정 관습에 따라 둘 이상의 언어 요소를 하나로 묶어서 이르는 말에는 ‘한자 부수’ 이외에도 통사적 단위가 축약되어 하나의 형태로 실현되는 경우나 동양인명에서 관습에 따라 성과 인명이 별개의 단어임에도 하나의 단어처럼 처리되는 경우가 있다. 통사적 단위가 축약되어 하나의 형태로 실현되는 ‘게’(것이), ‘내’(나의), ‘-단다’(다고 한다), ‘-래도’(-라 하여도) 등은 뜻풀이에서 각

각의 형태가 통합된 과정을 밝혔으므로 형태 분석에서는 따로 분석해 보이지 않았고 동양인명의 ‘성+인명’의 분리는 고유명사를 분석하지 않는 태도에 따라 별도로 분석해 보이지 않은 것이다.

(26 마)와 같이 ‘직간접’과 같이 둘 이상의 단어가 혼성(混成, blending)을 거쳐 만들어진 단어의 경우, 본고에서 채택하고 있는 형태 중심의 분석 방식으로는 이 단어를 구성하고 있는 ‘직접+간접’의 단어 구성을 보이기 어렵기 때문에 형태 분석을 하지 않는 태도를 취하였다. 이들을 형태 단위로 분석한다면 ‘직+간접’이나 ‘직간+접’으로 분석해야 할 터인데 전자의 분석의 경우 ‘직(직접)+간접’과 같이 표시하는 방법이 없는 것은 아니나 ‘직(직접)’의 처리가 타당한지에 대한 입증이 어렵고 ‘직간+접’의 분석하는 경우 ‘접’에 대한 형태론적 처리 기준이 명확하지 않다는 점에 어려움이 있다. 향후 이에 대한 형태론적 논의가 진전되어 이들을 합리적으로 분석하는 방안이 연구자들에게 공유된다면 이들도 형태 분석의 대상이 될 수 있을 것이나 현재의 <고려대 한국어대사전>에서는 이러한 혼성어들의 형태 분석을 보류하고 이들을 단일어에 준하는 형태로 보는 입장을 취하였다.²³⁾

(26 바)의 방언과 비표준어도 공식적인 표준어 중심의 형태 분석을 보인다는 <고려대 한국어대사전> 형태 분석의 일반 원칙에 따라 형태 분석의 대상에서 제외하였다. 방언과 비표준어의 경계를 확인하는 것이 쉽지는 않지만 적어도 <고려대 한국어대사전>에서 방언으로 표시하고 비표준어로 설정하여 표준어와 구별한 요소들에 대해서는 표준어와의 대응 관계만을 중시하고 특별한 경우를 제외하고는 별도의 형태 정보나 뜻풀이 정보를 하지 않는 원칙에 따른

23) 경우는 다르지만 ‘두동실, 아사삭, 데구르르’ 등의 부분 중첩어의 경우도 기본적으로는 여기 ‘동실, 아삭, 데굴’ 등에서 파생된 것으로 볼 수 있으나 단일어에 준하여 처리하는 단어에 포함된다.

것이다.

(26 사)에서 제시한 『역사』 분야 전문어 중에서 현대적 쓰임을 확인하기 어려운 고문헌상의 용어들에 대해서는 형태 분석을 보류하는 입장을 취하였다. <고려대 한국어대사전>뿐만 아니라 현재 출간되어 있는 대부분의 국어사전에서 『역사』 분야 전문어의 처리는 매우 어려운 측면이 있다. 『역사』 분야 전문어는 다른 분야의 전문어에 비해 비정상적으로 표제어수가 많은데, 그 이유는 『역사』 분야 전문어가 다른 전문어들에 비해 학술적 혹은 분야적 전문성을 띠고 선정되었다기보다는 현대 국어에 쓰이지 않는 수많은 사어(死語)들이 『역사』 분야 전문어로 흡수되었기 때문이다. 공시적인 표제어를 담는 현대 국어 사전에서는 고유어 사어(死語)들은 ‘옛말’로 처리되어 체계적으로 배제되는 것이 일반적인데 한자어의 경우는 그러한 과정이 철저하게 관철되지 못하고 있는 실정이다. <고려대 한국어대사전>에서는 『역사』 분야 전문어를 ‘『역사학』 분야 전문어’와 ‘역사적 단어’로 나누어 학술 용어로 선택되기 어려운 ‘역사적 단어’들을 사어화하기 위한 절차의 하나로 이들에게는 별도의 형태 분석을 제공하지 않았다.²⁴⁾

24) 『역사』 분야 전문어와 ‘역사적 단어’의 구별에 대한 논의는 김양진(2006)의 논의를 참조할 수 있다. 김양진(2006)에서는 현행의 역사전문어들을 [한국사]와 [동양사], [서양사]로 구획하고 [한국사]의 경우는 철저히 ‘옛말’에 해당하는 단어류와 『역사학』에서 사용하는 용어류를 구별해야 함을 주장한 바 있다. 이 논문에서는 [동양사]의 경우는 [한국사]와 연결된 범위 안에서는 [한국학]과 같은 기준으로 사어(혹은 유행어) 처리가 이루어지고 [한국사]와 연결되지 않은 단어들의 경우는 [서양사]의 학술용어들과 함께 현대 국어의 학술 용어로 인정받을 수 있을 것으로 보았다. 즉 ‘문예부흥, 명예혁명, 차카타이 한국, 보불전쟁, ...’ 등은 현대어에서 『역사』 혹은 『역사학』 학술용어로 인정받을 수 있는 단어들이지만 ‘전옥서, 승지, 안찰사, ...’ 등은 교육적 목적이나 교양적 필요가 있는 경우가 아니라면 현대 국어로 선정될 필요가 없는 옛 단어들에 불과하다.

(26 아)에 제시된 『민속』 전문어나 『음악』, 『무용』 전문어 중에서 『민속』에 준하는 일부 단어들 중에는 전국적이고 일반적으로 통용되는 용어가 아닌 특정 지역에서만 통용되는 단어들이 있다. 예를 들어, ‘탈춤’이나 ‘타령’, ‘사위’ 등은 일반적으로 한국어 화자라면 누구나 알 수 있는 단어들이고 표준어로 처리된 『민속』, 『음악』, 『무용』 분야의 전문어이지만 (26 아)에 제시된 단어들은 특정 지역 이외에서는 사용되지 않는 엄밀히 말하면 방언에 준하는 단어들이다. 따라서 <고려대 한국어대사전>에서는 이들을 방언에 준하여 형태 분석을 하지 않는 태도를 취하였다.

3. <고려대 한국어대사전> 형태 분석 정보의 특징

<고려대 한국어대사전> 형태 분석 정보의 유용성은 표제어의 가장 미시적인 형태 단위까지 다분지로 분석한 후 각 형태 단위들의 형태론적 범주를 명시적으로 보인 것에서 출발한다.²⁵⁾ 분석에 따른 유용성은 분석 결과 자체의 1차적 유용성과 사전의 다른 정보들과의 관계 속에서 보이는 확장된 유용성으로 나누어 볼 수 있겠다.

우선 다분지 분석의 경우 개별 표제어의 가장 미시적인 형태 단

25) 기존 국어사전들은 대부분 이분지의 단어 구성을 붙임표 ‘·’를 표제어에 보이는 방식으로 단어 구성을 보이고 있다. 이는 전통적으로 ‘직접 성분 분석’으로 이르던 것으로 단어의 형성이 이루어지는 최종 단계에서의 분석만을 보여주는 방식이라 하겠다. 송재영(2008ㄱ/2008ㄴ)에서는 ‘이원 분석’이라 하였는데 ‘이원(二元)’의 의미를 특정할 수 없으므로 <고려대 한국어대사전>에서는 형식적인 분절을 의미하는 ‘이분지’라는 용어를 택하였다. 이에 상응하는 ‘다분지’의 개념은 배주채(2006)에서 ‘최종 성분 분석’이라는 개념으로, 소개한 것과 같은 것으로 송재영(2008ㄱ/2008ㄴ)은 이를 ‘다원 분석’이라 하였으나 <고려대 한국어대사전>에서는 이를 단순히 형식적으로 나누었다는 의미를 강조하기 위해서 ‘다분지’라는 용어를 채택한다.

위의 분석을 보인다는 1차적 유용성에서 더 나아가 동일한 형태를 공유한 표제어들 간의 확장 양상이나 표제어 분석 마디에 따라 달라지는 뜻풀이 정보, 표제어 내부의 위치 변화에 따른 형태소들의 문법 범주 변이 등을 보여줄 수 있다는 확장적 유용성을 가지고 있다.

다음으로 형태론적 범주의 명시는 공시적인 형태/형태소 목록의 추출이나 한국어 단어의 내부 구성 요소들의 자격을 보여주는 1차적 유용성 외에 한국어 접사 목록의 정밀화나 이형태 개념의 확장, 불완전 어근 목록의 확인 등 한국어 형태론에 유용한 많은 기초 자료를 제공한다.

아래에서 <고려대 한국어대사전>의 형태 분석 정보가 가지는 장점을 간단히 정리하였다.

- 〈1〉 다분지 분석을 통한 표제어 구성의 최대 미시적 분석
- 〈2〉 형태 분석을 통한 접사 목록의 정밀화
- 〈3〉 자소 단위의 분석 방식
- 〈4〉 공시적 분석 가능성을 고려한 이형태 개념의 확장
- 〈5〉 계열 관계와 통합 관계를 모두 고려한 형태 분석
- 〈6〉 불완전 어근의 도입

3.1. 다분지 분석을 통한 표제어 구성의 최대 미시적 분석

기존의 일반적인 국어사전들에서처럼 표제어의 단어 구성에 대해서 불임표(-)를 통한 이분지 분석을 보이는 것만으로는 표제어의 완전한 구성을 보여주는 데 한계가 있다.²⁶⁾ 다분지 분석 방식은 이러

26) 김창섭(1992)에서는 국어의 파생어가 맞춤법 등의 이유로 소위 ‘불임표’를 통하여 보여지기 어려운 경우 따로 자리를 마련하여 형태소 분석의 모습을 보여줄 필요를 언급한 바 있으나 기본적으로 표제어를 한 번만 분석한다는 이

한 한계의 상당 부분을 극복하는 데 기여할 수 있다.

