

L1, L2 스페인어 접속법 사용에 대한 변이주의적 접근*

황지희

계명대학교

황지희(2023), L1, L2 스페인어 접속법 사용에 대한 변이주의적 접근, 이베로아메리카연구, 34(3), 57-86.

초록 스페인어의 법(mood)은 크게 직설법과 접속법으로 나뉘며, 이에 관한 연구에서는 통사적이거나 의미적인 요소들이 두 법의 대조와 관련이 있다고 주장됐다. 제2언어(L2)로서의 스페인어 연구에서는 L2 학습자들이 스페인어 법 대조 습득에 많은 어려움을 겪으며, 이에 따라 법 대조는 스페인어 학습의 후반기에 습득되는 언어 구조라고 받아들여 왔다. 이러한 주장과 가정들은 주로 접속법의 형태, 통사, 의미, 또는 화용적인 자질에 초점을 맞춘 기존의 연구들을 통해 확인됐다. 그러나 최근의 사회언어학 연구들은 실제 언어 사용 환경에서 발견되는 스페인어의 법 대조 현상에는 이보다 더욱 복잡하고 다양한 요인들이 작용한다는 것을 보여주었다. 본 연구는 이러한 논의를 기반으로 하여 L1과 L2 스페인어 법 대조를 변이주의(variationist) 방법론을 사용하여 분석했다. 특히, 일부 선행 연구에서 상반되는 결과가 도출된 동사의 형태 규칙성에 대한 변수를 고려하여 분석을 진행했다. 스페인어 작문 코퍼스인 CEDEL2에서 추출한 L2 스페인어 학습자와 L1 스페인어 원어민 화자의 작문 데이터를 비교한 결과, 규칙 변화 동사가 종속 명사절에서 접속법으로 더욱 자주 사용되며, 주절이 부정문으로 작성되거나 종속절 동사가 저빈도 동사 그룹에 속할 때 접속법이 더욱 자주 사용된다는 경향이 확인되었다. 본 연구는 변이주의 접근법을 통해 L2 스페인어 법 대조에 대한 새로운 통찰력을 제공하고자 시도했다는 점에서 의의가 있다.

핵심어 스페인어 접속법, L2 스페인어 법 대조, 작문 코퍼스, L2 변이주의, 동사 형태 규칙성

* 이 논문은 2022년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-과제번호)(NRF-2022S1A5B5A17038767)

I. 들어가는 말

본 연구는 변이주의 사회언어학(Variationist Sociolinguistics)을 활용하여 다양한 모국어 배경(L1 background)을 가진 L2 스페인어 화자가 종속 명사절에서 스페인어의 직설법(indicative)과 접속법(subjunctive)의 대조를 어떻게 습득하고 사용하는지에 대해 새로운 관점에서 분석을 제시하는 것을 목표로 한다. 스페인어의 법(mood)은 직설법과 접속법으로 구분되는데, 선행 연구들에서는 두 법의 대조에 통사적 또는 의미적 요소들이 밀접한 관련이 있다고 주장됐다. 그러나 비교적 최근의 사회언어학 연구에서는 더욱 다양한 요소들이 직설법과 접속법의 구분에 영향을 미친다는 것을 실험을 통해 실증적으로 분석하고 있다. 현재까지 한국에서 진행된 제2언어로서의 스페인어 습득 연구에서 이러한 변이주의적 및 사회언어학적 방법론의 활용이 부족했던 점을 고려하여, 본 연구에서는 대형 코퍼스를 통해 L2 화자들의 작문 자료를 수집하고 분석하여 기존 연구에서 다루어지지 않은 부분을 보완하는 것은 물론이고 통계 분석을 통해 본 연구의 보편성 역시 확보하고자 한다.

