

# 21세기 세종계획 관용표현 전자사전 구축에 대하여

박만규<sup>0</sup> 이선웅 나윤희 이광호  
관동대학교 서울대학교 서울대학교 서울대학교  
mgpak@mail.kwandong.ac.kr anhhung@chollian.net  
younhee420@hanmail.net cathol7@hanmail.net

## On the development of a computational lexical database of idiomatic expressions in the framework of 21st Sejong Project

Man-Ghyu Pak Sun-Woong Yi Yun-Hee Na Kwang-Ho Lee  
Dept. of French Linguistics & Literature, Kwandong Univ.  
Dept. of Korean Linguistics & Literature, Seoul National Univ.  
Dept. of French Linguistics & Literature, Seoul National Univ  
Dept. of Korean Linguistics & Literature, Seoul National Univ.

### 요 약

본고는 올해 처음 시도하는 세종계획 관용표현 전자사전 구축에 관한 글이다. 본 전자사전이 완성되면 관용표현의·총체적 정보(형태, 통사, 의미, 화용 정보)를 수록하는 최초의 업적이 될 뿐만 아니라 실제 언어 자료에서 흔히 볼 수 있는 관습적 표현까지 모두 포괄하는 4만 표제어의 대규모 사전이 될 것이다. 본 사전에서는 관용표현의 형태·통사적 구성과 그 분포적 속성뿐 아니라, 관용표현이 가지는 논항의 존재 유무, 구조, 조사 통합 양상, 그리고 고정명사에 대한 수식어 제약, 어휘적·통사적 변형 양상, 선어말어미 제약, 어말어미 제약, 문장 유형 제약 등이 수록된다. 또한 각 논항의 의미역과 선택제약에 관한 정보, 그 외 다양한 의미·화용 정보, 어원·표기 정보 등도 담기게 된다. 본고에서는 그러한 정보의 표기 양식을 하나하나 명시적으로 설명할 것이다.

### 1. 머리말

본 논문은 [21세기 세종계획 전자사전] 개발의 틀 내에서 올해부터 구축하기 시작한 관용표현 전자사전의 언어학적·사전학적 의의와 특성을 분석·제시함을 목적으로 쓰여진다. 본 사전은 우선 그 규모에 있어 4만 표제어에 이르는 대규모 범용 전자사전이라는 성격을 갖는데, 이는 현대 한국어 사용자의 관용표현에 관한 언어학적 지식을 총체적으로 표상하기 위함이다. 또한 본 사전은 음운, 형태, 통사, 의미, 화용 정보 등 모든 언어적 속성을 모두 낱말이 체계적으로 정밀하게 기록하려는 시도라는 특징을 갖는다. 현재 국어의 관용표현을 담은 본격적인 사전이 드물다는 점, 더구나 전자사전의 형태로 구축된 것은 전무한 실정이라는 점을 고려하면 그 시도 자체가 획기적이라고 할 수 있다.

본 사전의 또 하나의 특징은 특정 이론에 편중되지 않고 이론 중립적인 성격을 갖는다는 점이다. 이는 다양한

자연어 처리 프로그램에 지원이 가능하도록 하는 실용적 목적 이외에도, 본 사전의 자료가 논쟁적인 문제들에 대해 다양한 이론적 검증의 수단 내지 자료체가 되게 하기 위함이다. 이러한 목적으로 인해 비록 언어학적으로 엄밀한 의미의 관용표현이 아니더라도 한국어에서 관습적으로 고정된 표현(conventionalized expression)을 될 수 있는 한 많이 포함시킬 것이다. 다른 범주의 사전(예를 들어 명사 사전, 용언 사전 등)에서 용례 제시라는 암시적 방식으로밖에 제시될 수 없는 이 같은 부류에 관한 언어학적 기술이 제외되는 일을 막기 위해서이다.

