

# 한국어 구문 해석을 위한 동사 ‘가다’의 명사 결합 관계 연구

박건숙

상명대학교 국어국문학과

서울시 종로구 홍지동, 우:110-020

sook@klipl.com

## Noun Link Relation Research Of Verb ‘-Kata (가다)’ for Korean Syntactic Analysis

Keon-Sook Park

Department of Korean Language & Literature

Sangmyung University

### 요약

본 논문에서는 한국어 구문 해석을 위해 동사 중심의 구문 틀 정보를 구축하고, 나아가 결합 빈도가 높은 명사와의 결합 관계를 하나의 네트워크로 구성하는 구문 해석의 방법을 제안한다. 동사 중심의 구문 틀과 명사의 의미 자질은 구문 해석에서 아주 중요한 역할을 하는 것으로, 구문의 비문 여부를 가리는 데 도움을 준다. 그러나 명사의 의미 자질은 경계가 모호하여 구문의 적격성(wellformedness)을 가리기에는 부족한 점이 많다. 따라서 동사와 명사의 결합 관계를 이용하면 구문의 의미적 적격성을 좀 더 명시적으로 가릴 수 있다. 한국어에서 기본 동사이고, 초등학교 교과서에서 사용된 빈도가 아주 높은 동사 ‘가다’를 가지고 구체적으로 구문 틀 정보와 결합 명사의 의미 자질 및 결합 관계를 정리하였다.

### 1 머리말

자연어처리에 대한 연구는 크게 형태소 분석, 구문 분석, 의미 분석 그리고 대화(또는 담화) 분석으로 이루어져 왔다. 이러한 연구의 기본은 형태소 분석인데, 형태소 분석은 학문적인 연구 성과는 물론이고 실제 맞춤법 검사기 등에서 실용적으로 응용되고 있어 최근에는 이를 바탕으로

구문이나 의미, 대화 분석에 대한 연구가 상당히 진행되고 있다. 그러나 구문이나 의미, 대화 분석은 연구 성과에 비해 실제 응용 사례가 거의 없는 실정이다. 이것은 형태소 분석이 대부분 형태소 규칙을 삽입하여 정보를 구축하여 정보의 양이 다소 제한적인데 비해, 구문과 의미 분석을 위한 통사 및 의미 정보는 그 양이 막대하여 구축하는데 어려움이 많기 때문이다. 특히 대화 분석은 언어 외적인 정보 또한 분석에 많은 영향을 주기 때문에 대화 상황까지 일종의 시뮬레이션을 거쳐 구축해야 하는 실정이다.

최근까지 구문 해석은 주로 조사의 격 정보를 이용하거나, 서술어의 자릿수를 이용하여 분석하고, 구문의 적격성을 가리기 위해 명사의 의미 자질 정보를 사용하고 있다. 그러나 의미 자질은 경계가 모호할 뿐만 아니라 모든 명사에 의미 자질을 부여하는 데에 어려움이 많아 구문 해석의 한계로 지적되고 있다.

본 논문은 실제 응용 가능한 구문 해석 정보를 구축하기 위한 하나의 방법을 제시하고자 하는데, 한국어의 구문을 결정하는데 중요한 역할을 하는 동사를 중심으로 구문 해석을 시도할 것이다. 분석된 구문이 의미적으로 부적합한 구문인지를 가리기 위해서는 동사를 중심으로 구문이 취하는 명사의 의미 자질을 통해 결정한다. 그러나 문법

<sup>1</sup> 한국어의 형태소 규칙은 당/홀소리 규칙, 모음 조화 규칙, 불규칙 용언의 활용 규칙 등이 있다.

규칙과 동사의 자릿수 정보를 이용하는 분석은 한계가 있고 모든 명사의 의미 자질을 결정하여 이것을 구문 해석의 정보로 사용한다는 것 역시 한계가 있으므로 이것을 극복하기 위해, 동사가 사용되는 용례를 분석하여 그 동사와 빈번하게 결합하는 구체적인 명사 및 명사의 격 정보를 정리하여 구문 해석의 정보로 사용하고자 한다.

본 논문에서는 사용 빈도가 높은 동사 '가다'를 중심으로 초등학교 국어 '읽기' 교과서에서 예를 뽑아 구문 유형과 명사와의 결합 관계를 정리하였다. 이와 같이 '가다'를 선정한 이유는, 사용 빈도가 높은 기초 어휘여서, 구문의 유형이 다양하고 의미도 기본 의미에서 많이 확장되어 결합하는 명사의 의미 자질이 다양하기 때문에, 많은 구문 해석 정보를 제공할 수 있기 때문이다.

## 2 연구 개요

구문 해석은 형태소 분석에 비해 많은 구문 정보가 필요하고 또한 정확한 구문 해석을 위해서는 의미 정보까지 필요하다. 구문 해석의 방법은 크게 구 구조 문법과 문맥 의존 문법으로 나뉘는데, 이것은 의미를 그다지 고려하지 않은 방법이라서 구문의 적격성을 검사하기에는 부족한 점이 많다. 이러한 내용을 아래에서 살펴보겠다.

### 2.1 구문 해석의 정의

논문 [2]에 의하면 구문 분석은 자연어의 분석 과정의 하나로, 일반적으로 통사 규칙에 따라서 문장 내에서 각 형태소가 가지는 역할, 혹은 상호 관계를 분석하는 것이다. 간단히 말해서 하나의 문장이 구문적으로 적합한 지를 검증하는 것이다. 다음의 예를 보자.

(1) \*나는 사과가 먹다.

위의 문장을 형태소 분석을 통해 분석하면 다음과 같다.