스페인어의 접속법은 주로 명사절, 형용사절, 부사절에서 활용된다. 명사절에서 활용되는 접속법의 경우, ‘주절 + *que* + 종속절’의 구조에서 종속절의 동사가 접속법으로 나타나게 된다. 기존에 이루어진 연구들과 스페인어를 외국어로 배우는 학습자들을 위한 교재 및 교과서들에서는 주절 동사의 의미에 따라 종속절 동사가 직설법 또는 접속법으로 결정되어 둘 중 하나의 형태로 쓰여야 한다고 설명한다. 실제로 많은 구문이 이러한 문법적 내용으로 설명될 수 있는 것은 사실이지만, 같은 주절 동사를 지님에도 불구하고 종속절에서 직설법과 접속법을 모두 활용할 수 있는 구문들 역시 존재한다. 즉, 주절 동사의 의미라는 단순한 기준으로 종속절의 동사가 직설법 또는 접속법으로 의무적으로 구분하여 사용되어야 한다는 설명은 이보다 더욱 복잡하고 다양한 형태로 실생활에서 발견되는 스페인어의 법 대조(mood distinction) 현상을 이해하기에는 무리가 있다. 이러한 문제점을 해결하기 위해서는 더욱 다양한 관점에서 법 대조를 분석할 필요가 있으며, 언어학적 변이(variation)를 보이는 현상인 스페

인어의 법 대조를 분석하기에는 변이주의 방법론이 적합할 것이라고 기대할 수 있다. 또한, 기존에 이론적으로 또는 경험적으로 알고 있었던 스페인어의 접속법에 대한 지식이 대량의 언어 데이터의 통계적인 분석을 통해 실증적으로 검증이 되는지 또는 새로운 상반되는 결론이 도출되는지를 확인하는 것은 L2 스페인어 접속법 연구에 필요한 과정이라고 판단된다.

본 연구에서는 스페인어 작문 코퍼스인 CEDEL2(*Corpus Escrito del Español como L2*, L2 Spanish Written Corpus) (Lozano 2022)에서 제공되는 L2 스페인어 학습자들의 작문 데이터와 L1 스페인어 원어민들의 작문 데이터 중 종속 명사절이 활용된 구문들을 추출하여 살펴본다. 종속 명사절 동사의 법을 대조적으로 사용하는 데 영향을 미칠 것으로 예상되는 다양한 언어적, 그리고 언어외적 요인들을 통계 분석 툴(tool)인 Rbrul(Johnson 2009)을 활용하여 분석한다. 이를 통해, 스페인어의 종속 명사절에서 사용되는 동사의 법을 예측하는데 어떠한 요인이 얼마만큼의 확률로 상관관계를 맺는지 그 함의를 파악하는 것이 본 연구의 주요 내용을 구성한다.

II. 스페인어의 법: 직설법 vs. 접속법

스페인어의 법(mood)은 크게 직설법(indicative)과 접속법(subjunctive)으로 구분된다. 전통적으로, 일부 학자들은 직설법과 접속법의 대조는 통사적으로 결정된다고 주장했으며(Bello 1847/1982; Lakoff 1968), 다른 학자들은 법의 선택은 의미적으로 이루어지는 언어 현상이라고 설명했다(Guitart 1982; Terrell and Hooper 1974). 그러나 Silva-Corvalán(1994), Gudmestad(2010, 2012)와 같은 일련의 사회언어학 연구에서는 스페인어의 법 대조(mood distinction)에 영향을 미치는 요소에는 기존의 연구들에서 주장한 통사적, 의미적 요인들뿐만 아니라 동사의 형태 규칙성(form regularity), 시제(time reference), 실험 과업의 종류(type of task) 등과 같은 더욱 다양한 요인들이 포함된다고 지적한다.

스페인어의 접속법은 주로 명사절, 형용사절, 부사절에서 활용된다. 접속법

을 통해 전달되는 가장 두드러지는 의미는 비실제성(irrealis) 및 불확실성(uncertainty)으로 알려져 있다(cf. Haverkate 2002). 본 연구의 초점이 될 명사절의 경우, 아래 예문에서 볼 수 있듯이 주절 동사의 의미나 종속절 내용의 (비)실제성과 (불)확실성에 따라 접속사 *que*를 통해 연결되는 종속절에 내포된 동사의 법이 직설법 또는 접속법으로 결정된다고 주로 설명된다.