### 2. 예비적 고찰

#### 2.1. 정의

관용표현<sup>1)</sup>에 대한 기존의 일반적인 정의는 "둘 이상

1) 기존 연구에서는 관용어, 관용구, 관용어구, 속어, 익은말, 익힘말 등이 쓰여 왔으나 세종계획 전자사전에서는 관용표현으로 통일하여 사용한다.

의 단어가 필수 공기 관계에 놓이면서, 의미상 그 결합 요소들 각개 의미의 단순한 합이 아닌 제3의 의미를 지니게 된 특수 표현" 정도로 요약된다. 다시 말해 통사적 고정성과 의미적 불투명성을 필수적으로 갖고 있어야 관용표현으로 인정될 수 있다. 그러나 때때로 특정한 화용적 상황 하에서 형식적으로는 고정적이지만 의미적으로는 그다지 불투명하지 않은 표현도 많이 사용되는데, 광의적인 관점에서 보면 이렇게 관습적으로 굳어진 표현까지 포괄할 수 있다. 사실 이론적으로 후자의 의미로 관용표현을 정의한 논의는 아직 찾아볼 수 없으나, 전자사전에 이러한 표현의 의미 정보를 수록하지 않으면 언어 자료의 생성(예컨대 영한 번역기에서의 한국어 문장 생성)은 상당한 정도로 제약될 수밖에 없다.

예를 들어 다음과 같은 흔히 쓰는 표현을 보자.

- (1) 가. 사람 살려!, 맹견 주의, 일동 기립!  
 나. {어디가 어디인지, 뭐가 뭔지} 모르겠다, 알다가도 모르겠다  
 다. 이번 사건을 {중심으로, 구실로}, 이와 {함께, 별도로}

(1가)는 일반적인 관용표현의 '의미적 불투명성'을 거의 갖고 있지 않다. 이는 일종의 기호의 언어화라고 할 수 있을 것이다. 이를테면 소변(小便)을 보지 말라는 곳에 '가워'가 그려져 있는 것과 같은 것이다. 곧 어떤 전형적 상황에서 반드시 쓰이는 관습적 표현인바, 이러한 표현의 의미 정보 역시 언어 생성의 관점에서 관용표현 전자사전에서 다루어 줄 필요가 있는 것이다. (1나)는 (1가)보다는 의미가 불투명한 경우이다. 앞에 있는 '어디'와 '뭐'는 '내가 있는/가는 곳', '내가 닦친 상황/사물' 정도의 뜻을 갖는데, 그 뜻을 밝혀 말하는 일은 거의 없고 주로 '어디가 어디인지, 뭐가 뭔지'와 같은 전형적인 고정 표현을 사용한다. (1다)는 '중심으로/구실로/함께/별로로'라는 부사어가 논항을 요구하는 특이한 경우로서 이론적으로 이 표현 뒤에 '하여' 정도의 용언을 보충할 수 있다고 하여도 (1다) 표현 자체의 독자성을 보장해 주지는 못한다. 일반적으로 어떤 사전이든 용언과 보문 명사에만 논항 정보가 수록된다는 점을 고려할 때, 이번 관용표현 전자사전에서의 기술은 실제적인 정보 제시를 가장 중시한다고 할 수 있다.

## 2.2. 수록 항목 설정

국내의 주요 사전(국립국어연구원 편 1992, 한글학회

편 1992, 김민수 외 3인 편 1992), 관용어 사전(박영준·최경봉 1995, 조선어학과연구실 1984)에서의 모든 관용표현 등재어, 그리고 세종계획 전자사전 개발의 총책임자인 홍재성 교수가 개인적으로 작성한 목록을 1차 대상으로 삼고 '장난이 아니다, 게임이 안 되다, 골프의 골 자도 모른다, 졌다 졌어' 등의 미등재 관용표현을 최대한 수집하여 수록한다. 광의의 관점에서의 관용표현 항목은 3년 동안 각종 텍스트에서 수집한다.