(1a) 나+는 사과+가        먹+다  
 명사+조사 명사+조사 동사+어간+어미

한국어 형태소 분석은 한 어절 단위로 분석하기 때문에, 어절간의 상호 관계에 대한 분석은 불가능하다. 따라서 논문 [5]에서 제시하는 것처럼, 형태소 정보 이외에 동사 '먹다'가 두 개의 명사를

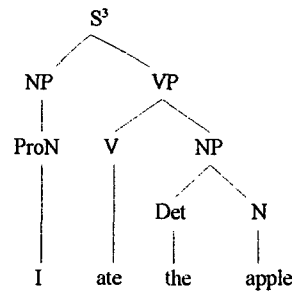
선택할 수 있는 동사이고, 그러므로 'N<sub>1</sub>이 N<sub>2</sub>를'의 형태를 취한다는 정보가 더해지면, 예문 (1)은 구문적으로 적합한 문장이 아니고, 무엇이 잘못 쓰였는 지를 지적할 수 있다.

결국 형태소 분석이 워드 프로세서에서 '맞춤법 검사기'로 응용되어 온 것과 비교하여, 구문 해석은 '구문 검사기' 또는 '구문 교정기'로 응용이 가능하다.

### 2.2 구문 해석의 방법

한국어에 대한 구문 해석의 방법은 크게 문맥 자유 문법 (context free grammar)과 문맥 의존 문법 (context sensitive grammar)이 있다.

[6]에 의하면 먼저 문맥 자유 문법은 초기 단계에서부터 널리 연구되어 온 문법 형태로, 생성언어학과 컴퓨터 과학에서 자연 언어 처리를 다루는데 많이 사용되어 왔다. 이것은 구 구조 (phrase structure)를 얻는 방법으로 문맥 자유 문법에서 사용되는 문법 규칙은 수형도에 나타나는 어떤 구문 기호가 다른 어떤 기호로 다시 구성될 수 있는가를 정의한다. 다음은 문맥 자유 문법의 한 예이다.<sup>2</sup>

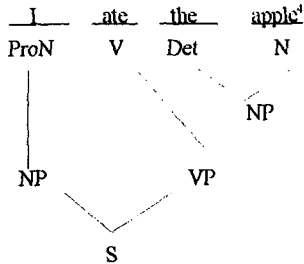


S	→	NP VP
NP	→	ProN
VP	→	V NP
NP	→	Det N
ProN	→	I
V	→	ate
Det	→	the
N	→	apple

<sup>2</sup> 아래의 규칙을 다시쓰기 규칙이라고도 한다.

<sup>3</sup> 김영택 (1994)에서는 이 수형도는 단어의 열을 입력으로 하여 주어진 문법 규칙에 따라 입력된 문장의 구조를 찾는 과정의 하나로, 수형도의 진행에 따라 top-down 구조 또는 하향식 구조라 한

문맥 자유 문법이 각 문장의 구성 성분을 바탕으로 문장을 파악하는데 반해, 문맥 의존 문법은 각 요소, 즉 단어와 단어의 지배 및 의존 관계를 문장을 파악한다.



주로 위와 같은 방법으로 구문 분석이 시도되어 왔고, 특히 한국어에 적합한 구문 분석을 위해 많은 노력이 있어 왔다. 그러나 문법 정보를 중심으로 한 구문 해석은 많은 한계가 있어 의미 자질에 대한 연구가 이루어지고 있지만, 이것 역시 어순이 바뀌거나 중의성을 가진 문장을 처리하는데 많은 어려움이 있어 아직 실용적인 성과는 미비한 실정이다.

2.3 분석 모형

자연어처리를 위한 구문 해석을 연구하기 위해 동사 ‘가다’의 예를 뽑아서 구문 해석의 기본이 되는 구문 틀을 정리하고, 각 구문이 취하는 명사의 의미 자질도 정리하겠다. 구문 틀 유형은 논문 [5]에서 동사의 의미 연구를 위해 제시한 구문 유형을 바탕으로 정리하겠다. 논문 [5]에서는 동사가 선택하는 명사의 수뿐만 아니라 명사의 격 정보를 포함하여 구문 유형을 나누고, 각 구문 유형별로 선택하는 명사의 의미 자질을 부여하는데, 이러한 명사의 수·격·의미 자질을 하나로 묶어 정리하는 방법은 자연어처리의 구문 해석에서 아주 유용하다. 따라서 구문 해석을 위한 구문 틀의 정보 구축은, 동사가 선택하는 명사의 수·격·의미 자질을 묶어 정리하겠다. 그리고 동사 ‘가다’와 결합 빈도가 높은 구체적인 명사 및 명사의 격 정보를 정리하여 이것을 하나의 네트워크로 연결하겠다. 논문 [3]에서는 미국 Princeton 대학의 WordNet 을 소개하고 있는데, 이 WordNet 은 명사 중심의 망인데 비해, 이 논문

서 구축하고자 하는 네트워크는 동사와 명사의 결합 관계를 중심으로 한 것이다.

구문 틀 형식은 동사가 필수 성분만 취하는 기본형을 제 1 유형으로, 필수 성분 외의 다른 성분을 취하는 확장형은 제 2 유형, 제 3 유형 등으로 나누어 다룬다.

동사 ‘가다’가 취하는 명사의 의미 자질은 각 구문 틀을 중심으로 정리한다. 먼저 N<sub>1</sub>의 의미 자질은 크게 [생물]과 [무생물]로 나누고, [생물]은 [사람], [동물], [식물]로, [무생물]은 [구체물], [추상물]로 나누어 제시한다. ‘가다’는 이동의 의미를 나타내는 동사이므로 N<sub>2</sub>의 의미 자질은 [구체적 장소]와 [추상적 장소]로 나누어 제시한다. 그리고 분석 대상이 되는 구문의 비문 여부를 가릴 수 있도록 각 구문 틀 유형에 대표적인 명사의 의미 자질을 연결시킨다.