- (1) a. Sé que tienen/*tengan razón. 나는 그들이 옳다는 것을 안다.
 b. Quiero que coman/*comen más. 나는 그들이 더 먹기를 원한다.
 (Perez-Cortes 2020, 150)

위의 (1a)는 주절의 *saber* ‘알다’ 동사가 발화의 순간에 화자가 사실이라고 단정하고 있는 내용을 표현하고 있음을 나타내며, 종속절의 ‘그들이 옳다’는 것에 대한 확실성을 나타내고 있기 때문에 *que* 이하의 동사는 직설법인 *tienen*의 형태로 쓰여야 한다고 설명할 수 있다. 반면, (1b)의 경우, 화자가 아직 실제로 이루어지지 않은, 즉 실제성이 결여된 ‘그들이 더 먹는 것’이라는 내용이 추후에 이루어질 것을 원하고 있는 상황이기 때문에 종속절의 동사는 접속법인 *coman*의 형태로 사용되어야 한다고 볼 수 있다.

앞서 살펴본 예문 (1)과 같이 주절에서 사용된 동사의 어휘적인 의미로 인해 종속절에 표현되는 내용의 (불)확실성과 (비)실제성이 결정되고, 이에 따라 종속절 동사의 법 역시 의무적으로 선택되는 것으로 보이는 구문들도 있지만, 아래와 같이 주절 동사와는 관계없이 종속절에서 직설법과 접속법이 모두 사용될 수 있는 경우도 존재한다.

- (2) a. El jefe les dice a los empleados que trabajan bien.
 사장은 직원들에게 그들이 일을 잘 한다고 말한다.
 b. El jefe les dice a los empleados que trabajen bien.
 사장은 직원들에게 일을 잘 하라고 말한다.
 (Perez-Cortes 2020, 150)

(2)의 두 예문의 경우, 주절에 쓰인 동사는 *decir* ‘말하다’로 동일하지만,¹⁾ 종속

절의 동사가 직설법과 접속법 중 어느 형태로 쓰이는가에 따라 전체 문장이 표현하는 의미가 전혀 다르게 된다. 종속절에 직설법이 사용된 (2a)의 경우, 사장이 ‘직원들이 일을 잘 한다’는 사실을 보고하는 의사소통적 기능을 수행한다면, 접속법이 사용된 (2b)는 ‘직원들이 (앞으로) 일을 잘 할 것’을 간접적으로 명령하는 기능을 수행한다고 볼 수 있다.

한편, Igualada Belchí(1989) 역시 예문 (2)와 마찬가지로 주절의 동사가 일종의 트리거(trigger)가 되어 자동으로 종속절의 법을 결정하지 않는, 즉 종속절에 직설법의 사용과 접속법의 사용이 모두 문법적으로 가능한 구문을 보여준다.

- (3) No dudo que {viene/venga}. 나는 그 사람이 올 것임을 의심하지 않는다.

(Igualada Belchí 1989, 661)

예문 (3)의 주절 동사 *dudar*는 ‘의심하다’는 뜻을 갖고 있는 동사로, (3)과 같이 부정문에서 사용되면 ‘의심하지 않다’, 즉 확실성을 표현하게 된다. 이때, 주절에서 드러난 확실성이 트리거로 작용하여 자동적으로 종속절에서 직설법을 사용하는 것이 적합하다고 설명되는 경우가 많은데, 예문에서 보이는 것처럼 직설법뿐만 아니라 접속법의 사용 역시 문법적으로 가능하다. 저자는 이 교호성을 종속절 내용에 대한 인정 또는 믿음의 정도의 차이로 설명한다. 종속절에서 직설법이 사용된 *No dudo que viene*는 대용형(proform)을 활용하여 *Lo acepto*로 바꾸어 쓸 수 있으며, 이는 ‘나는 그것을 인정한다/믿는다’의 의미를 나타낸다. 반면, 접속법을 사용한 *No dudo que venga*는 *Lo acepto como posible* ‘나는 그것이 가능하다고 인정한다/믿는다’를 표현하는 문장으로, 종속절에 직설법을 사용한 문장보다는 인정 또는 믿음의 정도가 덜하다는 것을 보여준다. 정리해

1) 익명의 심사자께서 지적하신 대로, 예문 (1)에 사용된 *decir* 동사가 (1a)에서는 ‘decir 1: 약한 명령의 말하다’로, 그리고 (1b)에서는 ‘decir 2: 의사소통의 말하다’로 사용되어 서로 차이를 보이는 것으로 제시해야 한다고 볼 수 있다. 그러나 본 연구에서는 (1a)와 (1b) 예문에서 *decir*가 궁극적으로는 공히 ‘말하다’라는 행위를 표현하고 있다는 점을 고려하여 이를 서로 다른 주절 동사인 것으로 구분하지 않았으며, 데이터에 포함된 다른 주절 동사 역시 같은 기준으로 분석하였음을 밝힌다.

