

표제어 선정의 제 문제

정 희 원

(국립국어연구원 학예연구사)

1. 머리말

사전이란 일정한 언어의 어휘들을 표제항으로 선정하여 철자로 표시하고 그 각각의 항목에 대해 필요한 정보(발음, 어원, 문법범주, 정의, 용례, 관련 어휘 등)를 체계적으로 제시해 주는 텍스트이다. 여기에서 일정한 원칙에 따라 선정되어 제시된 표제어들의 전체 구조는 흔히 사전의 거시구조(lexicographic macro-structure)라 하고 각 표제어에 대해 일정한 순서로 분류, 배열되어 제시된 다양한 언어 정보가 이루는 각각의 사전 항목은 미시구조(lexicographic micro-structure)라 하는데, 이에 따르면 사전은 거시구조와 미시구조의 통합 체계로서의 텍스트라 할 수 있다. 이상적인 사전에서는 거시구조와 미시구조가 일정한 기준에 따라 체계적으로 설정되어 통일성을 유지해야 할 것이다. 따라서 국어 사전 편찬에서 정확한 뜻풀이 못지 않게 중요한 것이 체계적이고 일관된 기준에 따른 표제어 목록을 제시하는 일이다.

새 사전 편찬을 위해 표제어를 선정하기 위해서는 일단 그 후보가 될 수 있는 어휘의 목록이 있어야 하는데, 기존 사전이 있는 경우에는 거기에 수록되어 있는 말들이 새로 편찬하고자 하는 사전의 자료로서 우선 조사·검토되어야 할 것이다. 이를 위해서 표제어부에서 가장 먼저 한 일은 기존 사전 중의 6개 사전¹⁾을 표제어별로 카드화하는 작

1) 대상 사전은 국어대사전(금성출판사, 1991), 우리말큰사전(한글학회, 1992), 새우리말큰사전(삼성출판사, 1986), 국어대사전(민중서림, 1982), 조선말대사전(북한 사회과학원, 1992), 조선말사전(북한 사회과학원, 1962)이다. 같은 표제어별로 풀이 내용을 하나의 카드에 붙이고 이에 대한 정보(관련 어휘, 수록 상황, 단어의 속성 등)를 전산 입력하여 활용하고 있다. 표제어 선정은 이 기초 자료를 통해 이루어진다.

업이었다. 다음에는 이미 카드화되어 있는 6개 사전의 표제어 목록을 검토하는 작업을 해 왔다. 이는 6개 사전 표제어 비교 카드를 검토하여 새 사전의 표제어로 수록할지 여부를 판단하는 것으로, 규범에 맞지 않는 것을 구분해 내고 언어학적으로 잘못 판단된 것이나 우리 사전의 성격과 맞지 않는 단어를 선별해 내는 것이다. 이 과정에서 일반어와 전문어를 구분하고, 북한어·방언·고어 표시를 한다. 또 단어의 형태를 분석하고 형태론적으로 관련되는 단어나 동의어를 함께 묶는 작업도 병행한다. 이 밖에 뜻풀이할 때 필요한 사항, 예를 들면 예상되는 원고 분량을 표시해 주거나 관련되는 단어 중 직접 뜻풀이할 단어를 지정해 주기도 한다. 94년 말 현재까지 전체 약 63만 표제어 중에서, 7항에서 8항까지 27만 어휘를 검토하여 일차적인 표제어 목록을 확정하였다. 그러나 이렇게 만들어진 목록이 새 사전의 최종적인 표제어 목록은 아니다. 선정자들이 직접 판정하기 어려운 표제어에 대해서는 유형별로 모아서 지속적으로 검토하여야 한다. 다음에는 실제 표제어 선정 과정에서 나타나는 몇 가지 문제점과 그 처리 방안을 항목별로 살펴보겠다.