동사—명사의 네트워크는 동사와 높은 빈도로 결합하는 구체적인 명사를 찾아, 각 명사가 취하는 동사의 유형을 살펴 본다. 동사와 명사가 서로 높은 빈도로 연결되는 것은 하나의 네트워크로 연결하는데, 이렇게 구축된 동사—명사의 네트워크는 구문 해석에서 의미 자질의 모호성이나, 조사 생략 등의 문제를 해결하는 데 도움을 준다.

이것을 하나의 모형으로 제시하면 <그림 1>과 같다.

<그림 1>의 단계 1에서 단계 3까지는 형태소 분석을 중심으로 한 과정으로, 본 논문에서는 단계 4에서 단계 7까지의 과정을 중심으로 연구하였다. 단계 4에서는 입력된 구문에서 동사의 구문 틀을 탐색하여 구문과 일치하는 구문 틀을 결정한다. 그리고 단계 5에서 결정된 구문 틀의 격에 따른 명사의 의미 자질을 확인하는데, 이때 의미 자질이 일치하지 않으면 입력된 구문은 의미상 비문임을 알 수 있다. 의미 자질이 일치하면 단계 6에서 해당 동사와 결합된 명사의 결합 관계를 탐색한다. 동사—명사의 결합 관계가 일치하면 입력된 구문은 구문 구조뿐만 아니라 의미상으로도 적절한 구문임을 알 수 있다.

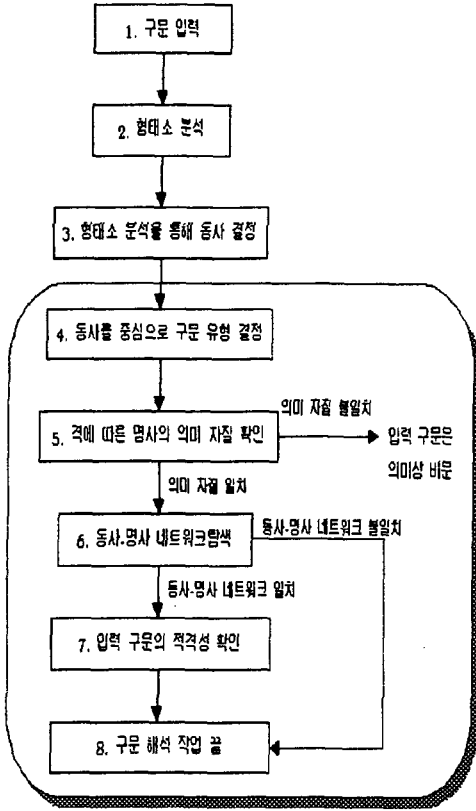
다.

<sup>4</sup> top-down과는 반대인 bottom-up 또는 상향식 구

조이다. (김영택:1994)

(제 10회 한글 및 한국어 정보처리 학술대회)

<그림 1> 동사를 중심으로 한 구문 해석의 모형



- ④ (어떤 직책으로) 자리를 옮기다
- ⑤ (소식이나 연락 등이) 전달되다
- ⑥ (물건이 어떤 사람에게) 옮겨지다
- ⑦ (관심이나 짐작 등이) 미치다
- ⑧ (시선 따위가) 미치다
- ⑨ (금이나 주름 따위가) 생기다
- ⑩ (어떤 일에 노력이나 손질이) 들거나 필요하다
- ⑪ (일정한 값에) 이르다
- ⑫ (어떤 정도나 차례에) 해당하다
- ⑬ (시간이) 경과하다
- ⑭ (어떤 상황이) 계속되거나 유지되다
- ⑮ (어느 때에) 이르다

타동사

- ① (어떤 곳을) 향하여 움직이다
- ② (어떤 장소 위를) 움직이다
- ③ 어떤 목적을 위하여 떠나가다
- ④ (일정한 시간을) 지속하다

국어 사전에서는 ‘가다’의 여러 가지 의미를 상세히 기술하고 있다. 따라서 동사의 의미를 확인하거나, 한국어가 아닌 다른 언어 사용자가 동사의 사용 범위나 의미에 대해서 학습하는 데 아주 효과적이다. 그러나 구문 정보는 ‘자동사’와 ‘타동사’의 분류 정도로만 기술되어 있어 ‘가다’의 구문 구조나 기능에 대한 설명은 아주 부족하다. 결국 자연어처리를 위해서는 좀 더 구문적으로 정리된 사전이 필요하다. 이러한 구문 정보를 중심으로 기술한 것이 사전 [10]으로, 여기에서는 동사를 중심으로 동사의 구문 정보를 상세히 정리하여 제시하고 있다. 사전 [10]에서 제시하는 동사 ‘가다’에 대한 구문 정보는 다음과 같다.

3 동사 ‘가다’의 구문 연구

‘가다’의 구문 틀 정보를 구축하기 위해, 먼저 동사 ‘가다’에 대한 사전적 의미 및 정보를 살펴보고 구문 해석을 위해서는 어떤 정보가 유용한 지에 대해 살펴보겠다.

3.1 ‘가다’의 사전적 분류

동사 ‘가다’는 사전에 따라 [8]과 같이 주어만을 필수적으로 취하는 자동사로 분류하거나 [4, 7, 9]와 같이 자동사와 타동사의 기능을 모두 가지는 동사로 분류한다. 여기에서는 가장 최근에 나온 사전 [7]의 내용만 정리한다. 다음은 사전에 정리된 ‘가다’의 품사 및 의미 내용이다.

