

종합국어대사전의 표제어

윤 용 선

(국립국어연구원 학예연구사)

1. 선정 방향

사전에 있어서 표제어 선정은 단순히 뜻풀이할 단어를 제시하는 것이 아니라 사전의 골격을 형성하고 사전의 성격을 구현하는 일이다. 그래서 잘 짜여진 표제어 목록은 좋은 사전을 만드는 전제 조건이 된다. 중요성에 비례해 일 자체도 어려울 수밖에 없다. 특히 대사전과 같이 그 규모가 방대할 때는 그 어려움은 더해진다.

우리가 새로이 사전을 만든다 하여 현재보다 수준이 몇 단계 진일보한 그런 사전을 편찬하려는 것은 아니다. 그간 축적된 국어학계의 연구 역량을 토대로 기존 사전을 종합하고 정제함으로써 한 차원 높은 사전으로 나아가기 위한 결산의 의미가 더 강하다. 여기에 시급한 현실적 욕구를 해결하기 위한 목적을 함께 갖는다.

표제어 선정도 마찬가지이다. 실제 우리의 역량은 이상적인, 외국의 유명 사전에 비견할 만한 사전을 만들기에는 미흡하다. 모든 것을 다하겠다고 나서는 것은 만용일 것이다. 그래서 우리는 시급한 현실적 욕구를 충족시키는 데 선정의 중점을 두고 있다. 시급한 현실적 욕구란 언어 생활의 표준을 제시하는 것이다. 사실 좋은 사전을 갖는 것보다 언어 생활의 표준을 제시하는 것이 사회적으로는 더 의미있고 가치있는 일일 것이다. 언어 생활의 표준이란 한마디로 바르고 고운 말이다. 어느 말을 어떻게 말하고 써야 옳은지를 일반 국민들이 쉽게 알 수 있도록 하는 것이 표준을 제시하는 일일 것이다. 이를 위해 어문 규범이 있고, 국어 순화, 화법 표준화 등의 일이 필요한 것이다.

그런데 규범이 있고 아무리 좋은 사업과 운동을 벌인다 해도 그 내용이나 결과가 일

반인들에게 알려지지 않는다면 그 의미는 반감되고 말 것이다. 국어 사전은 언어에 대한 모든 정보를 담고 있고 일반인들이 쉽게 접할 수 있는 것이어서 제도 효과를 충분히 살릴 수 있을 뿐만 아니라 그 자체가 쉽게 이해할 수 있는 구체적인 규범이다. 사전의 이러한 효용성은 표제어를 통해 드러난다. 어느 말이 표준어인지 어떻게 쓰는 것이 바른 표기인지를 표제어를 통해 확연히 보여 줄 수 있기 때문이다. 그래서 우리는 표제어가 한글 맞춤법, 표준어 규정, 외래어 표기법에 맞는지 검토하고, 모호하거나 애매한 부분은 명확히 해석하고, 문제되는 것은 심의하여 표준 문제에 대해서는 믿고 따를 수 있는 표제어 목록을 제시하는 데 주력하고 있다. 우리의 표제어 선정 작업은 광범위한 어휘를 대상으로 한 표준어 사정과 현행 어문 규범의 자세한 해설 작업의 성격을 갖는다 할 것이다.

2. 선정 원칙

종합국어대사전은 역사적으로는 우리말을 표기한 모든 문헌에 나오는 말, 지역적으로는 우리 사회는 물론 북한과 해외동포 사회에서 쓰고 있는 말 중 한국어로 수용 정착된 말을 수록함을 원칙으로 한다.

여기서, '말'은 우리가 생활하면서 쓰고 말하는 모든 말을 포괄한다. 흔히 국어 사전은 언어 사전이니 언어적인 것만을 수록하여야 한다는 지적이 있기도 한데 우리가 일상 생활에서 전문어, 일상어를 구분하거나 고유 명사를 따로 가려 쓰지는 않는다. 오히려 고유 명사나 전문어의 의미를 알고자 사전을 찾을 때가 더 많다. 새 사전이 국민 언어 생활의 편의를 위한 것이라면 당연히 백과사전적이라든가 고유 명사나 전문가가 포함되어야 한다.

