

《조선말대사전》의 거시구조에 대하여

한영균 *

1. 서론

1.1. 이 글은 평양 사회과학원 언어학연구소에서 편찬하고 1992년 사회과학출판사에서 간행한 《조선말대사전》의 거시구조를 이루고 있는 기본 단위의 검토를 통해 표제항 설정 기준을 확인하고, 아울러 이 사전의 편찬에 적용된 원칙과 그 원칙의 실제 적용 양상을 비교하여 사전의 충실성을 검토하는 것을 목적으로 한다. 이 작업은 확장형 한국어 사전으로서의 《조선말대사전》이라는 텍스트를 대상으로 한 사전학적 분석을 위한 기초 작업의 한 부분이다.

1.2. 사전(辭典)은 한 사회 안에서 언어생활의 규범과 실상을 직접적으로 보여주는 자료다. 이런 점에서 북한에서 편찬된 국어사전은 날로 이질화되어 가는 남한과 북한의 언어 실태를 비교하는 자료로써, 나아가 통일을 대비한 언어 정리 작업의 기초자료로 활용될 수 있다. 특히 본 연구에서 검토의 대상으로 삼

* 울산대학교 국어국문학과 교수

은 《조선말대사전》은(앞으로는 이 사전이라고 부른다) 두 가지 면에서 주목할 만하다.

첫째, 이론적 측면에서 이 사전은 《조선말사전》, 《문화어사전》, 《현대조선말사전》 등 일련의 사전들을 편찬해 온 경험과 언어정리의 성과가 결집된 것이어서 현 시점에서 북한 사전학계의 이론적 수준을 보여주는 대표적인 사전이라고 할 수 있다.¹⁾

둘째, 울림말의 선정에 있어서는 실제 간행된 문헌에서 사용된 어휘를 수집하여 정리하였음을 표방하고 있고(일러두기 2쪽), 또한 약 100만 어절의 문어 텍스트에서 추출된 어휘 빈도를 병기하고 있기도 하다. 이는 이 사전이 가능한 한 북한에서의 언어 생활의 실태를 반영하려 애썼음을 의미하는 바, 북한 사회의 실제 언어 사용 양상을 상당 부분 반영하고 있으리라고 기대할 수 있다.

본 연구에서는 이러한 점에 착안하여 이 사전의 편찬에 적용된 이론의 수준을 확인하고, 또한 이 사전이 북한사회의 실제 언어 사용 양상을 얼마나 잘 반영하고 있는가를 확인하는 데에 초점을 둔다.²⁾ 여기에는 앞에서 언급한 바 이 사전이 통일을 대비한 언어 정리 작업의 기초자료로서 지니는 가치를 평가하는 작업에 도움을 주고자 하는 의도도 포함되어 있다. 현재 국립국어연구원에서 진행 중인 《종합국어대사전》 편찬에도 이 사전이 중요한 참고자료의 하나로 채택되어 있는데,³⁾ 본고와 같은 작업을 통해서 《종합국어대사전》의 성공적인 편찬에도 기여할 수 있기를 기대한다.

-
- 1) 《조선말대사전》의 사전사적 의의에 대해서는 大江孝男(1993) 및 송천석(1993)에 자세히 다루어진 바 있다.
 - 2) 大江孝男(1993 : 8)에서는 이 사전의 기본 목적이 “북한 내의 언어 상황을 분석해 보여준 다기보다, 현실의 언어 상황을 규정하는 규칙을 가르치고, 어! : 방향으로 이끌고 나가려고 하느냐에 관한 지침을 알려주는 사전”일 것이며 “이 사전에 나타난 방침을 토대로 해서 신문, 잡지나 문학작품 같은 데서 쓰이는 문장을 분석하면 북한에서 목표로 하는 언어 상태가 어떤 것인지 알 수 있을 것”이라고 지적한 바 있다. 본 연구는 그와 같은 지적의 타당성을 검증하는 한 과정이 될 것이다.
 - 3) 《종합국어대사전》의 편찬은 기본적으로 남북한의 통일에 대비한다는 목표를 가지고 출발한 것이며, 이러한 목표의 달성을 위해서 남한과 북한에서 간행된 대사전류의 정보를 가능한 한 충실히 반영하여 편찬한다는 것이 사전편찬의 기본 방침인 것으로 알려져 있다. 이 사전이 간행된 지 몇 해 되지 않음에도 불구하고 사전의 기본적 성격 및 내용에 대한 검토가 유례없이 집중적으로 이루어진 것도 바로 이러한 사정에 연유한다.

1.3. 어떤 사전에 대한 사전학적 평가를 위해서는 거시구조에 대한 검토와 함께 표제항에 대한 문법정보, 뜻풀이의 정확성 및 그 배열, 음운정보, 어휘정보 등 미시구조에 대한 검토가 병행되어야 할 것이다. 그러나 한 편의 논문에서 그러한 문제를 한꺼번에 다루기는 어렵다. 따라서 이 글에서는 우선 거시구조 상의 기본 문제만을 개괄적으로 다룰 것인데, 특히 필자가 직접 문화어 자료를 분석한 결과 얻어진 자료와의 비교를 통해서 일러두기(2쪽)에서 내세우고 있는 실제 문헌에 사용된 언어 자료를 바탕으로 한다는 편찬의 기본 방침이 얼마나 충실히 실현되었는가를 검토하는 데에 일차적 목표를 둔다. 이러한 관점에서의 분석을 위해서 본 연구에서는 북한에서 비교적 최근에 간행된 문어 자료에서의 언어 단위의 사용 양태와 이 사전의 거시구조 및 뜻풀이에서의 용법·용례와 비교하는 방법을 취한다.

2. 《조선말대사전》의 거시구조 개관

2.1. 표제항의 수와 품사별 비율

2.1.1. 머리말에 따르면 이 사전에 수록된 표제어는 33만에 달하는 것으로 되어 있다. 그러나 우리가 분석 대상으로 삼은 현대어에 국한하면 표제어의 총수는 293,073항목이고, 관용구·속담이 29,013개 항목 수록되어 있다. 표제어 중 주표제어는 256,140 항목이며, 파생어류로 이루어진 부표제어가 36,933 항목이다.

