

세종 의존명사/대명사/수사 전자사전의 정보표상 구조

방성원 경희대학교 국어국문학과
호정은 경희대학교 국어국문학과
김종인 서울대학교 불어불문학과
jikim@inisolution.com

The Representational Structure of Lexical Informations of Korean non-autonomous nouns in the Sejong Electronic Dictionary

Seong-Won Bang, Kyunghee Univ.
Jeong-Eun Ho, Kyunghee Univ.
Jong-In Kim, Seoul National Univ.

요 약

세종전자사전이 궁극적으로 범용전자사전을 지향한다는 점에 비추어 볼 때, 텍스트 자동 분석과 생성, 정보 검색 및 자동 번역 등에 활용될 데이터베이스로서의 전자사전은 자연 언어 어휘의 내적 구조와 기능방식에 관한 정보들, 가령 음운·통사·의미·화용적 가치와 실현 조건 등에 관한 정보들을 체계적이고도 정교하게 담고 있어야만 한다.

의존명사, 대명사, 수사 범주에 속하는 언어 단위들은 단일 명사와 구별되는 어휘·통사적 속성들을 지니며, 사전의 기술 구조에는 그 정보 값들을 체계적으로 명시화할 수 있는 정보 항목과 표상 구조가 설정되어야 한다. 가령 의존명사처럼 통사·의미적 자율성을 지니지 않는 언어 요소의 경우, 어휘 관계 정보보다는 인접하는 여타 언어 단위들과의 호응관계나 결합제약 조건들이 더 중요한 정보일 수 있다. 본 사전이 체언사전의 하위사전으로 별도로 구축되는 것은 단일어 사전에서 그러한 정보들을 효과적으로 표상하기 어렵기 때문이다. 그러나 본 사전은 실제적으로는 체언사전에 통합되어 운영된다는 점에서 이중적 지위를 누린다고 하겠다.

1. 서론

세종전자사전의 핵심적인 목표는 현대 한국어 어휘부의 언어정보를 상세하고도 체계적으로 기술하는 것이다. 이 야심적인 목표는 사전 구축의 지향성에 관한 선언적인 차원을 넘어 사전의 거시구조 설계와 미시정보 기술 방식을 결정짓는 바탕이라 할 수 있다. 이미 알려져 있듯이 세종전자사전은 실질적으로 품사별 하위사전들로 구축되고 있다. 이 하위사전 분화라는 선택은 바로 어휘부 단위들 간에 정보의 내재적 특성이 상이하기 때문에 이루어진 것이다. '의존명사/대명사/수사' 사전이 체언사전의 하위사전으로 별도로 구축되는 것도 그러한 이유에서이다.

위의 언어 범주들은 지금까지의 사전 기술의 관례상 단일어 기술의 모형에 맞추어 그 정보값이 표상되었으며, 그러한 기술 방식이 이 범주들의 언어적 속성을 충분히 담을 수 없었다는 점은 짐작하기에 어렵지 않다. 비록 단일한 품사 범주로 분류되는 것은 아니지만 위의 범주들은 다른 요소들에 의존하거나 호응, 또는 공기하는 특성을 지니며, 그런 점에서 자립적인 단일어와 구별되는 공통점을 갖는다. 요컨대, '의존명사/대명사/수사' 사전은 이 범주들이 갖는 요소들 간의 결합 제약 및 관계 정보를 체계적으로 표상하기 위한 목적에서 설정된 하위사전으로 정의할 수 있다.

2. '의존명사/대명사/수사' 사전 구축의 기본 방향

1) 세종전자사전은 대규모의 범용적 어휘부 데이터베이스로 요약될 수 있다. 세종사전의 이러한 성격은 무엇보다도 어휘부 정보 기술상의 상세함과 체계성으로 대변된다. 이것은 세종전자사전이 사전의 거시논리구조 설계상에서나 정보의 미시기술에 있어서 현대 언어학의 연구 성과를 반영하고자 한다는 것을 의미한다. '정보 처리는 근본적으로 언어 처리'라는 관점에서 보면, 가령 인간의 언어 조작 원리를 반영하는 정보처리기를 구축하려고 할 때 가장 시급히 충족되어야 할 과제들 중 하나는 어휘부의 내재적 지식에 대한 총체적이고 정밀한 기술이라 할 수 있다.

2) '의존명사/대명사/수사' 사전의 구축 작업에서 어휘의 언어 정보를 상세화하고 체계적으로 표상하는 것만큼 중요하게 고려한 것은 거꾸로 '언어처리는 결국 정보처리를 거친다'라는 전산처리의 관점이다. 이는 전자사전의 응용성과 적용성에 관련된 실천적 문제이자 세종사전의 전체 구조의 일관성에도 관련된 문제이기 때문이다. 우리가 가령 품사 지위의 설정에서부터 정보 항목 설정 및 표상화 방식에 이르기까지 관련 분과와의 협의 및 조정을 거친 것은 그러한 이유에서이다.