- | | |
|------------------------|-----------------|
| (27) 가. 꽃맞이 [+꽃+맞-이] | 나. 꽃-맞이 |
| 꽃맞이굿 [+꽃+맞-이+굿] | 꽃맞이-굿 |
| 꽃맞이굿하다 [+꽃+맞-이+굿-하_다] | 꽃맞이굿-하다 |
| 꽃맞이하다 [+꽃+맞-이-하_다] | 꽃맞이-하다 |
| (28) 가. 가지급금 [=假+支給-金] | 나. 가-지급금 (假支給金) |
| 지급-금 [+支給-金] | 지급-금 (支給金) |
| 가-지급 [=假+支給] | 가-지급 (假支給) |
| (29) 가. 염주나무 [+念珠+나무] | 나. 염주-나무 (念珠--) |
| 개염주나무 [=개+念珠+나무] | 개-염주나무 (-念珠--) |

위 예에서 (27가), (28가), (29가)는 <고려대 한국어대사전>의 형태 분석 정보를 보인 것이고 (27나), (28나), (29나)는 일반적인 사전 표제어의 제시 방식을 보인 것이다. (27가)는 ‘꽃맞이’를 중심으로 한 단어의 확장 양상을 명시적으로 보여주고 있다. 또한 다분지 분석 방식은 (28)의 ‘가지급금’처럼 여기에 접두사와 접미사가 모두 결합한 경우 (28나)와 같은 2분지 구성보다 훨씬 유용한 방식이다. 즉 ‘지급금’에 ‘가-’가 붙은 것인지, ‘가지급’에 ‘-금’이 붙은 것인지, 더 나아가 ‘지급’을 중심으로 앞뒤에 결합한 ‘가-’와 ‘-금’이 ‘가지급’이나 ‘지급금’의 그것과 동일한 것인지 등을 보이기에는 (28가)의 방식이 적절하겠다. (29)도 마찬가지인데, 만일 사용자가 (29나)

분지의 원칙에서 벗어난 것은 아니다. 송재영(2008-)에서는 다분지의 분석을 보이는 국어사전의 예로 <조선어사전>(문세영, 1938)의 예를 들고 있으나 이는 주로 접두사 표제어에서 하위의 무표제어로 접두사 파생어의 예들을 묶음으로 처리할 때 나타나는 예외적인 처리에 불과하며 기존에 출간된 사전 중에서 다분지의 분석 방식을 일관성 있게 제시한 사전은 아직 없다고 보아야 할 것이다. <표준국어대사전>(1999)에서 어원란을 통하여 이러한 다분지의 예를 보이기는 하였으나 기본적으로 통시적 단어 분석을 제한된 것이고 전체 복합어 표제어를 대상으로 한 것도 아니다.

의 정보만 가지고 있다면 ‘개염주나무’에서 ‘염주나무’의 단어 구성을 보기 위해 다시 ‘염주나무’를 찾아야 하는 불편을 감수해야 한다. 이에 비해 (29가)의 정보를 가진 사용자라면 ‘개염주나무’의 형태 분석 정보만으로 해당 표제어의 구조를 쉽게 이해할 수 있다.

물론 이러한 다분지 방식의 형태 분석 제시는 궁극적으로 단어 구성의 계층 관계를 보이고 형태 의미 및 어휘 범주를 확인하고 연결해 줄 수 있는 방안으로까지 확장되어야 할 것이다.²⁷⁾ 그럼에도 불구하고 <고려대 한국어대사전>은 그 일차적 수준으로 다분지의 형태 목록을 일관성 있게 제시하였다는 점에 큰 의의를 두고 있다. 특히 이를 통해 단어의 확장 양상이나 단어 구성의 다양한 모습을 표제어와 함께 동시에 눈으로 직접 확인할 수 있다는 점에서 실용적, 학술적 가치는 충분하다고 판단된다.

3.2. 접사 목록의 정밀화에 따른 형태 분석

<고려대 한국어대사전>의 형태분석 작업은 한정된 표제어를 대상으로 한 것이 아니라 사전의 표제어 전반을 대상으로 한 작업이다. 대사전 규모의 어휘들에 대한 체계적인 형태분석을 위해서는 접사 목록의 전반적인 재정비가 필요하다. <고려대 한국어대사전>에서는 이러한 전체 복합어의 형태 분석의 기준이 될 공시적인 접사의 목록을 선별하고 이들을 통해 전체 복합어의 형태 분석을 새로이 하였다. 나아가 실제 형태 분석된 결과물에서 이들 접사가 일

27) 이에 대한 방법론은 송재영(2008ㄱ/2008ㄴ)에서 상세하게 논의된 바 있다. 다만 일반인 사용자를 대상으로 하는 국어사전에서 어디까지 이러한 복잡한 형태 정보를 보여 줄 것인지의 문제는 심각하게 논의될 필요가 있다. <고려대 한국어대사전>에서는 표제어의 형태 자체의 분석에 한정된 선까지만 제시하고 그 안에 내재된 다양한 계층 정보, 의미 정보, 형태 정보, 어원 정보 등의 세부 사항은 향후 구축될 전자 사전의 구성 내부에서만 보이기로 하였다.

정 정도 이상의 생산성을 가지고 있는지 등을 고려하여 접두사와 접미사 목록을 폐쇄 목록으로 새로이 작성하여 형태 분석 전체 결과를 재조정함으로써 형태 분석의 신뢰를 높일 수 있었다. <고려대 한국어대사전>에 접사 폐쇄 목록의 설정 과정에서, 기존 사전에서 접사로 보지 못했지만 접사로 보아야 할 예들이 새로 접사 목록에 추가되었고 기존 사전들에서 접사로 처리되었으나 접사로 보기 어려운 예들은 의존명사, 보조사, 불완전어근 등의 다른 형태 범주로 처리되어 전체 접사 목록에서 배제되었다.

다음은 <고려대 한국어대사전>에 추가된 몇몇 접두사의 예이다.

(30) <고려대 한국어대사전>에서 추가된 접두사의 예

가. 꽃-[=꽃] ㉠ 일부 명사 앞에 붙어, ‘맨 처음으로 된’ 또는 ‘맨 위에 뜬’의 뜻을 더하는 말. ㉠국/~등/~물/~잠/~소금/~소주/~다지.

나. 환-[=환] ㉠ 일부 명사 앞에 붙어, ㉠소리/~수작.

다. 객-[=客] ㉠ 일부 명사 앞에 붙어, ㉠소리/~식구/~술가락.

라. 희-[=稀] ㉠ 일부 명사 앞에 붙어, ㉠가스/~갈색/~금속.

국어사전에서의 접두사 설정 기준에 대해서는 유현경(1999)과 김창섭(2000)을 참고할 수 있다. 두 논문 모두 접두사의 설정 기준으로 위치성 이외에 ‘의미의 투명성’과 ‘조어의 생산성’을 들고 있다. 유현경(1999)에서는 이 두 조건 이외에 ‘어근의 불규칙성’과 ‘의미의 확장’이라는 부수적 조건을 들고 있고 김창섭(2000)에서는 ‘원 단어와의 유연성에서 멀어질 것’을 접두사 설정의 조건으로 들고 있다.

유현경(1999), 김창섭(2000)에서의 ‘의미의 투명성’은 의미하는 바가 분명하지 않다. 유현경(1999)에서는 ‘의미의 투명성’을 ‘실질 형태소처럼 구체적이고 실질적인 의미가 아니라 결합하는 어근의

의미에 따라 변화하는 추상적이고 투명한 의미’를 말한다고 서술하고 있는데 이러한 정의로는 파생 접(두)사의 의미와 굴절 접사의 의미 간의 차이를 명확히 구별하기 어렵다.

이 두 논문에서 말하는 ‘조어의 생산성’ 역시 접두사 설정의 확실한 기준이라고 하기 어렵다. 경우에 따라서는 ‘없애다’의 ‘-애-’나 ‘일으키다’의 ‘-으키-’처럼 유일한 결합형을 만드는 접사도 얼마든지 가능하기 때문이다.

유현경(1999)에서 제시한 ‘의미의 확장’과 김창섭(2000)에서 제시한 ‘원 단어와의 유연성에서 멀어질 것’의 조건은 기본적으로 같은 내용으로 볼 수 있다. 실질형태소와 접사를 구별하는 기준으로서의 ‘의미의 확장’은 ‘원 단어와의 유연성에서 떨어진 의미의 확장’을 뜻하기 때문이다. 하지만 원 단어로부터의 ‘의미의 확장’이 일반 합성어들에서도 얼마든지 가능하다는 점을 고려한다면(밤낮, 치맛바람, 꽃미남) 이러한 의미 확장이 접사의 설정 기준이 될 수 있을지는 불분명하다. <고려대 한국어대사전>에서는 이에 대한 대안으로 ‘범주 의미의 획득’을 제시하였다. ‘범주 의미’란 ‘명사화, 부사화, 자동사화, 타동사화, 형용사화’와 같은 문법 범주 의미와 ‘강조, 강세, 반복, 순환, 지소, 지대, 으뜸, 머금, 최초, 최종, 사람, 사물, 사건, 척도, 유사, 흡사, 긍정, 부정(멀시)...’ 등등의 어휘 범주 의미를 포괄한다.²⁸⁾

유현경(1999)에서는 “한자어 접두사가 고유어와 더 많이 결합하게 된다든지, 의미적으로 별다른 이유 없이 특정한 어근과는 결합하지 않는다든지 하는 특성”을 말하는 ‘어근의 불규칙성’을 별도의 조건으로 덧붙이고 있는데 ‘한자어+고유어’ 방식의 합성어나 선택

28) 이에 대한 자세한 논의는 지면상 <고려대 한국어대사전>의 접사 설정과 관련한 별도의 논문에서 따로 다루도록 하고 본고에서는 이상의 개략적인 소개로 그친다.