2. 표제어 선정의 실제 문제들

표제어를 어떤 기준에 따라 선정할 것인가 하는 문제는 새 사전의 목표와 성격이 어떤 것인가와 직접적인 관련이 있다. 우리가 목표하는 새 사전은 규범적인 언어 사전을 지향하고 있으며 고어는 물론 개화기 이후에 쓰인 국어 어휘를 총망라하여 표제어로 실을 것을 목표로 하고 있다. 또한 남북한 통일을 대비하여 분단 후 이질화된 어휘나 북한에서 새로 만들어진 어휘 등 북한어(소위 문화어)도 실을 것을 계획하고 있다. 따라서 실제 선정 작업도 이러한 목표에 따라 행해진다.

2.1 단어와 구의 판별 문제

국어 사전에서 표제어의 단위는 물론 단어가 될 것이다.²⁾ 즉 특정한 어휘가 단어이면 사전에 표제어로 등재되겠지만 그것이 구적인 구성이라면 표제어로서는 사전에 오르지 않을 것이다. 그러나 실제 편찬자의 입장에서 보면 구와 단어를 판별하는 일이 가장 어

2) 일반어의 경우는 단어만 수록하는 것이 원칙이다. 그러나 문법 형태의 경우는 형태소 단위도 표제어가 될 수 있고, 전문어· 고유 명사· 준말 등은 단어보다 큰 단위(구)도 표제어로 실릴 수 있다.

려운 문제가 된다. 기존 사전들에서도 구와 단어의 구별 기준에 관해서 가장 많은 불일치를 보이고 있다. 정확한 규범의 제시를 목표로 하는 우리 새 사전의 경우는 이 판별이 띄어쓰기에 대한 기준으로 작용할 것이기 때문에 특히 바른 단어 목록의 제시가 요구된다.

특히 둘 이상의 단어가 통사적으로 결합하여 구성되는 합성어는 구와의 경계가 분명하지 않다. 이의 판별을 위해서 몇 가지 원칙을 설정하여 놓고 이 원칙에 따라 합성어와 구를 판별하는데, 선전자 개인이 결정하기 어려운 어휘에 대해서는 따로 모아서 심의하는 절차를 거쳐야 한다. 합성어와 구의 판별에 대한 기존의 연구에서는 어순이 고정적이거나 내적 확장이 불가능한 것, 의미의 변화가 뚜렷한 것 등을 합성어로 판별할 것을 제시해 왔다. 이를 원용하여 우리의 표제어 선정 작업은 아래와 같은 기준³⁾을 설정하여 구와 합성어를 판별하고 있다.

2.1.1 명백한 통사적 구조의 형태는 표제어로 올리지 않는다.

명백하게 통사적인 구조를 형성하고 있는 것은 구로 인정하여 단어 목록에서 우선 제외한다. 일반적으로 'X+서술어'의 구조를 가진 유형이나, '수식어+X'의 구조를 가지는 것들은 우선 제외된다. 'X+서술어'의 통사 구조를 바탕으로 생성된 형태는 특히 한자 합성어의 경우에 매우 많이 나타난다. 예를 들어서 '능력 발휘, 문제 제기, 협상 제의, 임무 완수, 규제 완화' 등 많은 예들이 '목적어+술어'의 통사적 구조를 지니고 사전에 표제어로 실려 있다. 이들은 한자어 특유의 문 구성 원리에 따라 지극히 생산적으로 등장하는 형태들로서, 이들의 대부분은 조사나 어미를 적절히 사용하여 정상적인 국어 문장으로 바꿀 수 있으므로 모두 문장이나 구로 처리할 수 있다. '수식어+X'의 구조, 즉 '관형어+명사'나 '부사어+동사'의 구조로 형성된 유형도 단어 목록에서 제외된다. 예를 들어서, '일반 사회, 외부 압력, 적극 개입, 자진 사퇴' 등의 형태들은 원칙적으로 표제어 목록에서 제외된다. 그러나 실제 표제어 선정 과정에서 이 기준을 일률적으로 적용하기란 매우 어렵다. 특히 어떤 합성 형태가 특정 전문 분야에서 특수한 의미를 가지고 사용된다면, 이 기준에 의해 제외시킬 수는 없을 것이다. 예를 들어서 '경제계획, 경제교류, 형식논리, 거시(미시)정책' 등의 표제어는 특정 전문 분야에서 특수한 학술적 용어로 사용되는 낱말이므로, 비록 구적인 구성을 지니지만 표제어로 등재해야 할 것이다. 따라서 명백한 통사적 구조를 형성하고 있는 예들이라도 표제어 선