여기서 속담, 연어, 합성어와 구별하여 항목을 설정하는 몇 가지 기준에 대하여 언급할 필요가 있다. 관용표현과 가까운 다른 언어 형식으로 첫 번째로 들 수 있는 것은 연어(連語: collocation)인데, 연어는 적어도 어느 한 구성 요소는 자신의 중심 의미를 그대로 지니고 있는 반면에 관용표현은 모든 구성 요소의 합이 새로운 의미를 창조하기 때문에 어떤 요소도 자신의 중심 의미를 온전히 갖고 있는 일이 없다는 점에서 두 개념을 구분할 수 있다.

관용표현과 가까운 언어 형식으로 두 번째로 들 수 있는 것은 속담(俗談)이다. 속담은 많은 경우 관용표현과 구분이 용이하지 않을 만큼 관용표현과 통사·의미적으로 공통성을 갖고 있지만, 일반적으로 다음과 같은 속성에서 구분이 가능하다.

우선 속담은 의미적으로 자족적인 구조, 즉 스스로 완전한 구조를 가지고 있다. 따라서 논항을 가지지 않는다. 또한 속담은 전체가 완전히 고정되어 있는 언어형식이라고 할 수 있다. 따라서 통사적 제약에 있어 속담은 관용표현보다 더욱 많은 제약을 가지고 있다. 예컨대 수동화, 의문화, 대명사화, 분열문화, 어순 이동(permutation), 부정 등의 조작을 모두 불허한다. 속담에서 가장 많이 사용되는 시제는 현재인데 이른바 초시 간접 현재라고 불리우는 것이다. 이는 속담이 담고 있는 충칭성과의 관련을 맺고 있는 것으로 따라서 시제적 변이를 허용하지 않는다. 반면에 관용표현의 경우에는 이러한 제약이 있는 것은 아니다. 한편 속담이 전하는 의미 내용은 대체적으로, 보편적 진리 내지 현실 및 세태의 풍자이거나 또는 윤리적·교훈적 성격이 강한 것들이다. 특히 수사학적 관점에서 보면 은유가 주로 쓰이고, 의인법, 반복법, 대조법, 과장법 등도 많이 나타난다.

관용표현과 관련된 언어 형식으로 세 번째로 들 수 있는 것은 합성어(合成語)이다. 둘 이상의 언어적 요소가 결합하여 제3의 의미를 만들 수 있는 합성어가 관용표현

과 구별되는 이유는 합성어가 단어의 자격으로 어휘부에 등재된 요소로서 내적 비분리성, 변형 불가능성 등 단어가 가지는 일반적인 특성을 갖고 있기 때문이다. 예컨대 '꿀맛', '독 안에 든 쥐'나 '코가 비뚤어지게'와 같은 표현은 관용표현의 서술어가 어미 제약을 갖는 경우라고 볼 수도 있으나 변형이 불가능한 고정적인 표현이면서도 일반적인 명사와 부사로 쓰일 수 있다는 점에서 각각 합성명사와 합성부사로 봄이 더 타당하다(띄어쓰기는 관행적인 것으로 간주한다). 그러나 단어와 구의 명백한 판별 기준이 아직 알려져 있지 않기 때문에 합성어와 관용표현 사이의 경계선도 불분명한 면이 없지 않다.

언어 형식 분석의 관점으로는 그 결합력에 따라 '합성어>관용표현>기능동사>연어>일반자유구문'으로 볼 수 있는데, 관용표현 전자사전에서는 각 경계선을 엄밀하게 구획하는 일보다는 전자사전의 범용성, 실용성을 우선하여 다소의 잉여적 수록을 허용할 것이다. 다시 말해 이론적으로는 관용표현과 구분되는 연어, 합성어, 속담 등도 필요에 따라서는 관용표현 사전에 잉여적으로 수록될 수 있다는 뜻이다. 이를테면 어느 쪽에 속해 있는지 판단이 불분명한 경우에는 혹시 합성명사 전자사전, 합성부사 전자사전, 연어 전자사전, 속담(아직 세종계획 전자사전 작업에 들어 있지 않다)에 들어갈 수 있는 항목이라도 관용표현 사전에서 다룰 것이다.

