

# 소규모 학습자 코퍼스에 기반한 중국인 한국어 학습자의 구문 특징 연구

고결(군산대)

## 〈목 차〉

- |                       |                  |
|-----------------------|------------------|
| 1. 머리말                | 4. 학습자의 구문 오류 분석 |
| 2. 학습자 코퍼스의 구축과 구문 표지 | 5. 맺음말           |
| 3. 학습자의 구문 분석         |                  |

## 1. 머리말

본고는 중국인 한국어 쓰기 코퍼스를 이용하여 중국인 한국어 학습자의 구문 실현 양상을 기술하고 구문 특징을 밝히는 데 목적이 있다.<sup>1)</sup> 중국어와 한국어는 형태론적 유형성에서 고립어와 교착어라는 차이를 보인다. 그러나 형태론적 유형성뿐만 아니라 통사적인 차원, 특히 구문 구조에서도 차이를 보인다. 구문 구조와 관련하여 가장 눈에 띄는 두 언어의 차이는 어순이다. 한국어는 문장 성분이 문장 안에서 자유롭게 이동할 수 있지만 기본적으로는 ‘주어-목적어-서술어’ 어순을 취하는 반면 중국어는 어순이 고정적이며 기본적으로 ‘주어-서술어-목적어’ 어순을 취한다.

1) 구문(構文)이란 글의 짜임이나 문장의 구성이라고 할 수 있다.

구문형성 방식에서도 두 언어는 다르다. 특히 주어, 목적어, 부사어와 같은 문장 성분의 형성 방식에서 차이를 보인다. 중국어는 각 성분의 통사적 자격을 표시하는 조사가 존재하지 않는 반면,<sup>2)</sup> 한국어는 조사를 이용하여 성분의 통사적 자격을 분명히 한다.<sup>3)</sup> 구문을 종결하거나 구문과 구문을 연결하는 방식도 다르다. 중국어는 문장의 종결과 연결이 단어와 어순으로 표현되는 반면, 한국어에서는 어미 형태소가 이 역할을 맡는다.

한국어 구문 형성의 특징은 구문을 형성하는 핵심 역할을 하는 문법형태소, 곧 조사와 어미에 있다. 따라서 중국인 한국어 학습자가 작성한 한국어 구문에서 조사와 어미의 쓰임이 어떻게 나타나는지를 면밀히 살필 필요가 있다. 단어와 문법 형태소를 결합하여 문장을 만들거나 이해하는 능력을 구문 능력이라 할 때, 한국어 구문 능력을 높이기 위해서는 문법 형태소를 적절히 사용하는 법을 필수적으로 배워야 한다.

한국어 학습자의 쓰기 자료를 바탕으로 한 구문 분석 작업은 학습자가 한국어 구문의 특징을 이해하고 한국어 문장을 정확하게 쓰는 데 도움을 준다. 또한 중국인 한국어 학습자가 작성한 구문 유형의 경향성과 문법 형태소 사용에서의 오류 경향성을 파악함으로써 한국어 교수법의 체계화를 꾀할 수 있다.

최근 컴퓨터의 발달과 보편화로 코퍼스 자료가 더 광범위해지고 코퍼스에의 접근성도 높아졌다. 2000년대에 들어서면서 한국어 학습자 코퍼스에 관한 연구가 시작된 후부터 학습자들 코퍼스에서 보이는 언어적 특징을 밝히고 이에 기초한 교육과정 및 지도 방안이 개발되어 왔다. 특히 외국인 학습자들을 대상으로 하는 구문 연구가 새로운 국면에 접어들게 되었다. 고석주(2004)에서는 한국어 학습자 코퍼스의 구축이 오류 분석에 대한 중요성을 살펴보았다. 고석주(2004)에 의하면 코퍼스를 이용하여 오류 분석을 객관적으로 진행할 수 있다고 했다. 이승연(2007)에서는 한국어 학습자 코퍼스의 구축과 연구가 언어 교육을 위해 매우 중요한 부분이며 효율적인 방법이라고 했다. 이전에

---

2) 중국어에 존재하는 조사는 구조조사, 동태조사, 어기조사 세 가지뿐이다. 중국어의 조사는 문장의 구조 관계나 부가 의미를 나타내고 한국어 조사처럼 문장 성분의 통사적 자격을 표시하지 못한다.

3) 한국어의 격조사는 때로 생략되기도 하는데 생략 기제가 분명하지 않아 한국어 학습자들이 성분 구조를 학습하는 데 큰 어려움을 겪는다.

중국인 학습자에 대한 구문 연구는 용언이나 조사에 대한 연구가 대부분이었다.<sup>4)</sup> 최근 들어 코퍼스를 이용함으로써 학습자의 구문과 관련된 총체적인 특징과 교육 방안에 대한 연구도 가능해졌다. 또한 코퍼스를 이용한 구문 연구는 객관적이어서 구문 교육에 더 효과적이다.<sup>5)</sup> 예컨대 코퍼스 분석 결과를 통해 한국어 구문의 어순 특징, 문장의 호응 관계, 서술어의 선택 제약 등의 특징을 학습자에게 제시함으로써 학습자가 자신의 구문 형성 방식과 한국어의 구문 형성 방식의 차이를 더 직관적으로 느낄 수 있다. 이런 면에서 학습자 코퍼스를 이용한 구문 연구는 학습자 코퍼스에서 나타난 언어학적인 특징을 파악할 수 있으며 교육적인 측면에서는 한국어를 정확하게 사용하도록 할 수 있다.

본고에서 중국인 학습자의 구문 특징을 밝히기 위한 연구 방법은 아래와 같다.

첫째, 중국인 한국어 학습자의 작문 자료를 수집해서 정리한다. 정리하는 과정에서 중국인 학습자의 작문 자료를 텍스트 형식으로 저장한 후 구문을 부호화한다. 주로 기본문형, 구문 오류를 표지할 것이다.