적 합성어 형성이 얼마든지 가능하다는 점에서 이러한 방식의 어근에 대한 제약이 접(두)사에 대한 설정 기준이 될 수 있을지는 분명하지 않다.

<고려대 한국어대사전>에서는 이상의 논의를 포함한 기존의 견해들을 참고하여 접두사의 설정 기준으로 ‘어기에 대하여 접두 위치에 있을 것’, ‘어기에 범주 의미를 추가할 것’, ‘어기와 분절 가능할 것’ 등으로 단순화하고 이에 따라 국어의 모든 복합어를 대상으로 새로이 접두사를 설정하였다.²⁹⁾ <고려대 한국어대사전>에서는 이에 따라 (30) 이외에도 접사의 계열 관계 등을 고려하여 ‘갈-, 감-, 골-, 내-, 된¹-, 된²-, 드²-, 뜬-. 뻘-, 땃-, 쇠²-. 예²-, 애¹-, 어리-. 졸¹-. 줌-, 진²-. 진¹-, 해²-. 햅-. 홉-, 경(硬)-, 단(短)-, 동(同)-, 목(木)-, 선(鮮)-, 아(亞)-, 악(惡)-, 운(聞)-, 전(全)-, 중¹(中)-, 중(曾)-, 평(平)-’ 등의 목록을 <표준국어대사전>의 접두사 목록에 추가하였다.

<고려대 한국어대사전>에서 추가된 접미사의 경우도 이와 마찬가지로 이다.

(31) <고려대 한국어대사전>에서 추가된 접미사의 예

가. -바리[-바리] 𠵿^ㄹ 일부 명사나 어근 뒤에 붙어, ‘그러한 사람’의 뜻과 앞잡는 뜻을 더하여 명사를 만드는 말. ㄱ^ㄱ군~/테~/벗~.³⁰⁾

29) ‘어기의 접두 위치에 있을 것’의 조건은 통사적 비자립성을 의미하는 것으로 관형사와 구별하기 위한 조건이다. ‘어기에 범주 의미를 추가할 것’의 조건에서의 ‘문법 범주 의미’와 ‘어휘 범주 의미’는 굴절 접사의 ‘통사적 기능 의미’와 구별하기 위한 것이다. ‘어기와 분절 가능할 것’의 조건은 역사적 접사류들, 즉 공시적으로 분석 가능하지 않은 접사적 요소들과 공시적 접사류를 구별하기 위한 조건으로 제기된 것이다. 이에 대해서도 <고려대 한국어대사전>에서의 접두사와 접미사 설정 기준에 대한 다른 논문에서 좀더 구체적으로 논의하도록 하겠다.

30) 이 접미사에 의해 파생되는 단어로는 ‘감바리, 꿈바리, 데통바리, 뒤듬바리, 뒤툼바리, 샘바리, 악바리, 애바리, 트레바리, 하바리(下-)·…’ 등이 있다.

- 나. -쇠⁶[-쇠] ㉠ ㉡ 일부 부정적인 의미를 지니는 명사나 어근에 붙어, ‘그러한 속성을 지닌 사람’의 뜻과 얹잡는 뜻을 더하여 명사를 만드는 말. 주로 남자를 가리킨다. ㉠곤~/구두~/꼴떡~/덜렁~/돌~/밥~/알랑~/텡~. ㉡ 일부 장소를 나타내는 명사 뒤에 붙어, ‘거기에서 살거나 일하는 사내’의 뜻과 얹잡는 뜻을 더하여 명사를 만드는 말. ㉠개땅~/관~/마당~/바닥~. ㉢ 무당들의 말에서 일부 ‘성씨(姓氏)’를 나타내는 은어 뒤에 붙어, 그것이 ‘성씨(姓氏)’임을 강조하는 말. ㉠나구~/신줄~.
- 다. -잖다[-잖(<_지+안+하)_다] ㉠ 일부 용언 어간 뒤에 쓰여, 그 말을 부정하거나 가치를 비하하는 의미를 덧붙이는 말. ㉠갈~/달갑~/시답~/적~.
- 라. -잖다[-잖(<_하_지+안+하)_다] ㉠ 일부 서술성 체언이나 서술성 어근의 뒤에 쓰여, 그 말을 부정하거나 가치를 비하하는 의미를 덧붙이는 말. ㉠가당~/변변~/시원~/편~/하~.

<고려대 한국어대사전>에서는 접미사에 대해서도 심혜령(1999), 구본관(2000) 등의 기존의 견해들을 참고하여 접미사의 설정 기준으로 ‘어기에 대하여 접미 위치에 있을 것’, ‘어기의 문법 범주를 지배할 것’, ‘어기에 범주 의미를 추가할 것’, ‘어기와 분절 가능할 것’ 등의 기준을 제시하고 이를 기준으로 국어의 전체 복합어를 대상으로 하여 <표1>에서 보인 접미사 목록을 확정하였다. (31 가~라)의 예는 기존 사전에서 접미사로 인정하지 않은 예 중에서 <고려대 한국어대사전>에서 접미사로 인정한 예의 일부를 보인 것이다.³¹⁾

<고려대 한국어대사전>에서는 접사의 동음이의를 구분한 것도 형태 분석에 적극적으로 활용하였다. 예를 들어 여기와의 형태 결합의 과정에서 접미사 ‘-이’의 다의성이 다음과 같이 확인된다.

31) 접미사 ‘-바리’에 대해서는 김양진·정연주(2008)에서 상세히 설명한 바 있으므로 참고하기 바란다. 접미사 ‘-쇠’는 기존 연구에서는 부분적으로 접사로 논의된 바는 있으나 국어사전에는 포함되지 못한 것이다. ‘-값다’, ‘-참다’에 대해서는 논란의 여지가 있으나 공식적으로는 특정 용언 어기에 붙어 그 어기의 의미의 반대되는 의미를 나타내는 접미사로 재구조화되었다고 보았다.

(32) 접사 ‘-이’

접사 (파생어수)		뜻풀이	용례	참고	품사
-이 ²⁶	27	①자음으로 끝나는 사람 이름 뒤에 붙어, 그 사람을 자신과 동년배나 아랫사람으로 대하는 뜻을 더하는 말. ②일부 수사에 붙어, ‘그만큼의 사람’의 뜻을 더하여 명사를 만드는 말. 주로 열 이하의 적은 수에 결합한다. ③일부 한자어 어근 뒤에 붙어, ‘그러한 속성을 가진 동물’의 뜻을 더하여 명사를 만드는 말.	①영순이/성철이/영진이/최강철이/배영환이. ②셋이/여섯이/우리 둘이는 서로 사랑한다. ③송충이/호랑이		명사 파생
-이 ²⁷	412	①일부 타동사 어간 뒤에 붙어, ‘그렇게 하는 일’의 뜻을 더하여 명사를 만드는 말.	별이/풀이/고기잡이/봄맞이/꽃꽂이/가슴앓이.		명사 파생
-이 ²⁸	851	일부 상태성 형용사의 어간 뒤에 붙어, 그 상태의 척도의 뜻을 더하여 명사를 만드는 말.	깊이/높이/길이/넓이.	-기 ²⁹ {2}	명사 파생
-이 ²⁹	1,256	①겉모양이나 속성을 더하는 일부 명사나 어근 뒤에 붙어, ‘그러한 모양이나 속성을 지닌 사람’의 뜻을 더하여 명사를 만드는 말.	애꾸눈이/절름발이/육손이/곰배팔이/똥똥이/명청이.		명사 파생
-이 ³⁰	74	일부 명사를 거듭 합친 말에 붙어, ‘그것 하나하나마다’ 또는 ‘그곳 전체에’의 뜻을 더하여 부사로 만드는 말.	날날이/일일이/집집이/짹짹이/간간이/울울이.		부사 파생
-이 ³¹	50	일부 형용사 어간이나 상태성 어근의 뒤에 붙어, ‘그러한 상태로’의 뜻을 더하여 부사로 만드는 말. 자음형 어간이나 어근 뒤에만 붙는다.	적이/많이/깨끗이/뚜렷이/깊숙이/따뜻이.	-리 ¹⁰ {1}, -히 ⁶ {1}	부사 파생

김창섭(1992)에서 이미 형용사·동사로부터의 명사화 접미사, 형용사로부터의 부사화 접미사, 명사로부터의 부사화 접미사, 그리고 사람 이름의 일종의 축소 명사화라는 여러 기능의 접미사들을 별개의 표제어로 독립시켜야 할 필요가 있음을 언급한 바 있다.

<고려대 한국어대사전>에서는 사람 이름의 일종의 축소 명사화의 접미사를 ‘-이26’으로, 동사로부터의 명사화 접미사를 ‘-이27’로, 형용사로부터의 명사화 접미사를 ‘-이28’로, 명사 혹은 명사성 어근으로부터의 명사화 접미사를 ‘-이29’로, 명사 혹은 명사형의 반복형으로부터의 부사화 접미사를 ‘-이30’으로 형용사 혹은 상태성 어근으로부터의 부사화 접미사를 ‘-이31’로 각각 별개의 표제어로 나누어 보인 것이다. 접미사 ‘-이’를 이와 같이 여러 가지 동음이의로 구별한 까닭은 각 접미사로부터 만들어지는 단어들의 범주 혹은 의미 부류가 전혀 다른 종류로 실현된다는 사실을 고려한 것이다.