3) 요컨대, '의존명사/대명사/수사' 사전 분과 역시 타 분과들과 마찬가지로 현대 언어학의 이론적 요구와 효율적인 전산처리의 요구를 동시에 반영해야 한다는 이중의 과제를 안고 있다고 할 수 있다. 가령, "계단-식, 현대-식", "연구-차, 인사-차" 등에서 '식'과 '차'의 품사범주를 설정하는 경우에, 어근/접사 사전분과와의 협의를 거쳐 접사 범주 대신 의존명사 지위를 부여하고 상기 표현들을 합성어로 처리한 것이 그러한 예이다. 이는 기술대상이 되는 표제어를 목록화하는 분과간 협의 과정에서 제기된 문제로서, 접사 목록을 최소화하는 것이 처리적 관점에서 유리하다는 판단을 받아들여 결과적으로는 전통적인 언어학적 판단을 유보한 실례이다. 왜냐하면 접사는 결합 생산성이 매우 높아서 처리 부담이 증가할 수 있기 때문이다. 이는 일반 인쇄사전에서의 기술 방식과 큰 차이를 나타내는 좋은 예이다.

3. 의존명사의 사전 처리 방안

본 사전 작업에서는 의존명사를 일반 의존명사(이하 '의존명사')와 단위성 의존명사(이하 '분류사')로 하위 분류하고, 각각 별도의 하위 품사 범주로 다루어 두 범주 각각의 고유한 어휘·통사적 특성이 정보 항목과 표상 구조를 통해 분명하게 나타나도록 하였다.

여기에서는 의존명사 목록을 선정한 기준과 실제 목록을 제시하고, 의존명사의 통사적 특징을 반영해 주는 사전 논리 구조 및 정보 항목의 특징을 보여주고자 한다.

3.1 목록 선정 기준 및 규모

사전에 등재될 표제어들은 우선 언어학적 기준에 입각하여 범주를 정의하고, 그 정의에 따라 1차 등재 가능 목록에서 기술 범위를 한정하는 방식으로 선정되었다. 1차 등재 가능 목록은 <표준국어대사전>, <연세한국어사전>의 의존명사 표제어와 세종계획 국어 기초자료 구축 분과에서 구축한 기초 통계 자료의 의존명사 목록을 대상으로 하였다. 그리고 다음의 기준으로 표제어를 선정하였다.

① 의존명사를 '선행 요소로 관형어를 필수적으로 요구하여 명사구 구성을 이루는 명사의 하위 유형'으로 정의하고, 관형어의 필수성을 표제어 선정의 으뜸 기준으로 삼는다.

② 의존명사 중에는 일반 명사와 동형어인 경우가 많다. 일반 명사와 의존 명사의 구별이 곤란한 경우, (1ㄱ)과 같이 일반 명사의 의미가 문법적인 의미로 전이되어 통사적으로 반드시 관형어를 요구하는 경우는 의존명사의 목록에 포함한다. 그러나 (1ㄴ)과 같이 일반 명사와 동일한 의미를 가지고 있으며, 자립적으로 쓰일 수 있는 경우는 의존명사의 목록에서 제외한다.

(1) ㄱ. 마당, 모양, 사이, 새, 셈, 참, 한, 터
; 의존명사

ㄴ. 때, 자리, 작정, 주제, 차례, 찰나, 처지, 날, 다음, 동시, 형편 ; 명사

③ 의존명사 중에는 선행 관형어로 관형사나 관형절을 취하지 못하고 명사 수식 성분만을 취하는 경우도 있다. 특히 수량/시간 표현 명사구를 선행어로 취하는 경우(내, 말, 초, 나절, 남짓), 이름과 결합하는 경우(씨, 군, 양, 형, 웅, 편, 저, 역), 나열의 명사구와 결합하는 경우(들, 등, 등등, 등속, 등지)가 있는데, 이들은 반드시 명사 수식어를 요구한다는 점에서 의존명사의 목록에 포함한다.

④ '놈', '거'와 같은 비표준어 또는 준말도 표제어로 선정한다.