[표 1] 《조선말대사전》 주표제어의 품사별 구성

품사별	명사	대명사	수사	동사	형용사	부사	관형사	토씨	감동사	합계
빈도	212,147	105	183	13,829	11,905	15,058	71	576	368	254,242
비율	82.82%	0.04%	0.07%	5.40%	4.65%	5.88%	0.03%	0.22%	0.14%	99.26%

2.1.2. 이 사전에 수록된 주표제어의 품사별 분포는 [표 1]과 같다.⁴⁾

표에 제시된 것 이외에 접사류(앞붙이/뒤붙이)가 350개, 한자성구류가 595개로 각각 0.14%, 0.23%를 차지하는데, 특기할 것은 주표제어 중 950개 가량은 아무런 문법 범주 표시가 없다는 점이다(빈도의 총합계가 주표제어 총수와 다른 것은 그에 기인한다).⁵⁾ 이는 일러두기 4-②에서 밝히고 있는 바 “한 글자안에 들여다 배열한 일부 올림말이나 공고한 단어 결합 및 성구, 속담에는 품사표시⁶⁾를 하지 않았다.”는 편찬의 원칙에 따른 것이라고 할 수 있는데, 실제로는 일러두기에서 밝히고 있는 유형의 것들 이외에도 품사표시를 하지 않는 것들이 있다 (2.1.4. 참조)

2.1.3. 이 사전의 표제어 배열에서 주목되는 것은 부표제어의 등재 내용이다. 앞에서 밝힌 바와 같이 이 사전에는 총 36,933 개의 부표제어가 등재되어 있는데, 그 중 문법 범주 표시가 되어 있는 30,858개 중에서 단 하나의 예⁷⁾를 제외하고는 모두 “동사”인 것이다. 이는 “한 글자안에 들여다 배열한 일부 올림말(부표제어)”에는 품사표시를 하지 않았다는 앞서 일러두기의 설명을 고려할 때, 동사파생의 경우에만 품사표시를 한 것을 보여준다.

이렇게 부표제어에 품사표시를 하지 않은 것은, 대부분의 파생어를 주표제어로 등재한 것과 밀접한 관계가 있는 것으로 보인다. 이 사전에는 제시된 표제어에 해당 표제어의 내적 구성을 보여주는 정보를 담지 않고 있어서 주표제어 중 단일어와 합성어, 파생어의 비율을 간단히 산출하기는 어렵다. 그러나 부표제어를 분석한 결과 비유적으로 풀이된 ‘X+갈-’, ‘X+없-’형 파생어와 용어류를 제외

-
- 4) [표 1]의 비율은 주표제어 256,140개 중 해당 품사의 어휘가 차지하는 비율이다.
 5) [표 1]의 품사별 구성표를 이용함에 있어서 유의할 부분이 있다. 이 사전에서는 하나의 올림말에는 하나의 품사만을 인정하고 있기 때문에(일러두기 3쪽), 통계에도 그러한 사실이 반영되어 있다. 따라서 이 분포표는 말 그대로 이 사전의 표제항에 주어진 문법정보에서의 품사별 분포라고 할 수 있을 뿐 실제 어휘의 품사별 분포라고 하기는 어렵다는 점이다. 이 통계는 남한의 대사전에서의 품사 분포에 대한 검토와의 비교 자료로서, 그리고 우리가 이 글에서 검토 대상으로 삼고 있는 문법형태소(도씨)와 함께, 소위 접사류(앞붙이, 뒤붙이 등)의 비중이 어느 정도인가를 파악하는 기본 자료로서만 의미를 지닌다.
 6) 일러두기에는 “품사표시”라는 용어를 쓰고 있으나, “품사표시”라기보다는 문법범주라는 용어가 옳을 것이다. “앞붙이, 뒤붙이, 성구” 등도 포함되는 것이기 때문이다
 7) 그 하나의 예는 다음의 것이다.
 감회 [명] 고맙게 여기어 기뻐하는것. 감회하다 [동](타) 감회하다 [형] 고맙고 기쁘다. 感
 喜

하면 모두 “-하다, -되다”에 의한 파생어임을 볼 수 있었던 바, “-거리다, -이다, -대다, -스럽다, -답다” 등의 용언파생접사에 의한 파생어와 명사, 관형사, 부사 파생어는 모두 주표제어로 등재하는 것을 원칙으로 하고 있었다는 결론을 얻을 수 있었다.⁸⁾

2.1.4. 주표제어 중 문법적 특성을 표시하지 않은 항목들은 다음과 같은 몇 가지 유형으로 나뉜다.

- ① 노아: 노아의 방주 기독교에서, 노아의 홍수때 노아와 그의 가족이 만들어 타고 하는 네모난 큰 배. (=)노아의 족배. 노아의 상자배.
 뉴턴²: 뉴턴의 제3법칙 뉴턴에 의하여 발견된 고전역학의 기초로 되는 법칙의 하나. 작용(F1)과 반작용(F2)은 크기가 서로 같고 한 직선우에 있으며 방향은 반대로 되는 법칙 곧 $F2 = -F1$ 이다. 작용과 반작용은 작용점이 서로 다른 물체에 있으므로 크기가 같고 방향이 반대이면서 한 직선우에 있다 할지라도 서로 비키는 힘으로 되지 않는다.
- ② 감태²: 감태같다 (머리가) 검고 윤기가 나다. | 춘향을 잡아 대들아래에 내리치니 군노사령들이 벌떼같이 달려들어 감태같은 춘향의 머리채를 휘휘 칭칭 잡아 쥐고 동맹이쳐서...<고전소설 《춘향전》> §
- ③ 남태령: 서울이 무섭다니까 남태령부터 긴다 ㄱ서울.
- ④ 늘 쓰는 말 <언어> 사람들의 일상생활에서 늘 쓰는 어휘부류. 모든 사람들에게 또는 해당 활동분야의 사람들에게 자주 쓰이는 생활적이며 적극적인 어휘들이 이에 속한다.
 겸용종 소 <축산> 두가지 이상의 생산성을 가지는 소. 생산물가운데서 어느것이 기본인가에 따라 젓 고기 겸용종 소, 고기 젓 겸용종 소, 부림 고기 겸용종 소, 젓 고기 부림 겸용종 소로 나뉜다. 조선소는 부림 고기 겸용종 소이다.
- ⑤ 구충같다 <아득하게 높음>을 비겨 이르는 말.
- ⑥ 골로¹ (말체) <고것으로>의 준말. |더 커야지 골로 될하나? §
- ⑦ 가라사대 존경토 <사>가 들어간 옛날말 동사 <가라다>의 이음형. 남의 말을