특히 ‘-이26’와 ‘-이29’의 구별은 ‘-이26’이 조음(調音)의 기능을 기본으로 하고 새로운 의미를 추가하지 않는 화용적 요소에 가깝다면 ‘-이29’는 ‘...하는 사람 또는 동물’이라는 어휘 의미를 추가하는 어휘적 접사라는 점에 기반한 것이다. <고려대 한국어대사전>은 형태 분석의 과정에서 ‘-이26’~‘-이29’까지의 예들과 ‘-이30’, ‘-이31’의 예들이 갖는 의미들을 일일이 고려하였다. 다만 사전 간행본 형태 분석란에서는 가시적인 편의를 위하여 이러한 구별을 별도의 표지로 보이지는 않았다.

또한 기존 논의에서 애매하게 뭉뚱그려졌던 접사 파생어의 문제를 분명히 하는 데에도 이러한 형태 분석의 결과를 반영하였다. (33)의 예에서 보듯이 ‘돌’에 대해 (33가)부터 (33라)까지 합성어와 파생어의 구분을 엄격히 하여 표제어를 형태분석 하였다.³²⁾

32) 이에 대한 자세한 논의는 김양진(2004)의 논의를 참고할 것.

(33) ‘돌’ 결합형의 합성어와 파생어

- 가. 돌¹: 명사 ‘돌[石]’의 합성어 요소. 예) ‘돌부처’류: 돌다리¹ 등.
 나. 돌²: ‘돌다’의 어근 ‘돌-[廻]’의 합성어 요소. 예) ‘돌팔이’류: 돌 판죽, 돌매, 돌퇴 등.
 다. 돌-: 어근 ‘돌[溝<도랑]’의 합성어 요소. 예) ‘돌미나리’류: 돌다리², 돌둑, 돌창, 돌피 등.
 라. 돌-: 접미사 ‘돌-’. 야생의, 질 나쁜. 예) ‘돌배’류: 돌가시나무, 돌가자미, 돌갈매나무, 돌감, 돌감람나무, 돌능금, 돌단풍, 돌돔, 돌마자, 돌마타리, 돌말, 돌매화나무, 돌목상어, 돌바늘꽃, 돌배¹, 돌부채, 돌뽕나무, 돌삼, 돌상어, 돌씨¹, 돌지치, 돌콩, 돌팔 등

그동안 충분히 다루어지지 못한 고유어 접두성 어근 ‘생-’과 한자어 접두사 ‘생(生)-’ 등의 문제도 전체적으로 표제어의 일관성을 고려하여 처리하였다.

- (34) 가. 생(<새양<생강): 합성어로 처리. 예) 생쥐[+생+쥐], 생각시, 생 쪽매듭, 생차, 생토끼 ...
 나. 생(<西洋): 고유어화한 어근 합성어로 처리. 예) 생철[±생(<西洋)+철], 생목, 생무명, 생사, 생철 ...
 다. 생(生)-: 한자어 접두사 ‘生-’에 의한 파생어로 처리. 예) 생과일 [=生+과일], 생가죽, 생나무, 생김치, 생고생, 생고집, 생과부, 생부모, 생지옥, 생땅, 생이별, 생살, 생매, 생낮 ...

(34 가)의 ‘생’은 일부 연구에서는 접사처럼 처리된 것이지만 ‘생강’ 혹은 ‘새양’의 준말로서의 ‘생’이 존재함을 들어 모두 합성어로 처리하였다. 간혹 ‘생각시’와 같은 경우 한자어 접두사 ‘생(生)-’으로 처리된 경우도 있었는데 이는 이두식 표기에 불과하므로 형태 정보에서 바로잡았다. (34 나)의 ‘생(<西洋)’의 경우는 독립된 표제어로 삼지는 않았으나 ‘범주 의미’로 볼 만한 추상적 의미가 없다고 보아 불완전 어근으로 처리하였다. 다만 ‘양(洋)-’, ‘호(胡)-’를 접

두사로 처리한 것과 비교하면 일관성이 문제가 될 것이나 ‘당(唐)-, 왜(倭)-’ 등을 어휘적 요소로 처리한 것과의 중간적 범주로 처리한 것이다. 이들 접사성에 대한 문제는 좀더 천착이 필요하다. (34 다)의 한자어 접두사 ‘생(生)-’은 유현경(1999)에서 ‘의미의 확장’의 예로 제시된 것인데 <고려대 한국어대사전>에서는 이들의 접사성을 막연히 의미의 확장에 따른 다의성의 확보에 있다고 보지 않고 이들이 새로이 획득한 의미가 ‘범주 의미’로 평가된다고 보았기 때문에 이를 접사로 평가한 점에 차이가 있다.

그 외에 <고려대 한국어대사전>에서는 형태 정보의 일관성을 고려한 분석의 과정에서 국어의 ‘접두사+접미사’ 결합형의 목록을 다음과 같이 확인할 수 있었다.

(35) ‘접두사+접미사’ 구성 표제어의 예

군담[=군-談], 내뜨리다[=내-뜨리_다], 덧간[=덧-間], 덧배기[=덧-배기], 막치[=막-치], 만간[=만-間], 먹지[=먹-紙], 민짜[=민-짜], 본형[=本-形], 부쇠[=副-쇠], 생꾼[=生-꾼], 생짜[=生-짜], 악처[=惡-處], 양아치[=洋-아치], 외동이[=와-동이], 중쇠[=中-쇠], 쪽지[=쪽-紙], 찰기[=찰-氣], 푼내기[=푼-내기], 헛내기[=헛-내기],

마땅히 접사가 어기가 아닌 요소에 부착할 수 있는가 하는 질문이 있을 수 있으나 <고려대 한국어대사전>의 형태 분석은 학문적 질문을 던지기 위한 절차가 아니라 국어의 형태론적 질서를 ‘있는 그대로’ 보여주는 데 초점을 둔 과정이었음을 고려할 때 (35)와 같은 예들에 대한 학문적 해석은 이후에 연구자들을 통해 좀더 진지하게 논의될 부분이다.³³⁾

33) 국어에 ‘접두사+접미사’ 구성의 단어가 존재함에 대해서는 이미 김민수(1980) 등의 선행 연구에서 여러 차례 언급된 바 있다. <고려대 한국어대사전>에서는 접미사 형태는 접미사로, 접두사 형태는 접두사로 형태 분석을 일관성 있

3.3 자소 단위의 형태 분석

<고려대 한국어대사전>에서는 준말, 불규칙 용언의 분석이나 ‘사이시옷’의 처리 등에 일관성을 기하기 위해 자소 단위의 형태 분석을 허용하였다. 아울러 형태가 자소 단위로 분리될 경우, 그 의미나 기능을 분명히 하기 위해 형태소를 괄호 안에 넣어 밝힘으로써 형태소 단위의 검색이 가능하도록 처리하였다. 특히 앞 절에서 언급한 형태 보존 원리를 준수하기 위해 사이시옷을 살리되 선행 형태로부터 분리하여 사이시옷의 선행 형태가 분명히 드러나도록 처리하였다. 이를 통해 단순히 형태소 단위의 검색뿐 아니라 형태 단위 검색의 용이성 및 정확성을 추구할 수 있었던 점도 큰 수확이라고 할 수 있다.

(36) 자소 단위의 형태 분석과 형태소 분석의 실제

나가다[+ㄴ(나)_아+가_다], 다가오다[+다ㄱ(다그)_아+오_다], 떠나다[+ㅌ(뜨)_어+나_다], 바빠[+바빠(바쁘)-이], 슬피[+슬피(슬프)-이], 짜내다[+ㅈ(짜)_아+내_다], 차오르다[+ㅈ(차)_아+오르_다], 커지다[+ㄱ(크)_어+지_다], 터놓다[+ㅌ(트)_어+놓_다]

자소 단위의 형태 분석 방식이 가장 유용하게 적용된 것이 불규칙 용언의 활용형을 포함하는 표제어들의 형태 분석에서이다. (17)에서 간략히 소개한 예를 유형별로 다시 정리하면 다음과 같다.

- (37) 가. ㄷ불규칙 용언 : 걸음 [+걸(걷)-음], 걸리다 [+걸(걷)-리_다]
 나. ㅅ불규칙 용언 : 지은이 [+지(짓)-은+이],
 다. ㅂ불규칙 용언 : 가까이 [+가까(가깝)-이], 도움 [+도우(돕)-ㅁ],
 도와주다[+도우(돕)_아+주_다], 아름다움 [±아름-다우(답)-ㅁ],
 아름다워지다 [±아름-다우(답)_어+지_다]

게 적용하는 과정을 통해 이러한 결과가 나온 것이므로 이에 대한 이론적인 처리의 문제는 추후의 연구 과제로 남는다.