보면, 예문 (3)의 종속절에 사용된 직설법과 접속법의 차이를 확실성의 정도의 차이로 설명하는 것은 가능하지만, 확실성을 기준으로 두고 직설법과 접속법이 상보적 또는 배타적 분포(complementary distribution) 관계에 있는 것은 아님을 알 수 있다.

위의 (2-3)와 같은 구문들은 주절 동사의 의미가 종속절의 법 선택에 영향을 미친다는 설명을 통해 이해하기에는 어려움이 있다. 그러나 스페인어를 외국어로 학습하는 제2언어 화자들(L2 learners)에게 가장 빈번하게 제공되는 설명은 앞서 본 (1)과 같은 예문을 활용하여 주절 동사의 종류에 따라 종속절에 직설법 또는 접속법을 대조적으로 사용해야 한다는 것에 그치는 경우가 많다. 다시 말하면, 예문 (2-3)과 같이 직설법과 접속법이 모두 활용될 수 있는 구문에서 학습자들이 각 법의 출현이 어떠한 문맥적 차이를 수반하는지에 대해 이해할 수 있도록 도움을 줄 수 있는 적절한 설명이 부족한 현실이다. 특히 예문 (3)과 같은 경우는 스페인어 교재나 교실 수업 내에서 다루어지지 않는 경우가 대부분이기 때문에 L2 화자들에게 직접적으로 노출되지 않아 종속 명사절에서의 접속법 교호 여부조차 알지 못하는 학습자들이 많다.

III. L2 스페인어 법 대조 습득의 변이성(variability)

L2 학습자들이 스페인어 법 대조 습득에 많은 어려움을 겪고, 이에 따라 법 대조는 스페인어 학습의 후반기에 습득되는 언어 구조(late-acquired linguistic structure)라는 사실은 기존 연구들을 통해 잘 알려져 있다(cf. Stokes 1988). 학습자들이 스페인어에 노출되는 기간이 길어지고 스페인어를 공부한 시간이 늘어남에 따라, 그들은 다양한 경로를 통해 스페인어의 법 대조를 서로 다른 수준으로 습득하게 된다. 교실에서 배우는 내용은 물론, 원어민과의 의사소통이나 각종 미디어를 통해 명시적으로(explicitly) 또 암묵적으로(implicitly) 직설법과 접속법의 사용 양상을 습득하게 되는 것이다. 습득의 경로가 다양한 만큼 그 결과 역시 변이성을 보이는데, 이렇게 학습자의 목표어(target language) 생산

(production)에서 목표어 유사(target-like) 양상과 그렇지 않은(non-target-like) 양상이 발견되는 것을 제 I 변이(Type I variation)라고 지칭한다. 또한, 스페인어의 법 대조는 이러한 제 I 변이 외에도 특정 언어 형태에 대해 원어민 사이에서 발견되는 변이성을 다루는 제 II 변이(Type II variation) 역시 보여준다. Gudmestad(2012)에 따르면, 스페인어 원어민들에게 있어서 직설법과 접속법의 대조는 가변적(variable)이고 두 형태 중 직설법의 사용 비율이 점점 더 증가하는 방향으로 변화하고 있으며, 이러한 변이적인 법의 선택은 다양한 요인의 영향을 받는다는 것이 여러 사회언어학 연구들을 통해 밝혀져 왔다.

이렇게 제 I 변이와 제 II 변이 두 가지 모두를 보이는 언어 현상인 스페인어의 법 대조가 어떻게 습득되는가를 연구하는 데 있어서 사회언어학적, 특히 변이주의적(variationist) 분석이 중요한 역할을 할 것이라고 기대할 수 있다. 그러나 대부분의 선행 연구는 학습자들이 법의 대조를 문법적으로 올바르게 이해했는가, 즉 문맥에 따라 종속절에 직설법과 접속법을 구분하여 선택할 수 있는가에 초점을 맞춰 왔기 때문에 실제로 학습자들이 두 형태를 사용하는 양상에 대한 분석은 충분히 이루어지지 않았다. 따라서 본 연구는 스페인어의 L2 화자들이 직설법과 접속법을 활용하는 데에 영향을 미칠 수 있지만 기존에 충분히 다루어지지 않았던 요인 및 변수들을 중심으로 학습자들의 작문 데이터를 양적으로 분석하고자 한다.