8) 단적으로 주표제어의 부사 15,000여개 중에서 ‘-하니, -스래, -히, -이’형 부사가 5,800여개로 전체의 1/3이 넘는다. 이들이 모두 용언에서 파생된 것이며, 그 뜻풀이에서 “X하게”형으로 간단히 처리된 것을 감안하면 과연 주표제어로 일일이 등재할 필요가 있는가 하는 의문을 제기할 수 있다. 또한 ‘X+갈-’, ‘X+없-’형 파생어도 모두 부표제어로 처리된 것은 아니고, 주표제어로 등재된 경우가 적지 않다. 이 경우 파생어 처리의 일관성 문제가 제기될 수 있는 것이다. 이 사건의 파생어 처리 방식에 대해서는 별도의 글을 준비 중이다.

인용할 때 <말하되>, <말하기를>의 뜻을 옛날투로 이르는데 높이는 뜻이 있다. 주로 이 형태로만 쓰인다.

cf 가^ㅍ [동] <가다>의 이음형인 <가아>의 준말.

①, ②, ③의 예는 문법 범주 표시를 하지 않는 가장 전형적인 경우의 예인데, ①의 예에서의 “노아, 뉴턴”은 용어류를 등재하기 위해서 사용되었을 뿐 자체에 대해서는 뜻풀이를 하지 않으며, ②의 예에서의 “감태” 역시 “감태같다”는 부표제어를 등재하기 위한 것이고, ③의 “남태령”은 관용구를 등재하기 위한 것이다. 이렇게 다른 항목을 실기 위해서 등재한 항목의 경우, 모두 표제어 뒤에 콜론(:)을 써서 표시하고 있다.

④는 용어류이면서도 ①과 같은 방식을 사용하지 않고 용어 자체를 표제항으로 등재한 것인데, 다른 표제어들과는 달리 띄어쓰기를 하고 있어서 구분된다. “늘 쓰는 말”의 경우에는 ①의 방식으로 등재할 만한 적절한 명사류가 없다는 점이 그 배경일 수 있겠으나, “검용종 소”의 경우에는 “검용종”이 독립된 주표제어로 등재되어 있음에도 그 하위항목으로 등재하지 않고 따로 주표제어로 삼은 것이다.

⑤의 “구충같다”는 ②의 “감태같다”와 형태론적 구성상 차이를 찾기 어려운 데, 처리방식이 다른 점이 주목된다. 즉 선행요소인 “구충”이나 “감태”의 의미를 정확히 알기 어려우며,⁹⁾ “같다”가 붙어서 하나의 용언으로 인식되는 것인데 그 처리방식이 다른 까닭을 알 수 없는 것이다. 다만 “감태같다”의 경우 뜻풀이에 있어서 다른 “X+같다”와 달리 비유적 용법으로 해석하지 않고 본래적 의미로 풀이했다는 점을 지적할 수 있다. 부표제어로 등재되건 ⑤와 같이 주표제어로 등재되건 “X+같다”형의 풀이는 “<……을 비겨 이르는 말>”와 같은 형식을 취하는 것이 일반적이기 때문이다.¹⁰⁾

9) ②의 “감태같다”의 “감태”에는 어깨번호가 붙어 있어서 “감(海苔)”을 의미하는 “감태1”과 구분하고 있다. 이는 이 “감태”의 의미를 파악하기 어려운 데에서 기인한 것으로 생각된다.

10) 부표제어에서의 예를 들면 다음과 같다.

거미줄같다 ① <사방으로 이리저리 많이 늘어놓은 상태>를 비겨 이르는 말. ② <다치면 끊어질만큼 힘이 약한 상태>를 비겨 이르는 말.

⑥ ⑦은 주표제어 중에서 문법 범주를 표시하지 않는 것의 또다른 유형이라고 할 수 있다. 이 사전에서는 “준말”의 경우도 그것이 명사이면 문법범주 표시를 해주는 것을 원칙으로 하고 있다.¹¹⁾ ⑥과 같은 처리는 특히 구어체(말체)인 경우에 적용한 것으로 판단된다. ⑦의 예는 동사의 활용형을 표제어로 등재했다는 점이 주목되는 것인데, 이러한 유형의 예가 많지 않아서 정확히 단언하기는 어려우나 “기5”의 경우와 달리 문법범주를 표시하지 않은 것은 “가라사대”라는 한정된 형태로만 쓰이기 때문이 아닌가 한다.

2.2. 학술용어의 수용

2.2.1. 이 사전에는 총 38,228개의 학술용어가 수록되어 있다. 이는 명사 211,187 항목 중 18.1%에 달하는 적지 않은 비율이다. 이 사전의 일러두기에 따르면, 모두 46개 분야의 학술용어를 수록하고, 표제어 다음에 (())로 표시하였다고 밝히고 있다(일러두기 3쪽).

2.2.2. 학술용어의 영역별 수록 항목 수는 다음과 같다.

건설 : 1,839	경제 : 987	고고 : 380	금속 : 992	기계 : 1,547
기상 : 381	광업 : 629	농학 : 907	력사 : 793	론리 : 121
립학 : 437	무용 : 209	문학 : 217	문예 : 170	물리 : 2,063
미술 : 615	방직 : 542	법학 : 671	수리 : 416	수산 : 539
수학 : 1,418	심리 : 65	생리 : 995	생물 : 2,401	자동차 : 287
잠학 : 226	전기 : 999	전자 : 774	지리 : 791	지질 : 1,030
천문 : 461	철학 : 336	측산 : 575	출판 : 393	체신 : 587
체육 : 864	해양 : 521	화학 : 2,995	화학공업 : 257	약학 : 567
언어 : 1,498	연극 : 77	영화 : 259	운수 : 790	음악 : 1,314
의학 : 2,686				

거부기잔등같다 <마르거나 거칠어져서 갈라터진것>을 두고 이르는 말.

