

제1장

자연언어의 기본 속성과 제약

한국어나 영어와 같은 자연언어는 모국어 화자면 누구나 자유롭게 구사할 수 있다는 것은 주지의 사실이다. 이 장에서는 모국어 화자가 이러한 언어능력을 가질 수 있게 하는 가장 중요한 언어의 기본적 속성인 창조성(creativity)과 규칙 지배성(rule-governed)에 관해서 살펴본다. 또한 이러한 두 가지 기본적인 속성을 설명하기 위한 생성문법(generative grammar)의 기본 목표를 고찰해 봄으로써, 본 서에서 제시하는 한국어 구구조문법(KPSG: Korean Phrase Structure Grammar)의 방향성을 알아본다.

1.1 창조성

인간은 누구나 이전에 들어본 적이 없는 새로운 문장을 말하거나, 이해하는 능력을 가지고 있다. 예를 들어 설령 우리는 (1)과 같은 문장을 들어 본 적이 없더라도 이 문장을 아무런 어려움 없이 이해할 수 있다.

(1) 이 책은 한국어를 제약기반문법의 관점으로 분석한 새로운 시도이다.

한국어 화자들은 누구나 이러한 언어 능력을 가지고 있다.

이와 같은 모국어에 대한 언어 능력을 어떻게 설명할 수 있을까? 가장 간단한 가설은 우리가 모든 문장들을 기억하고 있다는 것이다. 그러나 이러한 가설의 문제점은 문장의 수가 무한하다는 사실에서 쉽게 찾을 수 있다. (2)와 같은 문장을 생각해보자.

- (2) a. 철수는 밥을 먹고 또 먹고 했다.
- b. 철수는 밥을 먹고 또 먹고 또 먹고 했다.
- c. 철수는 밥을 먹고 또 먹고 또 먹고 또 먹고 했다.
- d. 철수는 밥을 먹고 또 먹고 또 먹고 또 먹고 또 먹고 했다.
- e. ...

위의 예문에서처럼 ‘또 먹고’를 반복하면 얼마든지 더 긴 문장을 만들어 낼 수 있다. A라는 사람이 평생 동안 ‘또 먹고’를 반복하더라도 B가 거기에다 간단하게 ‘또 먹고’를 한 번 더 반복하면 더 긴 문장을 만들어 낼 수 있다.

이와 같이 문장의 수가 무한하다는 것은 무엇을 의미하는 것일까? 무한한 수의 문장을 이해할 수 있고, 무한한 수의 문장을 만들어 낼 수 있다는 것은 우리가 기본적으로 언어 사용에 관한 창조적(creative) 능력을 가지고 있다는 것을 의미한다. 모국어 화자라면 무한한 수의 정문(well-formed sentence)의 문장을 만들어 낼 수 있는 창조적 능력을 가지고 있는 것이다.¹

그렇다면 어떻게 이러한 창조성을 설명할 수 있을까? 인간의 뇌의 기능을 고려할 때 인간이 기억할 수 있는 자료에는 한계가 있다. 이러한 능력의 한계는 우리가 기억할 수 있는 정보의 양이 제한되어 있다는 것이며, 무한수의 문장을 기억할 수 없다는 것을 의미한다. 남이 이야기 해 준 것을 모두 그대로 기억해 두었다가 필요할 때마다 그것들을 하나씩 꺼내서 말하는 것

¹ 일반적으로 ‘정형(well-formedness)’이라는 개념은 문법적으로 올바르다는 것을 의미하며, ‘수용가능성(acceptability)’이라는 개념은 의미 혹은 화용론적으로 적합하다는 것을 의미한다.

이라고 볼 수 없다. 제한된 기억력을 가진 우리가 어떻게 무한한 수의 문장을 이해할 수 있으며 말할 수 있을까? 가장 설득력 있는 대답은 ‘모국어 화자는 유한수의 규칙을 가지고 무한수의 문장을 만들어 낼 수 있는 문법적 능력(grammatical competence)을 가지고 있다’라는 것이다. 인간만이 가지는 이러한 내재적으로 타고난 언어적 능력이 우리로 하여금 언어를 배울 수 있게 하고 언어를 사용할 수 있게 한다.

1.2 규칙 지배성

모국어 화자가 무한수의 문장을 만들어 낼 수 있는 능력이 있지만, 이들 문장은 모국어 화자들이 정문으로 받아들이는 문장이어야 한다. 문장은 의미를 가진 소리들(sounds) 혹은 단어들의 집합으로 구성되어 있지만, 모든 단어들의 집합이 문법적이지는 않다. 예를 들어 ‘잘, 원빈은, 공을, 정말, 찬다’와 같은 다섯 단어들을 결합할 수 있는 가능 수를 수학적으로 계산해 보면 $5! = 120$ 가지나 된다. 하지만 아래 예문에서 보여주는 것과 같이 모든 결합이 정문이 되는 것은 아니다.