본격적으로 본 연구의 주요 내용을 살펴보기에 앞서, 기존의 L2 스페인어 접속법 습득 연구는 어떻게 진행되어 왔는가를 간단히 정리해보고자 한다. 우선, 접속법의 교육과 관련한 내용을 다룬 Farley(2001)의 경우, L2 스페인어 학습자들이 종속절 동사의 법을 기준으로 주절에 사용된 동사가 불확실성(*Yo dudo que...* ‘나는 ...라는 것을 의심한다’)을 나타낼지 또는 확실성(*Es verdad que...* ‘...라는 것은 사실/진실이다’)을 나타낼지를 구분하여 선택할 수 있는지를 밝혀내고자 했다. 이때, (입력) 처리 교수법((Input) Processing Instruction, 이하 PI)과 의미 중심 출력 교수법(Meaning-based Output Instruction, 이하 MOI) 중 어떠한 교육 방법이 실험 과업에 대한 학습자들의 수행 능력에 더욱 긍정적인 영향

을 주는지를 중점적으로 살펴보았다. 결론적으로, MOI보다는 PI를 기반으로 한 수업에 참여했던 학습자들이 더 높은 정답률을 보였기 때문에 저자는 접속법과 같이 복잡한 구조를 가르칠 때는 “(a) 한 번에 한 기능/용법씩, (b) 의미에 초점을 두고, (c) 문장 단위에서 시작하여 점차 담화 단위로, (d) 구어(oral)와 문어(written) 입력(input) 모두를 활용하여, (e) 학습자들이 입력 받은 내용을 직접 활용해보도록 하며, (f) 학습자의 입력 처리 전략을 염두에 두는 (VanPatten 1993, Farley 2001, 289에서 재인용)” 입력 중심의 교수법이 효과적임을 강조했다. 이 연구는 효과적인 접속법 교육 방법을 제시했다는 점에서 의의가 있지만, ‘주절의 확실성+종속절의 직설법’과 ‘주절의 불확실성+종속절의 접속법’이라는 일종의 전형적인 패턴을 유일한 정답으로 설정하고 이를 기준으로 L2 화자들의 학습 정도를 파악했다는 한계를 보인다. 다시 말하면, 앞서 살펴보았던 예문 (3)과 같이 주절에서 확실성이 드러난다고 하더라도 종속절에 접속법이 사용된 구문(*No dudo que venga*. ‘나는 그 사람이 온다는 것을 의심하지 않는다’)들은 Farley(2001)의 기준으로는 오답이자 비문법적인 문장으로 취급되었다는 점에서 아쉬움이 남는다.

한편, Collentine(2010)은 Collentine(2003) 이후로 진행된 L2 스페인어 접속법의 습득과 교육 관련 연구들을 총정리했다. 보편문법(Universal Grammar, UG)의 관점과 심리언어학적(psycholinguistic) 관점에서의 연구들을 비교 분석한 이후, 학습 환경의 측면에서 입력 중심, 출력 중심, 그리고 과업 중심 연구들을 두루 살펴보았다. 주제별 대표 연구들의 내용 정리를 통해 발견한 기존 연구의 도전 과제(research challenges)를 제시했는데, “L1의 역할과 L1 전이, 담화 화용론 중심의 접근, 음운론 중심의 접근, 접속법 사용에 영향을 미치는 문법-어휘적 특징, 해외 연수의 영향과 역할(Collentine 2010, 47)”을 특히 앞으로의 연구들에서 다루어야 할 주제로 꼽았다. 주목할만한 점은, 비록 저자가 온라인 스페인어 코퍼스인 Corpus del español(Davies 2016-)에서 분석한 2,085,990개의 동사 형태 중에서 오직 7.2%만이 접속법 형태로 사용(Collentine 2010, 39) 되었을 정도로 접속법은 실제 스페인어 사용 환경에서 굉장히 미미한 역할을

하고 있지만, L2 스페인어 습득이라는 분야에서 접속법은 여전히 중요한 주제라고 결론을 내린 부분이다. 이는 접속법이 지닌 통사적, 의미적, 화용적 복잡성이 L2 스페인어 습득 과정을 이해하는 데 중요한 통찰(insight)을 제공할 수 있기 때문이라고 볼 수 있겠다.