11) 그 예를 하나만 들어 둔다.

골창¹ [명] <고랑창>의 준말. ㄱ ~ 밑바닥에 시꺼멓게 쌓인 흙.

2.2.3. 이 사전에서의 학술용어 수용 양상에 대한 검토에서 드러나는 특성과 문제점을 정리하면 다음과 같다.

첫째, 사회주의적 인식체계를 반영하고 있다는 점이다.

그것을 잘 보여주는 것이 이 사전에는 종교·민속·신화 등과 관련된 분야의 구분이 없는 점이다.¹²⁾ 정치·사회·교육 등과 관련된 분야도 별도로 설정하지 않고 있는데, 이는 사회·정치적 용어를 학술용어로 다루지 않았기 때문이라고 할 수 있고, 이 역시 북에서의 인식론적 특성을 보여주는 것이며, “임학, 방직, 운수, 잠학, 축산” 등이 포함된 것도 과학기술의 실용화를 강조하는 체제적 특성과 사전의 인식교양적 기능¹³⁾을 강조하는 태를 반영한 것이라고 할 수 있는 것이다.

둘째, 학술용어의 수용 기준이 분명치 않다는 점을 들 수 있다.

예를 들어 이 사전에는 적지 않은 군사용어가 수록되어 있는데 이에 대해서는 아무런 구분이 없으며, ‘문예’라는 영역을 ‘문학·미술·연극·영화’과는 별도로 설정하고 있는데 그들 사이의 구분 기준¹⁴⁾ 등도 분명치 않은 것이다.

셋째, 이 사전에서는 총 38,228개의 학술용어 중 1.59%에 달하는 607개 항목이 일러두기에 언급되지도 않은 25개 영역의 학술용어로 구분되어 있다. 그러한 영역과 빈도는 다음과 같다.

건축 : 6	경공업 : 2	공예 : 1	광물 : 2	광산 : 2
광학 : 2	교예 : 18	교육 : 126	농업 : 2	동물 : 1
력학 : 24	류리 : 8	림업 : 5	민속 : 367	산림 : 1
선박 : 5	수의 : 1	식물 : 6	잠업 : 1	종교 : 9
철도 : 1	화학공학 : 1	예술 : 5	원자 : 1	인쇄 : 10

이들 중에서 일부는 일러두기에 언급된 46개 영역으로 재분류될 수 있는 것

12) 그렇다고 해서 해당 분야의 어휘를 신지 않은 것은 아니다. 영역범주에 포함하지 않았을 뿐, 뜻풀이 속에서 어떤 영역에서 사용하는 것인가를 밝히고 있다.

13) 사전의 인식교양적 기능이란, 인민의 의식수준을 높이고, 과학과 문화를 발전시키는 기능을 의미한다. 북한의 사전은 규범사전이건 확장형 사전이건 이러한 기능에 이바지하는데에 큰 의미를 부여한다.

14) ‘생리’와 ‘생물·의학’의 경우도 마찬가지라고 할 수 있다.

들도 있고(광물, 광산, 립업, 삼림 등), 일러두기에 미처 언급되지 못한 것일 가능성이 있는 것들도 있지만(교육, 민속 등), 이들의 존재는 이 사전이 치밀한 통제 아래에 편찬된 것이 못됨을 시사한다.

3. 《조선말대사전》의 거시구조 구성상의 몇 문제

3.1. 표제항 선정 기준 설정 문제

3.1.1. 한 사전의 거시구조는 ① 거시구조의 단위(표제항) 설정 ② 표제어의 배열이라는 두 요소에 의해서 결정된다. 그런데 두 번째 문제는 다의어(polysemy)와 동음이의어(homophony)의 구분 및 그 사전적 처리라는 어휘론 및 사전편찬학의 오랜 문제와 관련을 가지고 있는 것이어서 본고에서 본격적으로 다루기 어렵다. 여기서는 주로 첫째 문제를 중심으로 하되 주로 합성어의 등재와 관련하여 이 사전에서의 표제항 선정상의 문제점을 검토해 보기로 한다.

3.1.2. 이 사전에서는 표제어의 선정 기준을 “오늘날 널리 쓰이고있거나 지난날 쓰이었던 단어와 공고한 단어결합 및 성구와 속담” 그리고 “우리 말 문법구조의 특성을 반영하여 앞붙이, 뒤붙이, 토”로 한다고 규정하고 있다(일러두기 1)-①).

이러한 올림말 설정 기준은 이른바 합성어의 표제항으로서의 등재 여부를 결정하는 문제와 관련하여 다음과 같은 두 가지 어려움을 남긴다.

첫째, 한국어의 형태론적 특성상 문어에서의 “단어”에 대한 규정은 정서법(orthography)과 밀접한 관계를 가지게 된다는 점이다. 정서법에서의 ‘단어’의 기준이 바로 표제항의 단위 설정에 영향을 줄 수 있기 때문이다. 특히 규범적 성격을 지닌 사전인 경우 사전의 표제항 설정은 정서법에 반영된 띄어쓰기 단위와 밀접한 관계를 가지게 된다. 사전의 사용자가 사전의 활용에 응용할 수 있는 최소한의 정보가 띄어쓰기이기 때문이다.

둘째, ‘공고한 단어결합’과 ‘공고하지 않은 단어 결합’을 구별한 이론적 준거를

제시할 수 있어야 한다는 점이다.¹⁵⁾ 《문화어문법규범》에 의하면 ‘단어’는 띄어쓰기에 의해 구분하도록 되어있다.¹⁶⁾ 그러나 많은 경우 단순한 단어와 단어의 결합에서도 “공고한 단어결합”과 마찬가지로 띄어쓰기를 하지 않기 때문에, 띄어쓰기를 바탕으로 해서 단순한 단어와 단어의 결합과 “공고한 단어결합”을 구분하는 것은 불가능하며, 이의 구분을 위한 별도의 장치가 마련되어야 한다.