둘째, 중국인 학습자의 구문 특징을 명확하게 파악하기 위해서 각 표지 항목을 통계화한다. 중국인 학습자의 기본문형을 통계화하고 한국어 기본문형이 수집된 자료에서 어떻게 사용되고 있는가, 빈도가 가장 높은 기본문형은 어떤 것인가 등을 살펴본다. 그리고 구문 형성상의 특징을 밝히기 위해서 중국인 학습자가 사용하는 문장 구조도 살펴본다. 학습자의 단문과 복합문의 사용 빈도, 복합문을 쓸 때 안긴 문장과 이어진 문장의 사용 빈도, 학습자의 어미 사용 빈도를 통계화한다.

셋째, 구문 표지를 한 후에 중국인 학습자의 작문 자료에서 나타난 오류의 양상 및 원인을 분석한다. 오류 분석을 하기 전에 오류 유형을 먼저 나눠야 한다. 주로 문법 오류를 분석할 것이다.

---

4) 이 분야에 대한 연구는 김충실(2009) 등이 있다.

5) 구재희(2007)에서는 학습자에게 올바른 문장 사용을 빠른 시간 안에 익히게 하는 효과적인 한국어 학습 방법으로 구문 교육을 강조한다.

## 2. 학습자 코퍼스의 구축과 구문 표지

문장의 구조를 밝히는 데에는 구문 표지가 유용하다. 구문 표지를 통해 문장 구조에 대한 자세하고 총체적인 분석이 가능하다. 더 나아가 학습자들이 한국어로 문장을 만들 때 어떤 종류의 기본문형을 많이 사용하고 구문 형성에서 핵심 역할을 하는 조사와 어미의 사용 양상도 밝힐 수 있을 것이다.

중국인 한국어 학습자들의 구문 표지를 하기 위해서 먼저 학습자들이 쓴 작문 자료를 수집하여 규모가 작은 학습자 코퍼스를 구축하였다. 중국 산둥성에 있는 연대(煙臺) 노동대학교(魯東大學校)에서 한국어를 전공하는 48명의 2학년 중국인 학습자들의 작문 자료를 수집하였다. 작문은 모두 420편이고 총 5만 어절이다. 학습자들의 연령은 20대의 성인 학습자들로서 한국어의 수준은 중급 정도이다.

자연언어처리 분야의 구문분석에 대한 연구는 일반적으로 형태소분석을 마친 후에 입력문장의 문법적인 구조를 분석한다(안광모·서영훈, 2011). 따라서 일반적으로 구문 표지는 형태적 표지, 문법적 표지, 기능적 표지를 모두 포함한다. 중국어는 한국어와 달리 고립어라서 통사적 자격을 표시하는 조사가 존재하지 않는다는 사실과 한국어와의 어순 차이를 고려해서 본고에서는 중국인 학습자의 구문 특징, 특히 학습자의 기본문형과 어순 관계, 호응 관계에 초점을 맞추었다. 따라서 구문 표지를 할 때 주로 기본문형 표지와 오류 표지를 취할 것이다. 오류 표지는 주로 문법 오류 표지를 할 것이다. 자세한 구문 표지와 부호는 아래와 같다.

〈표1〉 구문 표지의 분류와 사례<sup>6)</sup>

표지 유형		분류	표지 부호	사례
문형 표지	기본 문형	NP이 V	S1	일생 중에는 작고 큰 <u>기억들이</u> 남는다.
		NP이 Adj	S2	내가 좀 <u>설렘했다</u> .
		NP이 NP이다	S3	<u>그가</u> 가족 중에 제일 빠른 사람이다.
		NP이 NP을 V	S4	<u>친구들이</u> 밥을 먹고 나서 게임을 했다.
		NP이 NP이 아니	S5	한 눈에 반한 것이라고 말해도 <u>과연이</u> 아니다.

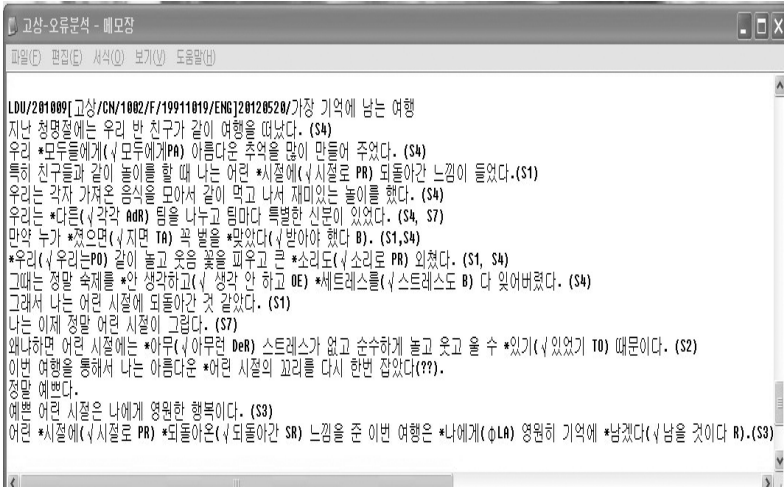
6) 구문 표지의 분류와 사례는 21세기 세종계획 구문 분석 말뭉치를 구축할 때의 기준을 참고해서 변형했다.

		다/되다			
		NP이 NP을 NP으로 V		S6	그 때 내가 꼭 이 여자를 내 여자 친구로 만들기로 마음을 먹었다 <sup>7)</sup> .
		NP이 NP이 Adj		S7	등산을 시작한 지 얼마 지나지 않아서 모두들 힘이 없고 많이 피곤해 보였다.
오류 표지	문법 오류 표지	조사 오류	조사의 생략	PO	다가오는 대학교 <u>입시(√입시가 PO)</u> 우리를 귀찮게 한다.
			다른 조사로의 대체	PR	그 순간은 기쁘기는(√기쁘기가 PR) 더 말할 나위가 없었다.
		어미 오류	선어말어미오류	T	사장님께 아주 미안하다고 생각합니다. 다른 사람에게 이 일을 <u>부탁하기(√부탁하시기H O<sup>8)</sup>)</u> 를 바랍니다. 나는 동아리에서 탈퇴하고 열심히 공부하기로 <u>정한다(√정했다TO<sup>9)</sup>)</u> .
			종결어미오류	E	쓰기가 한국어 능력시험에서는 굉장히 중요해요. 그런데 어떤 학생들이 이것을 소홀히 해요. 그래서 우선 *하는(√할 DR) 일이 *학생들을(√학생들이 PR) 쓰기에 대한 좋은 태도를 가지게 하는 *것이다(√거예요 EnR).
			연결어미오류	C	여행 때 소매치기를 <u>당하거든(√당하면CR)</u> 누구에게 도움을 구하지요?
전성어미오류	D	저에게 가장 기억에 남는 여행은 가족들과 임기시에 <u>가는(√간 DR)</u> 것이다.			