그러나 ①의 일반 국어학적 정의에 입각하여 의존명사의 목록을 확정짓는 데에 몇 가지 문제가 있다. 관형어의 유형 중에는 명사구 수식어도 포함되는데, 명사구만을 수식어로 취하는 형태 중에는 기존에 접사로 처리된 것들이 상당수 포함될 가능성이 있기 때문이다. 실제로 ③에서 제시된 예들을 의존명사로 처리하였으므로, 다음과 같이 통사적 단위인 명사 확장 구에 결합하는 접사, 소위 통사적 접사들도 결합 제약의 면에서 의존명사와 동일한 문법적 특성을 갖는 것으로 파악하게 된다.¹⁾

1) (3)의 예들은 띄어쓰기의 면에서 의존명사와 구별되어 왔지만, 띄어쓰기는 다양한 용례에서 그 일관성을 찾기가

- (3) ㄱ. 한 개에 천 오백 원짜리 사과
- ㄴ. 이 주일에 한 번쯤 만난다.
- ㄷ. 1996년 8월 하순께부터
- ㄹ. 만 원어치

따라서 다음의 기준이 본 사전 작업의 의존명사 목록을 선정하는 데에 추가되었다.

⑤ 기존 사전에서 접사로 처리된 목록 중에서 명사 확장 구에 결합하는 통사적 접사들을 의존명사의 목록에 포함한다.²⁾

위의 기준 중에서 ②와 ⑤는 의존명사가 일반명사, 접사와 경계가 모호한 경우를 판단하기 위한 것이다. 이러한 범주 경계의 모호성은 의존명사의 정의 자체가 명사성과 의존성을 바탕으로 하기 때문이다. 이러한 범주 경계의 모호성으로 인해 의존명사에 대한 기존의 여러 연구와 사전 처리에서도 목록 선정의 면에서 차이가 나타났다. 본 사전 작업에서는 의존명사의 의존성, 즉 결합 제약 특성을 중요하게 봄으로써, 의존명사의 선정 규모가 기존 사전 처리에 비해 확대되었다. 다음의 예들이 확정된 의존명사 목록이다.

■ 의존명사 목록 (총 148개)

거리, 것, 격, 겸, 길, 김, 간, 나름, 녀석, 년, 녘, 노릇, 놈, 대로, 동안, 둥, 듯, 등, 따름, 따위, 탄, 때문, 마당, 만, 만치, 만큼, 무렵, 바, 바람, 법, 분, 뻘, 뽕, 성, 셈, 시, 식, 양, 외, 이래, 이상, 자, 적, 점, 조(調), 족족, 중, 즘, 지, 지경, 짝, 쪽, 차, 참, 채, 척, 체, 축, 측, 터, 턱1, 통, 투, 판, 편(便),

어려우므로 의존명사와 접사를 구별해 주는 신뢰성 있는 기준이 되지 못한다. 실제로 '쯤, 짜리, 어치' 등의 통사적 접사들은 의존명사와 접사의 구별이 모호함으로써 띄어쓰기의 오류가 빈번한 예들이기도 하다.

2) ⑤의 기준은 세종 전자 사건의 분과간 협의에 따른 결과이다. 현재 접사 분과에서는 접사의 범위를 단어 파생 접사로 제한하고 소위 통사적 접사들은 접사가 아닌 다른 범주로 보고 있다. 또한 전자 사건의 처리 방침이 접사의 목록을 줄이는 방향으로 진행되면서, 다음 i), ii)의 예와 같이 기준에 접미사로 처리된 형태 중에서 동일 의미로 그 결합의 단위가 통사적 단위(명사 확장 구)에서 단어 단위까지 나타나는 예들도 의존명사로 처리하게 되었다. 그리고 ii)와 같이 새 단어가 되는 예는 의존명사의 정보 항목 중 합성어로 제시하기로 하였다. 이는 일반적인 인쇄 사전, 국어학 연구에서 파생어로 본 것과 다른 처리 방식이다.

- i) 답안지를 보니, 틀린 대투성이다
- ii) {흠진흠{먼지{피땀...}투성}의

품, 한 (→ 2000년 68개 기술 항목)

가랑, 가운데, 간(刊), 간(間), 거, 겨를, 결, 경, 곧, 군(君), 꺾, 께, 꿀, 끝, 나머지, 나위, 나절, 남짓, 내, 내내, 놉, 덕, 데, 들, 듯이, 등등, 등속, 등지, 탄, 리(理), 마련, 만, 말(末), 맛, 망정, 말, 모양, 백, 별, 사이, 상, 새, 생, 수, 십상, 씨, 양, 어치, 여지, 역, 웅, 이내, 이외, 이전, 이하, 이후, 일수, 저, 전, 정도, 조(條), 줄, 짓, 짜리, 체, 짬, 초, 치1, 치2, 탓, 턱2, 투성이, 티, 편(篇), 편(編), 하, 형(兄), 형(型), 형(形), 후 (→ 2001년 80개 기술 항목)

3.2 사전의 논리 구조 및 정보 항목

'의존명사·대명사·수사 사전'에서 정보 표상은 기본적으로 체언 사건의 방식을 따르되, 각 범주별로 고유한 정보를 표시하기 위한 정보 구획들을 추가로 설정한 것이다. 여기에서는 의존 명사의 통사적 특징을 표시하기 위해 중요한 정보 항목, 또는 일반 명사와 변별되는 정보 항목의 내용과 표상 방식 및 그 의의를 설명하려고 한다.³⁾

@pos=[nb] : 품사범주

일반 의존명사의 품사 정보는 'nb'로 표기한다.