3.1.3. 공고한 단어결합의 구분 기준에 대해서는 이 사전보다 상당히 이른 시기에 간행된 것이지만 정순기·리기원(1984 : 38-47)을 참조할 수 있다. 정순기·리기원(1984)에서는 형태적으로 구분되지 않는 단순한 단어결합과 공고한 단어결합을 구분하는¹⁷⁾ 기준으로 ① 조선어의 고유한 서사적 특성과 어음적 특성 ② 조선어의 유형학적 특성의 두 가지를 제시하고 있는 것이다.

정순기·리기원(1984)에서 제안하고 있는 기준은 기본적으로 규범사전이 아닌 실용적 목적을 위한 뜻풀이사전을 위한 것이다. 즉 “사전리용자들이 어휘에 대한 구조적인 분석을 하지 않고도 해당 울림말을 단번에 찾아 그 뜻을 알수 있게 하는 편리를 도모해줄 수 있도록 하기 위한” 것이다. 이는 결국 북한 문어의 서사규범에서 단어와 구분되지 않는 단어결합, 성구속담 등을 사전 표제어로 등재하는 문제를 이론적으로 합리화하기 위한 설명이라고 할 수 있는데, 여러 가지 경우를 동원해서 설명하고 있지만 기본적으로 “의미”를 바탕으로 한 설명이 주류를 이룬다. 즉 ‘로력문제, 로동일수, 교원집단’ 등과 같이 “구태여 사전의 울림말로 올려서 그 뜻을 풀이하지 않아도 그 단어의 구성요소들을 통하여 뜻이 명확히 파악되는 단어들은 따로 올리지 않고(《사전편찬이론연구》 43쪽)” “하

15) 명시적으로 드러나 있지는 않지만, 이 사전에서는 자립적 단위의 결합 중에서 ① 직접구 성성분(Immediate Constituent)들의 의미를 바탕으로 전체 구성의 의미를 파악할 수 없는 경우 ② 이른바 학술용어(전문용어)에 해당하는 경우의 두 가지를 “공고한 단어결합”으로 분석하고, 그것을 표제항으로 채택하는 것으로 판단된다. 이는 북한의 어휘 연구에서의 “공고한 단어결합”에 대해서는 고신숙(1989) 참조.

16) “문화어 문법 규범 제1장 2항: 조선어의 글에서는 단어를 단위로 하여 띄어쓰는 것을 원칙으로 하되, 자모를 소리마디단위로 묶어쓰는 특성을 고려하여 특수한 어휘부류는 붙여 쓰도록 한다.” 참조.

17) 정순기·리기원(1984)에서는 이를 단어결합, 성구속담, 일부 표현적 단위라고 지칭하고 있다.

나의 전일적 의미로 통합된” 것들은 표제어로 수록한다는 것이다. 이는 앞의 주 15)에서 지적한 《조선말대사전》에서의 “공고한 단어결합”의 처리기준과 크게 다르지 않다.

3.1.4. 이러한 의미를 바탕으로 한 표제항 선정은 자칫하면 사전 편찬자의 자의에 의해서 표제항이 선정되기 쉽다는 문제를 안고 있다.¹⁸⁾

한 예로 ‘하늘같이’와 ‘구름같이’의 경우를 들 수 있다. ‘하늘같이’는 이 사전에서 “① 우리름이 아주 높거나 은혜를 느낌이 아주 크게 ② (의탁하여 믿는데 있어서) 아주 든든하게 또는 크게”라는 풀이와 함께 ‘하늘’의 부표제어로 등재되어 있다. 이는 ‘하늘’의 은유적 의미 확대가 ‘-같이’와의 결합에 기인한다고 해석함을 의미한다. 한편 ‘구름같이’는 ‘구름’ 항의 성구 중에 ‘구름같이 사라지다’가 등재되어 있을 뿐 독립된 표제어로는 인정되지 않았다.¹⁹⁾ 그러나 “공고한 단어결합”에 대한 앞에서의 기준에 의할 때 ‘하늘같이’와 ‘구름같이’의 구성요소 사이의 의미관계가 서로 다른 성격을 지니는 것으로 분석하기는 어려울 것이다. 결국 ‘하늘같이’와 ‘구름같이’ 중에서 ‘하늘같이’만이 표제어로서의 자격을 가지게 된 기준은 사전 편찬자의 자의적 판단 이외에는 없는 것이다.

이를 더 확대하여 ‘명사+부사’ 합성어에 대한 처리를 보면 이 사전에서의 표제항의 선정이 뚜렷한 기준 아래에서 이루어진 것이 아님을 더욱 확실하게 확인할 수 있다.

18) 기존의 연구에서는 의미론적 기준 이외에 사전의 표제항 선정에 적용될 수 있는 기준에 대해서 언급된 것을 보기 어렵다. 필자의 검토에 따르면 대체로 ①통사적 관계 특히 언어구조(collocation structure)에 대한 참조를 통한 표제항 사이의 평형성 유지 ②동일한 구성요소를 지니고 있는 형태들의 구성 형식에 대한 참조를 통한 표제항의 평형성 유지 ③단어와 단어 사이의 의미관계를 구성하는 틀 즉 단어장(word field)을 통한 표제항의 포괄성 확보 등을 들 수 있을 것이다. ①과 ②는 자립형식의 결합으로 이루어지는 복합 단위들의 어휘화 여부를 판단하는 기준으로, ③은 “대사전(Unabridged Dictionary)”의 편찬에 있어서의 표제항 사이의 균형성 확보에 유용한 기준으로 적용될 수 있을 것으로 판단된다. 이 문제에 대해서는 따로 글을 준비 중이다.