3) 이는 국립국어연구원에서 사전 편찬 사업의 기초 작업으로 용역 의뢰한 "국어사전에서의 합성어 처리에 관한 연구"(김광해 김동식, 1993)에서 제시한 기준과 같다.

정의 첫 단계에서는 유형별로 모아서 다른 기준들의 적용 가능성 여부를 따져 보아야 한다.

2.1.2 불투명한 성분을 포함한 형태는 표제어로 수록한다.

많은 복합 형태들 중에는 구성 성분 중의 하나가 자립적인 형태로 분리되지 않으므로 써 마치 접사처럼 사용되는 경우가 있는데, 이런 형태들은 결합된 각각의 성분으로부터 의미를 쉽게 산출해 낼 수 없으므로 따로 표제어로 등재해야 할 것이다. 예를 들어 보면, ‘옷갓, 난장, 날담비, 널방석, 선소리, 도배반자’ 등 명사들과 ‘궐나다, 귀나다, 마들다, 태먹다, 축나다’ 등 많은 동사(또는 형용사)들이 있다. 이들 복합 형태들은 공시적으로 널리 쓰이지 않는 단어나 어근적인 성격을 가진 명사를 포함하고 있다. 즉 이들은 결합된 형태 전체의 의미 또는 이들을 구성하고 있는 구성 성분 중 일부의 의미가 불투명하기 때문에 해당 항목에서 각각 뜻풀이를 해 주어야 한다. 그러나 의미의 불투명성 여부는 사람에 따라 달라질 수 있는 것이므로 실제 적용에서는 많은 어려움이 따른다.

2.1.3 의미의 변화를 수반하는 형태는 표제어로 수록한다.

이는 흔히 의미의 특수화라 하여 합성어와 구를 구분 짓는데 있어서 가장 중요한 기준으로 지적되어 오던 것이다. 일반적으로 복합 표현의 의미는 그들을 구성하고 있는 단순 표현들의 의미의 합으로 해석된다. 그러나 단순 표현들이 결합하여 복합 표현이 되는 과정에서 특정한 의미를 갖게 되는 경우가 있는데, 이들은 합성어가 된다. 예를 들어 두 개의 단어 X, Y의 의미를 각각 a, b라고 했을 때 이것이 결합된 X+Y의 의미가 a+b+α로 해석된다면, 즉 +α의 의미가 형성되었다면 X+Y의 형태를 합성어로 인정한다는 것이다. 이렇게 +α의 의미가 덧붙여진 형태의 정확한 의미를 알기 위해서는 그 각각의 표제어에 대해 구체적이고 특수한 정보를 추가해야 할 것이다. ‘끈붙다, 패차다, 막놓다, 놀아먹다, 구워막다’ 등 많은 예들이 이 유형에 해당하는데, 위에서 제시한 두 번째 기준(“불투명한 성분을 포함한 형태는 표제어로 수록한다”)과 뚜렷이 구분하기는 어렵다. 이 기준은 특히 기존 사전에 수록된 많은 어휘 항목 중에서 의미의 변화를 수반하지 않은 예들을 구로 처리하여 표제어에서 제외하는 데 유용하게 사용할 수 있다. 예를 들어서 단순히 구성 형태들 간의 결합 의미 외에는 별다른 의미를 가지지 않는 ‘개구리알, 가방꾼, 가방고리, 개간공사’ 등의 표제어는 이 기준에 따라 표제어 목록에서 제외될 수 있다.