@prt=[] : 조사 제약

의존명사 뒤에 결합하는 조사의 제약을 명시한 것이다. 의존명사 중에는 특정 조사와 결합이 제약되는 경우가 매우 많다. 이러한 후행 격조사의 결합 제약은 의존명사를 격조사의 결합이 자유로운 '보편적 의존명사'와 특정 조사와의 결합만으로 제약되는 '제한적 의존명사'로 분류하는 근거가 된다는 점에서 의존명사의 특징을 설명하기 위한 중요한 정보 항목이 된다.⁴⁾

<예> @form=[나위] ... @prt=[~가 (없다)]
 @form=[바람] ... @prt=[~에]

@comb_aj=[] : 형용사-자유 구성

일반 명사의 기술에서 표제어와 의미적인 연관성을 가지며 빈번히 공기하는 형용사를 기재하는 항목이지만, 의존명사의 기술에서는 주로 표제어와 결합 제약을 갖는 형용사가 기재된다.

<예> @form=[어지] ... @comb_aj=[~가 있다|없다]

@comb_v=[] : 동사-자유 구성

형용사-자유 구성처럼 표제어와 결합 제약을 갖는 동

- 3) 의존명사 사전의 논리 구조는 뒤의 덧붙인 상세 기술의 예를 참조하기 바란다.
- 4) 보편적 의존명사와 제한적 의존명사의 분류에 대해서는 남기심·고영근(1987) 참조.

사가 기재된다.

@form=[검] ~ @pre_s=[S을 ~]

<예> @form=[거를] ~ @comb_v=[~을 주다]

@comb_idda=[~이다] : '-이다' 자유 구성

의존명사가 '-이다'와 결합하여 특정한 구문을 이루는 경우를 기술한 것이다. 이는 의존명사가 서술화의 기능을 갖게 되는 경우를 보여주는 정보 항목이다.

<예> @form=[노릇] ~ @comb_idda=[~이다]

@pre_d=[] : 한정사 제약

의존명사에 선행하는 관형사 수식의 제약을 기재한 것이다. 의존명사에 선행하는 관형사 수식의 제약이 없는 경우<예①>, 특정한 관형사와만 결합하는 경우<예②>, 관형사 수식이 불가능한 경우<예③>가 다음과 같은 방식으로 표상된다.

- <예>① @form=[것] ~ @pre_d=[]
- ② @form=[덕] ~ @pre_d=[그 ~]
- ③ @form=[씨] ~ @pre_d=[#]

@pre_n=[] : 수식어-명사

의존명사에 선행하는 명사적 수식어의 결합 제약을 기술한 것이다. 명사구의 수식에 일정한 제약이 내재하지 않는 경우<예①>와 명사구 결합에 일정한 제약이 있는 경우<예②>, 명사구의 수식이 불가능한 경우<예③>에 대한 정보를 모두 기술한다. 또 명사구 결합에 일정한 제약이 있는 경우에는 명사구 뒤에 '-의'가 필수적으로 결합하는 경우<예②-ㄱ>, '-의'가 수의적인 경우<예②-ㄴ>, '-의'가 나타날 수 없는 경우<예②-ㄷ>를 모두 표상한다.

- <예>① @form=[것] ~ @pre_n=[]
- ② ㄱ. @form=[여지] ~ @pre_n=[N-의 ~]
- ㄴ. @form=[덕] ~ @pre_n=[N(-의) ~]
- ㄷ. @form=[씨] ~ @pre_n=[N ~]
- ③ @form=[셈] ~ @pre_n=[#]

@pre_s=[] : 수식어-문장

의존명사에 선행하는 보문 수식어의 결합 제약을 기술한 것이다. 의존명사 중에서 보문 수식이 불가능한 경우<예①>, 특정한 유형의 보문과 결합하는 경우<예②>를 기술하며, 결합 가능한 보문의 유형은 세종 전자 사전의 보문 기술 원칙에 따라 가능한 보문의 형태를 모두 기재하였다. 즉 '-기/는/다는/라는/냐는/은/있을/던/있던/을'의 보문 형태가 의존명사에 선행하는 가능한 보문 유형으로 기재되었다.

- <예>① @form=[씨] ~ @pre_s=[#]
- ② @form=[때문] ~ @pre_s=[S기 ~]

통사정보 C 구획에서 기술되는 표제어의 선행 수식어 제약 @pre_d, @pre_n, @pre_s는 의존명사의 특징을 파악하는 데 매우 중요하다.⁵⁾ 이러한 결합 제약의 특징은 의존명사의 고유한 통사적 특징을 드러내는 것으로서, 전산 처리뿐만 언어학적으로도 중요한 정보가 된다.