코퍼스에 대한 표지는 수작업<sup>10)</sup>으로 하고 이분지 구조를 기본 원칙으로 삼아 분석했다. 문장 구조의 성분 관계는 기본적으로 결합된 격조사에 의존하지 만 조사가 생략되거나 보조사가 결합된 경우에는 서술어의 구문 틀에 넣어 판단 후 분석했다. 학습자 코퍼스는 아래와 같이 구축됐다.

- 7) 예문은 S6은 S4에도 포함되어 있는데 S6의 특징을 보여 주기 위해서 필요한 부분만 밑줄을 그었다. 구문 표지할 때 두 가지 기본문형은 다 표시됐다.
- 8) HO는 존대 어미의 생략을 나타낸다.
- 9) TO는 시제 어미의 생략을 나타낸다.
- 10) 구문 표지 도구는 반자동 구문분석기가 존재하지만 코퍼스를 교정해야 하는 부분이 많고 정확성을 보장하기 위해서 구문 하나하나를 분석하여 수작업으로 진행하는 과정이 꼭 필요하다.

〈그림1〉 학습자 코퍼스 구문 표시의 화면



### 3. 학습자의 구문 분석

#### 3.1. 기본문형의 분석

한국 국내에서 기본문형을 설정하는 문제에 대한 기존 연구는 많은 편이다. 그 중 남기삼·고영근(2006), 이익섭(2006) 등이 대표적이다. 먼저 남기삼·고영근(2006:234)에서는 기본문형을 다음과 같이 네 가지로 설정하였다.

- (1) ㄱ. 무엇이 어찌한다. (주어-서술어1)
- ㄴ. 무엇이 어떠하다. (주어-서술어2)
- ㄷ. 무엇이 무엇이다. (주어-서술어3)

(1)은 각각 동사문, 형용사문, 명사문을 나타낸다. 모든 한국어 문장의 기본 형식은 기본적으로 (1)의 문형으로 환원될 수 있다. 그러나 (1)의 문형은 한국어 학습자의 구문 문형 특징을 밝히기에는 지나치게 단순화되어 있다. (1)의 구조에는 목적어나 보어, 필수적 부사어 등 문장 형성에서 꼭 필요한 성분이

포함되어 있지 않기 때문이다.

한편 이익섭(2006:7)에서는 기본문형을 다음과 같이 일곱 가지로 제시하였다.

- (2) 가. 무엇이 어찌한다. (주어-자동사)
- 나. 무엇이 어떠하다. (주어-형용사)
- 다. 무엇이 무엇이다. (주어-명사이다)
- 르. 무엇이 무엇을 어찌한다. (주어-목적어-타동사)
- ㅁ. 무엇이 무엇이 아니다/된다. (주어-보어-서술어)
- ㅂ. 무엇이 무엇을 무엇으로 삼는다/여긴다. (주어-목적어-보어-서술어)
- ㅅ. 무엇이 무엇이 어떠하다. (주어1-주어2-서술어)

이익섭(2006)에서는 서술어가 요구하는 필수적 성분을 기본 문형에 명시하여 (1)보다 세분하였다. 본고에서는 중국인 학습자들의 구문 특징을 연구 대상으로 삼는다. 중국인 학습자는 한국어 문장 구조에서 요구되는 필수 성분을 명확히 이해하고 정확히 사용할 필요가 있다. 이에 따라 기본문형에 필수 성분이 포함된 (2)의 분류 방식을 좇기로 한다.

중국인 한국어 학습자의 기본문형의 분석은 두 가지 측면에서 이루어진다. 첫째, 위에서 설정한 한국어 기본문형이 학습자 코퍼스에서 어느 정도의 빈도로 사용되고 있는지를 분석한다. 둘째, 중국인 학습자의 문장 구조를 대상으로 문장 형성에 참여하는 조사와 어미의 사용 양상을 분석한다.

먼저 학습자의 코퍼스에서 사용된 기본문형의 양상을 살펴보기로 한다. 이런 작업을 통해 중국인 학습자가 자주 쓰는 기본 문형을 확인할 수 있다. 이와 함께 한국어 모국어 화자의 기본문형 사용 빈도<sup>11)</sup>와 비교하여 중국인 한국어 학습자의 기본문형 사용의 특징과 문제점을 찾을 수 있다. 중국인 학습자의 기본문형 사용과 관련하여서는 기본문형의 빈도가 분석 대상이 된다. 48명 학생의 작문의 문장은 모두 4525개이고 분석 대상이 되는 문형은 모두 8212개이다.<sup>12)</sup>

---

11) 한국어와 중국어는 모두 '주어, 서술어, 목적어'를 가지고 있지만 중국어에는 어미, 조사가 존재하지 않으므로 모국어 화자와 중국인 한국어 학습자의 기본 문형의 사용 빈도를 비교하면 그 특징과 문제점을 찾을 수 있다.

중국인 한국어 학습자들의 기본문형 사용 빈도는 다음과 같다.

〈표2〉 중국인 한국어 학습자의 문형 사용 빈도

순위	유형	빈도	백분율(%)
1	NP이 NP을 V	2031	24.7%
2	NP이 Adj	1135	13.8%
3	NP이 V	572	7.0%
4	NP이 NP이다	387	4.7%
5	NP이 NP이 아니다/된다	83	1.0%
6	NP이 NP이 Adj	67	0.8%
7	NP이 NP을 NP으로 V	34	0.4%

중국인 한국어 학습자의 기본문형 사용 빈도 중 1, 2, 3위는 ‘NP이 NP을 V’, ‘NP이 Adj’, ‘NP이 V’ 순이다. 한편 한국인 모국어 화자의 문형 빈도에 대한 조사는 노은희(1999)에서 이루어졌다. 중국인 한국어 학습자의 기본문형과 한국인 모국어 화자의 기본문형을 비교하면 그들의 차이를 찾아낼 수 있다. 노은희(1999)에서는 서술어를 중심으로 다섯 가지 기본문형을 설정했다. 자동사 술어형(N이 V), 타동사 술어형(N이 N을 V), 형용사 술어형(N이 Adj), 명사 술어형(N이 N이다) 등이고 각 문형의 사용 비율을 비교하면 아래와 같다.