- 라. ㅎ불규칙 용언 : 까만색 [+까마(까맣)_ㄴ+色], 까매지다 [+까매
(+까맣_아)+지_다], 노래지다 [+노래(+노랑_아)+지_다]
 마. 르불규칙 용언 : 놀리놓다 [+놀르(누르)_어+놓_다], 빨리 [+빨
르(빠르)-이]
 바. 우불규칙 용언 : 퍼내다 [+포(푸)_어+내_다],
 사. 여불규칙 용언 : 그리하여 [+그리+하_어], 해내다 [+해(+하_
어)+내_다]
 아. 애불규칙 용언 : 그래서 [+그래(+그러_어)+서], 이래도 [+이래
(+이러_어)+도]

<고려대 한국어대사전>의 형태소 분석 과정에서 불규칙 용언과 접미사의 결합이 형태론적 조건에 따라 다양한 실현을 보이고 있음이 확인되었다. ㄴ불규칙 용언의 경우, 결합하는 접미사가 부사 파생의 ‘-이’인 경우는 형용사 어기만 결합하며 어기의 ‘ㄴ’이 탈락하는 방식으로 실현되지만 결합하는 접미사가 명사 파생의 ‘-이’인 경우는 어기가 형용사인 경우는 ‘더위, 추위’에서처럼 어기가 ‘더우-, 추우-’ 형으로 실현되었고, 어기가 동사인 경우는 부사 파생의 경우처럼 어기의 ‘ㄴ’이 탈락하는 양상을 띠었다.(녕마주이 등) 또 접미사가 형용사 어기에 결합한 경우도 접사의 의미가 ‘척도’가 아니라 ‘사람’의 의미일 때에는 어기의 ‘ㄴ’이 그대로 남겨지는 등(싱겁이 등) 다양한 특징이 드러났다.³⁴⁾

기존 학교 문법에서 불규칙으로 처리하지 못한 ‘그러다, 이러다, 저러다’의 활용 양상을 ‘애’불규칙 용언으로 불러야 할 만한 것임을 밝힌 것도 주요한 관찰이다. 기존의 ‘여’불규칙 용언과 관련한 논의에서 ‘하여’와 ‘해’의 처리가 불명확하게 처리되어온 것들도 ‘애’불규칙의 설정에 따라 “‘하다’가 ‘여’ 불규칙 활용과 ‘애’ 불규칙 활용을 보인다”는 식으로 재조정될 필요도 있을 것이다. ‘빨리’의 경우

34) 이에 대해서는 김양진·정연주(2008)의 논의를 참조할 것.

<고려대 한국어대사전>에서와 같은 형태 분석을 통해 ‘빠르-’의 이 형태 ‘빨르-’에 접미사 ‘-이’가 결합한 구성임을 보인 것도 기존 사전에서 충분히 보이지 못한 것이다.

<고려대 한국어대사전>에서는 또 사이시옷을 포함하는 합성어의 경우 사이시옷을 선행하는 체언의 일부로 보되 자소 단위 분석 방식을 따라 ‘ㅅ’ 형태를 여기로부터 구별하여 보이는 태도를 취했다.

- (38) 가. 나뭇가지 [+나무ㅅ+가지], 오랫동안 [+오래ㅅ+동안]
 나. 깃발2 [+旗ㅅ+발], 뒷산 [+뒤ㅅ+山]
 다. 숫자 [+數ㅅ+字], 찻잔 [+茶ㅅ+盞]

(38)과 같은 처리는 종이 사전에서 전자 사전으로 전환될 경우 형태 단위 검색의 용이성이나 효율성을 고려한 처리 방식이다. 나아가서 ‘ㅅ’을 별개의 형태소로 분리해 보는 방안에 대해서도 고려해볼 수 있을 것이다.³⁵⁾

마지막으로 국어사전의 표제어 중에는 불규칙 용언의 경우에서처럼 둘 이상의 언어 단위가 줄어들어 자소 단위의 분석 방식을 적용하여도 그 온전한 형태를 보존하기 어려운 예들이 다수 있다.³⁶⁾ 이들은 부득이 표제어의 표면형을 인정하여 1차로 형태 분석을 하되 일부 표면형에 ‘()’를 두어 그 안에서 형태소 분석의 과정을 제

35) 소위 사이시옷 ‘ㅅ’을 별개의 형태소로 볼 수 있다는 견해는 임흥빈(1981)에서의 ‘통사적 과격요소’, 최호철(1991)의 ‘의미상의 도열 상태’를 나타내는 요소 등의 언급에서 암시를 얻을 수 있고 김양진(2005ㄱ)에서는 구체적으로 사이시옷 ‘ㅅ’을 ‘선행체언의 총칭화’같은 의미 기능을 통해 형태소로 평가할 수 있음을 언급한 바 있다. 이에 대해서는 앞으로도 지속적으로 논의될 필요가 있다.

36) 이는 전통적인 형태론적 개념에서 복면 형태소에 해당하는 것들이다. ‘것이’의 줄어든 풀인 ‘게’, ‘터이다’의 줄어든 풀인 ‘테다’, ‘-시어요’의 줄어든 풀인 ‘-세요’ 등이 이에 해당한다.

시하는 것으로 처리하였다.

- (39) 가. 보내오다 [+보내(+보내_어)+오_다]
 나. 가려내다 [+가려(+가리_어)+내_다]
 다. 봐주다 [+봐(+보_아)+주_다]
 라. 에워싸다 [+에워(+에우_어)+싸_다]
 마. 돼먹다 [+돼(+되_어)+먹_다]
 바. 띄다[+띄(+뜨-이)_다]
 사. 글썸다 [+글썸(+글썸+이)_다]

국어의 ‘-어/-아’활용형과 ‘-이’계열 접미사들은 선행 어간이 모음으로 끝나는 경우 다양한 방식으로 축약형을 이루고 있다. ‘-어/-아’계열 활용형의 다양한 특성에 대하여는 김양잔·정경재(2008)에서 정리하여 보고한 바 있으나 이들의 형태론적 처리에 대해서 아직 합리적인 대안이 마련된 것은 아니다. <고려대 한국어대사전>에서는 우선 이들을 형태 단위로 분리하기 어려운 유형으로 보고 형태 분석란에서 표면형을 그대로 보이고 내부 형태 구조를 ‘()’ 안에 보이기로 절충하였으나 이에 대한 좀더 체계적인 처리 방안이 마련될 필요가 있다.

3.4. 공시적 형태 분석을 고려한 이형태 개념의 확장

<고려대 한국어대사전>에서 공시적 분석, 분석의 일관성 등의 요건에 따라 이형태 개념을 확장한 것도 중요한 성과이다. 흔히 이형태는 용언의 활용이나 체언에 조사가 결합하는 환경에서만 설정하는 것이 일반적이었다. 예를 들어, ‘먹-으니’, ‘가-니’의 ‘-으니’와 ‘-니’의 관계나 ‘도우-면’, ‘돕-고’의 ‘도우-’와 ‘돕-’, ‘나무가’와 ‘생각이’의 ‘이’와 ‘가’ 등이 대표적인 이형태로 설정되던 것들이다. 하지

만 공시적인 형태 분석의 대상을 한국어 복합어 전체로 확장하면서 이러한 요소들은 합성어나 파생어의 형태 분석의 결과에서도 여러 가지 유형으로 확인할 수 있게 되었다.

이형태는 일반적으로 서로 음운론적으로 관련 있는 형태가 음운론적 혹은 형태론적 환경에서 상보적으로 나타나는 둘 이상의 형태에 서로에 대해서 규정하는 관계를 나타내는 데 사용되는 용어이다. 다음과 같은 유형은 이러한 기준을 적용해 볼 때 역사적으로는 형태론적 흔적이나 음운 변화형이 화석형으로 남겨진 것이지만 공시적으로는 이형태의 범위에 포함될 성질의 것들이다.

- (40) 가. 암캐[=암+캐(개)], 수강아지[=수+강아지(강아지)], 머리카락[+머리+카락(가락)], 살코기[+살+코기(고기)] …
 나. 줍쌀[+줍(조)+쌀], 줍씨[+줍(조)+씨], 찹쌀[=찹(찰)+쌀], 휘-싸다[=휼(휘)-싸_다], 휘-쓸대[=휼(휘)-쓸_다] …

(40 가)는 역사적으로 ‘ㅎ’ 곡용어가 합성어나 파생어를 이루면서 남겨진 어형들이다. (40 나)는 이와 반대로 ‘ㅂ’계 어두 자음군을 가진 어형의 ‘ㅂ’이 받침 없는 선행 어기에 남겨진 어형이다. 이들은 공시적으로도 남겨진 부분의 분석된 의미는 분명하지만 형태론적으로 화석적이라는 점에 차이가 있는 어형들이다. 따라서 ‘개, 강아지, 컷, 키와, 닭, 탕나귀, 툄찌귀, 돼지, 병아리, 카락, 코기, 줍, 찹, 휘…’ 등을 공시적으로 화석화된 별개의 불완전 어근으로 설정하는 것보다는 이들을 각각 ‘개, 강아지, 컷, 기와, 닭, 당나귀, 돌찌귀, 돼지, 병아리, 가락, 고기, 조, 찰-, 휘-’ 등의 형태론적 이형태로 보는 것이 타당하다고 평가하였다.

이러한 태도는 아래의 ‘아드님’류의 단어들에도 확장되었다.

(41) ‘아드-님’류

- 가. 아드님 [+아드(아들)-님], 나날이[+나(날)+날-이], 소나무[+소(솔)+나무]
 나. 다달이 [+다(달)+달-이], 무더위[+무(물)+더위(땀)-이], 미단이 [+미(밀)+단-이]
 다. 고살 [+고(골)+살], 부삽 [+부(불)+鋤], 화살[+화(활)+살], 무쇠 [+무(물)+쇠]
 라. 꼬집다 [+ㄷ(끝)+집_다], 무좀 [+무(물)+좀],

(41 가~라)에 제시된 단어들은 각각 후행어의 어두 자음이 ‘ㄴ, ㄷ, ㅅ, ㅈ’인 경우 선행어의 어말 ‘ㄷ’음이 탈락한 경우이다. 이러한 ‘ㄷ’ 탈락은 공시적으로는 이미 생산적이지 않고 통시적인 변화를 거친 것으로 평가되고 있지만 그 양상은 일정한 음운론적 환경이 확인되는 것이므로 공시적으로 처리하는 데에도 큰 문제가 없다고 평가된다. 따라서 ‘-님, 날, 나무, 달, 더위, 단이, 살, 삽, 활, 쇠, 집다, 좀’ 등의 앞에 나타나는 ‘아드, 나, 소, 다, 무, 미, 고, 부, 화, 무, ㄷ, 무’ 등은 각각 ‘아들, 날, 솔, 달, 물, 밀, 골, 불, 활, 물, 끝, 물’의 이형태로 보는 것이 이들을 화석형의 불완전 어근으로 해석하거나 분석하지 않고 통합형으로 처리하는 태도보다는 좀더 합리적인 태도로 평가된다.