19) 이는 ‘하늘’의 부표제어로 ‘하늘같다’가 등재되어 있는 반면, ‘구름같다’는 부표제어가 되지 못한 것도 관계있는 것으로 생각된다. 그런데 ‘뚝구름’, ‘뚝구름같다’ 등도 이 사전에는 표제항으로 등재되지 않았다. 결국 ‘허황한 존재’라는 의미로의 ‘구름’은 ‘◇ 구름같이 사라지다’ 이외에는 관련 표제항이 없는 것이다. 이는 사전 표제항 선정에서 어휘군의 의미 즉 단어장(word field)이 고려되지 않았음을 방증하는 것이라고 할 수 있다.

1992년분 《노동신문》의 사설에서 확인된 ‘명사+부사’의 형식으로 이루어진 합성어에는 ‘가슴뜨거이, 감동깊이, 긍지높이, 불길높이, 신심드높이, 심장깊이, 합성드높이’ 등이 있다. 이들 합성어의 의미가 이들을 구성하고 있는 요소들의 의미의 복합으로 파악될 수 있는 것인지 아니면 별개의 의미를 형성한 것인지를 판정하는 일은 쉽지 않다. 그러나 이들 결합체의 통사적 구성은 제각기 다르다고 할 수 있다. 따라서 이들을 표제어로 선정할 것인가 여부는 이러한 두 측면에 대한 고려를 바탕으로 하여야 한다. 즉 이들의 통사적 구성 및 의미관계를 바탕으로 판단되어야 하는 것이다. 그런데 이 사전에서는 이들 중 ‘가슴뜨거이’만이 주표제어로 등재되어 있다. 이는 ‘가슴뜨겁다’가 존재하지 않는다고 판단했기 때문이라고 할 것이다. 그러나 이러한 기준에 의한다면 ‘심장깊이’가 주표제어로 등재되지 못할 까닭이 없다. ‘심장깊다’라는 형태도 존재할 수 없는 것이기 때문이다.²⁰⁾ 나아가서 ‘감동깊이’와 ‘심장깊이’를 비교해보면 양자의 구성 양식과 구성요소 사이의 의미관계는 전혀 다르다. 따라서 ‘감동깊이’를 ‘가슴뜨거이’나 ‘심장깊이’와 달리 처리하기 위해서는 그 처리기준이 명시적으로 드러나야 할 것이다. 그러나 이 사전의 어떤 부분에서도 그러한 기준은 찾아보기 어렵다.

3.1.4. 결국 이 사전에서는 띄어쓰기가 이루어지지 않은 단어결합을 표제항으로 선정할 것인가의 여부를 결정하는 데에 의미론적 기준 이외의 준거를 찾아내지 못한 상태에서, 사전 편찬자의 자의에 의해 표제어의 선정이 이루어졌다고 할 수 있다. 이러한 양상은 이 사전의 편찬에 적용된 거시구조의 단위 선정 방법은 면에서는 《조선말사전》, 《문화어사전》, 《현대조선말사전》 등의 편찬 당시의 이론적 수준에서 별로 발전하지 못한 상태에 머물러 있다고 할 수 있는 것이다.

3.2. 문법 형태소의 처리

3.2.1. 한국어와 같이 심한 굴절(Conjugation)을 보이는 언어를 대상으로 한 사전의 경우, 그 사전이 규범성을 지향하는 사전인가 아니면 실용성을 중시하는

20) 여기서 ‘가슴뜨겁다’, ‘심장깊다’의 존재 가능성은 논의하지 않는다. 그것은 별개의 문제이기 때문이다. 여기서의 논의의 주안점은 ‘명사+부사’ 복합체의 처리에 놓인다.

사전인가에 따라서 문법형태소들의 이형태 및 복합형에 대한 처리방법이 달라지게 된다.

규범성을 지향하는 사전이라면 사전의 규범적 기능을 가능한 한 거시구조의 단위 설정에 반영하려 할 것이고, 그것은 바로 표제항 선정에 반영될 것이다. 즉 한 문법형태소가 여러 이형태를 가지고 있는 경우, 사전의 규범적 기능을 극대화하기 위해서는 현실 자료에서 확인되는 형태를 모두 표제항으로 등재하기보다 이론적 분석에 바탕을 둔 대표적 형태를 표제항으로 삼아야 할 것이고, 아울러 표제항이 되는 단위 사이에도 체계상의 일관성이 유지되어야 한다. 또한 문법형태소의 복합형이 다양하게 나타나는 경우에도, 그들 모두를 거시구조의 단위로 반영하기보다는 문법형태소들 사이의 결합 유형을 분석하고, 그들 바탕으로 가능한 한 단순한 형태를 중심으로 한 거시구조의 단위 설정이 이루어 질 것이다. 반면에 실용성을 지향하는 사전이라면, 어느 정도 차이는 있겠지만, 가능한 한 현실 자료에서 출현하는 문법형태소의 이형태 및 복합형태소들을 포괄적으로 표제항으로 제시하고 그들 각각의 용법·용례를 설명하는 방법을 택할 것이다.

따라서 문법형태소의 표제항 선정 기준은 이론적 분석의 수준을 보여줄 수 있을 뿐 아니라, 사전의 지향점을 보여주는 자료가 될 수 있는 것이며, 뜻풀이에서 실제 출현할 수 있는 이형태의 용법·용례에 대해서 얼마나 자세히 그리고 정확하게 기술하고 있는가는 바로 사전의 충실성을 측정하는 기준이 될 수 있는 것이다.

밝힌 바와 같이 본고의 관심사 중의 하나는 문법형태소 특히 조사 및 어미의 처리가 어떻게 이루어져 있는가 하는 점이다. 문법형태소의 처리는 이 사전의 편찬에 적용된 이론적 분석의 수준과 함께 사전의 충실성을 확인하는 기준의 하나가 될 수 있다고 판단하였기 때문이다. 그러나 분석의 결과는 그 어느 면에서도 기대치에 미치지 못한다.