〈표3〉 중국인 학습자와 한국인 모국어 화자의 기본문형 사용 빈도의 대조

순위	분류	문형	백분율(%)
1	중국인 학습자의 기본문형	NP이 NP를 V	24.7%
2		NP이 Adj	13.8%
3		NP이 V	7.0%
4		NP이 NP이다	4.7%
5		NP이 NP에/에게 (NP을) V	3.3%
1	노은희(1999) <sup>13)</sup>	NP이 NP을 V <sup>14)</sup>	38.23%
2		NP이 NP이다	15.48%
3		NP이 V	12.40%
4		NP이 Adj	8.84%
5		NP이 NP에/에게 (NP을) V	6.16%

12) 문형은 서술어를 중심으로 추출했다. 내포문을 안은 문장이나 이어진 문장의 경우 서술어가 둘 이상인 경우가 있다. 따라서 문형 수와 문장 수는 일치하지 않는다.

위의 표를 보면 중국인 학습자와 한국인 모국어의 기본문형 사용 빈도는 모두 ‘NP이 NP을 V’이 가장 높다. 그 다음으로 중국인 학습자의 기본문형은 ‘NP이 Adj’와 ‘NP이 V’가 2, 3 위이고 한국인 모국어 화자의 기본문형은 ‘NP이 NP이다’, ‘NP이 V’가 각 2, 3 위이다. 이런 특징을 보이는 이유는 중국어 기본문형의 사용과 관계가 있다고 생각한다. 베이징 언어학원 문형 연구팀(1989)에 따르면 중국어의 기본문형의 사용 순위는 ‘NP이 (NP을) V<sup>15)</sup>’, ‘NP이 Adj’, ‘NP이 NP이다’ 순이다.

이와 같이 중국인 학습자와 한국인 모국어 화자의 기본문형의 사용 빈도를 비교하면 몇 가지 특징이 나타난다. 첫째, 문장의 기본 성분인 주어, 서술어, 목적어 세 가지를 다 가지고 있는 ‘NP이 NP을 V’의 빈도가 가장 높다. 중국인 학습자의 경우 이 문형의 사용 빈도가 가장 높은 것은 언어 습득과 학습의 측면에서 설명할 수 있다. 한국어와 중국어에서 모두 ‘NP이 NP을 V’ 문형의 사용 빈도가 가장 높다. 그리고 한국어와 중국어 문장을 구성하는 주요성분이 모두 ‘주어, 술어, 목적어<sup>16)</sup>’이어서 중국인 학습자가 중국어 습득과 한국어 학습 과정에서 이 문형에 가장 자주 노출되어 익숙해졌다는 점을 고려하면 이 문형의 사용빈도가 높게 나타난 이유를 알 수 있다.

둘째, 중국인 학습자가 두 번째로 자주 사용하는 기본문형은 ‘NP이 Adj’인 반면 한국인 모국어 화자가 두 번째로 자주 사용하는 기본문형은 ‘N이 N이다’이다. 중국어의 서술어는 동사와 형용사 두 가지밖에 없다. 동사 서술어형 외에 두 번째 기본문형은 형용사 서술어형이다. 그러므로 중국인 학습자는 ‘NP이 Adj’인 문형을 상대적으로 자주 사용한다.

셋째, ‘NP이 V’의 중국인 학습자의 사용 빈도는 한국인 모국어 화자의 사용 빈도의 순위와 마찬가지로 3위이다. 이 유형의 빈도가 높은 이유는 중국어에는 자동사의 개념이 없지만 한국어를 처음 배우는 학습자들에게 일반적으로

13) 노은희(1999)에서는 구어와 문어를 나눠서 문형의 사용 빈도를 통계했다. 본고에서는 구어와 문어를 구분하지 않고 총체적인 사용 빈도를 살펴보았다.

14) 노은희(1999)에서 ‘NP’ 대신 ‘N’으로 표시되어 있는데 이때의 ‘N’은 명사구를 가리키므로 ‘NP’로 통일했다.

15) 중국어에서는 자동사와 타동사가 구별되지 않는다.

16) 중국어 문장의 주요성분은 ‘주어, 서술어, 목적어’인 반면에 한국어 문장의 주요성분은 ‘주어, 서술어, 목적어, 보어’이다.

로 단어를 외우면서 품사까지 외우게 하기 때문인 것 같다.

### 3.2. 구문 형성상의 특징

문장은 주어-서술어 관계가 한 번만 이루어진 홑문장(단문) 뿐만 아니라 주어-서술어의 관계가 두 번 이상 이루어진 겹문장(복합문)도 있다. 한국어 문장을 홑문장(단문)과 겹문장(복합문)으로 나누고 복합문은 다시 이어진 문장과 안긴문장으로 나뉜다. 안긴문장은 명사절로 안긴문장, 서술절로 안긴문장, 관형절로 안긴문장, 부사절로 안긴문장, 인용절로 안긴문장이 있다. 이어진 문장은 대등적으로 이어진 문장과 종속적으로 이어진 문장으로 나뉜다. 부사절로 안긴문장과 종속적으로 이어진 문장은 구별이 어려운 경우가 있다. 본고에서는 어미 형태를 기준으로 나누었다<sup>17)</sup>.

구문 분석을 통해 학습자가 어떤 문장 구조에 익숙한지 알 수 있고 문장을 만들 때 어떤 문제점이 존재하는지 찾을 수 있다. 중국인 학습자들이 작문에서 사용하는 문장 구조의 유형과 빈도는 아래와 같다.