@rep=[] : 반복 구성

해당 어형이 반복되어 일종의 구문을 이루는 구성을 기술한 것이다. 특정한 형태로 반복 구문을 형성하는 의존명사 구문의 유형이 이 정보 항목을 통해 기재된다. 동사 간의 일치 정보가 필요한 경우에는 동일한 지표를 사용하여 해당 정보를 표상하였다.

<예> @form=[대로] ~ @rep=[Vi을 ~ Vi]

(지칠 대로 지쳤다)

이상에서 의존명사의 통사적 특징을 나타내는 데 특별히 유용한 정보 항목을 중심으로 그 의미와 표상 방식을 설명하였다. 이러한 정보 항목은 여타의 정보 항목과 함께 의존명사의 고유한 어휘·통사적 특성을 자세하게 명시해 준다.

4. 분류사의 사전 처리 방안

4.1 목록 선정 기준 및 규모

앞서 언급한 바대로 본 사전에서는 의존명사 중 수량 명사와 결합하는 단위성 의존명사를 분류사라 하여 의존 명사와 별개로 다룬다. 일반적으로 분류사라 하면 결합하는 명사를 의미적으로 부류화하는 범주를 가리킨다. 물론 국어의 단위성 의존명사가 모두 이처럼 분류사 기능을 하는 것은 아니다. 선행하는 명사와 비교적 분명한 의미 호응을 보이는 '마리, 개, 채, 권, 명, 그루, 송이, 켈레, 별' 등과 달리, '근, 자, 흙, 킬로그램, 와트' 등은 각각 무게, 길이, 부피, 전력 등 특정 의미를 띠지만 그렇다고 해서 선행 명사를 특별히 제약하지는 않는다. 그럼에도 불구하고 단위성 의존명사를 따로 분류사 범주로 설정한 것은 이들이 일반 의존명사와 다른 통사·의미 특성을 가지고 있기 때문이다. 이러한 특성은 사전에서 별도의 정보 항목을 두어 그 특성을 표기하

5) 서정수(1996)에서는 의존명사의 선행어 결합 정보로 관형사, 용언의 관형형(-을, -은, -는), 체언의 관형형(-의), 체언 자체, 용언의 명사형(-기)의 결합 제약을 들고, 이에 준하여 의존명사의 유형을 분류한다.

게 된다.

분류사의 표제어 목록은 <표준국어대사전>, <연세한국어사전>의 의존명사 표제어와 세종계획 국어 기초자료 구축 분과에서 구축한 기초 통계 자료의 목록을 1차 등재 가능 목록으로 잡고, 몇 가지 기준으로 기술 범위를 한정하여 약 400여 개의 목록을 선정하였다.

사전에 등재할 표제어 선정에 가장 문제가 되는 부분은 동형어와 다의어의 구분이다. 본 사전에서는 일반명사 중 '가마니', '그릇', '잔'과 같은 용기 명사류는 분류사 표제어로 선정하여 명사와 이중범주어로 처리하고, '정거장', '군단', '사단'처럼 '수/양'의 개념이 내재되지 않고 단순히 수사나 수관형사와 결합하는 것은 가산명사로 보고 분류사의 지위를 주지 않았다. 그러나 전자사전의 효율성을 고려하여 사전 미등재어 가운데 '사람, 나라' 등 단위명사로 자주 사용되는 것은 표제어에 포함하였다.

또한 액수를 나타내는 화폐 단위 명사 중 '달러'처럼 명사로 흔히 사용되는 어휘 역시 명사와 이중범주어로 다루었다.

<예> 은행에 가서 달러 좀 바꿔 와야겠다.

분류사에는 다양한 전문 분야의 도량 단위가 포함된다. 이들 목록을 전문 분야에 따라 재점검하고, 그 하위 분류를 세밀화할 필요가 있다. 이러한 하위 분류는 cl_sub라는 정보 항목을 두어 기술한다.

4.2 사전의 논리 구조 및 정보 항목

분류사 사전의 정보 표상은 기본적으로 체언 사전의 방식을 따르면서, 분류사의 고유한 정보를 나타내기 위해 다음과 같은 정보 구획들을 추가로 설정하였다.

@pos=[cl] : 품사범주

분류사의 품사 정보는 'cl'로 표기한다.