‘가다’ [7]

자동사

- ① 옮겨 움직이다
- ② 지향점을 가지고 나아가다
- ③ 복무하거나 종사하기 위하여 다니다

‘가다’ [10<sup>5</sup>]

자동사 :  $N_0, S_2$ -러  $N_1$ -(에+에게+로+에게로)  $V$ [이동]

1.  $N_0$ = 인물,  $N_1$ = 장소 · 인물
2.  $N_0$ = 교통,  $N_1$ = 장소 · 인물

<sup>5</sup> 홍제성 (1997)에서는 동사 구문을 특징짓는 데 자동사 · 타동사 구문의 구별을 중요시하고, 여기에 동일한 의미 해석-또는 규칙적으로 차이가 지는 의미 해석-을 가지면서 두 구문의 교체가 가능한 자타 양용 구문을 별도의 구문 유형으로 설정한다. 그러나 이 같은 범주화에 앞서 ‘일반(자유) 동사, 숙어 동사, 기능 동사, 보조 동사’의 네 가지 동사 유형을 구분한다.

(제 10회 한글 및 한국어 정보처리 학술대회)

$N_0 S_2$ -러  $N_1$ -에  $V$ [이동]

$N_0$ = 인물,  $N_1$ = 모임

$N_0 N_1$ -에  $V$

$N_0$ = 인물,  $N_1$ = 단체

$N_0 N_1$ -로  $V$

1.  $N_0$ = 교통·길,  $N_1$ = 장소

2.  $N_0$ = 추상(이야기, 회의, 풍조),

$N_1$ = 추상(방향, 데, 쪽)

3.  $N_0$ = 인물,  $N_1$ = 저승·하늘 나라

:

타동사 :  $N_0 S_2$ -러  $N_1$ -을  $V$

:

숙어자동사 :  $N_0 C_1$ -이  $V$

예) 이런 옷은 이제 한물이 갔어.

:

숙어타동사 :  $N_0 C_1$ -을  $V$ ..

예) 왜 굳이 가시밭길을 가려고만 하니?

:

[10]은 동사가 취하는 구문을 유형별로 정리하고 각 유형에서 취하는 명사의 의미 자질을 상세하게 정리하였다. 특히 숙어 동사에 대한 설정은 아주 특이한데, 동사 구문 중에서 동사의 의미가 지나치게 확장되어 일반적으로 선택하는 명사의 유형에서 벗어난 명사를 선택하는 경우 또는 관용적으로 굳어진 구문을 숙어 동사로 제시하고 있다. 예를 들어, '한물 가다', '맛이 가다'와 같은 구문으로, 동사가 취하는 보통의 의미 자질로는 구문 해석이 어렵기 때문에, 이러한 동사 구문을 하나로 묶어 처리하는 것은 아주 새롭고 독특한 방법이라 하겠다. 그러나, 각 동사의 구문 유형과 명사의 의미 자질에 대한 정보의 양이 아주 많아 구문 해석의 정보로 활용하기에는 부담이 크다. 또한, 빈도가 높은 구문 유형이나 특별히 강한 결합 관계를 보이는 명사의 의미 자질에 대한 정보가 전혀 없어, 구문 해석을 위한 정보를 선택하는 데는 어려움이 많다.

결국 국어 사전은 구문 해석의 기초적인 자료는 제공하지만 표제어의 의미를 중심으로 이루어져서 통사 정보가 거의 없고, 구문 사전은 너무 방대한 양의 정보를 가지고 있어 정보를 선별해야 하는데, 빈도나 결합 관계의 정도에 대한 정보가 전혀 없어 정보를 선별하는데 어려움이 많다. 따라서 구문 해석에 사용 가능한 구체적인 정보를 구축해야 하는데, 다음 장에서 이것을 정

리하겠다.

3.2 동사 '가다' 구문 구조의 분류

본 논문에서는 대부분의 사전에서 '가다'를 자동사와 타동사로 구분하는 것과는 달리, '가다'의 자동사 유형만을 기본형으로 삼고 그 이외의 것은 기본형에서 확장된 형태로 분류한다. 그 이유는 자동사에 비해 '가다'의 타동사 용례 빈도가 현저하게 낮고 목적격 조사 '-을/를'에 오는 명사의 자질 역시 보통 '가다'와 결합하는 명사의 자질과는 다소 성격이 다르기 때문이다. 따라서 필수 성분만이 오는 '가다'의 기본형을 중심으로 필수 성분 외에 규칙적으로 나타나는 임의적인 성분을 중심으로 분류한다. 이때 주어가 생략된 구문이라도 대부분 추측이 가능하거나 생략되어도 어색하지 않은 문장이므로 모두 이 분류에 포함시켰다. 여기에서 제시하는 구문 유형은 논문 [5]에서 제시한 동사가 선택할 수 있는 명사의 수에 관한 구문 유형을 따른 것이다. 논문 [5]에서 제시하는 구문 유형은 동사의 선택 제한을 밝히는 과정에서 동사가 선택할 수 있는 명사의 의미자질과 제한 받는 명사의 의미자질에 대해 살펴보고 있어, 자연어처리에서 단순히 기계적이고 획일적인 구문 해석에서 벗어나 인지적인 정보를 제공할 수 있는 중요한 방법을 보여주고 있다.

이렇게 분류된 동사 '가다'의 구문 유형은 다음과 같다.<sup>6</sup>

제 1 유형 (기본형)  $N_1, V$

: -(이)가 가다

제 2 유형 (확장형)  $N_1, (N_2)$ 에  $V$

: -(이)가 -에(에게) 가다

<sup>6</sup> 다음은 신현숙 (1984)에서 제시한 동사가 선택할 수 있는 명사의 수에 관한 세 가지 구문 유형이다.

$N1, V - N1$ 이

$N1$ 이 ( $N2$ 에서)...