'우리말을 표기한 모든 문헌'이라 함은 시간적으로 고어도 포괄하는 동시에 借字表記¹⁾나 한문으로 기록된 말도 수용한다는 뜻이다.²⁾

- 1) 물론 차자 표기의 후론 결과가 학계에서 일반화된 경우에 한한다. 차자 표기는 주로 固有名詞表記, 口訣, 吏讀, 鄕札을 가리키는 뜻으로 쓰나 여기서는 '雨雷', '烏賊魚'와 같이 고유어를 비롯한 음의 한자로 옮긴 取音表記나 音譯語도 포함하는 뜻으로 이 말을 쓴다.
- 2) 한문으로 기록된 문헌에 우리말이 있느냐고 반문할는지 모른다. 그러나 한자어도 우리말의 일부이고 한문 문헌에는 과거 일상 생활에 흔히 쓰던 한자어가 포함되어 있을 수 있다. 물론 언해 문헌에 나오는 한자어가 비로소 우리말로 굳어진 한자어라 하겠지만 그 양이 제한되어 있어 이것만으로는 충분하지 않다. 오히려 단어로 굳어진 한자어가 한문 문헌에만 나올 가능성도 있다. 지금은 사정이 다르나 과거에는 충분히 우리말의 자격을 갖추었던 것으로 보아 수록하고자 하는 것이다.

'말'의 수록 단위는 단어가 된다.旬는 단어의 연속체로 단어를 알면 생산적으로 만들어 낼 수 있는 것이므로 旬를 표제어 단위로 하지 않음은 당연한 일이다. 다만 예외가 존재하는데 문법 형태, 고유 명사, 전문어, 융합형은 수록 단위가 旬 이하나 이상이 될 수도 있다. 문법 형태는 기본적으로 그 단위가 형태소이기 때문에 문법 형태는 자연 형태소 단위가 표제어가 된다. 고유 명사와 전문어는 언어적으로는 旬이더라도 그 개념이 단일하다면 표제어가 될 수도 있다. 예를 들어 '플레밍의 왼손 법칙'은 명백히 旬라고 볼 수 밖에 없지만 하나의 원리를 나타내는 말이고 현실적으로 필요하기 때문에 사전에 수록되어야 한다. 이런 필요성을 감안해 전문어와 고유 명사는 약간의 예외를 허용하고 있다. 융합형이란 단어의 연결에서 음절수 줄이기가 일어나 의존 요소로 재구조화된 것을 말한다. 즉 기원적으로는 두 개 이상의 단어였지만 그 결합이 공고해져 공식적으로는 더 이상 분석할 수 없는 것을 말한다. 엄격한 의미에서 이들을 단어로 보기는 어렵겠지만 그 속성이 단어와 별반 다를 바가 없으므로 표제어로 수록한다. 이상을 제외하고는 모두 단어 단위를 표제어로 수록한다. 그러면 관용구와 속담은 수록하지 않느냐고 반문할 수도 있다. 새 사전은 기존 어느 사전보다 관용구와 속담을 적극적으로 수록할 것이다. 이들이 현실적인 효용성이 높기 때문이다. 관용구와 속담이 부표제어적인 속성을 갖지만 표제어라고는 할 수 없다. 하나의 표제어에 부속된 뜻풀이의 성격을 갖는 것이므로 표제어 수록 단위와 상충되지 않는다.