@cl_sub=[] : 분류사 하위부류정보

분류사는 그 사용 영역 및 결합되는 명사 부류, 지각 속성에 따라 하위분류되어야 하겠지만, 본 사전에서는 먼저 척도와 수량으로 대분류하였다. 척도와 수량으로 대분류하는 기준은 결합하는 명사와의 의미적 관련성 여부에 따른 것이다. 즉 수량은 정도의 차이는 있지만 다음처럼 결합하는 명사에 대한 일정한 어휘적·의미적 제한이 따른다.

<예> 사물-개, 시체-구, 사람-명, 두부-모, 김-튀, 오징어-축, 말-필, 생선-손, 의류-벌

반면 척도는 결합하는 명사에 대하여 이러한 어휘적·

의미적 제한이 따르지 않는다. 예컨대 '자'는 길이를 잴 수 있는 명사는 어느 것이나 결합 가능하며, '섬'도 그 속에 들어가는 대상과 관계가 없다. 또한 시간이나 횟수를 나타내는 분류사는 특정한 명사와 의미적인 관련성을 별로 맺지 못한다는 점에서 척도에 포함시켰다.

하위분류 정보는 27개 항목으로 세분화하였으며, '시각'은 '시간'으로, '거리'는 '길이'로 각각 통합하며, '전기량', '일량' 등은 '양'으로, '킬로바이트'는 '크기'로 분류하였다. 본 사전의 분류사의 하위 부류와 대표적인 예들을 보이면 다음과 같다.

- 굵기(thk): 예) 텍스, 서클러밀, 번수
- 길이(lnt): 예) 자, 치, 센티미터, 킬로미터
- 날개(uni): 예) 명, 마리, 그루, 포기, 개, 권
- 넓이(sq): 예) 평, 마, 제곱미터
- 농도(dst): 예) 가밀, 도, 피피엠, 프루프
- 무게(wgt): 예) 관, 근, 냥, 돈, 킬로그램
- 묶음(grp): 예) 컬레, 벌, 손, 접, 축, 패, 톳
- 밝기(lmm): 예) 럭스, 포트
- 부분(div): 예) 가지, 조각, 쪽, 토막
- 부피(blk): 예) 되, 말, 홉, 병, 모금, 가마니
- 비율(rat): 예) 퍼밀, 프로, 프로밀
- 세기(str): 예) 밀리쿼리
- 속도(vct): 예) 밀리갈, 밍스, 보, 비피에스
- 시간(tim): 예) 시, 분, 초, 월, 년, 세기
- 압력(prs): 예) 킬로바, 토르, 토리첼리, 파스칼
- 액수(sum): 예) 원, 푼, 전, 달러, 엔, 프랑
- 양(amt): 예) 칸, 쿨롬, 패럿
- 열량(kal): 예) 칼로리, 큐, 킬로칼로리
- 온도(tmp): 예) 도, 디그리데이
- 저항(rst): 예) 킬로옴, 메그옴, 음향옴
- 정도(dgr): 예) 정도
- 종류(knd): 예) 종류, 갈래, 부류
- 차례(odr): 예) 차례, 세
- 크기(lrg): 예) 포인트, 메시, 폰
- 폭(wdt): 예) 폭, 너비
- 횟수(fqn): 예) 회, 바퀴, 번, 선
- 힘(pwr): 예) 켈빈, 크로커다일, 킬로볼트, 암페어

@rel_n=[] : 관련 명사 (related noun)

이 항목에는 분류사와 결합 가능한 대표적인 명사 정보를 기술한다. 분류사 관련 명사는 두 가지로 구별될 수 있다. 즉 '홉, 되, 말', '밀리미터, 센티미터, 미터, 킬로미터'처럼 해당 분야의 관련 도량 단위 정보를 표기하는 것과, '모금, 뿔, 옹금, 아름'처럼 단위 의미를 실현하기 위해서는 특정 신체 부위와 관련을 맺는 어휘를 표기하는 것을 나누어야 한다. 첫번째 관련어 정보가 어휘관계 구획에 표기되는 것이라면 두번째 정보는 결합가능 관련어 정보로 별도 기술된다.

<예> 되의 관련어: (도량 단위) 톳, 홉, 말, 가마니

되와 결합가능 관련어: (선행명사) 쌀, 보리

이 정보 항목에 기술하는 것은 바로 두번째 정보가 되는데, 관련 명사의 성격에 따라 다음과 같이 기술한다.

① 특정 어휘부류가 명시되는 경우는 부류자질을 기술한다.

<예> @form=[그루] ... @rel_n=[나무]
@form=[송이] ... @rel_n=[꽃]

② 자질로 기술하는 것이 쉽지 않을 경우 대표적인 결합 어휘를 ()안에 나열한다.