〈표4〉 중국인 학습자의 문장 구조 빈도<sup>18)</sup>

문장 구조			빈도	백분율 (%)	사례	
홑문장			3877	47.2%	2년 전의 청도 여행은 가장 기억에 남는 여행이다.	
겹문장	이어진 문장	대등적 이어진 문장	-고	1464	17.8%	친구와 같이 오사광장에서 놀고 바닷가에서 사진을 찍었다.
			-지만	425	5.2%	
			기타 연결어미	338	4.1%	
	종속적 이어진 문장	-아/어서	1123	13.7%	처음에 기차를 타서 기분이 정말 기뻐다.	
		-(으)면	143	1.7%		
		기타 연결어미	342	4.2%		

17) 부사절로 안긴문장과 종속적으로 이어진 문장은 위치 이동의 용이성과 주어의 통제 여부 등에서 차이가 난다. 논란의 여지가 있겠지만 본고에서는 일단 '-어서, -니까, -(으)면, -(으)면서' 등의 연결어미가 있는 문장은 종속적으로 이어진 문장으로, '-게, -도록, -듯이, -리수록' 등이 있는 문장은 부사절로 안긴문장으로 보고 통계화했다.

18) 모어 화자의 문장 구조 양상이 어떻게 되는지와 비교해서 설명하면 설득력이 훨씬 더 있는데 선행 연구를 살펴봤는데 모어 화자의 문장 구조 양상에 대한 연구가 없어서 비교 연구가 이루어지지 못했다.

안긴 문장	명사절로 안긴문장	168	2.0%	우리는 마음을 놓고 자연의 아름다움을 느끼게 된다.
	서술절로 안긴문장	39	0.4%	내가 고등학교 때 성격이 내향적이었다.
	관형절로 안긴문장	214	2.6%	그때 나는 이 세상에는 이처럼 더 아름다운 경치가 없다고 생각했다.
	부사절로 안긴문장	57	0.7%	우리는 기분이 좋게 등산하기 시작했다.
	인용절로 안긴문장	22	0.3%	아저씨가 우리에게 아이스크림이 무엇이나고 물었다.

분석 결과 중국인 학습자가 사용하는 문장의 구조는 세 가지 특징을 가지고 있다. 첫째, 홀문장보다 겹문장의 사용 빈도가 더 높다. 학습자가 어미를 습득하면서 어미를 가지고 복잡한 문장을 만드는 경향이 있다. 둘째, 겹문장 중 안긴문장보다 이어진 문장을 더 많이 사용한다. 학습자가 한국어를 습득할 때 대등적 연결어미와 종속적 연결어미에 대한 이해가 전성어미나 부사형 어미, 인용형 어미보다 깊은 것으로 해석된다<sup>19)</sup>. 셋째, 연결어미 ‘-고’의 사용 빈도가 가장 높고 그 다음에 ‘-아/어서’를 많이 사용한다. 한국어에는 종속적 연결어미가 대등적 연결어미보다 수가 훨씬 많고 의미도 다양한데 학습자 문장에서 종속적 연결어미를 선호하는 특징은 보이지 않는다.

#### 4. 학습자의 구문 오류 분석

앞에서 구문 오류에 관한 표지는 주로 문법 오류를 분석한다고 했다. 문법 오류는 다시 조사의 오류와 어미 오류로 나뉘었다. 오류 유형별 빈도는 아래와 같다.

〈표5〉 중국인 학습자 오류의 빈도

순위	오류 유형		빈도	백분율(%)
1	조사 오류	조사의 대치(PR)	682	30.0%
2		조사의 생략(PO)	481	21.3%

19) 표를 통해서 알 수 있듯이 중국인 학습자의 대등적인 연결어미와 종속적인 연결어미의 사용은 다소 많고 부사형 어미와 인용형 어미로 만든 안긴 문장의 백분율은 적다고 볼 수 있다.

3	어미 오류	연결어미 오류(C)	251	11.1%
4		전성 어미 오류(D)	199	8.8%
5		선어말 어미 오류(T)	172	7.6%
6		종결어미 오류(En)	60	2.6%

위의 표에서 볼 수 있듯이 조사에 관한 오류의 빈도가 높다. 그 이유는 중국어에는 조사가 존재하지 않기 때문이다. 그 중에서 조사의 대치 오류(PR)가 가장 높다. 그 이유는 학습자가 조사의 의미와 기능을 충분히 인숙하지 않아 발생하기 때문이다. 어미 오류 중에서 연결어미의 오류(C)가 가장 높다. 그 이유는 한국어 학습에서 연결 어미가 차지하는 비중이 많다는 원인도 있고 연결어미의 문법적인 차이도 많기 때문이다. 가장 중요한 원인이 학습자가 연결 어미에 대한 인식이 부족하고 특히 유사한 연결 어미를 쉽게 혼동할 경우가 많기 때문이다.

#### 4.1. 조사 생략의 오류(PO)<sup>20)</sup>

조사 생략의 오류(PO)는 조사를 써야 할 자리에 조사를 쓰지 않고 비우는 경우이다. 한국어는 교착어로서 조사를 사용하는 경우가 많은 반면에 중국어는 독립어라서 조사가 존재하지 않는다. 중국어에 조사가 없다는 점뿐만 아니라 한국어에서 조사 생략이 빈번한 점 그리고 조사가 언제 생략되는지 규칙화하기 어려운 점 때문에 조사 생략의 오류는 중국인 학습자들에게 빈번하게 발생하는 오류이고 쉽게 고치지 못한다. 특히 ‘이/가’ ‘을/를’ ‘은/는’과 같은 조사에서 조사 생략의 오류가 많이 나타났다. 다음은 조사 생략 오류의 예들이다.

- (3) ㄱ. 다가오는 대학교 입시(√입시가 PO) 우리를 귀찮게 한다.  
 ㄴ. 내가 가장 기억에 남는 아르바이트(√아르바이트는 PO) 그 과일장있다(√과일 공장 아르바이트였다 R).  
 ㄷ. 이렇게 자신(√나 자신을 PO) 바꿔서 성공감이(√성취감이 LR<sup>21)</sup>) 크다.

20) 한국 사람은 구어에서 조사를 자주 생략한다. 본고에서 제시한 문장은 중국인 학습자의 쓰기 자료를 연구 대상으로 삼았기 때문에 구어가 아닌 문어에 속한다.