'한국어로 수용 정착된'이라 함은 외국어를 배제한다는 뜻이다. 현실적으로 외래어와 외국어를 구분하는 것이 쉬운 일은 아니지만 대응하는 우리말이 있는 외래적인 말 중 현실적인 효용성이 없는 것은 신지 않는다. 또 신는 경우라도 대응하는 우리말을 밝혀 일반인들이 가급적 우리말을 사용하도록 유도한다. 이런 취지는 순화어에도 그대로 적용된다. 순화의 결과를 적극 수용하여 순화한 말이 널리 사용될 수 있도록 하려 한다. 오표기와 비표준어도 같은 취지에서 표제어로 수록한다. 잘못된 말을 신지 않는 것보다 실어 놓고 그것이 잘못되었으며 바른 것은 무엇인지를 알려 주는 것이 바른 언어 생활을 유도하는 데 더 효과적이기 때문이다.

관련되는 단어가 여럿일 경우 어느 것을 기본적인 것으로 보느냐도 표제어 선정의 중요한 과제다. 동의 관계에 있는 말들이나 표준어와 비표준어, 바른 표기와 오표기, 순화한 말과 순화된 말, 우리말과 외래어 등의 관계에서 어느 것을 기본적인 것으로 선택하고 어느 것을 참고하게 하느냐에 대한 판단과 처리는 새 사전의 표준성을 살리는 효과적인 방법일 뿐 아니라 표제어 전체의 체계화를 위해서도 필요한 작업이다. 우리는 여러 단어가 문제될 때 현행 규범에 맞는 표준형을 기본 표제어로 삼는 것은 당연한 일이고 그것이 불투명할 때는 고유어를 최우선하고 한자어, 외래어 순으로 기본 표제어를

잡는다. 또 일상 생활에 많이 쓰이는 것을 우선하는데 예를 들어 본말과 준말, 경어와 평어가 있을 경우 일률적으로 본말이나 평어를 기본 표제어로 잡는 것이 아니라 널리 쓰이는 것을 기본 표제어로 잡는다.

이렇게 선정된 표제어의 표기는 한글로 한다. 영문 약어라도 그 표기는 한글로 한다. 현행 어문 규범을 따라야 함은 당연한 일이지만 오표기나 북한어는 예외로 한다.³⁾ 특히 북한어의 표기는 북한식의 표기를 그대로 따른다. 이는 북한의 문헌을 보거나 북한 말과 우리말을 비교할 때 편리하게 하기 위해서이다. 그래서 '노인'은 '로인'으로 실려 '노인'항을 참고하게 한다. 북한어인가에 대한 판단은 일차적으로 북한 사전에만 나온 단어가 검토 대상이 된다. 북한 문헌을 조사해 나온 말이라도 우리 사전에 실려 있지 않다면 그 역시 검토 대상이 된다. 그런데 수록 상태만으로는 북한어를 판별하기 어렵다. 북한 사전에는 우리로서는 ㄱ이라고 판단되는 예가 다수 실려 있어 이를 그대로 북한어라고 보기 어렵다. 우리는 북한어를 북한 사회에서만 쓰는 말, 형태는 같더라도 북한 특유의 용법을 보이는 말, 북한식 표기를 보이는 말로 규정하고 있다. 북한 사전에만 실려 있더라도 우리 입장에서 ㄱ로 판단되고 또 앞의 범위에 포함되지 않는다면 그 말은 수록하지 않는다. 사전 편찬의 태도에서 기인한 차이를 반영할 필요는 없기 때문이다.

기존 사전에서는 표제어를 '-'를 이용하여 분절해 주고 있다. 이렇게 분절하면 의미를 파악하는 데 도움이 된다는 장점이 있다. 그러나 단어에 따라서는 음절 단위로 분절되지 않아 분절 표시가 불가능한 것(예: 개구리)도 있고, 또 그 분절이 애매한 것(예: 회로에탁, 동서남북)도 있어 결과적으로 전체 표제어가 일관되게 처리되지 못한다는 단점도 있다. 우리로서는 아직 어느 것이 더 효용성이 있고 필요한 일인지 결정 내리지 않았다. 다만 작업 과정에서는 후에 넣을 경우에 대비해 표제어를 둘로 분절하고 그 표시를 해 주고 있다. 이 과정에서 분절의 주요 기준은 의미 파악에 도움이 되게 한다는 것이다. 이는 엄격한 언어학적 분석과 우리의 분절 결과가 다를 수도 있음을 암시한다. 만일 학술적인 것과 실용적인 것이 상충된다면 우리는 실용성을 우선한다. 사전은 쓰는 사람이 편리하게 하는 것이 우선하기 때문이다.