2.1.4. 전문 용어나 명칭은 따로 기준을 설정하여 수록 여부를 결정한다.

복합형이 명칭일 때는 모두 단어로 판정한다. 단체의 명칭, 사회적·역사적 사건의 명칭, 전문 분야에서 사용되는 학술 용어 등의 결합 형태들은 고유 명사와 동일한 기능을 하는 것으로서 합성어의 자격을 충분히 가진다. 단일한 개념을 형성하고 있고, 내부적인 확장이 불가능한 점 등 일반적인 합성어의 조건을 모두 만족시킨다. 우리의 새 사전은 고유 명사나 전문용어 등을 모두 표제어로 반영하는 백과사전적인 성격을 가지므로 이들은 모두 표제어로 등재된다. 다만 이들을 어느 정도까지 사전 표제어로 수록하는가 하는 점이 문제로 남게 된다. 이러한 성격을 가지는 표제어에 대해서는 일차적인 선정 과정에서 전문어나 고유 명사의 표시를 해 주게 되면 다시 전문어 분과로 넘어가 각 영역별로 담당자가 심의하여 사전 표제어로 등재할지의 여부를 결정하게 된다.

2.2 표준어 결정의 문제

우리 사전의 가장 큰 목표는 국민 언어 생활의 표준을 제시한다는 것이다. 언어 생활의 표준을 제시하는 데는 여러 가지 방법이 있겠지만 사전 편찬을 통한 방법이 가장 효과적일 것이다. 현재 우리에게는 표준어 규정과 한글 맞춤법, 외래어 표기법, 표준 발음법 등의 어문규정이 있지만 그것은 언어학적 규칙이어서 일반인이 쉽게 이용하기는 어렵다. 실제로 많은 사람들이 바른 표기법이나 발음, 표준어를 잘 모를 때에는 규정보다는 국어 사전을 찾아보게 될 것이다. 그러나 기존 사전들을 비교, 검토해 보면 사전마다 달리 처리된 낱말들이 많이 발견되어 사전 이용자들이 오히려 혼란을 겪고 있는 실정이다. 이는 사전에 따라 규범을 따르지 않고 독자적인 기준에 의해 표기법이나 표준어를 정한 경우도 있지만, 규정에 따라서는 해석에 따라 의견이 다를 수 있는 것들이 있는데 그 결과가 각 사전에서 표제어가 달리 처리되는 결과로 나타나기도 한 것이다. 우리 사전이 추구해야 할 중요한 목표의 하나가 표준적인 단어 목록을 제시하는 것이므로 혼동을 빚고 있는 단어에 대해서는 현행 어문 규범을 철저히 분석하여 적용하되, 규범의 적용이 모호하거나 규정 자체가 미비한 부분에 대해서는 그에 대한 구체화와 보충이 필요할 것이다.

2.2.1 사이시옷

기존 사전의 표제어들을 검토하다 보면 사이시옷이 사전 간 달리 처리된 예들이 상당히 많이 발견된다. 사이시옷은 두 개의 명사가 결합되어 합성 명사가 구성될 때, 아래 명사의 첫소리가 된소리로 나는 등 발음의 변화가 생기는 경우가 있는데, 그 된소리로