이상의 예들은 비록 역사적으로 형성된 단어들이지만 공시적으로 분석이 가능하고 공시적인 형태 이론의 입장에서 충분히 설명 가능한 것들로 평가된 것이다. 그러나 <고려대 한국어대사전>에서는 사전 표제어의 구성 중 이러한 설명이 가능하지 않은 역사적 단어의 경우 분석하지 않았다. 이는 공시적으로 활성 상태인 형태만을 대상으로 형태소를 추출한다는 장점을 갖는다.³⁷⁾

이에 따라 <고려대 한국어대사전>에서는 자음형 어기에 결합했

37) 역사적 부류에 대해서는 별도의 처리를 거쳐야 할 것이다.

을 때 비분절형으로 나타나는 접미사들이 모음형 어기에 결합했을 때, 마치 분석되는 것처럼 보이는 예들을 분석하지 않는 것으로 일관되게 처리하였다.

- (42) 가. 뜨덤뜨덤[±뜨덤±뜨덤], 도담도담[±도담±도담], 발맘발맘[±발맘±발맘], 자밤자밤[±자밤±자밤], 주섬주섬[±주섬±주섬]…
 나. 보암보암[±보암±보암], 죄암죄암[±죄암±죄암](썸썸/*죄엄죄엄/*썸썸), 쥐엄쥐엄[±쥐엄±쥐엄], 기엄기엄[±기엄±기엄], 띄엄띄엄[±띄엄±띄엄], 쉬엄쉬엄[±쉬엄±쉬엄], 이엄이엄[±이엄±이엄] ; 도리암직하다[±도리암직-하_다], 보암직하다[±보암직-하_다], 하엄직하다[±하엄직-하_다] …

(42 가)와 같이 접미사 ‘-암/-엄’이 자음형 어기에 결합했을 때 비분절형 접미사들은 모음이 결합했을 때에도 분석하지 않는다. 자음형 어기 뒤에서 분석할 수 없는 접미사를 모음형 어기에서 분석한다는 것은 형태론적인 일관성을 지키지 못하는 것이다. 만약 모음형 어기 뒤의 접미 형태소 ‘-암/-엄’이 형태론적으로 가시적이라면 (42 가)의 예들을 ‘뜯엄뜯엄, 돈엄돈엄, 뽕엄뽕엄, 잡엄잡엄, 좇엄좇엄’으로 표기하지 못할 이유가 없다. 그러나 공시적으로 이들을 (42 가)와 같이 연철하여 표기하는 이유는 이들이 공시적으로 형태소로 분리되기 어렵기 때문이다. ‘-암직’, ‘-엄직’에 대한 분석의 경우 자음형 어기의 뒤에서 나타나는 사례가 없지만 이러한 논리에 따라 별도로 분석하지 않는 원칙에 따랐다(‘-음직’, ‘-ㅁ직’류는 공시적으로 분석하였다).

이러한 원칙은 ‘-으스름/-으스레’류의 분석 태도에도 일관성 있게 반영되었다.

- (43) 가. 불그스름하다[±불그스름-하_다], 거무스름하다[±거무스름-하_다], 기르스름하다[±기르스름-하_다], 말그스름하다[±말그스름-하_다], 배스름하다, 비스름하다, 알브스름하다, 파르스름하다 ; 가느스름하다, 둥그스름하다, 허여스름하다, ; 구부스름하다, ; 희읍스름하다, 해읍스름하다 …
- 나. 푸르스름하다[±푸르스름-하_다], 누르스름하다[±누르스름-하_다], 노르스름하다[±노르스름-하_다] …

(43 나)의 예들은 일견 ‘푸르-스름-하_다’로 분석될 수 있을 듯하지만 (43 가)의 예를 볼 때 분석하지 않는 것이 타당하다. ‘불그스름, 거무스름, 기르스름, 말그스름, …’ 등에서의 ‘-으스름’은 ‘노랗다, 파랗다, 그렇다’ 등에서 어간의 일부로 융합된 ‘ㅎ(<ㅎ-)’과 같이 처리되어야 할 것이다. 이들은 일정한 의미를 공유하고 형태론적으로 일정한 유형을 이루기는 하지만 공시적인 입장에서는 독립된 형태소로 평가될 수 없다.

만약 (43 나)의 예들에 근거해서 여기 ‘푸르-’와 접미사 ‘-스름’을 분리하는 입장을 따른다면 (43 가)의 예들에서도 접미사 ‘-으스름, -으스레’와 불완전 어근 ‘불그, 거무, 기르, 말그, 배, 비, 알브, 파르, 가느, 둥그, 허여, 구부, 희읍, 해읍’ 등이 분리되어야 할 것이다. 결과적으로 ‘푸르-, 누르-, 노르-’의 세 어형을 ‘-스름-’으로부터 분리하기 위해 ‘불그, 거무, 기르, 말그, 배, 비, 알브, 파르, 가느, 둥그, 허여, 구부, 희읍, 해읍’ 등의 불완전 어근을 설정해야 하는 부담을 가져야 한다는 것이다. <고려대 한국어대사전>에서는 ‘불그스름, 거무스름, 기르스름…’ 등과의 동형성을 고려하여 ‘푸르스름, 누르스름, 노르스름’을 형태 분석하지 않는 분석 태도를 유지하였다.

이는 ‘-스름’, ‘-스레’의 문제에 한정되지 않고 ‘-레(불그레, 부끄레…), -무레(불그무레, 거무레, 시크무레, 달크무레…), -드레, -시레, -주그레, -주레, -지레, -질구레, -찌레, -프레/푸레(희습푸레, 어습푸

레)’ 등 색채어나 기타 형용성 어간에 결합하여 새로운 불완전 어근을 이루는 접사적 요소들은 물론 일반적인 역사적 단어형에 모두 해당된다.³⁸⁾

3.5. 계열 관계와 통합 관계를 모두 고려한 형태 분석

<고려대 한국어대사전>의 형태 분석의 또다른 장점은 모든 형태 분석이 계열관계와 통합관계를 함께 고려하여 이루어졌다는 점이다.

(44) 가. 커지다 [+크(크)_어+지_다]

나. 커다랗다 [±커-다랗_다]

(45) 가. 밝아지다 [+밝_아+지_다], 쏟아지다 [+쏟_아+지_다], 넘어지다 [+넘_어+지_다], 이루어지다[+이루_어+지_다], 주어지다 [+주_어+지_다] …

나. 길다랗다 [+길-다랗_다], 굵다랗다 [+굵-다랗_다], 굵다랗다 [+굵-다랗_다], 깊다랗다 [+깊-다랗_다], 높다랗다 [+높-다랗_다], 되다랗다 [+되-다랗_다], 두껍다랗다 [+두껍-다랗_다], 작다랗다 [+작-다랗_다], 잔다랗다 [±잔-다랗_다], 좁다랗다 [+좁-다랗_다] …

‘-지다’는 용언 어간과 직접 결합하는 일이 없이 항상 연결어미 ‘-어’를 매개하여 합성어를 이루지만 ‘-다랗-’은 반드시 용언 어간에 직접 결합하거나 용언 어간으로부터 변화된 일부 불완전 어근에만 결합하는 특성을 지닌다. ‘커지다’와 ‘커다랗다’의 ‘커’는 일견 동일한 형태 정보를 가지고 있을 듯하나 이러한 계열관계와 통합관계를 고려할 때 ‘커지다’는 ‘[+크_어+지_다]’로 분석되는 데 비해 ‘커다

38) ‘너무, 도로, 마주, 겨우’ 등과 같은 계열에 속하는 ‘되우, 외우’ 등의 ‘-우’ 등도 이러한 유형에 해당하나 <고려대 한국어대사전>(초판, 2009)에서는 이러한 부분이 철저히 지켜지는 못했다. 재판에서는 마땅히 수정되어야 할 것이다.

랴다’는 ‘[±커-다랴_다]’로 분석된 것이다. 계열관계와 통합관계의 어느 한쪽만을 참고할 때는 일견 동일한 형태로 보이는 형태 단위들도 계열관계와 통합관계를 동시에 고려할 때 그 모순의 충돌을 해결하기 위한 방안의 마련이 검토될 수 있다. <고려대 한국어대사전>에서는 이러한 방안들이 구체적으로 실현되어 있다는 점에서 단순한 형태 분석의 일관성을 뛰어넘는 정밀성을 지니게 되었다고 자인한다.

3.6. 불완전 어근의 인정

<고려대 한국어대사전>에서는 완전 해석에 따른 형태 분석 원리에 따라 모든 형태소를 분석함으로써 공시적 형태소 목록인 자립어, 접두사, 접미사, 어미로 설명할 수 없는 역사적 결과로서의 불완전 어근의 목록을 따로 확인할 수 있었다. 이들은 비록 제약적인 환경에서 폐쇄 목록으로 나타나는 것이 <고려대 한국어대사전>에서는 이들 불완전 어근에 대해서도 최대한 동일한 형태에 대해서 동일한 의미를 부여하기 위해 노력하였다.

- (46) 가. 마땅찮다 [±마땅-찮_다] (사람이 일이나 현상 따위가) 알맞지 않거나 마음에 달갑지 않다.
 나. 마땅하다 [±마땅-하_다] (일이나 현상이 어찌하여야) 이치로 보아 당연하다.
 다. 마땅히 [±마땅-히] ① 그렇게 하는 것이 이치로 보아 당연하게.
- (47) 가. 마뜩이 [±마뜩-이] 제법 마음에 들어 좋게.
 나. 마뜩잖다 [±마뜩-잖_다] (일이나 현상 따위가) 별로 마음에 달갑지 않다.
 다. 마뜩하다 [±마뜩-하_다] ((주로 ‘못하다’, ‘않다’와 함께 쓰여)) (사람이 무엇이) 제법 마음에 들어 좋다.