현행 맞춤법은 띄어 쓰는 것이 원칙이지만 붙여 쓰는 것도 허용하는 경우가 있다. 고유 명사와 전문어가 이에 해당되는데 그 대상이 되는 말이 구체적으로 열거되어 있지

3) 오표기나 북한식 표기를 보이는 모든 단어를 다 실는다는 뜻은 아니다. 단독형을 싣고 그것이 쓰인 복합어나 파생어는 실지 않는다. 예를 들어 '녀성'은 실지만 복합어인 '녀성복', '녀성문제'는 표기가 다르다 해서 수록하지 않는다.

않아 일반어까지 붙여 쓰는 일이 왕왕 있다. 새 사전은 이를 고려해 그 대상이 되는 단어를 ‘ˆ’를 이용해 밝혀 주려 한다. 조항대로 하나의 단일 개념이지만 언어적으로는 句로 보아야 하는 전문어, 고유 명사의 목록을 밝혀 주는 셈이다. 그러나 일반어에는 이 표시가 들어갈 수 없다. 규범 자체가 허용하고 있지 않기 때문이다.

표제어는 한글 맞춤법의 자모순에 의해 배열된다. 문제되는 것은 동음 이의어와 고어인데 고어는 컴퓨터에서의 자모 순서에 따라 배열한다. 동음 이의어는 나름대로의 기준⁴⁾을 적용하여 일관되게 처리한다.

3. 선정 과정

새 사전에 수록할 표제어를 선정하기 위해서는 우선 기존 사전을 검토하는 작업이 선행되어야 할 것이다. 우리의 작업이 무에서 유를 창조하는 완전히 새로운 작업이 아닌 바에야 우리에게 주어진 기존의 업적들을 검토해 장점을 취하고 단점을 보완해야 하기 때문이다. 다행히 우리에게는 선학들이 노력하여 이루어 놓은 자료가 적잖이 있다. 최초의 대사전이라 할 “큰사전”을 위시하여 대사전 수준의 국어 사전이 10여 종 있다. 우리는 이 중에서 체제나 기술이 체계적이고 좋은 사전이라고 평가 받는 6개 사전을 선정해 검토하는 일을 하였다. 민중서림에서 출간한 “이희승 국어대사전”, 삼성출판사의 “새우리말큰사전”, 맞춤법이 바뀐 이후에 나온 금성출판사의 “국어대사전”, 한글학회가 편찬한 “우리말큰사전”⁵⁾의 우리 사전 4종과 북한에서 편찬한 1960년의 “조선말사전”, 1992년의 “조선말대사전” 2종이 그것이다. 혹자는 최초의 대사전인 “큰사전”이 빠진 점을 지적하기도 하는데 “큰사전”의 내용은 후대의 사전들이 많은 부분 답습하고 부족한 부분을 보완해 놓았기 때문에 최근의 것을 검토하는 것이 효율적이라는 판단에서 일단 제외한 것이다. 물론 참고 자료로는 당연히 이용한다.

-
- 4) 동음 이의어는 아래의 배열 기준을 단계적으로 적용하여 결정한다.
- 어휘 형태를 문법 형태에 우선한다.
 - 고유어, 한자어, 외래어의 순으로 배열한다.
 - 일반어를 전문어보다 우선한다.
 - 체언-용언-보조 용언-관형사-부사-감탄사-어근-어간-조사-어미-접사 순으로 배열한다.
 - 기타 사용 빈도가 높은 것을 우선한다.
- 5) 한글 맞춤법이 바뀐 이후에 나온 사전이지만 ‘우리말큰사전’은 현행 맞춤법을 따르지 않고 있다. 한글학회 자체의 규범에 따라 편찬된 사전이어서 현행 맞춤법과는 상충되는 경우가 적지 않다. 이용자들의 주의가 요망된다.