된 것을 정확하게 표기에 반영하기 위해서 그 사이에 ‘ㅅ’을 끼워 넣기로 한 것이다. 한글 맞춤법 규정 제30 항에서는 사이시옷이 고유어끼리 결합한 합성어 또는 고유어와 한자어가 결합한 합성어에만 적용되는 것임을 분명히 하고 있다. 즉 한자어 사이에는 사이시옷을 붙이지 않는 것이 원칙인데, ‘곳간(庫間), 셋방(貰房), 숫자(數字), 차간(車間), 뒀간(退間), 횡수(回數)’의 6개 단어만을 예외로 인정하여 사이시옷을 쓰도록 한다는 것이다. 우리의 새 사전은 현행 어문 규범에 충실한 사전이므로 이 규정을 엄격히 적용하여 사이시옷을 표기하게 될 것이다. 규정을 충실히 적용하므로 한자어와 한자어가 결합된 표제어에 대해서는 위에 제시한 6개의 예외를 제외하고는 일률적으로 사이시옷을 붙이지 않게 된다. 따라서 기존 사전에 있는 ‘전셋방(傳貰房), 내과(內科), 장미과(薔薇科)’ 등의 표제어는 각각 ‘전세방, 내과, 장미과’로 표기될 것이다. 고유어와 고유어, 고유어와 한자어로 구성된 형태의 경우에는 사이시옷을 쓸 것인가를 판단하기 위해서는 단어와 구의 판별, 정확한 발음의 결정이 선행되어야 한다. 그러나 이들은 개인에 따라 판정에 차이가 생겨날 수 있다. 합성어와 구의 판별은 위에서 제시한 기준에 따라 어느 정도 해결될 수 있을 것이다. 발음의 경우에는 표준 발음법 규정에 따라 먼저 발음 표시를 한 후에 사이시옷 표기를 결정해야 할 것이다. 이러한 과정을 거쳐서도 해결하기 어려운 문제들은 유형별로 모아서 원내외의 검토나 심의를 받게 될 것이다.

2.2.2 접사 ‘-이/-히’의 표기

이 표기 역시 발음 현상에 의존하므로 판단이 개별적이어서 문제가 된다. 한글 맞춤법 제6 장 제51 항에서는 “부사의 끝음절이 분명히 ‘이’로만 나는 것은 ‘-이’로 적고, ‘히’로만 나가거나 ‘이’나 ‘히’로 나는 것은 ‘-히’로 적는다”고 되어 있다. 그리고 각 항목에 ‘-이’ 또는 ‘-히’로 표기해야 할 말들을 나열해 놓았다. 그런데 이 예들 중에 포함되지 않은 단어들로서 사전 간에 표기 차이가 나는 것들이 있다. 이 규정을 보다 구체적으로 풀이해 보면 다음과 같다. ‘이’로만 소리나는 것에 해당되는 예들은 접미사 앞의 어근이나 어간이 다음과 같이 네 가지로 나누어진다. (1) ‘-하다’가 붙는 어근의 끝소리가 ‘ㅅ’인 경우(깨끗하다→깨끗이, 반듯하다→반듯이, 의젓하다→의젓이). (2) 비불규칙 형용사의 어간인 경우(곱다→고이, 대견스럽다→대견스러이, 날카롭다→날카로이). (3) 규칙적인 활용을 하는 형용사의 어간인 경우(많다→많이, 적다→적이, 헛되다→헛되이). (4) 같은 말이 겹쳐진 첩어로서, 뒤에 ‘-하다’가 오지 못하는 경우(겹겹→겹겹이, 일일→일일이, 튼튼→튼튼이). ‘이’나 ‘히’로 소리나는 것에 해당되는 예들은 모두 ‘-하다’가 붙는 어근이 앞에 올 경우뿐이다(솔직하다→솔직히, 나른하다→나른히). 다만 끝소리가 ㅅ인 어근은 언제나 ‘이’로만 소리나므로 제외된다. 그런데 기존 사전들

의 표제어를 비교해 보면 이 예에 명시되지 않은 단어들 중에서 사전에 따라 달리 표기한 예를 찾아볼 수 있다. 예를 들어 '깊숙이/깊숙히, 뻑뻑이/뻑뻑히, 자욱이/자욱히'의 경우가 있는데 이런 예에 대해서는 규정을 적용한 후에 해결되지 않는 예에 대해서는 따로 정리하여 검토하여야 할 것이다.