3.2.2. 우선 문법 형태소의 표제항 선정과 관련된 문제를 검토하기로 한다. [표 1]에서 볼 수 있듯이, 이 사전에는 576개의 토씨(조사와 어미)가 수록되어 있다. 이 중에서 조사로 분류될 수 있는 것은 서술격 조사 ‘-이’의 경우를 포함하여 총 132개였고, 나머지는 444개는 어미에 해당하는 것이었다.²¹⁾

그런데, 우리가 북한 문어자료를 직접 분석한 결과는 이와는 상당한 차이를

보인다. 본 연구를 위해서 분석한 북한문어자료는 노동신문사설 1992년분 174건, 장편소설 《꽃피는처녀》, 《피바다》, 그리고 1991년에 간행된 《조선문학》 소재의 단편 및 중편소설 51편으로 50만 어절 남짓한 분량이었는데, 거기서 얻어진 조사의 이형태 총수는 1,271개에 달하며, 어미의 이형태 총수는 3,495개에 달했다.

결국 조사의 경우 북한에서 간행된 문어자료 50만어 어절의 분석에서 얻어지는 형태의 10.4%만이, 어미의 경우는 12.8%만이 표제어로 등재되어 있음을 확인한 것이다.

이는 기본적으로 이 사전에서의 문법형태소 수록의 원칙이 대표형을 중심으로 하는 것이었고, 복합조사나 복합어미는 특별한 경우를 제외하고는 수록하지 않는 것을 원칙으로 한 데에서 기인한다고 할 수 있다. 문제는 이러한 원칙이 사전의 어느 부분에서도 명시적으로 드러나지 않고 있을 뿐 아니라, 원칙의 적용이 자의적이어서 어느 정도까지 그러한 원칙을 수용하고 있는지를 판단하기 어렵다는 점이다. 결국 우리가 3.2.1에서 언급한바, 사전의 편찬이 규범성을 지향한 것인지 아니면 실용적 목적을 위한 것인지를 판단할 만한 뚜렷한 근거를 찾기 어려울 뿐 아니라, 3.1.4에서 검토한 어휘 단위의 처리에서와 마찬가지로 문법형태소에 대한 처리 역시 현대 북한 문어에서 출현하는 형태를 전체적으로 포괄하지도 못하면서, 또한 이론적 분석의 기준도 뚜렷하지 않은 양상을 확인할 수 있는 것이다.

단적인 증거로 1992년의 《노동신문》 사설에서 확인되는 형태들조차 사전에 수록되지 않은 것을 지적할 수 있다. [표 2]와 [표 3]은 각각 남한의 학교문법에 따른 분석 결과 얻어진 조사와 어미의 이형태 중 빈도수가 높은 25 종류를 보인 것인데, 이 중에서 조사 '-인, -에서는, -에는, -께서는' 등과 어미 '-니다, -여

21) 이 사전에서는 조사와 어미를 구분하지 않는데, 특히 남한의 학교문법에서 서술격 조사로 처리하는 '-이'의 활용형을 체언의 용언형으로 다루면서 용언의 활용형과 같이 처리하고 있다. 따라서 여기서의 조사와 어미 구분은 용례를 바탕으로 할 수밖에 없었는데, 예문이 하나밖에 없어서 조사 혹은 어미 한편으로만 처리되었지만, 실제로는 조사·어미 양쪽에 해당할 수 있는 예들도 적지 않다. 다음의 것이 그 예 중의 하나다.

-거니[도] 맺음을 나타내는 풀이토의 하나. 어떤 사실을 당연한 사실로 미루어 인정할 때 <반말> 말차림으로 쓰인다. /그대는 정녕 아름다운 조선의 딸이거니.

야, -였다' 등은 이 사전의 표제어로 등재되어 있지 않음을 확인할 수 있다. 이들이 표제항으로 등재되지 않은 것이 이들이 모두 복합조사이거나 복합어미이기 때문이라고 할 수도 있다. 그러나 이 사전에는 '-이든지, -이든가'와 같은 서술격 조사의 활용형이나 '-ㄴ가, -ㄴ답니다, -여야지'와 같은 복합어미가 표제항으로 등재되어 있다. 결국 문법형태소의 표제항 선정 역시, 어휘형태소의 표제항 선정의 경우에서와 마찬가지로 뚜렷한 기준이 없이 자의적으로 이루어졌다는 지적이 가능한 것이다.

[표2] 노동신문 사설 조사 빈도(부분)

	조 사			
	형태	발생횟수	빈도	누적빈도
1	을	14039	18.211	18.211
2	의	13409	17.394	35.605
3	과	6865	8.905	44.510
4	에	5482	7.111	51.621
5	은	5085	6.596	58.217
6	를	5010	6.499	64.716
7	이	4569	5.927	70.643
8	으로	3157	4.095	74.738
9	에서	2813	3.649	78.387
10	와	2305	2.990	81.377
11	이다	1820	2.361	83.738
12	는	1367	1.773	85.511
13	로	1160	1.505	87.016
14	가	1154	1.500	88.516
15	인	1072	1.391	89.907
16	도	877	1.138	91.045
17	이며	552	0.716	91.761
18	에서는	492	0.638	92.399
19	나	472	0.612	93.011
20	서	462	0.599	93.610
21	에게	410	0.532	94.142
22	께서는	358	0.464	94.606
23	에는	288	0.374	94.980
24	께서	285	0.370	95.350
25	까지	283	0.367	95.717

[표3] 노동신문 사실 어미 빈도(부분)

	어미			
	형태	발생횟수	빈도	누적빈도
1	ㄴ	10672	20.711	20.711
2	고	5343	10.369	31.080
3	는	4826	9.366	40.446
4	게	3748	7.274	47.720
5	ㄴ다	3139	6.092	53.812
6	여	2515	4.881	58.693
7	다	2252	4.370	63.063
8	며	1543	2.994	66.057
9	여야	1493	2.897	68.954
10	은	1431	2.777	71.731
11	기	1323	2.567	74.298
12	운	948	1.840	76.138
13	르	770	1.494	77.632
14	도록	672	1.304	78.936
15	켜	588	1.141	80.077
16	야	559	1.085	81.162
17	켜야	527	1.023	82.185
18	지	498	0.966	83.151
19	ㄱ으로써	461	0.895	84.046
20	으며	461	0.895	84.941
21	어야	450	0.873	85.814
22	였다	310	0.602	86.416
23	시였다	308	0.598	87.014
24	신	289	0.561	87.575
25	ㅏ	243	0.472	88.047

3.2.3. 문법형태소의 처리가 만족스럽지 못하다는 것은 뜻풀이 부분을 보면 더욱 잘 드러난다. 만약 이 사전에서 대표형을 중심으로 한 기술방법을 택했기 때문에 이형태들에 대한 배려가 소홀한 것이었다면, 풀이말에서 그를 보완하기 위한 노력이 있었을 것이다. 그러나 실제 문법형태소에 대한 풀이말을 살펴보면 전혀 그렇지 못함이 확인된다.