이들을 검토하기 위해 우리는 같은 단어별로 사전의 기술을 모아 비교할 수 있는 자료를 만들었다. 우리는 이를 '표제어별 비교 카드'(이하 '비교 카드'라고 한다)라고 부른다. 사전별로 수록하고 있는 표제어가 달라 비교 카드가 약 60여만 장 되는데 우리가 검토한 사전 중 최대 어휘수를 가진 사전이 약 40만 단어인 점을 감안하면 기존 사전들이 표제어 선정면에서 상당히 차이 난다는 점을 알 수 있다. 이 차이는 구와 단어의 판단이나 고유 명사, 전문어, 외래어, 방언의 수록 범위가 달라 일어나는 것이 많은 부분을 차지하고, 같은 단어라도 표기가 달라 별개의 카드가 된 것, 북한 특유의 어휘나 표기를 보이는 것 역시 많은 부분을 차지한다. 그런데 이 차이에는 상당한 것도 있으나 잘못된 것도 없지 않다. 우리의 표제어 선정은 6개 사전에서 차이 나는 부분을 새 사전의 성격에 맞게 조정 수용하는 일이 주가 된다.

표제어 비교 카드는 여러 사전을 비교하는 데는 요긴한 자료가 되나 표제어 전체의 골격이나 단어 상호 간의 관련을 파악하는 데는 부족한 점이 있다. 이 문제를 해소하고 표제어 전체를 일목연하게 보기 위해 비교 카드의 주요 정보를 전산 입력했다.

이 작업은 단어 상호 간의 관련을 밝히고 표제어 전체를 단어의 속성에 따라 분류하는 데 도움을 준다. 즉 동의 관계나 파생 관계에 있는 단어는 무엇인지, 오표기나 방언, 북한어라면 바른 표기나 대응하는 표준어는 무엇인지를 입력하고 해당 표제어의 속성, 즉 전문어, 고유 명사, 북한어, 방언, 특수어, 오표기, 고어, 외래어인가를 구분하도록 하였다. 이를 입력하여 전산 처리한 결과물과 카드를 바탕으로 하여 표제어 선정 작업을 하게 된다.

여기에 어떤 단어가 실제 쓰이는지 보여 주는 용례 자료도 표제어 선정에 도움이 되겠지만 우리가 입력한 6천만 어절에 나오는 어휘는 사전에 수록된 단어수에 비해 형편 없이 적기 때문에 실제적인 도움이 되지 못한다. 입력 자료로부터 찾아낸 단어는 차후 전체 표제어 목록과 비교하고 미등재 어휘를 색출하는 데 활용될 것이다.

표제어 선정 작업은 일차적으로 어떤 단어가 있을 때 그 단어가 우리 사전에 수록될 말인가를 결정하는 작업이다. 표제어 선정 원칙에 따라 표제어를 일일이 검토하는데 한정된 인력으로는 모든 단어를 충실히 검토하기 어렵다. 그래서 일단 기존 사전의 처리를 참고하고 있다. 풀이 분량이 많거나 6개 사전 모두에 실려 있는 말이라면 일단 수록한다. 우리의 기존 사전들이 서로를 참고했기 때문에 잘못이 그대로 답습되는 일도 없지 않지만 일단 6개 사전에 모두 올라 있는 단어를 삭제한다는 것은 현실적으로 매우 어렵다.