2.2.3 접사의 선택

형태 결합에 자주 이용되는 접사의 형태를 규정하고 그 교체 조건을 명확히 하면 표제어의 표기에 유용하게 이용할 수 있다. 표준어 규정에는 명백히 조건이 제시되어 있는데도 그 해석에 따라 각기 다른 형태를 표준어형으로 잡는 예가 있는데, 접미사 '-장이/-쟁이'가 대표적인 경우이다. 표준어 규정 제9 항을 보면 기술자에게는 '-장이', 그 외에는 '-쟁이'가 붙는 형태를 표준어로 한다고 하고, '미장이, 유기장이, 멧쟁이, 소금쟁이' 등의 예를 들어 놓았다. 문제는 '-장이/-쟁이'에 혼란을 보이는 어떤 표제어가 있을 때 그것이 특별한 기술과 관련되는지 여부의 판단이 개인에 따라 다를 수 있다는 데 있다. 예를 들어서 '염장이/염쟁이, 점장이/점쟁이' 같은 예들은 시체를 염하거나 점을 치는데 있어서 특별한 기술이 필요하다고 인정할 때와 그렇지 않은 경우에 판단이 달라질 것이다. 또한 같은 어간에 '-장이'와 '-쟁이'가 결합하였을 경우에 뜻풀이가 달라지는 경우도 있을 수 있을 것이다.⁴⁾ 접사의 선택이 문제가 되는 경우에도 사전 간 비교를 통해 문제가 되는 예들을 모아서 일관성 있는 처리를 해야 할 것이다.

2.2.4 오표기나 비표준어의 등재

우리의 새 사전이 표준적인 단어 목록을 제시한다고 해서 오표기나 비표준어를 전혀 표제어로 신지 않는 것은 아니다. 오히려 해당 낱말이 규범에 어긋난다는 것을 적극적으로 알리기 위해서 제한된 범위 내에서 오표기도 신는 것을 원칙으로 한다. 그러나 잘못된 어형을 무한정 반영할 수는 없으므로, 우선 기존 사전들에 실려 있는 오표기나 비표준어들이 등재 대상이 된다. 1980년대에 들어 규범이 바뀌면서 달라지게 된 말들의 경우나 일반 언중이 틀리기 쉽다고 인정되는 형태는 되도록 많이 반영한다. 예를 들어서 '수도물/수돗물' 같이 사전 간에 달리 처리된 단어는 두 가지 어형을 모두 실어서 '수도물'이 '수돗물'의 틀린 어형임을 보여 줄 것이다. 다만 사전에 복합형이 많이 등재되어 있는 경우는 단독형만 오표기를 보일 것이다. 예를 들어서 '까스'와 '까스난로'

4) 예를 들어서 '그림장이'와 '그림쟁이'는 각각 다른 의미를 나타내는 말로 사전에 실려 있다.

는 사전에 모두 등재되어 있는데 이 경우에는 ‘까스’만을 표제어로 실어서 바른 표기 형태인 ‘가스’로 보내 주고 ‘까스난로’는 실지 않을 방침이다. 이는 표기만 다른 북한어의 경우에도 같이 적용된다.

2.3 표제어 전체의 균형 문제

표제어 상호 간에는 의미, 형태 등에서 밀접하게 관련된 것들이 있다. 이 대응 관계에서 기존 사전의 수록 어휘가 공백을 보일 경우 그것을 보충하느냐가 문제가 된다. 표제어 전체의 균형을 위해서는 공백을 메우는 것이 좋으나 그것이 실제 언어 현실을 왜곡할 위험이 있으므로 반드시 좋은 것은 아니다. 따라서 사전 편찬 과정의 미비함으로 인한 것과 실제 언어를 반영한 것인가를 판별하여 보충 여부를 결정하여야 한다.

2.3.1 복합 형태의 대응

복합적으로 구성된 형태의 경우에 그것을 구성하고 있는 요소가 의미상 상대되는 것 일지라도 복합형이 꼭 일대일로 대응되는 것은 아니므로 공백을 채워서 대응하도록 표제어를 정리할 필요는 없을 것이다. 다만 구성 방식이나 의미 면에서 대응 관계가 명확한 것이라면 경우에 따라서는 공백을 보충할 수 있다. 공백을 보충하는 경우에는 반드시 언어 자료에 의한 검증이 있어야 할 것이다. ‘남, 여’로 구성된 복합어의 사전 수록 상황을 살펴보자.