### 3. 관용표현 전자사전에 담길 주요 정보2)

#### 3.1. 구조적 속성

관용표현의 구조적 속성에 관해서는 우선 관용표현의 내부구조가 형태·통사론적으로 제시되어야 할 것이다. 또 관용표현은 그 전체가 하나의 언어 형식으로서 논항을 요구하는데, 전자사전에는 논항의 존재 유무, 논항의 형식(조사 통합 양상), 논항의 수, 논항의 구조(명사구/보문), 논항의 생략 가능성 등에 관한 정보 역시 수록되어야 한다. 다음은 그 예이다.

2) 아래 3.1.-3.6.에서 제시된 정보 이외에 중요한 정보로는 어원 정보, 준말/분뒤말 정보, 참조 정보 등을 들 수 있다. 어원 정보는 관용표현 구성 요소의 언어 종류를 밝히는 것으로서 '말문이 막히다'는 '<org>C0=ko+si(말뭉치)'로, '히트를 치다'는 '<org>C0=eng(hit)'로 표시된다. 준말 정보로서 '미역을 감다'는 '<var type="abb">먹을 감다</var>'와 같이 표시한다. 참조 정보는 같은 형태의 연어가 있음을 보여 주는 것으로서 예컨대 관용표현 '눈을 뜨다'는 언어로서도 존재하므로 '<see>col</see>'와 같은 정보를 준다.

- (2) 가. 형태적 정보: '몸에 배다'→  
 <idmMTag>몸/nn+에/jkb 배/vgi+다/eft</idmMTag><sup>3)4)</sup>  
 나. 표준 기본 구조: '눈이 벌췌다'→  
 <idmStr>C0 A</idmStr><sup>5)</sup>  
 다. 논항 표시:  
 (i) '몸에 배다'→  
 <caseFrame>N0 N1-이 Id</caseFrame>  
 (너는 그런 행동이 몸에 배었구나)  
 (ii) '목이 날아가다'→  
 <caseFrame>N0 Id</caseFrame>  
 (그 사형수가 목이 날아가다)  
 <caseFrame>N1-의 Id</caseFrame>  
 (그 사형수의 목이 날아가다)  
 (iii) '바람을 넣다'→  
 <caseFrame>N0 N1-에게 S-라고|자고 Id</caseFrame>  
 (나는 그에게 술집에 가자고 바람을 넣었다)  
 (iv) '꿈 깨'→<caseFrame>Id</caseFrame>(논항 없음)  
 (v) '꺼백 죽다'→  
 <caseFrame n=1>N0 N1 앞에서는 Id</caseFrame>  
 <caseFrame n=2>N0 N1-라면 Id</caseFrame>  
 <caseFrame n=3>N0 S1-면 Id</caseFrame>  
 라. 조사 삭제 정보:  
 (i) '몸을 더럽히다'→  
 <var type="JK\_drop">JK1=0</var>  
 (대격조사 삭제 불가)  
 (ii) '들통이 나다'→  
 <var type="JK\_drop">JK0=1</var>  
 (주격조사 삭제 가능)  
 마. 조사 변이 정보: '가슴이 찢리다'→  
 <var type="JK\_var">JK0=에</var>  
 (조사 '이'는 '에'로 교체 가능)

#### 3.2. 분포적 속성

(2나)의 기본 구조 제시가 관용표현의 하위범주, 즉 통사적 유형을 지적함이 주목적이라면, 이는 관용표현 내부구조의 각 어휘요소의 분포를 기술하기 위한 것이다. 관용표현의 중심어인 고정명사와 동사, 형용사, 부사 및 부정어의 분포를 지적함으로써, 예컨대 우리말에서 특정 고정명사를 기반으로 한 관용표현에 어떠한 것

- 3) 세종계획 전자사전에는 기술 대상이 되는 모든 범주에 형식화된 태그를 붙임으로써 용이한 전산 처리를 도모하고 있다. <idmMTag> 태그는 관용표현 내부의 형태론적 구성 정보를 앞뒤에 붙이는 것이다.  
 4) nn(보통명사), jkb(부사격 조사), vgi(자동사), eft(어말어미) 등은 세종계획 전자사전에서의 문법범주 약호이다.  
 5) C는 관용표현 내부의 상수항(고정명사구: Constant)를 뜻하고, 뒤의 지수는 문장 성분(여기서는 주어=0)을 보이기 위해 붙인 것이다. A는 형용사이다.