<예> @form=[축] ... @rel_n=[(오징어)]
@form=[자루] ... @rel_n=[(연필|볼펜|만년필|도끼|총)]

@flt=[] : 결합 가능 구조 (floating)

이 항목에는 분류사 및 수사어 결합 가능 구조에 관한 정보를 기술한다. 이 정보에는 분류사가 결합하지 않는 '아이 셋', '세 아이'와 같은 유형은 기술하지 않는다. 반면에 결합 명사를 명시할 수 없는 '1시', '20분'과 같은 유형은 기술한다. 또한 조사 '-의'의 삭제 가능 여부에 따라 별도의 하위 유형을 나누지 않고 삭제 가능하면 '-(의)-'로, 삭제 불가능하면 '-의-'로 기술하기로 한다. 그 결과 결합 가능 구조(floating)의 유형은 다음과 같이 3가지로 나누어진다.

<결합 가능 구조(flt)의 유형>

- ① n-mdq-cl (명사+수관형사+분류사)
<예> 나무 한 그루, 맥주 한 잔, 밤 한 틀
- ② mdq-cl-(의)-n (수관형사+분류사-(의)+명사)
<예> 한 잔(의) 술, 세 명의 아이
- ③ mdq-cl (수관형사+분류사)
<예> 한 시(=1시), 삼십 분(=30분)

<예> @form=[잔]
@flt=[n-mdq-cl;mdq-cl-(의)-n]
@form=[시]
@flt=[mdq-cl]

@cardord=[] : 기수/서수 정보(cardinal/ordinal)

이 정보 항목에는 분류사 앞에 결합하는 수관형사가 고유어만 가능한지 한자어만 가능한지, 또는 둘 모두 가능한지에 대한 수관형사 제약정보를 기술한다. 또한 기수만 가능한지, 서수만 가능한지 또는 둘 모두 가능한지와 관련한 정보를 기술한다.

예를 들어 '권'은 고유어 수량 표현이 오면 기수적 의미를 갖는다(한 권). 한자어 수량 표현과 쓰이면 서수적 의미를 갖는다(일 권). 한편 '회'는 이는 동일하게 한자어 수량 표현이 오더라도 기수((재)5회), 서수(5회)의 의미가 각

각 있다. 이러한 차이를 밝혀 주면 문장 내 분류사가 어떤 의미로서 쓰이고 있는지 구별할 수 있다.

5. 맺는말

'의존명사/대명사/수사' 사전 구축 작업은 2000-2001년도에 걸쳐 진행되어 금년말에 기술이 완료될 예정이다. 본 사전은 애초에는 각 하위 범주 개개의 언어학적 특성을 반영하는 별도 사전으로 구축될 예정이었으나, 상위의 체언 사전(단일어 사전) 분과의 요구를 수용하여 그 하위사전으로 설계되고 기술되었다. 이는 전체 사전 구조를 단순화하려는 의도 때문이었으나, 여기서 몇 가지 아쉬운 점이 남게 되었다. 가령, 하위 정보항목 중 선행 관형어 정보항인 <synt_c>의 @pre_s=[]는 단일어의 경우에는 결합 가능 정보를 나타내는 반면, 의존명사의 경우에는 선행요소의 필수성이 반드시 표상되어야 하는데도 현재의 설계구조에서는 그 차이가 명세되지 않는 문제가 남아 있다. 이같은 미비점들은 향후 세종사전의 단계적인 확장 및 수정 과정에서 개선되어야 할 것이다.

6. 참고 문헌

고영근(1989). 국어 형태론 연구, 서울대 출판부.
 김영희(1984). 한국어 셈술화 구문의 통사론. 탑출판사.
 남기심·고영근(1987). 표준국어문법론, 탑출판사.
 민현식(1998). 의존명사, 문법자료와 연구, 태학사.
 서정수(1996). 국어 문법, 한양대학교 출판원
 유동준(1983). 국어 분류사와 수량화. 국어국문학 89. 국어국문학회.
 임홍빈(1991). 국어 분류사의 성격에 대하여. 국어학의 새로운 인식과 전개. 민음사.
 채 완(1990). 국어 분류사의 기능과 의미. 진단학보 70. 진단학회.
 유형식(2001). 한국어 분류사의 범주화 기능 연구. 도서출판 박이정.
 홍재성 외(2000). 21세기 세종계획 전자사전 개발 분과 연구보고서