L2 스페인어 법 대조 습득을 변이주의적 관점에서 다루고 있다는 측면에서 본 연구와 가장 큰 유사성을 지니고 있는 Gudmestad (2006, 2012)는 일련의 언어적 요인(Gudmestad 2006, 2012)과 언어외적 요인(Gudmestad 2012)들을 설정하여 어떠한 요인이 L2 스페인어 화자의 접속법 선택 및 사용을 예측할 수 있는 주요 변수로 작용하는가를 살펴보았다. Gudmestad(2006)에서는 종속절에서 사용된 법(mood)이 서로 다른 두 문장을 제시한 후(*Quiero que viajamos en Italia, en España o en México* vs. *Quiero que viajemos en Italia, en España o en México* ‘나는 우리들이 이탈리아, 스페인, 또는 멕시코에서 여행하기를 원한다’), 실험 참가자들이 둘 중 어떠한 문장을 선호하는지 또는 두 문장 모두 적합하다고 생각하는지를 확인하였다. 이때, 종속절 동사의 형태 규칙성(form regularity), 문장 내용의 미래성(futurity), 주절 동사가 원망(desire) 동사인지 아닌지, 그리고 주절 동사가 감정(emotion)을 표현하는지 아닌지를 독립 변수로 설정하여 응답을 분석했다. 실험 참가자의 스페인어 실력(중급 vs. 고급)에 따라 결과의 차이가 있었지만, 두 그룹에서 공히 발견된 결과는 종속절에서 불규칙 변화를 하는 접속법 동사가 사용되었을 때 접속법에 대한 선호가 유의미하게 드러났다는 것이다. 이러한 결과를 바탕으로 저자는 후속 연구(Gudmestad 2012)에서 실험 참가자의 스페인어 실력과 실험 과업의 종류를 2개의 언어외적 변수로 설정하고, 동사의 형태 규칙성, 주절 동사 의미군(semantic category), 시제(time reference), 그리고 가상성(hypotheticality)이라는 4개의 요인을 언어적 변수로 설정했다. Gudmestad(2006)에서 L2 화자들의 언어 생산(production)을 확인하지 못하고 주어진 문장에 대한 그들의 선호도만을 파악하는 데 그쳤다는 한계를 보완하여 Gudmestad(2012)에서는 L2 화자들의 구어 생산(oral production)을 녹음한 후 전사(transcribe)하여 분석했다. 그 결과, 실험 과업을 제외한 다른

변수들은 학습자의 스페인어 실력에 따라 서로 다른 영향을 주는 것으로 나타났다. 특히, 동사의 형태 규칙성은 낮은 레벨의 학습자들에게는 유의미한 영향을 주지 못했지만 Level 4-5 그룹과 원어민 통제 그룹(control group)에서는 영향력을 보이는 변수로 작용했다는 점이 이 연구의 중요한 발견이라고 판단된다.

IV. 연구 방법

앞서 살펴본 선행 연구들에 대한 분석을 바탕으로 하여 스페인어의 L2 화자들이 명사절에서 직설법과 접속법을 사용하는 데 영향을 미치는 요인을 보다 광범위하게 파악하기 위해, 본 연구에서는 다양한 언어적(linguistic) 그리고 언어외적(extralinguistic) 변수를 설정하여 그들의 스페인어 생산을 분석하고자 한다. 특히, Collentine(1997) 등의 연구를 따라 접속법의 사용이 규칙 변화 동사(regular verbs)와 더욱 빈번하게 이루어지는지 아니면 불규칙 변화 동사(irregular verbs)와의 활용도가 더욱 높은지에 초점을 두고 하나의 독립 변수(dependent variable)로서 동사 형태의 규칙성이 통계적 유의성을 갖는지를 파악한다.