$N1, N2, V - N1$ 이  $N2$ 를

$N1$ 이  $N2$ 를 ( $N3$ 에)...

$N1, N2, N3, V - N1$ 이  $N2$ 를  $N3$ 에게

$N1$ 이  $N2$ 를  $N3$ 에게 ( $N4$ 에)...

<sup>7</sup> 운평어문연구소 (1992)에서, 조사 '-에'는 처소, 때, 진행 방향, 원인, 경우 등을 나타내는데 '가다'와 함께 쓰이는 경우는 대부분 '진행 방향'이나 '지향점'을 나타내는 부사격 조사로 쓰인다.

(제 10회 한글 및 한국어 정보처리 학술대회)

<표 1> 동사 '가다'의 구문 유형

동사 '가다'의 구문 유형		예 문	빈도	비율
제 1 유형	$N_1, V: -이(가) 가다$	적는 <u>가야겠어요.</u> 그의 학문은 <u>남이 갈수록</u> 깊어졌다. 나는 학교 공부 때문에 <u>갈</u> 수가 없었다.	49	15.2%
제 2 유형	$N_1, (N_2에) V: -이(가) -에 가다$ $N_1, (N_2에게) V: -이(가) -에게 가다$	영수가 <u>학교에 갑니다.</u> 장사꾼은 <u>관가에 가서 ..</u> 나는 <u>선생님께 갑니다.</u> 쉬는 시간에 <u>술환이에게로 갔다.</u>	119 7	39%
제 3 유형	$N_1, (N_2로) V: -이(가) -(으)로 가다$	거꾸로 <u>나라로 가는</u> 길 뭘, 숲 <u>속으로 갔다</u> 고? 자, <u>다음으로 가</u> 볼까?	87	27%
제 4 유형	$N_1, (N_2를) V: -이(가) -을(를) 가다$	길을 <u>가던</u> 한 소년이 이것을 보았습니다. 신문 <u>배달을 가</u> 야 할 텐데... 서울에서 살기 위해서 <u>이사를 가</u> 는 것이다.	23	7%
제 5 유형	$N_1, (S_2러) V: -이(가) -(하)러 가다$	할머니께 세배를 <u>드리러 가</u> 겠습니다. 우리 <u>반은 학교 뒷산에 청소하러 갔다.</u>	17	5.3%
제 6 유형	$N_1, (N_2에) (N_3에) V: -이(가) -에 -에 가다$	나는 <u>아침 일찍 학교에 갑니다.</u> <u>토끼는 벌써 숲 속에 가</u> 있지.	21	6.5%
합계			323	100%

제 3 유형 (확장형)  $N_1, (N_2로) V$

: -이(가) -(으)로<sup>8</sup> 가다

제 4 유형 (확장형)  $N_1, (N_2를) V$

: -이(가) -을(를)<sup>9</sup> 가다

제 5 유형 (확장형)  $N_1, (S_2러) V$

: -이(가) -(하)러 가다

제 6 유형 (확장형)  $N_1, (N_2에) (N_3에) V$

: -이(가) -에 -에 가다

동사 '가다'의 구문 유형과 유형에 따른 빈도 수는 위의 <표 1>과 같다. 실제 아주 다양한 구문 유형이 나오지만 빈도가 높게 나온 것만을 정리하였다. 각 예문에서 명사  $N$ 은 단일 밑줄 ( )로 동사  $V$ 는 이중 밑줄 ( )로 표시하였다.

<표 1>의 내용을 보면 동사 '가다'가 장소나 처소를 나타내는 조사 '-에'와 '-(으)로'와 결합하여 사용된 예가 66%임을 알 수 있다. 이것은 동사 '가다'가 장소 이동의 의미를 나타내는 동사임을 분명히 보여줄 뿐만 아니라, 구문 해석 과정에서 가장 우선하여 검사해야 하는 것임을 보

여준다.

다음 장에서는 '가다'와 결합하는 명사의 의미 자질을 제 1 유형과 제 2 유형, 제 4 유형을 중심으로 살펴보겠다. 이 세 가지 유형을 중심으로 살펴보는 이유는, 제 1 유형은 '가다'의 기본 구문 유형이고, 제 2 유형은 가장 빈도가 높은 구문 유형이고, 제 4 유형은 '가다'의 타동사적 성격을 나타내는 구문 유형이기 때문이다. 제 1 유형에서는  $N_1$ 의 의미 자질을 살펴보고, 제 2, 4 유형에서는  $N_2$ 의 의미 자질을 살펴보겠다. 제 2 유형에서 ( $N_2$ 에게)의 의미 자질은 빈도가 적으므로 본 논문에서는 생략한다.

3.3 '가다'와 결합하는 명사의 의미 자질

이 장에서는 동사 '가다'와 결합하는 명사 중 주어인  $N_1$ 의 의미 자질은 크게 [생물]과 [무생물]로 나누어 보고, 조사 '-에'와 결합하는 명사  $N_2$ 의 의미 자질은 [구체적 장소]와 [추상적 장소]로 나누어 보고자 한다.

명사의 의미 자질은 이미 자연어처리에서 상당히 중요한 개념으로 인식되어 명사의 개념 분류 및 구문 해석에서 많이 사용되고 있다. 의미 자질 정보는 경계가 분명하지 않고 자질을 어느 정도까지 분류해야 하는 지에 대한 문제가 있지만, 구문 해석은 물론 의미 해석까지 가능케 하는 아주 중요한 정보이다. 이처럼 구문 해석에서

<sup>8</sup> 앞 사전에서 조사 '로'는 방향, 수단, 기구, 원인, 신분 등을 나타내는데 '가다'와 함께 쓰이는 '로'는 '방향과 수단, 방법을 나타낸다.