- (3) a. 원빈은 공을 정말 잘 찬다.
 b. 원빈은 정말 공을 잘 찬다.
 c. *잘 원빈은 공을 정말 찬다.
 d. *원빈은 찬다 잘 공을 정말.

한국어 화자라면 누구나 위 다섯 단어를 가지고 (3a)나 (3b)와 같은 정문들을 생성해내는 문법적 능력을 가지고 있다. 비록 의식하지는 못하지만 한국어 화자는 특정 규칙에 따라(rule-governed), 위와 같은 정문을 만들어 내는 것이다. 문법의 기본 목표는 이처럼 모든 가능한 조합에서 정문이 되는 조합들만 생성해낼 수 있는 규칙들을 찾아내는 것이며, 이를 흔히들 생성문

법(**generative grammar**)이라 한다.²

(4) 생성문법:

유한한 수의 규칙을 가지고 무한수의 정문을 생성해 낼 수 있는 문법 오늘날 문법이 추구하는 궁극적인 목적은 위의 정의처럼 자연언어의 창조성과 규칙 지배성을 가장 잘 설명해 줄 수 있는 최적의 생성문법을 개발하는 것이다. 이러한 생성문법을 개발하기 위해서 1950년대 이래 학자들은 많은 노력을 기울여 왔다 (Chomsky(1957) 참조).

1.3 문법의 기본 목표 및 제약성

앞에서 논의한 것처럼 자연언어 문법의 기본 목표 중의 하나는 “생성문법”을 구축하는 것이다. 즉, 무한수의 정문을 만들어 낼 수 있는 한정된 수의 규칙을 발견하는 것이 그 기본 목표인 것이다. 이러한 기본 목표를 수행하는 데 있어서 지적인 바와 같이 이 규칙 수가 무한할 수는 없다. 우리의 인지 능력의 한계로 인하여, 가능한 간단하며 적은 수의 규칙을 선호하는 것은 당연하다 할 수 있다. 또한 목표하는 문법이 대상 언어 현상에서 발견할 수 있는 일반화(**generalizations**)를 설명할 수 있는 기술적 타당성(**descriptive adequacy**)을 가져야 한다.³

자연언어 기술에 있어서 이러한 기술적 타당성을 가진 문법을 구축하기 위한 전제조건 중의 하나는 적절한 **제약(constraint)**성을 포함하고 있어야 한다는 것이다. 예를 들어 (5)와 같이 정확한 구조적 제약성이 없는 규칙으로는 올바른 문법을 기술할 수 없다.⁴

² 여기서의 ‘생성문법’ 개념은 생성성(**generativity**)에만 국한해서 사용한다. 자세한 논의는 Chomsky(1957) 참조.

³ 이러한 기술적 타당성이 갖추어 졌을 때에, 다음 목표는 이 문법이 설명적 타당성(**explanatory adequacy**)을 가지고 있는지의 여부이다. 설명적 타당성은 제한된 언어자료로 어린이들이 어떻게 언어를 완벽하게 습득하는지를 설명할 수 있는 문법 이론이다.

⁴ 이러한 규칙이 왜 구조적 제약성이 없는지에 관해서는 다음 장에서 자세하게 다룬다.

(5) 동사는 바로 앞에 있는 명사와 존칭 정보에서 서로 일치해야만 한다.

이와 같이 ‘바로 앞’과 같이 다소 단순한 규칙으로는 아래 (6a)를 설명할 수 있을지 모르지만, 좀더 복잡한 구조를 가진 (6b)와 같은 문장의 일치 현상은 설명할 수 없다.

(6) a. 선생님이 오시었다.

b. 선생님이 철수를 만나고 나서, 영희와 함께 오시었다.

위의 (6b)에서 동사에 가장 가까운 명사는 ‘선생님’이 아니라 ‘영희’이기 때문에 이러한 문장은 (5)의 제약을 위반한다. 이러한 사실은 문법이 보다 엄밀하게 언어학적으로 제약되어야 한다는 것을 잘 보여준다.

문법에 있어서 제약의 중요성을 인식하면서 1980년대 이후 발전하기 시작한 제약기반문법(Constraint-Based Grammar)은 제약들(constraints)의 집합으로 이루어져 있고 이들 제약 간의 상호작용이나 통합(unification)과 같은 메커니즘이 문법 기술에 중심적인 역할을 한다는 것을 받아들인다. 오늘날에도 지속적으로 발전하고 있는 대표적인 제약문법으로는 어휘기능 문법(LFG: Lexical Functional Grammar), 일반화 구구조문법(GPSG: Generalized Phrase Structure Grammar), 핵어중심 구구조문법(HPSG: Head-driven Phrase Structure Grammar), 구문문법(Construction Grammar) 등을 들 수 있다. 이러한 제약기반이론들이 갖추고 있는 공통된 목표와 기준을 요약하면 다음과 같다.⁵

생성성(generativity): 문법이론은 유한수의 규칙으로 무한수의 정문의 문장을 생성할 수 있어야 한다. 즉, 문법이 만들어 내는 구조들이⁶ 정형인지 비정형인지를 구분할 수 있어야 한다.