한 예로 흔히 사용되는 과거시제 형태소 '-았-'에 대한 기술 내용을 인용하기로 한다.

-았 [토] 시간을 나타내는 끼움토의 하나. 용언에서 말줄기의 끝소리마디에 있는 모음이 밝은 모임일 때 그와 모음조화가 되면서 붙는다. 1) 이야기하는 순간을 기준으로 하여 과거에 있는 시간관계 곧 절대적시간을 나타낸다. / 이 회상기를 보았소? 2) 어떤 현상이나 사실을 지난날의 일로써 묘사하는 데 쓴다. / 비가 좀 오는 것 같았다. 3) 과거에 일어난 결과가 현재까지도 지속되고있음을 나타낸다. / 편지를 받았다. / 마개를 막았다. 4) 미래 또는 현재의 사실을 단정적으로 나타낸다. / 막을가 어떻게 할가? 에이 막았다. 그까짓것! 5) 미래에 있을 사실임에도 불구하고 이미 정해진 결정적인 사실임을 강조하는데 쓴다. / 비오는 꼴이 래일 야유회는 다갔네.

문법형태소에 대한 뜻풀이에는 최소한 ①이형태의 종류 ②이형태의 출현 환경 ③문법적 기능 ④의미 등 네 요소에 대한 설명이 기본적으로 요구된다고 할 수 있다. 그러나 위의 인용에서 볼 수 있듯이 가장 흔한 문법형태소 '-았-'에 대한 설명조차 기본적 요소에 대한 설명이 제대로 되어 있지 않음을 확인할 수 있다.

우리가 분석한 자료의 결과를 보면, '-았-'의 이형태는 어말어미 '-다'와의 결합형만 하더라도 '-았다, -켰다, -였다, -쓰다, -셨다, -셨다, -었다, -켰다'의 일곱 종류로 50만 어절의 자료에서만 총 17,514예에 달한다. 따라서 이 사전이 규범성을 지향하는 것이든 아니면 실용적 목적을 추구하는 것이든, 최소한 이러한 형태들의 출현 환경에 대한 설명이 필요하다고 할 수 있다. 그러나 이 사전의 뜻풀이에서는 단지 '-았-'이 모음조화에 의해 교체되는 형태라는 사실만 언급하고 있을뿐, 실제 그 이형태로는 어떤 것이 있는지, 그리고 그 구체적 출현환경이 어떠한지에 대해서는 전혀 언급이 없다. 결국 이 사전은 문법형태소에 관한한 표제어의 선정과 풀이말 어느 면에서도 충실하지 못한 사전이라는 지적을 피할 수 없는 것이다.

4. 결 론

4.1. 본 연구의 목적은 《조선말대사전》의 편찬에 적용된 이론의 수준을 확인하고, 또한 이 사전이 북한사회의 실제 언어 사용 양상에 대해서 얼마나 정확하고 풍부한 정보를 제공해 주는가를 검토하는 것을 주된 목적으로 하는 것이었다. 이제 지금까지의 논의를 요약함으로써 결론에 대신하고자 한다.

4.2. 2.2.1.에서는 부록의 고어·방언·이두를 제외한 현대어의 표제항들의 품사별 분포를 검토하였고, 2.2.2.에서는 전문용어(학술용어)의 수용양상을 검토하였다. 특히 전문용어의 수용이 체계적이지 못하다는 사실을 확인하였다.

4.3. 3.1.에서는 어휘형태소의 표제항 선정과 관련한 문제를 다루었다. 그 결과 《조선말대사전》의 편찬에서는 합성어류의 표제항 선정 여부를 결정하는 데에는 종래의 사전 편찬에서 택해왔던 의미론적 기준 이외에 별개의 장치를 도출해 내지 못함으로써, “의미적 단일성” 내지는 의미의 은유적 확대라는 기준에 의해 표제항으로의 선정 여부를 결정하였고, 그 결과 거시구조의 단위 설정이 사전 편찬자의 자의적 판단에 의존하는 결과를 낳았음을 확인하였다.

4.4. 3.2.에서는 문법형태소의 처리에 대하여 검토하였다. 그 결과 이 사전은 규범성을 지향하는 사전으로서의 역할도, 실용적 목적으로 활용할 수 있는 사전으로서의 역할을 담당하기에도 부족한 양상을 보임을 확인하였다.

참고문헌

- 고신숙(1987), 《조선어리론문법-품사론》, 평양: 과학백과사전출판사.
 국립국어연구원(1992), 《북한의 국어사전연구》 I.
 국립국어연구원(1993), 《북한의 국어사전연구》 II.
 김광해(1993), <올림말의 관련어휘처리>, 《새국어생활》 3권 4호, 국립국어연구원.
 남기심(1993), <‘조선말대사전’과 문법정보>, 《새국어생활》 3권 4호, 국립국어연구원.
 사회과학원 언어학연구소(1992), 《조선말대사전》 상·하.
 송천식(1993), <‘조선말대사전’의 성격>, 《새국어생활》 3권 4호, 국립국어연구원.

이상복(1993), <북한사전의 올림말>, 《새국어생활》 3권 4호, 국립국어연구원.

정순기·리기원(1984), 《사전편찬리론연구》, 평양 : 사회과학출판사.

大江孝男(1993), <사전의 역할에 관해서>, 《새국어생활》 3권 4호, 국립국어연구원