Gudmestad(2012)에서 정리한 것과 같이, 이전의 연구들은 동사 형태의 규칙성이 유의미한 변수로 작용한다는 것은 밝혔지만, ‘규칙 vs. 불규칙’과 ‘직설법 vs. 접속법’ 사이의 상관관계에 대한 세부적인 사항에 있어서는 서로 일치하지 않는 결과를 보여주었다. 즉, 일부 연구에서는 규칙 변화 동사가, 다른 연구에서는 불규칙 변화 동사가 접속법으로 더욱 자주 활용된다는 서로 상반되는 결론이 도출된 것이다(cf. Geeslin and Gudmestad 2008; Gudmestad 2006; Lubbers Quesada 1998). 이러한 상황에서 다양한 모국어(L1) 배경의 스페인어 학습자들의 작문에서는 어떠한 양상이 발견되는지를 분석하는 것은 제2언어 습득 연구에 기여하는 바가 클 것으로 예상된다.

1. 연구 질문과 가설

선행 연구 전반에 걸친 상이한 실험 결과와 그 분석에 비추어, 본 연구의 목적은 여러 연구에서 분석한 각 독립 변수들, 특히 동사의 형태 규칙성과 스페인어 법 구분 사이의 관계를 더욱 깊이 있게 이해할 수 있는 연구의 첫 걸음을 시작하는 것이다. 이러한 연구 목적을 달성하기 위해 본 연구의 내용 및 방법에 밑바탕이 되는 연구 질문(research question, 이하 RQ)과 이에 대한 가설(hypothesis, 이하 H)은 아래와 같다.

RQ1: 다양한 모국어 배경의 스페인어 학습자들의 명사절 법 선택에 영향을 미치는 요인은 어떠한 것들이 있는가?

H1: 영어를 모국어로 구사하는 스페인어 학습자를 대상으로 한 일련의 선행 연구들에서 발견된 바와 같이, 주절 동사의 시제, 종속절 동사의 빈도(frequency), 동사의 형태 규칙성 등의 언어적 요인들과 스페인어 숙련도(proficiency), 스페인어 학습 기간, 스페인어권 국가 거주 경험 등의 비언어적 요인들이 영향을 미칠 것이다.

RQ2: 동사의 형태 규칙성이 직설법과 접속법 사용에 통계적으로 유의미하게 작용한다면, 해당 변수들 간의 세부적인 상관관계는 어떠한 양상을 보이는가?

H2: 접속법에서 불규칙 변화를 하는 동사들 중 다수가 고빈도 어휘이기 때문에 양적으로 불규칙 변화 동사가 접속법으로 활용되는 토큰(token)이 많이 발견될 것이다. 또한 동사의 종류를 개별 요인(individual factor)으로 설정하여 통계 분석을 실시할 경우에도, 불규칙 변화 동사와 접속법이 공통적으로 지니고 있는 유표성(markedness; saliency)으로 인해 이 두 변수 간의 상관관계가 더욱 높을 것이다.

이러한 연구 질문에 기초하여, 본 연구는 ‘모국어(L1)에 부재한 제2언어(L2)의 특수한 언어 현상을 충분한 교육(instruction)이나 입력(input) 없이도 성인 L2 학습자가 습득할 수 있는가?’라는 더욱 본질적인 제2언어 습득에 관련된 질

문에 대한 대답을 시도할 수 있을 것으로 기대한다.

2. 코퍼스

앞서 제시한 연구 질문에 대한 대답을 모색하기 위하여 본 연구에서는 스페인어를 제2언어로 학습한 학습자가 생성한 작문 데이터를 수집하여 형성된 대형 데이터베이스(large database)인 CEDEL2(*Corpus Escrito del Español como L2*, L2 Spanish Written Corpus) 코퍼스를 활용하였다. 본 연구에서 사용한 CEDEL2 v.2는 그 이전의 버전인 CEDEL2 v.1의 데이터를 포함하여, 총 4,399명의 참가자들로부터 얻은 1,105,936 단어의 작문 데이터를 수집한 코퍼스이다. 작문을 작성한 L2 스페인어 학습자의 모국어 배경은 영어, 독일어, 네덜란드어, 프랑스어, 포르투갈어, 이탈리아어, 그리스어, 러시아어, 일본어, 중국어, 아랍어로 매우 다양하다. 또한, L1 스페인어 화자들과의 대조를 위해 다양한 스페인어권 출신의 모국어 화자들의 작문 데이터 역시 포함하고 있다.