이 있는지를 검색할 수 있게 하고, 부정어를 포함한 관용표현이 전체 관용표현에서 차지하고 있는 비중 등을 알 수 있게 할 것이다.

아울러 관용표현의 고정명사구를 수식하는 한정어의 제약 정보도 수록한다.

(3) 가. 관용표현 내부구조의 어휘 분포 :

(i) '말문이 막히다'→

<idmStrDis>C0=말문, V=막히다</idmStrDis>

(ii) '숨이 붙어 있다'→

<idmStrDis>C0=숨, V=붙다, Aux=있다</idmStrDis>

(iii) '밤낮을 가리지 않다'→

<idmStrDis>C1=밤낮, V=가리다, Neg=않다</idmStrDis>

나. 논항의 어휘 분포:

'답을 쌓다'→

<selRst>N0=인간 N1=인간|추상</selRst>

(철수는 공부와는 답을 쌓았다),

<selRst>N0=인간 N1>S것</selRst>

(철수는 공부하는 것과는 답을 쌓았다)

다. 한정어 제약:

'명예를 메다/지다'→

<CMd>무거운 C0</CMd>

('명예'를 수식하는 말로 '무거운'밖에 올 수 없다)

### 3.3. 변형적 속성

관용표현의 변형적 속성에는 자동-타동, 능동-피동, 주동-사동 용법 사이의 대응과 같은 구문 사이의 대응관계와 관계화 구성, 명령문으로의 변형 가능성 등이 있다.

(4) 가. 구문 사이의 대응

(i) 피동사 대응: '틀을 잡다'→

<pasStr>틀이 잡히다</pasStr>

('틀이 잡히다'가 가능)

'발품을 팔다'→

<pasStr>0</pasStr>

('발품이 팔리다'가 불가능)

(ii) 능동사 대응: '씨알도 안 먹히다'→

<actStr>0</actStr>

('씨알도 안 먹다'가 불가능)

(iii) 사동사 대응: '손발이 맞다'→

<causStr>손발을 맞추다</causStr>

('손발을 맞추다'가 가능)

(iv) 주동사 대응: '시간을 죽이다'→

<uncausStr>0</uncausStr>

('시간이 죽다'가 불가능)

(v) 자동사 대응: '꼬리를 드러내다'→

<intStr>꼬리가 드러나다</intStr>

('드러나다'는 '드러내다'의 대응 자동사)

(vi) 타동사 대응: '땅에 떨어지다'→

<traStr>땅에 떨어뜨리다</traStr>

('떨어뜨리다'는 '떨어지다'의 대응 자동사)

나. 관계화 구성 제약: '몽니가 굶다'→

<GrInfo type="rel">0</GrInfo>

('굶은 몽니'가 불가능)

다. 명령문으로의 변형 제약: '스타일을 구기다'→

<GrInfo type="Imp">0</GrInfo>

(명령형 '스타일을 구겨라'가 불가능)

### 3.4. 형태론적 제약

형태론적 제약에 관해서는 시제, 상, 서법 등의 선어말 어미 제약과 어말 어미 제약에 관한 정보가 수록된다.