<sup>9</sup> 앞 사전에서 조사 '을/를'은 동사 '가다'와 사용되는 경우, 동작이 이루어지는 장소를 나타내거나, 그 행동이 목표표 하는 대상을 나타낸다.

(제 10회 한글 및 한국어 정보처리 학술대회)

의미 자질은 구문의 비문 여부를 가릴 수 있는 정보를 주지만, 여전히 정계가 모호하다는 문제가 남는다.

따라서 동사 '가다'의 각 구문 틀이 취하는 명사의 의미 자질을 살펴보고, 각 구문 별로 의미 자질의 빈도를 정리하여 빈도가 높은 의미 자질에 우선 순위를 줌으로써 구문 해석의 효율을 높일 수 있다.

3.2.1 제 1 유형  $N_1$ ,  $V$ : -이(가) 가다

이 유형에서 주어에 오는 명사는 거의 대부분이 행동의 주체가 될 수 있는 사람이나 동물이 오는데, 이것은 나머지 유형에서도 마찬가지이다. 의미 자질은 [생물]은 '사람, 동물, 식물'로, [무생물]은 '구체물, 추상물'로 나누어 분석하였다.

<표 2> 제 1 유형  $N_1$ 의 의미 자질

$N_1$	[생물]			[무생물]		합계
	사람	동물	식물	구체물	추상물	
빈도수	37	4	0	2	6	49
비율	75.5%	8.1%	0%	4.0%	12.2%	99.8%

'가다'의 행위자가 되는 주어의 자질은 위의 표에서 보듯이 대부분이 [생물]의 자질을 가진다. 특히 [생물]이면서 사람인 것의 빈도가 높으므로 이러한 자질을 가진 명사가 가장 높은 결합 관계를 가진다고 할 수 있다. 무생물이 주어로 오는 경우는 빈도가 상당히 적는데 그 예는 다음과 같다.

(2) 프린터는 가다가 저절로 멈추고

- 달이 가고 해가 바뀌는 동안
- 전깃불이 잤다.
- 김치 맛이 잤다.
- 벽에 금이 잤다.

위의 '달이 가다'는 '해가 가다/ 날이 가다/ 겨울이 가다'와 같이 사용되는 것으로, 논문 [1]에서는 이것이 시간의 흐름 중 한 부분인 추상적인 것임에도 불구하고, 하나의 이동 가능한 개체로 보고, 그 개체가 '화자'를 기준으로 한쪽 방향에서 다른 방향으로 이동하는 심리적 구성이 가능하기 때문이라고 설명한다. 그리고 '금이 가다'는

마치 선이 한쪽 끝에서 다른 쪽 끝으로 이어지는 과정이 있었던 것으로 심리적으로 느껴 '장소 이동'의 흔적 같은 것을 인식하기 때문이고, '맛이 가다'는 상태의 변화를 '상태 이동'으로 본 것으로 모두 은유적 구성이라고 설명한다.

[10]에서는 이런 구문을 숙어 동사로 처리하지만, 본 논문에서는 모두 기본형으로 묶어 처리한다. 숙어 동사는 동사 중심의 개념이어서 명사의 의미 확장은 거의 배제되어 있다. 논문 [1]에서 제시한 것처럼 이동이나 상태 변화로 인식하는 것은 명사 '맛/ 달/ 금' 등에 대한 인지 작용인데, 숙어 동사는 동사를 중심으로 묶은 것이어서 인지 작용에 역행한다고 할 수 있다. 다음의 예를 보자.

(3) ㄱ. 달이 가다

- 달이 지나다.
- ㄴ. 맛이 가다.
- 맛이 변하다.
- ㄷ. 금이 가다.
- 금이 생기다.

예 (3)을 보면 명사에 대한 인지 작용으로 인해, '맛/ 달/ 금'이 주어로 오는 경우 결합하는 동사가 '가다/ 지나다/ 변하다/ 생기다' 등과 같은 이동 및 변화의 의미를 갖는 동사와 결합한다. 결국 이러한 구문을 숙어 동사로 묶는 것은 인지적인 관점에서 맞지 않을 뿐만 아니라 해석 관점을 흐리게 만드는 문제점이 있다. 따라서 위와 같은 구문 역시 제 1 유형으로 분류한다.

다음의 제 2 유형과 제 4 유형에서는  $N_2$ 의 의미 자질을 중심으로 살펴보겠다.

3.2.2 제 2 유형  $N_1$ , ( $N_2$ )  $V$ : -이(가)-에 가다

이 유형에서 진행 방향이나 지향점에 오는 명사는 거의 모두가 독립적인 건축물이나 일정한 공간을 차지하는 장소의 성격을 가진다.

구체적인 장소는 '학교/ 집/ 도서관/ 산/ 병원' 등과 같은 것부터 '민속촌/ 고장/ 기상청'까지 아주 다양하게 나타나고, 특히 '학교', '집' 등의 빈도가 아주 높게 나타난다.  $N_2$ 에는 대부분 구체적인 장소가 사용되지만, 이렇게 구체적인 장소가 아닌 추상적인 장소, 그러나 우리가 장소로 인식하고 사용하는 것도 나타나는 데 이것을 [추상적 장소]로 분류한다. [추상적 장소]는 대부분 '앞/

(제 10회 한글 및 한국어 정보처리 학술대회)

곳/ 옆/ 거기/ 근처'와 같은 단어로, 특정한 장소를 한정적으로 가리키는 말이 아니므로 [추상적 장소]로 분류한다. 그 밖에 '외가/ 음악회/ 만찬/ 천국/ 저승' 등도 여기에 속하는 것으로, 특히 '음악회'나 '만찬'은 '음악회를 여는 장소', '만찬을 하는 장소'의 의미로 음악회나 만찬이 일정한 장소의 의미를 가지는 명사임을 알 수 있다.<sup>10</sup> '천국'이나 '저승' 등은 상상이나 인간의 인지 속에만 존재하는 것이지만, 장소의 성격을 강하게 지니고 있다.