21) LR는 한국어에 없는 어휘로 대치하는 표지이다. 학습자 코퍼스에서 어휘 오류도 표지했는데 본고의 연구목적과 상관성이 없어서 앞에서 언급하지 않았다.

- ㄹ. 집에 돌아가는 것이 내 인생중(√인생에서 PO) 가장 중요한 소풍이다.
- ㄻ. 마음대로 놀고 \*가족들(√가족들과 PO) 함께 좋은 시간을 보냈다.
- ㄴ. 이제 내 인생관과 앞으로(√앞으로의 PO) 계획에 모두 도움이 된다고 여긴다.

위의 예문에서 보는 조사 생략 오류를 정리해서 표로 살펴보면 다음과 같다.

〈표6〉 조사 생략 오류 빈도

순위	오류 유형	격조사		빈도	백분율(%)
1	조사 생략 오류	주격조사	-이/가	154	32.0%
2		목적격조사	-을/를	129	26.8%
3		부사격조사	에	65	13.5%
4		관형격조사	의	51	10.6%
5		보조사	-은/는	40	8.3%
6		부사격조사	로	14	2.9%
7		공동격조사	와/과	13	2.7%
8		부사격조사	에게	9	1.9%
9		부사격조사	에서	6	1.2%

위의 표에서 볼 수 있듯이 중국인 학습자의 조사 생략 오류는 주격조사 ‘이/가’, 목적격조사 ‘을/를’, 보조사 ‘은/는’, 처격조사 ‘에/에서’, 공동격조사 ‘와/과’, 관형격조사 ‘의’ 등이 포함되어 있다. 그 중에 주격조사 ‘이/가’의 생략 빈도가 가장 높다. 조사 생략 오류는 조사가 없는 중국인 학습자들에게 빈번하게 발생하는 오류이다. 주격조사 ‘이/가’의 생략 빈도가 가장 높은 이유는 한국어와 중국어의 어순 차이로 발생한다. 일부 성분은 한국어에서는 주어로 사용되지만 중국어에서는 목적어로 사용되기 때문이다.<sup>22)</sup> 목적격조사 ‘을/를’의 빈도가 높은 이유는 한국어의 동사가 자동사와 타동사로 나뉘고 타동사 앞에는 목적격조사 ‘을/를’이 와야 하는데 학습자가 자동사와 타동사를 구별하지 못하고 또는 오류의 발생을 걱정해서 목적격조사 ‘을/를’을 생략하는 경우가 많다.

22) 예를 들어 ‘밖에 비가 오고 있다.’ ‘비’는 한국어에서는 주어인데 중국어에서는 목적어이다.

## 4.2. 조사 대치의 오류(PR)

조사 대치의 오류(PR)는 조사를 사용했으나 조사의 용법을 정확하게 사용하지 않은 경우이다. 조사의 의미와 기능을 충분히 이해하지 못했거나 품사를 정확하게 구별하지 못해서 발생하는 오류이다. 다음은 조사 대치 오류의 예들이다.

- (4) ㄱ. 나는(√내가 PR) 가장 기억에 남는 여행은 우리 반 친구(√친구들과 PO) 같이 태산에 간 여행이다.  
 ㄴ. 제주도는(√제주도의 PR) 자연(√자연의 PO) 아름다움은(√아름다움이 PR) 최고라고 생각했습니다.  
 ㄷ. 입기<sup>23)</sup>가(√입기는 PR) 여행지로 아주 좋은 곳이다.  
 ㄹ. 사람들은 살다보면 많은 기억을(√기억이 PR) 가슴에 남는다.  
 ㅁ. 아름다운 자연 풍경뿐만 아니라 제주도에 살고 있는 사람들의 친근감을(√친근감은 PR) 나에게 깊은 인상을 남겼습니다.  
 ㅂ. 지난 달에 우리반 학생들은(√학생들과 PR) 같이 즐겁게 놀았다.

위의 예문에서 보는 조사 대치 오류를 정리해서 표로 살펴보면 다음과 같다.

〈표7〉 조사 대치 오류 빈도

순위	오류 유형	조사 대치의 구체적인 분류		백분율 (%)
1	조사 대치 오류	보조사와 격 조사의 대치	주격조사 ‘이/가’를 보조사 ‘은/는’으로의 대치	23.1%
2			보조사 ‘은/는’을 주격조사 ‘이/가’로의 대치	13.4%
3			목적격조사 ‘을/를’ 보조사 ‘은/는’으로의 대치	4.3%
4		격조사 간의 대치	목적격조사 ‘을/를’ 주격조사 ‘이/가’로의 대치	6.1%
5			부사격조사 ‘로’를 ‘에’로의 대치	3.4%

23) 지명. 통사적으로 이 문장이 비문인 것은 아니다. 그러나 이 문장은 전후 맥락을 고려하면 ‘입기에 대하여’ 설명하는 내용이다. 이 경우 ‘주제’를 나타내는 ‘-은’을 사용하는 것이 한국어다운 용법이라서 오류로 처리했다.

6			부사격조사 ‘에게서’를 ‘에서’로의 대치	1.2%
7		보조사 간의 대치	보조사 ‘도’를 ‘은/는’로의 대치	0.4%

중국인 학습자의 조사 대치의 오류를 보면 보조사와 격조사의 대치 오류 빈도가 높다. 특히 주격조사 ‘이/가’와 보조사 ‘-은/는’의 대치 빈도가 가장 높다. 한국어를 배우기 시작할 때부터 주격조사 ‘이/가’와 보조사 ‘은/는’의 구별이 힘들기 때문이다. 위에서 알 수 있듯이 중국인 학습자에게는 ‘이/가’와 ‘은/는’의 의미와 기능에 대한 정확한 이해가 어렵다는 것을 알 수 있다. 다음으로 목적격조사 ‘을/를’과 주격조사 ‘이/가’와 보조사 ‘은/는’의 대치의 오류도 상당히 비율이 높다. 그 이유는 학습자가 한국어 품사 특히 자동사와 타동사의 구분을 정확하게 하지 못하기 때문이다. 격조사 간의 대치 오류 중에서 목적격조사 ‘을/를’을 주격조사 ‘이/가’로 대치하는 비율이 다소 높게 나왔다. 그 이유도 마찬가지로 학습자가 자동사, 타동사, 형용사의 품사를 구별하지 못하기 때문에 발생하는 오류이다. 부사격조사 ‘로’를 ‘에’로 대치하는 오류는 두 가지 조사의 구체적인 목적지와 방향을 표현하는 차이를 이해하지 못해서 발생하는 오류이다. 보조사 간의 대치 오류는 상대적으로 적게 나왔다. 보조사 ‘도’를 ‘은/는’으로 대치하는 오류는 보조사 ‘도’의 포함한다는 의미를 정확하게 이해하면 해결될 것이다.