(5) 가. 시제, 상, 서법 등의 선어말 어미 제약:

(i) '그릇 깨다'→<VRst>꺾</VRst>

('그릇 깨겠다' 형태만 가능)

(ii) '지다'→<VRst>엇</VRst>

('졌다' 혹은 '졌다 겹어' 형태만 가능)

나. 어말 어미 제약:

(i) '같은 값이다'→<VEndRst>면라면</VEndRst>

('같은 {값이면, 값이라면}' 형태만 가능)

(ii) '밤낮을 가리지 않다'→

<VEndRst>tend=고</VEndRst>

(주로 '밤낮을 가리지 않고'로 쓰임)

(iii) '놀고 있다'→<VEndRst>네</VEndRst>

('놀고 있네' 형태만 가능)

(iv) '머리에 피도 안 마르다'→

<VEndRst>ㄴ</VEndRst>

('머리에 피도 안 마른' 형태만 가능)

### 3.5. 문장의 유형과 관련된 정보

문장의 유형과 관련되어서는 명령문, 의문문으로만 쓰이는 관용표현임을 밝혀 주는 일과 의미적 부정의 문맥, 통사적 부정의 문맥을 밝혀 주는 일이 주된 작업이다. 이론적으로 말하면 명령문, 의문문으로만 쓰이는 관용표현임을 밝혀 주는 일은 앞서 언급한 어말 어미 제약에 포함된다. 그러나 관용표현 전자사전에서는 명시성(expliteness)을 높이기 위해 문장의 유형과 관련된 정보는 따로 밝혀 준다. 한편 의미적 부정의 문맥 정보는 관용표현의 형식이 어떻게 변형되든 언제나 부정적인 의미를 담기 위한 것임을 보이는 것으로서 실제 상당수의 관용표현에 적용될 수 있는 매우 중요한 정보이다. 그러나 통사적 부정의 문맥 정보는 예컨대 '사족을 못 쓰다'에서처럼 통사적으로도 언제나 부정문 형식을 유지하여야 하므로 부정의 형태로 제시된 표제어에 별도의 태그

를 사용하여 정보를 추가할 필요는 없다. 이미 표제어에 반영되어 있기 때문이다.

(6) 가. 명령문, 의문문 제약:

- (i) '냉수 먹고 속 차리다'→  
<SMod>명령문</SMod>  
('냉수 먹고 속 차려'와 같은 명령형으로만 쓰임)
  - (ii) '이걸 누구 코에 바르다'→  
<SMod>의문문</SMod>  
('이걸 누구 코에 바르냐?'처럼 의문형으로만 쓰임)
- 나. 의미적 부정의 문맥:
- (i) '씨알도 먹히지 않다'→  
<SVar>씨알도 안 먹히다<씨알이나 먹히겠  
나?</SVar>
  - (ii) '코빼기도 보이지 않다'→  
<SVar>코빼기도 못 보다<코빼기나 봤니?코빼기라  
도 보았으면 좋겠다</SVar>

3.6. 의미 정보

의미 정보로는 관용표현의 명사구 논항의 의미역 정보와 선택 제약 정보, 그리고 관용표현의 내포 의미, 레지스터에 관한 정보, 관용표현의 동의/반의 표현 정보, 강조 표현 정보 등이 수록된다. 그 외에도 반어적으로만 쓰이는 표현으로서만 관용구로 성립되는 화용적 제약이나 특수한 화용적 상황에서만 관용구로 성립되는 경우에 관한 정보를 수록한다. 아울러 전문어 영역 정보, 영어 대역 정보도 지적해 둔다.

(7) 가. 논항의 의미역 정보:

'들통이 나다'→  
<argStr>N0=THM, N1=EXP</argStr>  
(그 일이 철수에게 들통이 났다)

나. 논항의 선택 제약 정보:

'스타일을 구기다'→  
<selRst>N0=인물, N1=추상</selRst>  
(철수가 그 일로 스타일을 구겼다),  
'보통이 넘다'→  
<selRst>N0=인물|추상</selRst>  
(철수는 보통이 넘는다, 철수의 바둑 실력은 보통  
이 넘는다)

다. 내포 의미 정보:

'입을 놀리다'→  
<con>비하</con>,  
'혈안이 되다'→  
<con>경멸</con>

라. 레지스터 정보:

'안광이 지배를 철하다'→  
<reg>문어</reg>

마. 영어 대역 정보:

'머리가 돌다'→

<trans>be out of one's mind|be off one's  
head|be wrong in the head</trans>

바. 전문어 영역 정보:

'악화가 양화를 구축하다'→  
<domain>경제</domain>,  
'모자를 씌우다'→  
<domain>바둑</domain>

사. 동의/반의 표현 정보:

'포문을 열다'→  
<lr type="syn">물꼬를 트다</lr type>,  
'운전대를 잡다'→<lr type="ant">  
운전대를 놓다</lr type>

아. 기타 관련 표현 정보:

'귀에 거슬리다'→  
<lr type="rel">눈에 거슬리다</lr>

자. 강조 표현 정보:

'서슬이 퍼렇다'→  
<var type="emph">서슬이 시퍼렇다</var type=  
"emph">

차. 비속어 정보:

'입을 다물다'→  
<var type="vulg">추둥이를 다물다</var>

카. 화용적 제약:

너 잘났다, 꼴 좋다→  
<intention>반어</intention>,  
잘 먹고 잘 살아라→  
<intention>저주</intention>,  
두고 보자→  
<intention>악의적 다짐</intention>

3.7. 붙임 : 관용표현 전자사전의 기술 구조

<superEntry> - 최상위 표제어

<entry> - 표제어

<mntGrp> - 관리 정보 그룹

<dic> - 사전 이름

<wrt> - 작성자

<crd> - 최초 작성일

<mdd> - 최종 수정일

<note> - 메모

<headGrp> - 표제어 정보 그룹

<orth> - 표제어 표기 정보

<org> - 어원 정보

<idmType> - 관용표현 유형 정보

<see> - 참조

<note> - 메모

<cnstGrp> - 구성 정보 그룹

<idmMTag> - 관용표현 형태 분석 정보

<idmStr> - 관용표현 내부구조

<idmStrDis> - 관용표현 내부구조 분포 정보

<var type="emph"> - 강조형 정보

<var type="vulg"> - 비속어형 정보

<var type="abb"> - 생략형 정보

- <var type="hon"> - 존대형 정보
- <var type="JK\_drop"> - 격조사 생략 정보
- <var type="JK\_var"> - 조사 변이 정보
- <var type=etc> - 기타 변이 정보
- <sense n="1"> - 의미 갈래
- <semGrp> - 의미 정보 그룹
  - <trans> - 영어 대역 표현
  - <domain> - 의미 영역(전문 분야)
  - <reg> - 레지스터
  - <con> - 내포적 의미
  - <intention> - 의도적 의미
  - <lr type="syn"> - 동의적 표현
  - <lr type="ant"> - 반의적 표현
  - <lr type="rel"> - 기타 관련 표현
- <synGrp> - 통사 정보 그룹
  - <synt\_a> 통사 정보 A 구획
    - <frame n="x"> - 통사적 구성의 갈래
      - <caseFrame> - 격틀 정보
      - <synSem> - 통사-의미관계
        - <selRst> - 선택 제약
        - <thtRol> - 논항 구조
        - <eg> - 예문
        - <note> - 메모
    - <synt\_b> 통사 정보 B 구획
      - <SMod> - 문장의 유형
      - <SVar> - 통사적 변이형 정보
      - <CMd> - 수식어 제약
      - <VRst> - 용언의 시제, 상, 서법의 제약
      - <VendRst> - 용언의 어말 어미 제약
      - <synRel> - 통사적 관계
        - <pasStr> - 피동사 대응
        - <actStr> - 능동사 대응
        - <causStr> - 사동사 대응
        - <uncausStr> - 주동사 대응
        - <intStr> - 자동사 대응
        - <traStr> - 타동사 대응
        - <GrInfo> - 관계화 제약, 명령문 제약

등의 각종 문법 정보

#### 4. 남은 문제

이상으로 세종계획 관용표현 전자사전에서의 수록 정보를 실례와 더불어 간단히 설명하였다. 위에서는 사전적 기술이 비교적 쉬운 예들을 중심으로 설명하였지만, 현재까지의 작업에서는 미해결된 문제가 적지 않다. 이는 관용표현 형식의 다양성에 1차적으로 원인이 있을 것이다. 세종계획 전자사전에서는 의미 기술을 치밀하게 하면서도 형식적으로도 체계를 갖추도록 노력하는바, 관용표현에서의 미해결된 문제들은 주로 현상에 대한 이해의 어려움에서 비롯된 것이 아니라 형식적 기술에서의 어려움에서 비롯된 것이 대부분이다. 다음에 몇 가지 문

제를 지적하는 것으로 글을 마감하고자 한다.