<별첨>

```

<superEntry>
●수
<entry n=1> %dic=[채언/의존명사] %writer=[방성원]
  %crd=[2001/08/13] %mdd=[2001/08/13] %note=[]
  <toplevel> @form=[수] @pos=[nb] @see=[]
    <morph_a> @var[xs=; xd=; xx=] @abb=[] @lng=[]
      @str=[] @org={()} @symb=[]
    </morph_a>
    <morph_b> @hom=[E] @der=[(n);(v);(a);(av)]
      @comp=[(n);(v);(a);(av)] @metc=[(n);(v);(a);(av)]
      @img =[xs={}; xd={}; xx={}; xp={}]
    </morph_b>
    <froz> @idnp=[] @idna=[] @idnv=[] @idda=[]
      @prv=[] @idetc=[]
    </froz>
  </toplevel>
  <sense n=1> @cg=[결국 나는 산에서 혼자 구조대를
    기다리는 ~밖에 없었다;내년이 되면 상황이 좀
    나아질 ~도 있다] @trans=[] @domain=[] @reg=[]
    @con= [] @curs=[] @sem=[] @nm_sub=[]
    @cl_sub=[]@np_sub=[sem=;num=;ref=] @rel_n=[]
  <lr> @syn=[] @ant=[] @hyper=[] @hypo=[]
    @coord=[] @holo=[] @mero=[] @rel=[]
  </lr>
  <synt_a> @cl=[uni=;grp=;div=;qnt=] @prt=[~가|도
    (있다|없다);~밖에 없다] @av=[] @ds=[]
  </synt_a>
  <synt_b> @comb_aj=[~가 있다|없다] @magn =[]
    @comb_v=[] @comb_ida=[] @comb_n=[]
    @supv=[; cor_v={}] @max_n=[] @sel_res=[]
  </synt_b>
  <synt_c> @pre_d=[#]@pre_n=[#]
    @pre_s=[S는|있을|을 ~]
  </synt_c>
  <synt_d> @flt=[] @carord={car=;ord=} @etc= []
  </synt_d>
</sense>
<sense n=2> @eg=[나는 이제야 겨우 그녀를 잊을 ~
  있게 되었다;이 일은 도저히 혼자 할 ~ 없는
  일이다] @trans=[] @domain=[] @reg=[] @con=[]
  @curs=[] @sem=[] @nm_sub=[] @cl_sub=[]
  @np_sub=[sem=;num=;ref=] @rel_n=[]
  <lr> @syn=[] @ant=[] @hyper=[] @hypo=[]
    @coord=[] @holo=[] @mero=[] @rel=[]
  </lr>
  <synt_a> @cl=[uni=;grp=;div=;qnt=] @prt=[~가|도
    (있다|없다)] @av=[] @ds=[]
  </synt_a>
  <synt_b> @comb_aj=[~가 있다|없다] @magn =[]
    @comb_v=[] @comb_ida=[] @comb_n=[]
    @supv=[; cor_v={}] @max_n=[] @sel_res=[]
  </synt_b>
  <synt_c> @pre_d=[] @pre_n=[] @pre_s=[]
  </synt_c>
  <synt_d> @flt=[n-mdq-cl;mdq-cl-(의)-n;mdq-cl]
    @carord={car=kol;si;ord=si} @etc=[]
  </synt_d>
</sense>
</entry>
</superEntry>

```

<superEntry>

●회

```

<entry n=1> %dic=[채언/분류사] %writer=[조형실]
  %crd=[2001/09/11] %mdd=[2001/09/21] %note=[]
  <toplevel> @form=[회] @pos=[cl] @see=[]
    <morph_a> @var[xs=; xd=; xx=] @abb=[] @lng=[]
      @str=[] @org=[si(미)] @symb=[]
    </morph_a>
    <morph_b> @hom=[E] @der=[(n);(v);(a);(av)]
      @comp=[(n);(v);(a);(av)] @metc=[(n);(v);(a);(av)]
      @img=[xs=; xd=; xx=; xp=]
    </morph_b>
    <froz> @idnp=[] @idna=[] @idnv=[] @idda=[]
      @prv=[] @idetc=[]
    </froz>
  </toplevel>
  <sense n=1> @eg=[우리는 3~ 입장권을 샀다;이영화는
    하루에 4~ 상영된다] @trans=[] @domain=[]
    @reg=[] @con=[]@curs=[] @sem=[] @nm_sub=[]
    @cl_sub=[uni] @np_sub=[sem=;num=;ref=]
    @rel_n=[]
  <lr> @syn=[] @ant=[] @hyper=[] @hypo=[]
    @coord=[] @holo=[] @mero=[] @rel=[]
  </lr>
  <synt_a> @cl=[uni=;grp=;div=;qnt=] @prt=[] @av=[]
    @ds=[]
  </synt_a>
  <synt_b> @comb_aj=[] @magn=[] @comb_v=[]
    @comb_ida=[] @comb_n=[] @supv=[; cor_v=]
    @max_n=[] @sel_res=[]
  </synt_b>
  <synt_c> @pre_d=[] @pre_n=[] @pre_s=[]
  </synt_c>
  <synt_d> @flt=[n-mdq-cl;mdq-cl-(의)-n;mdq-cl]
    @carord={car=kol;si;ord=si} @etc=[]
  </synt_d>
</sense>
</entry>
</superEntry>

```