이렇게  $N_2$ 에 나타나는 명사의 의미 자질은 대부분 장소의 의미를 가지지만 구체적인 것에서 추상적인 것까지 그 범위가 넓고 다양하다. 이것을 정리하면 <표 3>과 같다.

<표 3> 제 2 유형  $N_2$ 의 의미 자질

$N_2$	구체적 장소	추상적 장소	합계
빈도수	100	17	117
비율	85.5%	14.5%	100%

앞에서 제시한 것처럼 명사 '학교'가 조사 '-에'와 결합하여 동사 '가다'와 연결된 예가 가장 많은데, 이렇게 빈도가 높은 명사를 선택해서 명사를 중심으로 '가다'와의 결합 빈도와 다른 동사와의 결합 빈도를 살펴보면 다음 <표 4>와 같다.

<표 4> '학교에( $N_2$ 에)가 결합하는 동사 및 결합 빈도

$N_2$ 에	연결 동사	빈도	비율
학교에	가다	12	50%
	기타 (나가다/ 다녀오다/ 다니다/ 오다 등)	12	50%
합계		24	100%

동사 '가다'와 결합 빈도가 높은 '학교'는, '학교에'를 중심으로 보았을 때도 '가다'와의 결합 빈도가 상당히 높다는 것을 알 수 있다. 여기에서 중요한 것은 동사 '가다'만을 중심으로 결합 관

계를 보는 것이 아니라, 명사를 가지고도 동사와의 결합 관계를 살핀다는 것이다. <표 4>에서 제시된 것처럼 '학교에'는 '가다' 이외에 결합하는 동사도 '가다'와 같이 대부분 이동의 의미를 나타내는 동사임을 알 수 있다.

3.2.3 제 4 유형  $N_1$ , ( $N_2$ 를)  $V$ : -(이)가 -을(를) 가다

이 유형은 마치 타동사가 목적어를 취하는 형태와 비슷한데, 사용된 예를 살펴보자.

(5) 길을 가던 한 소년이 이것을 보았습니다.

우리는 산으로 소풍을 가기로 하였습니다.

신문 배달을 가야 할 텐데...

토끼는 멀리 도망을 가 버렸습니다.

미란이가 다시 서울로 전학을 가게 되었다.

서울에서 살기 위해서 이사를 가는 것이다.

$N_2$ 에 오는 명사는 일정한 장소를 차지하는 것이 아니어서, 위에서 사용된 '에, -(으)로'의 조사로 바꿔 쓸 수가 없다. '소풍, 전학, 이사' 등은 구체적인 장소를 차지하지는 않지만 '소풍지/ 옮겨가는 학교/ 옮겨가는 집'이라는 일정한 장소와 그 장소로 '이동'하는 모습을 포함하고 있는데, 이렇게 명사 자체가 '장소 이동'의 의미를 지니는 것이 23개의 용례 중 16개로 약 70%를 차지한다. 이러한 명사가 ( $N_2$ 를)의 유형으로 '가다'와 결합된 빈도를 정리한 <표 5>를 보면 그 성격을 분명히 알 수 있다.

<표 5> 제 4 유형 명사 ' $N_2$ 를'과 결합 동사

$N_2$ 를	$N_2$ 를 $V$ 유형의 출현 빈도	' $N_2$ 를'과 '가다'의 결합 빈도	'가다'와의 결합 비율
소풍	4	4	100%
배달	2	2	100%
전학	3	2	66.7%
도망	4	1	25%
이사	8	3	37.5%
귀양	3	2	66.7%

이렇게 '장소 이동'의 의미 자질을 지니는 명사는 특히 '가다'와 강하게 결합하므로 명사를 중심으로 동사와 결합시키는 것도 언어처리에 효과적이다.

<sup>10</sup> 홍계성 (1997)에서는 이러한 명사의 의미 자질을 [모인]으로 분류하고 '동창회/ 음악회/ 개강 잔치/ 체육 대회' 등을 제시하고 있다.

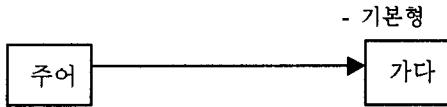


4 분석 결과

지금까지 동사 ‘가다’의 구문 유형을 정리하여 결합하는 명사의 의미 자질을 살펴보고, 각 구문 틀에서 높은 빈도로 결합하는 명사의 의미 자질에 대해서 살펴보았다.

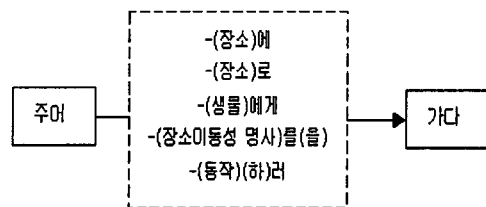
3.2 와 3.3 에서 정리한 것처럼 동사 ‘가다’의 구문 틀 유형과 각 구문 틀과 결합하는 명사의 의미 자질은 상당히 다양하지만, 높은 빈도로 나타나는 구문 틀과 의미 자질을 우선적으로 검사하면 구문 해석의 효율을 높일 수 있다. 먼저 제 1 유형의 구문 틀을 도식으로 나타내면 다음의 <그림 2>와 같다.

<그림 2> 자연어처리를 위한 ‘가다’ 구문 구조



제 1 유형을 제외한 확장형의 구문 틀과 명사의 의미 자질을 합친 구문 구조를 도식으로 나타내면 다음 <그림 3>과 같다.