⁵ 제약기반 이론의 특징에 관해서는 Pollard & Sag (1997) 참조.

⁶ 예를 들어 HPSG에서의 자질구조나 지배결속이론인 GB(Government and Binding)에서의 수형도.

표현성(expressivity): 문법 기술을 위해 사용되는 메타언어(meta language)가 충분한 표현력을 가지고 있어야 한다.⁷ 특정 문법이 사용하는 메타언어가 기술 대상 언어가 생성하는 모든 구조를 설명, 표현할 수 없다면 이는 문법으로서 적절치 않다는 것을 의미 한다.

경험적 타당성(empirical adequacy): 제약문법의 또 다른 중요한 기준은 경험적 자료들을 정확히 기술하여야 한다는 것이다. 모호하고 증명할 수 없는, 혹은 추상적이며 불확실한 원리나 공리(axiom)를 설정하는 문법이 아니라, 기술 대상 언어의 광범위한 자료들을 설명할 수 있는 명시적인 규칙(explicit rules)이나 원리를 가진 문법이어야 한다.

심리언어학적 적절성(psycholinguistic reality): 문장의 문법성을 판단하는 내재적 언어 능력(linguistic competence)을 설명할 뿐만 아니라 나아가서 언어의 이해나 발화와 같은 언어 수행(linguistic performance) 현상도 잘 설명할 수 있는 문법이론이어야 한다. 이러한 적절성을 충족시키기 위해 다음과 같은 조건들도 만족시켜야 한다.

- 비파괴성(nondestructiveness): 문법이 언어 구조를 변형하거나 파괴하는 Move- α 와 같은 이동 규칙을 설정하지 않는다. 문법 규칙이나 어휘 정보들은 언어 구조를 동시에 만족할 수 있도록 상호작용하는 제약들로만 존재한다.⁸
- 국부성(locality): 주어진 언어 구조의 정형성(well-formedness)을 판단할 때, 자신을 포함한 보다 큰 구조나 다른 구조와 비교해서 결정하

⁷ 술어논리, 자질구조, 구표식(phrase marker), PROLOG, LISP 등을 메타언어라 할 수 있다.

⁸ 따라서 제약기반이론에서는 의문사 이동(wh-movement), 접사 이동(affix-hopping), 핵어 이동(head-movement)과 같이 구조를 변경할 수 있는 규칙들이 존재하지 않는다. 그 주된 이유는 기존의 구조를 변경할 수 있는 언어 처리 모형이나 분석기를 구축하는 것이 어렵기 때문이다. 특히 우리가 언어를 처리하거나 이해하는 과정이 점진적(incremental)이며 통합적(integrated)인 과정이라는 사실은 구조를 변경하는 규칙을 설정하는 것은 적절치 않다는 것을 말해준다. 1.2 참조.

거나 혹은 관련지어 결정되는 것이 아니라, 그 구조 자체 내에서 관련 제약들을 모두 준수하였는지의 여부만 점검한다.

- **평행성(parallelism)**: 언어 정보는 음운, 통사, 의미, 화용 등 다양한 정보 계층(level)을 포함하고 있다. 전통적 변형문법에서 설정하는 T-모형과 같이 이들 정보 계층 사이에 위계나 순서가 존재하는 것과는 달리, 제약기반이론에서는 이러한 관련 언어 정보가 동일한 계층 내에서 수평적, 평행적으로 존재한다.
- **비독립성(nonautonomy)**: 다양한 언어 정보들이 서로 밀접한 상호 제약관계에 있다는 사실을 반영해야 한다. 비록 언어 기술이나 이해에 있어서 경우에 따라 한 가지 언어 정보만 필요할 수도 있지만, 많은 언어 현상은 여러 언어 정보 사이의 상호작용의 결과물이라는 사실을 반영해야 한다.⁹

본 서는 이러한 기본적인 목표를 준수하는 한국어 문법을 제시하는 데 주 목적이 있다. 즉 한정된 수의 규칙을 가지고 정문의 한국어 문장을 만들어 낼 수 있는 제약기반 생성문법(constraint-based generative grammar)을 구축하는 것이 본 서의 목적이다. 본 서에서 제시할 한국어 생성문법 KPSG는 이러한 생성성, 표현성, 경험적 타당성, 심리언어학적 적절성을 목표로 한다. 특히 이 문법은 표면구조(surface structure)를 중요시 하며 문법 정보 간의 상호 제약이 중요한 역할을 한다. 즉 형태, 통사, 의미, 화용 정보들이 서로 유기적인 관계를 맺는 한국어 생성문법이다.

⁹ 예를 들어 어순현상에서의 통사, 음운적 정보의 상호작용, 결속이나 영역(scope) 현상에서의 통사, 의미적 정보의 상호작용, 초점이나 주제 현상에서 음운적 화용적 정보들의 상호작용 등을 들 수 있다. 최근 10년간의 문법 연구는 이러한 문법 정보 간의 상호작용에 중점을 두고 있다고 해도 과언이 아니다.