여가(女家)	남가(男家)
여-교장	-
여-기사	-
여-직원	-
여-학생	남-학생

위의 예에서 볼 수 있듯이 ‘여교장, 여기사, 여직원’ 등의 표제어에 대응하는 ‘남교장, 남기사, 남직원’ 등의 낱말은 현재 사전에 등재되어 있지 않다. 이러한 표제어에 대해서는 대응형이 실제 언어에서 사용되고 있는지 조사를 통해서 명백한 공백이라고 인정되는 경우는 따로 목록을 만들어서 관리하고 있다.

복수표준어의 경우에는, 어느 한 쪽의 어형만이 복합어를 형성하는 경우가 있다. 예를 들어서 ‘가물’과 ‘가뭄’은 복수표준어인데, 복합어의 경우는 각각 ‘가물-칠, 가뭄-

해'의 예에서 달리 나타난다. 이런 경우에 대응하는 복합어를 따로 만들 필요는 없을 것이다.

2.3.2. 의성의태어

의성의태어는 일반적으로 어감이 다른 큰말, 작은말, 센말, 여린말 등을 관련 어휘로 규칙적으로 가지고 있다. 그런데 이런 관련 어휘가 사전에 따라 등재된 정도가 많이 다르게 나타난다. 북한 사전의 경우는 우리 사전에 비해 어휘의 공백을 최소화하려는 경향이 강하게 나타난다. 따라서 우리 사전에는 없는 의성의태어가 상당히 많이 수록되어 있는 것을 볼 수 있다. 이것이 실제 북한에서의 언어 현실을 반영하는 것인지 또는 공백을 채워 넣기 위한 인위적 조작의 결과인지는 명확히 판단하기 어렵지만, 일단 사전에 수록된 것들을 제외하기는 어렵다. 따라서 이미 사전에 있는 표제어들을 일부러 빼지는 않지만 거기에서 더 공백이 생기더라도 그 어휘를 보충하지는 않을 것이다. 예를 들어서 사전에 '개굴개굴'은 '개굴개굴'의 큰말로 실려 있다. 또한 '개굴개굴'은 그 부표제어로 '개굴개굴하다'를 가진다. 그러나 '개굴개굴'의 부표제어인 '개굴개굴하다'는 실려 있지 않다. 이런 경우에 '개굴개굴하다'를 따로 만들어 실지는 않겠다는 입장이다. 그러나 용례 추출을 위한 입력 자료로부터 '개굴개굴하다'가 실제로 쓰인 예가 발견된다면 표제어로 올릴 수 있다.

2.4 북한어의 판별

우리의 새 사전은 북한 지역에서 사용하는 어휘까지도 망라하여 표제어로 등재할 것을 목표로 하고 있다. 이는 북한에서 사용하는 어휘를 우리 말과 비교하여 어떤 점이 다른가를 보이고 일반 국민들이 북한의 말을 대했을 때 쉽게 찾아 볼 수 있는 자료를 제공하기 위해서이다.

북한어의 판정은 기존 사전의 어휘 중 우선 북한 사전에만 실려 있는 것을 대상으로 하지만, 북한 사전에만 실려 있다고 해서 모두 다 북한어라고 판정할 수는 없다. 우리 사전과 수록에 차이가 있는 것은 그 원인에 따라 언어적인 것, 사전 편찬 태도의 차이에 의한 것으로 나뉘기 때문이다. 원칙적으로 우리는 언어적인 차이를 보이는 표제어만을 북한어로 본다. 북한 사전에만 실려 있는 어휘들 중 상당수는 복합 형식으로 이루어져 있다. 따라서 북한어를 판별할 때는 북한 사전에만 실려 있는 단어 중 우리 사전에서는 구로 판정하여 수록하지 않은 어휘를 먼저 제거하여야 한다. 이 때의 구별 기준은 우리 사전에서의 단어, 구에 대한 판단 기준이다. 예를 들어서 '가로동갯, 가로수길,