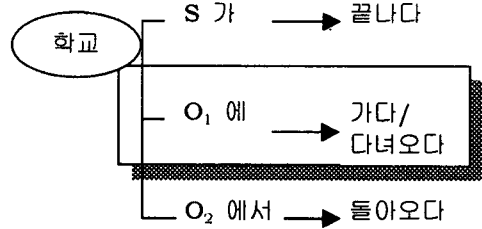
<그림 3> 자연어처리를 위한 ‘가다’ 구문 구조 - 확장형



위의 <그림 2, 3>은 아주 간략하게 구문 구조를 나타낸 것으로, 실제 응용 과정에서는 각 구문 별로 높은 빈도로 결합하는 명사의 의미 자질을 탐색하도록 해야 한다.

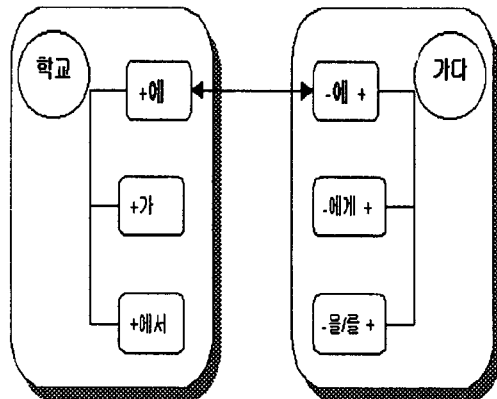
동사가 특정한 의미 자질의 명사를 취하는 것처럼 명사 역시 특별히 강하게 취하는 동사의 자질이 있다. 이 논문에서는 동사를 중심으로 강하게 결합하는 명사를 찾아, 역으로 명사가 문장에서 나타나는 격 형식에 따라 어떤 동사와 주로 결합하는 지를 하나의 망으로 연결할 수 있음을 설명하였다. <표 4>의 내용을 ‘학교’의 격 형식에 따라 정리하면 <그림 4>와 같다.

<그림 4> 명사 ‘학교’의 결합 관계



이렇게 정리한 구문 유형 및 동사-명사 결합 관계를 <그림 1>에서 제시한 구문의 적법성을 위해 사용하면 <그림 5>와 같다.

<그림 5> 동사 ‘가다’와 명사 ‘학교’의 결합관계



<그림 5>는 동사 ‘가다’와 명사 ‘학교’를 중심으로, 각각 강한 결합 관계를 보이는 경우를 네트 워크로 연결한 것이다. 이것은 동사 중심의 정보 구축에서 벗어나, 명사와 동사를 의미나 구문에서 동등하게 다룬 것으로, 동사 중심으로 일방적인 결합 관계가 이루어지는 것이 아니라, 명사와 동사의 상호 작용으로 결합 관계가 이루어진다는 것을 보여준다. 이런 방법은 동사는 물론 명사의 속성을 밝히는 데에도 많은 도움을 줄 수 있고, 특히 자연어처리에서 구문의 구조 및 의미 해석을 위한 구체적인 정보로서 유용하게 이용할 수 있다.

5 맺음말

지금까지 동사 '가다'의 구문 유형을 정리하여 각 구문에 빈번하게 결합되는 명사의 의미 자질을 살펴보고, 결합 빈도가 높은 구체적인 단어는 동사 '가다'와 네트워크로 연결을 하였다.

본 논문에서 제시하는 구문 해석의 방법은 기존의 연구에서 한국어 구문에 대한 전체적인 분석 모형을 제시하던 것과는 달리, 구체적인 동사를 중심으로 구문 해석 방법을 제시하였다. 사용 빈도가 높고 기초 어휘에 속하는 주요 동사를 중심으로 구체적인 정보를 구축하여 점차적으로 확장해 나가고자 하는 것이다.

실용 가능한 구체적인 정보 구축을 위해 동사의 구문 정보와 명사의 의미 자질 정보는 물론, 동사와 명사의 결합 관계를 중심으로 하나의 네트워크를 구축하였는데, 이러한 구체적인 결합 관계 모형은 구문 해석의 실제적인 정보로 사용할 수 있을 뿐만 아니라, 동사와 명사의 속성을 밝히는 데도 많은 도움을 줄 것이다.

물론 의미 자질이 정보가 방대하고 분류에 어려움이 있다고 지적한 것처럼 본 논문에서 제시한 방법 역시 아주 방대한 작업이다. 그러나 주요 동사를 중심으로 기본적인 정보 구축을 확장해 나간다면 구문 해석은 물론 의미 해석의 정보로도 유용하게 사용할 수 있을 것이다.

감사의 글

본 연구를 위해 아낌없는 지도와 도움 말씀을 주신 신현숙 교수님께 깊은 감사를 드립니다.

참고 문헌

- [1] 김미형, '가다'의 의미 연구, 상명여자대학 석사학위 논문, 1983.
- [2] 김영택, 자연 언어 처리, 교학사, 1994.
- [3] 문유진, "한국어 명사를 위한 WordNet의 설계와 구현", 한국정보과학회 논문지(C), 2권 4호, pp.437-445.
- [4] 신기철, 신용철, 새우리말 큰사전, 삼성출판사, 1983.
- [5] 신현숙, 한국어 학습자료 개발을 위한 동사의 의미연구, 상명여자대학 논문집, 1984.
- [6] 오길록, 최기선, 박세영, 한글 공학, 대영사, 1994.
- [7] 윤펽어문연구소, 그랜드 국어사전, 금성출판사, 1992.
- [8] 이희승, 국어 대사전, 민중서림, 1982.
- [9] 한글학회, 우리말 큰사전, 어문각, 1991.
- [10] 홍재성 외, 현대한국어동사구문사전, 두산동아, 1997.