### 4.3. 어미 오류(E)

어미는 조사와 같이 한국어의 문장 형성에서 높임법과 시제와 같은 문법 범주를 표시하고 문장을 종결하거나 연결하는 등 중요한 기능을 담당하고 있다. 위치에 따라 어미는 선어말 어미와 어말 어미로 나뉘고 어말어미는 기능에 따라 연결 어미, 종결 어미, 전성 어미로 나뉜다. 어미 오류에서는 다른 오류보다 어미를 생략하거나 다른 어미로 대치하는 경향이 두드러진다. 어미 오류의 구체적인 유형과 예문은 아래와 같다.

〈표8〉 어미 오류의 구체적인 유형과 오류 빈도

유형	구체적인 분류		사례	백분율	
어미 오류 (E)	선어 말 어미	존대어미오류 (H)	어미의 생략(HO)	사장님께 아주 미안하다고 생각합니다. 다른 사람에게 이 일을 부탁하기(√/부탁하시기 HO)를 바랍니다.	2.3%
			다른 어미로의 대치(HR)	선생님이 도와 주셔서 정말 감사 하셨습니다(√/감사했다 HR).	0.1%
		시제 어미 오류(T)	어미의 생략(TO)	나는 동아리에서 탈퇴하고 열심히 공부하기로 정한다(√/정했다 TO).	12.7%
			다른 어미로의 대치(TR)	오전 11시에 타산에 도착하고 먼 저 밥을 먹겠다(√/먹었다 TR).	4.2%
	종결 어미 오류(En)	어미의 생략(EnO)	산동성에 산이 많은데 예를 들어 태산, 노산(√/태산, 노산이 있다 EnO).	0.6%	
		다른 어미로의 대치(EnR)	쓰기가 한국어 능력시험에서는 굉장히 중요해요. 그런데 어떤 학생들이 이것을 소홀히 해요. 그래서 우선 *하는(√/할 DR) 일이 *학생들을(√/학생들이 PR) 쓰기에 대한 좋은 태도를 가지게 하는 *것이다(√/거예요 EnR).	6.3%	
	연결 어미 오류(C)	어미의 생략(CO)	드디어 원한다(√/원하는 DR) 대학교에 입학하(√/입학해서 CO) 축하하려고요.	6.3%	
		다른 어미로의 대치(CR)	여행 때 소매치기를 당하거든(√/당하면CR) 누구에게 도움을 구하지요?	16.3%	
	전성 어미 오류(D)	어미의 생략(DO)	엑스포는 첫 번째로 중국에서 열려서 중국인(√/중국인인 DO) 난 편하게 관람할 수 있으니까 영광이었다.	6.1%	
		다른 어미로의 대치(DR)	제에게 가장 기억에 남는 여행이 가족들과 임기시에 가는(√/간 DR) 것이다.	21.7%	

중국인 학습자의 작문에서 어미의 오류에서 보이는 특징은 세 가지이다. 첫째, 총체적으로 보면 전성 어미의 오류가 가장 높아서 중국인 학습자가 용

언의 활용에 많은 어려움을 겪고 있음을 알 수 있다. 이는 중국인 학습자가 한국어의 관형형 어미에 나타나는 시제의 복잡성과 명사형 어미 ‘-는 것’과 ‘-기’의 차이를 쉽게 이해하지 못하기 때문이다.

둘째, 어미의 대치 오류가 가장 많다. 전성 어미의 대치와 연결 어미의 대치 오류가 가장 많은 것으로 나타났다. 어미의 대치 오류가 많은 것은 한국어가 교착어이기 때문에 어미의 표현이 풍부하면서 문법적인 정확성을 필요로 한다는 특징이 있다. 그 중 가장 중요한 원인은 학습자가 어미에 대한 인식이 부족하고 특히 유사한 어미를 쉽게 혼동하는 경우가 많기 때문이다.

셋째, 어미 생략 오류는 대치 오류에 비해 낮은 편인데 그 중 시제 어미의 생략 빈도가 상대적으로 높다. 중국인 학습자가 시제 어미를 생략하는 것은 중국어에서는 시제를 단어로 표현하고 시제 어미는 존재하지 않기 때문이다.

## 5. 맺음말

지금까지 중국인 한국어 학습자의 구문 특징을 살펴보았다. 총체적으로 보면 중국인 학습자의 기본문형 사용이 모어 화자와 비슷한 것으로 나타났다. 이는 중국어와 한국어 모두 ‘주어, 목적어, 서술어’ 등 문장 성분을 가지고 있지만 중국어에는 조사와 어미가 존재하지 않기 때문에 중국인 학습자가 한국어 조사와 어미를 습득할 때 많은 어려움을 겪고 있는 것이다. 먼저 중국인 학습자의 기본문형을 분석해서 정리하면 아래와 같은 특징이 있다.

첫째, 기본문형의 사용 빈도가 높은 순서는 ‘NP이 NP을 V’ > ‘NP이 Adj’ > ‘NP이 V’ > ‘NP이 NP이다’의 순이다. 그 이유는 중국어의 기본문형도 한국어와 같이 ‘NP이 (NP을) V’, ‘NP이 Adj’, ‘NP이 NP이다’의 순서이기 때문이다.

둘째, 중국인 학습자가 사용하는 문장 구조를 보면 홀문장보다 겹문장의 사용 빈도가 더 높다. 겹문장 중 안긴문장보다 이어진 문장을 더 많이 사용한다.

중국인 학습자의 오류를 분석한 결과는 아래와 같다.