가죽가방, 가죽신발' 등의 표제어는 북한 사전에만 나타난다. 그러나 이런 낱말들은 우리말에서는 전혀 쓰이지 않거나 우리말과는 다른 뜻을 지니고 쓰이는 것이 아니라, 구적인 구성이므로 우리 사전에 등재되지 않은 예이다. 따라서 이들은 새 사전의 표제어로 등재되지 않을 것이다. 그러나 북한에서의 소위 말다듬기 사업의 결과로 만들어진 다듬은 말의 경우는 구적인 구성이더라도 사전 표제어로 등재할 수 있다. 예를 들어서 북한 사전의 표제어인 '가까운바다, 꽃피는때' 등은 명백히 구적인 구성이지만 각각 '근해(近海), 개화기(開花期)'의 다듬은 말이므로 북한어로 등재할 수 있다는 것이다. 그러나 이 문제는 아직 확정된 사항이 아니므로, 앞으로 더 논의되어야 할 것이다.

일반적으로 북한어인 표제어는 그 특성에 따라 (1)북한에서만 사용하는 단어나 북한에서 표준으로 인정한 단어('가두녀성, 가시어머니' 등), (2)북한 특유의 조어법을 보이는 단어('물들성, 걸그림' 등) (3)표기 규범이 달라 차이 나는 것('로인, 녀성' 등)으로 나눌 수 있다.

북한어라고 판단된 표제어는 뜻풀이를 따로 하지 않고 우리의 표준어와 대응시킨다. 우리 표준어와 대응하지 않는 단어는 독자적인 뜻풀이를 한다. 이 때 표기 규범이 우리와 다르면 뜻풀이에서 바른 어형을 제시한다.

3. 맺음말

지금까지 표제어 선정 과정에서 나타나는 몇 가지 문제점들을 유형별로 살펴보았다. 그러나 실제 표제어 선정 과정에서는 위에서 제시한 것 말고도 많은 문제점들이 있다. 예를 들어서 표제어의 본질 문제나 동음어/다의어의 구분에 대한 문제는 위에서 전혀 언급하지 않았지만 계속 문제가 되는 것들이다.

표제어의 선정은 기존 사전에 있는 각각의 표제어에 대해 위와 같은 절차를 거쳐서 일차적인 목록이 마련된다. 그러나 기존 사전의 표제어는 일차적인 검토 목록이 될 뿐이며 기존 사전에 등재되지 않은 어휘에 대해서도 지속적인 조사가 필요하다. 현재는 용례 수집을 위한 문헌 입력 자료로부터 기존 사전에 실려 있지 않은 미등재어를 수집하는 일을 광범위하게 하고 있다. 이렇게 수집된 목록에 대하여 또다시 새 사전 표제어로 적당한지 심사를 해야 할 것이다. 그러나 실제로 표제어 목록을 검토해 보면 위에 제시한 몇 가지 기준과 선정 지침에도 불구하고 선정자 개인이 판단하여 결정하기 어려운 표제어들이 있다. 현재는 이들을 문제 표제어로 분류하여 문제의 유형에 따라 각각 다른 구별 표지를 해서 따로 전산 관리하고 있다. 한편으로는 이러한 문제 표제어를 처리하기 위한 표준안을 마련하기 위한 연구도 병행하여 진행하고 있다.

그러나 표제어의 선정이 전체 목록을 한 번만 검토하고 마는 일회적인 작업은 아니다. 사전의 거시구조가 일관성 있는 체계를 이루기 위해서는 지속적인 검토가 필요하다. 따라서 특정 어휘를 사전 표제어로 실을 것인지의 문제는 사전이 나오기 직전까지 지속적으로 검토되어야 할 것이다.