(46), (47)의 ‘마땅’과 ‘마뜩’은 불완전 어근으로 결합하는 형태가 많지 않아서 단어 각각의 의미를 분명히 알기 어렵다. 그러나 이런 경우도 되도록 같은 형태소로 해석될 수 있도록 처리하였다.

<고려대 한국어대사전>에서는 2자어, 3자어 한자어의 형태 분석에서도 이러한 불완전 어근들을 적극적으로 해석하고 그 의미를 구체화하려고 노력하였다.

- (48) 가. 전자 [電磁] … 전기와 자기를 아울러 이르는 말. <원어>전자기.
 나. 전자기 [電磁氣]³⁹ … 전기와 자기를 아울러 이르는 말. <약어>전자(電磁).
 (49) 가. 전자계 [+電磁±界] … 전기장과 자기장이 서로 연관되어 함께 나타난 공간.
 나. 전자기력² [+電磁+力] … 전기나 자기에 의한 힘을 통틀어 이르는 말. <원어>전자기력.
 다. 전자기장 [+電磁+場] … 전기장과 자기장이 서로 연관되어 함께 나타난 공간. <원어>전자기장.
 라. 전자파¹ [+電磁+波] … 주기적으로 그 세기가 변하는 전자기장(電磁氣場)이 진공이나 물질 속을 전파해 가는 현상. 또는 그 파동. <원어>전자기파.
 마. 전자기학² [+電磁+學] … 전자기(電磁氣) 현상 전반에 관하여 연구하는 학문. <원어>전자기학.
 (50) 전자석 [±電+磁石] … 전류가 흐르면 자성(磁性)을 띠고, 전류가 끊기면 원래의 상태로 돌아가는 일시적 자석. <유의> 전기 자석¹(電氣磁石)

39) ‘전자기(電磁氣)’는 2.2.2절의 (26)에서 언급한 형태 분석에서 제외할 대상 중 둘 이상의 단어가 뒤섞여 하나의 단어로 된 경우에 속하기 때문에 따로 분석하지 않았다. 이에 따라 이 단어는 다음과 같은 복합어의 어기로 사용될 때 단일어로 분석된다. “전자기나팔 [+電磁氣+喇叭], 전자기력 [+電磁氣-力], 전자기장 [+電磁氣+場], 전자기파 [+電磁氣+波], 전자기학 [+電磁氣-學], 전자기혼 [+電磁氣+영.horn]”

(50)의 ‘전자석(電磁石)’은 한자 구성은 ‘전자(電磁)’ 혹은 ‘전자기(電磁氣)’와 유사한 것으로 보이지만 전혀 다른 구성을 갖는 예이다. ‘전자기’는 ‘전기’와 ‘자기’의 축약어로 ‘전자’의 원어에 해당하는 말로 “전기와 자기를 아울러 이르는 말”이고 따라서 별도의 형태학을 갖지 않는 말이지만, ‘전자석’은 “전류가 통하는... 자석”이라는 뜻으로 ‘전기 자석’과 유의 관계에 있는 말로 ‘자석’을 형태학으로 하는 말이다. 따라서 <고려대 한국어대사전>에서 ‘전자(電磁)’ 혹은 ‘전자기(電磁氣)’는 별도의 언어 단위로 분석되지 않고 단일어에 준하여 처리되지만, ‘전자석’은 ‘자석(磁石)’이라는 자립어의 분석을 통해 ‘전류가 흐르는’의 의미가 부여되는 ‘±電’이라는 불완전어근 형태소가 분리되므로 이 단어는 [±電+磁石]과 같이 분석되는 것이다.

이와 같이 <고려대 한국어대사전>에서는 형태 분석의 과정에서 접사로 볼 수 없는 1음절 한자어의 불완전 어근들에 정밀한 의미를 부여하기 위해 다양한 방식으로 노력하였다. 다만 이를 정밀화하고 이러한 형태 의미를 구체화하기 위해서는 <고려대 한국어대사전>의 1차적인 성과를 바탕으로 앞으로 좀더 치밀한 연구가 뒷받침되어야 할 필요가 있다.

4. 맺음말

이상에서 <고려대 한국어대사전>의 형태 분석 정보의 실제와 학술적 의의에 대해서 개략적으로 살펴보았다. <고려대 한국어대사전>은 한국어 대사전으로는 최초로 한국어 복합어 전체의 형태 분석 정보를 수록한 사전이다. 이 사전에는 현대 국어 표준어 복합어를 대상으로 하여 174,686개의 형태 분석 정보가 결합된 표제어를

제공하였다. 모든 복합어의 형태 분석은 공시적 기준을 바탕으로 다분지 방식으로 분석되었으며 계열 관계와 통합 관계를 철저히 고려하여 자소 단위의 최대 미시적 분석을 행하였다.

이를 바탕으로 <고려대 한국어대사전>에서는 약 11,000여 개의 고유어 형태소(자립 어근 7,000여개, 불완전 어근 4,000여개)와 117,000여 개의 한자어 형태소(자립 어근 110,840개, 불완전 어근 6,300여개), 14,000여개의 외래어 형태소(자립 어근 13,629, 불완전 어근 500여개)의 한국어 형태소 목록을 확정할 수 있게 되었다.⁴⁰⁾

각 품사별 고유어 형태소와 한자어 형태소, 외래어 형태소 및 접사별 결합 어기의 유형과 빈도, 각 개별 어근(완전 어근과 불완전 어근 포함)의 결합형의 유형과 빈도 등이 확인되고 이를 통해 국어 형태론 연구에 중요한 진전이 있을 것으로 예상된다. 나아가 이러한 자료를 통해 향후 한국어 형태소 사전 구축의 기반이 마련되었다는 것도 <고려대 한국어대사전> 형태소 분석 작업의 주요한 성과이다. 각 형태소 목록에 맞는 형태 정보들이 체계적으로 정리되고 일목요연하게 간추려진 형태소 사전을 기대해 본다. ◆

주제어 : <고려대 한국어대사전>, 형태 분석 정보, 형태 표지, 파생어, 합성어, 불완전어근, 형태소 목록, 단어 형성

40) 현재 <고려대 한국어대사전>(초판, 2009)의 개정판 작업이 진행 중이다. 각 형태소의 정확한 수치는 개정판 작업이 완료되면서 <고려대 한국어대사전 편찬백서>(2010, 예정) 등을 통해 공식적으로 제공될 예정이다.

참고문헌

- 구본관. 1998. “접미사의 사전적 처리,” 「새국어생활」 8-1. 국립국어연구원. 23-48.
- 김양진. 1999. 국어 형태 정보 연구. 고려대 박사학위논문.
- _____. 2004. “접두사 ‘쇠-, 돌-, 알-, 애-’의 사전 처리에 대하여,” 「한국사전학」 3. 한국사전학회. 5-25.
- _____. 2005ㄱ. “형태소와 의미,” 국어 연구와 의미 정보. 월인. pp.231~249.
- _____. 2005ㄴ. “일음절 한자어 어기의 형태론적 재해석,” 「어문논집」 52. 민족어문학회. 97-120.
- _____. 2006. “국어 중사전의 전문어 표제어 선정에 대하여 -역사전문어를 중심으로”, 한국사전학 7. 한국사전학회. pp.191-215.
- 김양진·정경재. 2008. “국어사전에서 ‘-아/-어’ 계열 어미의 문법 정보 기술에 대하여,” 한국사전학 12. 한국사전학회. 97-121.
- 김양진·정연주. 2008. “접미사 ‘-이’의 동음이의 구별 방안 - 형태론적 특성을 중심으로,” 한국사전학 12. 한국사전학회. 185-207.
- 김양진. 2009ㄱ. “특수어미 ‘-느니’의 형태통사론,” 한글 284. 한글학회. 73-105.
- _____. 2009ㄴ. “한자어의 형태 분석과 동음이의 처리 - 두 가지 ‘문학사(文學史)’를 중심으로,” 형태론 11-2. pp.377-405. 형태론연구회.
- 김창섭. 1992. “파생접사의 뜻풀이,” 「새국어생활」 2-1. 국립국어연구원. 72-88.
- _____. 1998. “접두사의 사전적 처리,” 「새국어생활」 8-1. 국립국어연구원. 5-22.
- _____. 2001. “한자어 형성과 고유어 문법의 제약,” 「국어학」 37, 177-195.
- 김홍규·강범모. 2000. 「한국어 형태소 및 어휘 사용 빈도의 분석 1」 서울: 고려대 민족문화연구원.
- _____. 2004. 「한국어 형태소 및 어휘 사용 빈도의 분석 2」 서울: 고려대 민족문화연구원.
- _____. 2009. 「한국어 사용 빈도」. 한국문화사.
- 도원영. 2003. “「민언사전」(가칭)의 특징,” 한국사전학 2. 한국사전학회. 85-109.
- 민경모. 2003. “현대 국어 형태 정보 주석 말뭉치의 주석 단위 표기 방안에 대하여,” 「언어 정보와 사전 편찬」 12-13. 연세대학교 언어정보연구회. 69-111.
- 배주채. 2006. “사전 표제어의 성분구조 표시에 대하여,” 「이병근선생 퇴임기

념 국어학논총」. 태학사.