첫째, 조사와 어미에 대한 오류가 가장 많이 발생한다. 특히 조사와 어미의 생략과 대치 오류가 가장 많다.

둘째, 조사 오류에서 조사 생략 오류가 조사 대치 오류보다 더 많다. 조사 오류는 조사가 없는 중국인 학습자들에게 빈번하게 발생하는 오류이다.

셋째, 어미 오류에서는 전성 어미의 오류가 가장 높다. 그중 전성 어미의 대치 오류가 가장 많고 연결 어미의 대치 오류도 많은 것으로 나타났다. 어미 생략 오류는 낮은 편인데 그 중 시제 어미의 생략 오류가 두드러진다.

본고에서는 소규모의 학습자 코퍼스를 구축해서 중국인 한국어 학습자의 구문 특징을 간단하게 살펴보았다. 그러나 코퍼스의 규모가 작아서 분석 결과의 정확도에 신뢰가 가지 않는다. 앞으로 더 많은 자료를 수집해서 규모와 정확성 및 신뢰성을 보완할 것이다. 또한, 본고는 학습자를 수준별로 나누어 연구를 진행하지 못한 점에서 한계가 있다. 이후에 학습자 수준별로 연구를 진행함으로써 좀 더 정밀한 분석을 추가로 보완할 것이다. 본고에서의 여러 가지 문제점은 앞으로의 과제로 남기기로 한다.

## 〈참고문헌〉

- 강범모·김홍규(2009), 『한국어 사용 빈도』, 한국문화사.
- 고석주(2004), 『한국어 학습자 말뭉치와 오류 분석』, 한국문화사.
- 구재희(2007), 『한국어 기본문형 교육 연구』, 이화여자대학교 대학원 박사학위논문.
- 김선화(2005), 『한국인 학습자에게서 나타나는 문형별 오용에 관한 연구』, 단국대학교 교육대학원 석사학위논문.
- 김윤자(2010), 『중등학교 영어 학습자들의 논항 구조 습득 연구: 여격 구문을 중심으로』, 한국교원대학교석사학위논문.
- 김충실(2009), 『중국인 학습자를 위한 한국어 '를' 구문 교수방법 연구』, 부산외국어대학교 박사학위논문.
- 김혜숙(1998), 『한국어 기본 문형 설정에 대하여』, 『국어국문학』.
- 남기심·고영근(2006), 『표준국어문법론 개정판』, 탑출판사.
- 노은희(1999), 『한국어 교재 개발을 위한 한국어의 문형 빈도 조사』, 『국어교육연구』 6.
- 박성숙(2010), 『외국인을 위한 한국어 어순교육 방안』, 경희대학교 교육대학원 석사학위논문.
- 송창선(2011), 『국어의 기본 문형 설정』, 『국어교육연구』.
- 신서인(2006), 『구문 분석 말뭉치를 이용한 한국어 문형 연구』, 서울대학교 대학원 박사학위논문.
- 안광모·서영훈(2011), 『구문분석의 중의성 해소를 위한 전처리 규칙의 확장』, 『한국엔터테인먼트산업학회논문집』.
- 음승민(2012), 『중국인 한국어 학습자의 구문 발달에 대한 양적 연구: 학습 단계 향상에 따른 문장의 구조 및 길이변화를 중심으로』, 고려대학교 교육대학원 석사학위논문.
- 이승연(2007), 『한국어 교육을 위한 학습자 말뭉치의 구축과 활용 연구』, 고려대학교 대학원 박사 학위 논문.
- 이익섭(2006), 『한국어 문법』, 서울대학교출판부.

조남호(2003), 『한국어 학습용 어휘 선정 결과 보고서』, 국립국어연구원.  
베이징언어학원문형연구팀(1989), 『現代漢語基本句型』, 世界漢語教學 제1호.

## 【국문초록】

본 연구는 규모가 작은 중국인 한국어 학습자 코퍼스를 구축하고 이에 기초하여 학습자의 구문 특징을 분석하는 데 목적이 있다. 한국어 학습자의 구문 특징을 분석하기 위해서 한국어 기본문형을 설정하고 학습자들의 기본문형 사용 양상을 분석했다. 이와 함께 중국인 학습자의 기본문형 사용 빈도와 한국인 모국어 화자의 기본문형 사용 빈도를 비교해서 논의하였다. 그 결과 중국인 한국어 학습자의 기본문형의 사용 빈도는 ‘NP이 NP을 V’> ‘NP이 Adj’> ‘NP이 V’> ‘NP이 NP이다’의 순으로 나타났다.

구문 관련 오류는 주로 세 가지 특징을 지닌 것으로 드러났다. 첫째, 조사와 어미의 생략 오류와 대치 오류가 많이 나타난다. 둘째, 조사 오류에서는 조사 생략 오류가 조사 대치 오류보다 빈도가 높았다. 어미 오류에서는 연결 어미의 대치 오류가 가장 많았다. 셋째, 어순 오류는 생각보다 많이 나타나지 않았다.

주제어 : 학습자 코퍼스, 구문 특징, 기본 문형, 구문 오류

【Abstracts】

## A study on syntactic characteristics of Chinese Korean-learner's writings based on small learner corpus

Gaojie

This study aims to analyze the students' syntax features based on Small Chinese Korean learner corpus. The basic sentence patterns were formed for these purposes and the distinctive features of using these basic sentence patterns were analysed. It was additionally examined how frequently Chinese learners and Korean native speakers use these basic sentence patterns respectively. As the result of this successive studies, Chinese learners most preferably use 'NP+NP+V' pattern and it was followed by 'NP+Adj', 'NP+V', and 'NP+NP' patterns.

Three major characteristics of the sentence errors are able to be indicated like the followings: First, the learners made most of the errors in using postpositions and endings. Second, the learners mainly omitted the postpositions rather than incorrect using of them. As for the endings, the adequate words were not used in most cases. Third, the errors in the arrangement of words did not happen so much than expected.

Key words : learner corpus, syntactic characteristics, basic korean sentence patterns, syntactical errors

이 논문은 2013년 6월 30일에 투고되었으며, 2013년 7월 26일에 심사 완료되어 2013년 8월 6일에 게재가 확정되었음.