#### (㉠) 보다 풍부한 화용적 의미 정보 수록의 문제

(㉡) 평서형으로 끝나는 일이 없는 관용표현을 평서형의 표제어 표기 문제: 예컨대 "{그런 법이/이러는 게/그렇다고 그냥 가는 게} 어디 있어?"라고 할 때의 '어디 있어?'가 현행 방식으로는 '어디 있다'로 제시되어야 한다. 그렇다고 '어디 있어'로 제시하는 것도 그리 쉽지 않다. '어디 있냐?', '어디 있니?', '어디 있습니까?' 등의 표현을 보면 알 수 있듯이 표제어에서 상대경어법 어미의 선택 문제 역시 간단한 문제는 아니기 때문이다.

(㉢) '국수를 먹다'는 "내가 국수 먹었다."와 같이 결혼식 하객이 주어가 될 경우는 어색한 표현이 되고, "국수 좀 먹자.", "국수 좀 먹여 줘.", "언제 국수 먹게 해 줄래?"처럼 '결혼해라'의 의미로 주변인들이 쓸 때 자연스런 표현이 되는데, 그렇다고 '국수를 먹이다'를 표제어로 할 수는 없으므로 통사 정보를 주기가 상당히 까다롭다.

(㉣) 고정명사에 'N-의', 'N-라는', 'N' 등의 수식어가 필수적인 경우의 격틀 제시 방식이 문제가 된다. 예컨대 '{충각, 학생, 실업자} 딱지를 떼다'와 같은 말의 격틀을 'N0 N1-의 Id'로 표시하면 N1의 의미역을 주는 곳에서 곤란을 겪게 된다. 가능한 대안은 'N 딱지를 떼다'처럼 표제어 자체에 수식어 정보를 주는 방법이다. 이 경우 관용표현 내부구조분포 정보는 '<idmStrDis>N=인물, C1=꿈, V=깨다/<idmStrDis>' 정도로 되겠지만, 형식적으로는 이해적이라고 할 수 있다.

(㉤) '눈에 흙이 들어가기 전에'와 같은 말은 어미 제약도 아니므로 형식적으로 보여 주기 어렵다. 또 '{내, 네, 그 사람} 눈에 흙이 들어가기 전에' 등이 가능하므로 합성부사로 보기도 어렵다. 전자사전에서는 관용표현의 기본 유형을 용언(동사, 형용사, 이다, 아니다, 있다, 없다 포함)의 '-다'로 끝나는 것으로 제한했기 때문이다.

(㉥) 관용표현의 경우 적절한 영어 대역 표현을 찾기가 매우 어렵다.

#### 5. 참고 문헌

- [1] 국립국어연구원 편(1999), 표준국어대사전, 두산동아.
- [2] 김민수·고영근·임홍빈·이승재 편(1992), 국어대

사전, 금성출판사.

- [3] 문금현(1996), 국어의 관용표현 연구, 서울대 박사 학위논문.
- [4] 박영준·최경봉 편(1995), 관용어 사전, 태학사.
- [5] 이기문(1980), (개정판)《속담사전》, 일조각.
- [6] 조선어학과연구실 편(1994), 조선어 관용구집(상·하), 동경외대 어학교육연구협의회.
- [7] 한글학회 편(1992), 우리말근사전, 어문각.
- [8] 홍재성(1993), 먹다 속어동사구문의 통사적 기술, 어학연구(서울대 어학연구소) 29권 3호.