- 서태길·김양진·이상혁·도원영·권오희 공역. 2008. 「사전 편찬의 원리와 실제」. 제이앤씨
- 성광수. 1992. “문법 형태소의 뜻풀이,” 새국어생활 2-1. 국립국어연구원. 89-100.
- 송재영. 2008ㄱ. 국어사전에서의 조어 정보 기술 방안 연구. 연세대 석사학위논문.
- _____. 2008ㄴ. “국어사전에서의 조어 정보에 대하여,” 한국사전학회 제13회 학술대회 발표논문집. 177-188.
- 심혜령. 1999. “접미사의 사전적 처리,” 사전편찬학 연구 9. 연세대학교 언어정보개발원. 151-181.
- 유현경. 1999. “접두사의 사전적 처리,” 사전편찬학 연구 9. 연세대학교 언어정보개발원. 183-203.
- _____. 2000. “사전에서의 동형어 구별을 위한 새로운 제안-구분자(distinguisher)의 사용에 대하여-,” 사전편찬학 연구 10. 연세대학교 언어정보개발원. 133-159.
- _____. 2002. “부사형 어미와 접속어미,” 한국어학 16. 한국어학회. 333-352.
- 이병근. 1989. “국어사전과 파생어,” 어학연구 22-3. 서울대학교 어학연구소. 389-407.
- 이상복. 1987. “국어 사전 편찬과 문법 형태소의 처리-조사와 연결어미를 중심으로,” 인문과학 59-88.
- _____. 1991. “형태소 복합법에 관련된 몇 가지 문제”, 「국어의 이해와 인식」(갈음 김석득교수 회갑기념 논문집). 229-246.
- 이현희. 2008. “사전 표제어의 형태분석을 위한 1음절 한자 처리 방안 연구,” 「우리어문연구」 32, 117-158.
- 정동환. “우리말 사전 처리에서 본 앞가지 뜻의 정리,” 건국 어문학 11-12. 343-373.
- 조재수. 1989. “국어사전에서의 비자립어 다루기 문제(Ⅰ),” 「애산학보 7」.
- 홍종선. 2001. “국어 단어 형태의 정보화”, 배달말 28. 배달말학회. 29-55.
- 홍종선·최호철·한정환·최경봉·김양진·도원영·이상혁. 2009. 「국어 사전학 개론」 서울: 제이앤씨.

【국문초록】

본고는 <고려대 한국어대사전>의 표제어부에 제시된 ‘형태 분석 정보’에 대한 소개와 그 효용성에 대해 논의하는 것을 목적으로 한다. 조어법이 발달한 한국어의 특성을 고려하여 <고려대 한국어대사전>에서는 표제어를 제시하면서 해당 표제어가 복합어(파생어, 합성어)일 경우 그 구성을 보여줄 수 있는 표지를 부착하여 형태 분석 정보를 함께 제시하였다. 이러한 형태 분석 정보는 기존 사전들에서 볼 수 없는 획기적인 정보 제공 방식으로 대사전 편찬 사상 표제어의 분석 단계를 다분지로 확장하였다는 점, 표제어 구성 형식들의 형태론적 범주를 명시화하였다는 점에서 실용적 가치와 학술적 가치를 모두 구현할 수 있게 되었다.

<고려대 한국어대사전>에 제시된 형태 분석 정보는 다음과 같이 구성된다. 우선 분석된 표제어 구성 성분들에 부착하는 형태 분석 표지가 있다. ‘완전 어근/자립어’의 경우는 ‘+’로, 불완전어근/비자립어는 ‘±’, 접두사는 ‘=’, 접미사는 ‘-’, 어미는 ‘_’의 표지를 부착한다. 예를 들어 명사와 명사의 결합인 ‘길가’는 ‘[+길+가]’로 불완전어근인 ‘오술’을 포함하는 ‘오술길’은 ‘[±오술+길]’로 형태 분석이 제시된다. 접사의 경우 접두 파생어는 ‘가지급[=假+支給]’, 접미 파생어는 ‘깊이[+깊-이]’로 분석하였다. 이 외에도 특이하게 어미를 포함하는 표제어도 적지 않은데, ‘밝아지다[+밝_아+지_다]’, ‘높다랗다[+높-다랗_다]’ 등으로 분석하였다.

현재 <고려대 한국어대사전>에는 총 386,889개의 표제어 중 174,686개의 표제어에 대한 형태 분석 정보가 제시되어 있으며, 이

중 접두 파생어는 8,518개, 접미 파생어는 77,413개, 접사를 포함하지 않은 합성어는 86,197개, 어미를 포함한 경우는 2,128개가 있다. 어종별로 고유어 포함 복합어 111,446개, 한자어 포함 복합어 62,501개와 기타 736개 등이 있다. 이렇게 분석된 형태 분석 DB는 언어 분석의 기초 자료로서 다음과 같은 의의를 가진다. 첫째, 한국어의 조어 구성을 체계적이고 전체적으로 조망할 수 있는 토대를 마련하였다. 둘째, 한국어의 공식적 언어를 대상으로 형태소의 목록을 추출하여 제시할 수 있게 되었다. 셋째, 언어 자료에 대한 통계 처리나 말뭉치 구축, 형태 분석 말뭉치, 신어, 번역 등 다양한 자연 언어 처리의 토대 자료가 마련되었다.

국어학이나 언어학적 관점에서도 형태 분석 정보는 큰 의의를 가진다.

첫째, 한국어의 형태/형태소의 총목록을 제시할 수 있는 기반을 마련하였다. 실제로 <고려대 한국어대사전>에서 단일어 동사는 2,077개, 단일어 형용사는 432개인데 비해 전체 동사의 수는 55,308개, 전체 형용사의 수는 12,431개이다. 즉 파생이나 복합의 과정을 통해 이루어진 어휘가 상당수라는 것인데, 이에 대한 국어학적, 언어학적 접근이 좀더 용이하게 체계적으로 이루어지기 위해서는 어휘 각각에 대한 형태 분석 작업이 필수적이라 하겠다.

둘째로 단어의 구조 및 의미에 대해 사용자가 정확하고 쉽게 확인할 수 있다는 장점이 있다. 기존 사전들에서 ‘다리미질’과 ‘다림질’은 ‘준말-본말’ 관계로 처리되어 왔다. 그러나 형태 분석 정보를 ‘다리미질[+다리미-질]’과 ‘다림질[+다리-口-질]’과 같이 제시함으로써 이들이 단순한 유의 관계의 어휘임을 볼 수 있었고 이것이 표제어의 뜻풀이에 반영될 수 있었다. 셋째로, 동일 음상의 표제어들을 구분할 수 있는 언어 정보들이 제시되었다. 형태 분석 정보를 제시

함으로써 사용자들은 동일 음성 표제어들의 뜻풀이를 구조적으로 이해할 수 있을 뿐 아니라 어휘들 간의 관계까지도 체계적으로 파악할 수 있게 되었다.

넷째, 동일 원어 정보의 표제어들도 형태 분석 정보를 통해 구분할 수 있게 되었다. 특히 원어 정보가 같은 한자어의 경우 유용한데, 동형의 한자라고 해도 단어의 구조에 따라 의미가 달라지거나 표제어의 어느 위치에 오느냐에 따라 의미나 문법적 지위가 달라지는 부분들까지 모두 보일 수 있게 되었다. 마지막으로 형태 분석 정보와 뜻풀이 정보가 서로 연계되어 표제어의 뜻풀이에 일관성을 기할 수 있게 되었다.

[Abstract]

The Morphological Analysis Information of <Korea University Korean Dictionary>

Kim, Ryangjin* · Lee, Hyeonhee**

This paper aims to introduce 'morph analysis information' included entry part in <Korea University Korean Dictionary> and to discuss about its usefulness.

We already know that Korean language has variety and advanced word formation rule. From this point of view, we presented morph analysis information about each entries which showed us structures of complex words entry including derivation words and compound words. Morph analysis information, which was not appear in existing dictionaries, extended analysis level of words structure and specified morphological categories of their constituents. As an epoch-making method, morph analysis information realized both practical value for normal users and academical value of professional users.

Morph analysis information is included of several morphological index in <Korea University Korean Dictionary>. Morphological indexes represents morphological status of words constituents. '+' represents words, '±' represents roots, '=' represents prefixes, '-' represents suffixes and '⌊' represents endings. For example, noun '길' plus noun '가' are presented '[+길+가]' that is compound word and root '오솔' plus noun '길' are presented '[±오솔+길]' that is

* Researcher Professor, Research Institute of Korean Studies, Korea University.

** Senior Researcher, Research Institute of Korean Studies, Korea University.

also compound word. Derivation words are presented such as '[=假 + 支給]' or '[+畛-이]'. Some words are included endings so they are presented such as '[+밖_아+지_다]' or '[+높-다랗_다]'.

At this time, the entries represented morph analysis information are totally 386,889 entries in <Korea University Korean Dictionary>. They are consisted of prefix derivation words 8,518, suffix derivation words 77,413, compound words excluded affix 86,197 and words included ending 2,128. Morph analysis information has several value. First, we prepare a groundwork for whole and systematic seeing about Korean word formations. Second, we can extract the whole morphemes lists concerning about modern Korean language. Third, we prepare the base for statistical treatments about linguistic data, corpus building, morph analysis corpus building, coining new words and translations.

Morph analysis information has several huge value in Korean linguistics or general linguistics. First, we can extract the whole morph/morphemes lists concerning about modern Korean language. In <Korea University Korean Dictionary>, simple verbs are 2,077 and simple adjective are 432, however, total verbs are 55,308 and total adjective are 12,431. It means Korean words are made by derivation or compounding word formation rules.

Second, user can identify word formation more easily and accurately. For example, hitherto, '다림질' is a shortened word of '다리미질'. However, we can know that two words are actually simple synonymous relation because we showed morph analysis information '[+다리마-질]' and '[+다리-ㅁ-질]'. These information is connected to meaning interpretations of '다리미질' and '다림질'.

Third, we presented linguistic information for classifying same phonemic words. User can understand about same phonemic words using word structures which are showed morph analysis information.

Forth, we can classify entries which have same original

information because of morph analysis information. Especially, morph analysis information is more useful in case of a word written in Chinese characters because Chinese characters are changed their meaning of grammatical status according to position in words. At last, we can describe a consistent meaning interpretations because of morph analysis information.

〈Key Words〉 <Korea University Korean Dictionary>, morph analysis information, morphological index, derivation words, compound words, complex words, root, morpheme lists, word formation

접수일 : 2009. 10. 31, 심사일 : 2009. 11. 11, 게재확정일 : 2009. 12. 3