표제어 선정의 미묘한 부분은 개인차가 있을 수 있기 때문에 담당자들이 서로의 경험과 문제점을 토의해 서로의 감각을 일치시키고, 최대한 구체화된 선정 지침을 만들어

일관되게 처리하려고 한다.⁶⁾ 표제어를 선정하다 보면 일군의 단어들은 하나의 유형을 형성하기도 하는데 이런 유형은 어떻게 처리하자는 것이 선정 지침의 골자가 된다. 그런데 그 선정 지침이 해당 단어의 일부만을 보고 결정하기 어려운 점이 있다. 예를 들어 부사화 접미사 ‘-이/히-’의 표기 문제는 어떤 규칙성이 발견되기도 하는데 극히 일부 예외가 발견되곤 한다(예: 자육이). 사전 전체의 해당 단어를 모두 모아 놓으면 그 예외는 더 생길 수도 있다. 이때 규칙성을 부여하고 일관되게 처리함으로써 독자들이 조금 더 쉽게 이해하고 예측하도록 하느냐, 아니면 사전에 있는 그대로를 살려 두느냐를 결정하여야 한다. 후자의 방법은 처리의 일관성을 해치고 문제의 해결이 되지 못한다는 단점이 있으며 전자는 언어 사실을 왜곡할 소지가 있다. 더구나 지침이 충분한 객관성을 갖지 못하고 작업의 편의를 위한 것이라면 그것은 더욱 문제가 될 것이다. 현재는 어느 한 쪽을 선택하지 않고 문제 되는 단어들을 유형별로 분류하는 일을 하고 있다. 일부분을 보고 내린 지침이 잘못될 경우는 오히려 그 잘못을 바로 잡기가 어렵기 때문이다. 이를 방지하기 위해 문제 되는 유형을 세분하여 선정 과정에서 분류한 다음 전체를 검토해 타당한 처리 방법을 강구할 생각이다. 이 단어들을 우리는 문제 표제어라고 부르는데 이렇게 모아진 문제 표제어들은 내부적으로 심의하여 그 해답을 찾고, 자체적으로 해결하기 어려운 것은 별도의 자문회의나 국어심의회 등의 검토 및 심의를 거쳐 결정하게 될 것이다.

새 사전은 언어 생활의 표준을 제시하는 것이 목적인 만큼 표제어 선정의 주요 활동도 규범을 해석하여 개별 단어에 적용하는 것이다. 문제 되는 어형(사전으로는 표기) 중 어느 것이 현행 어문 규범에 맞는 것인지 판단하는 일로 세심한 해석이 필요하다. 규정에 명시되어 있지 않은 경우는 규정의 근본 취지를 생각하고 예시된 단어로부터 규칙을 추론해 적용하기도 한다. 한마디로 현행 어문 규범을 재해석하여 어문 규범 해설서를 만드는 작업과 유사한 일을 하는 셈이다. 이런 과정에서 새로운 심의와 결정이 필요한 단어나 규칙이 생길 수 있는데 이 역시 문제 표제어로 분류하여 원내외의 검토 및 심의를 받게 될 것이다.

전체적으로 우리의 표제어 선정 작업은 크게 2개의 일로 나뉜다. 기존 사전에 수록된 단어들을 검토하여 다듬는 일이 그 첫째이고, 실생활에는 쓰이지만 사전에 수록되지 않은 미등재어를 찾아 표제어 목록을 보완하는 일이 그 두 번째이다. 미등재 어휘는 자체적으로 혹은 용역을 주어 전산 입력한 문헌 자료에 쓰인 어휘를 모두 찾은 후 이를 우리의 표제어 목록과 비교하여 미등재어를 찾고, 그 수록 여부를 결정하는 과정을 거치

6) 이 지침의 자세한 내용은 이 책에 실린 다른 글 “표제어 선정의 제문제”를 참조하기 바란다.

게 된다. 첫 번째의 작업은 앞서 말한 대로 기존 사전의 검토와 전산 입력, 개개 표제어의 수록 여부 판단, 표준성(규범성)의 검토, 문제 표제어의 분류와 심의라는 절차를 거쳐 결정된다. 이렇게 결정된 표제어 목록은 집필과 교열 과정을 거쳐 최종적으로 한번 더 검토될 것이다. 자세한 뜻풀이 과정에서 예상하지 못했던 문제점들이 발생할 수도 있기 때문이다. 이 과정에서는 단어들 사이의 관련, 기본 표제어의 선택, 부표제어의 설정 등이 집중적으로 검토될 것이다.