CEDEL2의 작문 데이터를 수집하기 위해 참가자들에게 주어진 작문 주제는 총 14개로, 각 내용은 다음과 같다: (1) 당신이 사는 지역, (2) 유명인에 대한 소개, (3) 최근에 본 영화에 대한 설명, (4) 작년 여름 휴가, (5) 미래의 계획, (6) 최근에 했던 여행, (7) 최근에 겪은 경험, (8) 테러(terrorism)가 세계에 미치는 영향, (9) 금연 구역 지정 관련 법에 대한 의견, (10) 성소수자 커플에 대한 의견, (11) 대마초(marijuana) 합법화에 대한 의견, (12) 이민(immigration)에 대한 의견, (13) 개구리 이야기, (14) 찰리 채플린 영화. 이 중에서 (1)–(12)번 주제는 기존 버전(v.1)에 이미 있었던 것으로 현재 버전(v.2)에서는 이 중 (2)번과 (3)번의 두 주제에 대해서만 계속해서 작문 데이터를 수집하였으며, 현재 버전에서 새롭게 추가된 (13)번과 (14)번 주제는 기존 주제와는 다소 다른 방식으로 작문 과정이 이루어졌다. 기존의 주제는 작문을 작성하는 참가자들이 해당 주제에 대해 자유롭게 글을 작성하는 자유 주제 작문이었던 반면 (13)번의 개구리 이야기는 그림책을 보고 그 내용을 스페인어로 표현하는 방식으로, (14)번의 찰리 채플린 영화는 찰리 채플린의 무성 영화 영상을 보고 그 내용은 스페인어로 요

약하는 방식으로 진행되었다. 이때, 모든 주제는 참가자의 스페인어 실력과 상관 없이 모두 선택 가능하도록 주어졌다. 이에 따라, 본 연구에서 역시 (1)번부터 (14)번까지 모든 주제에 대한 작문 데이터를 분석하였다.

또한, CEDEL2에서는 각 참가자들의 인적 사항과 언어적 배경(language background)과 관련된 다양한 정보들을 수집하여 작문 데이터와 함께 공개하고 있기 때문에 본 연구에 필요한 언어외적인 요인들을 선별하여 분석 분석에 포함할 수 있었다. 코퍼스에 공개된 자료를 간단히 나열하면, 학습자의 모국어; 학습자의 아버지의 모국어; 학습자의 어머니의 모국어; 가정에서 사용하는 언어(들); CEDEL2에서 자체적으로 실행한 스페인어 능력 검증 시험 점수(1-43 점); 스페인어 능숙도 수준(A1-C2); 스페인어 말하기/듣기/쓰기/읽기 각 영역에 대한 학습자의 자가 평가; 스페인어 외 다른 외국어 말하기/듣기/쓰기/읽기 각 영역에 대한 학습자의 자가 평가; 공인 스페인어 시험 점수 또는 레벨; 성별; 나이; 스페인어를 처음 접했을 당시의 연령; 스페인어를 공부한 기간; 스페인어권 거주 경험 (장소, 시기, 기간); 소속 교육 기관; 전공; 참가 당시의 학년의 총 20가지 정보로 구성되어 있다. 이러한 다양한 자료 중 어떠한 것을 본 연구의 데이터 분석에 활용했는지는 다음의 ‘IV.3. 데이터 코딩과 변인 설정’에서 자세히 설명하도록 한다.

3. 데이터 코딩과 변인 설정

분석에 필요한 변인을 설정하고 각 작문 데이터를 분석에 용이하도록 코딩하고자 우선적으로 L2 화자의 데이터와 L1 화자의 데이터를 따로 구분하여 2개의 하위 데이터 세트(subdataset)를 구성하였다. 논의의 편의성을 위해, 먼저 두 하위 데이터 세트에서 공통적으로 설정된 종속 변인(dependent variable)과 독립 변인(independent variable)의 첫 번째 카테고리인 언어적 변인들을 알아본 후 각 하위 데이터 세트에서 서로 다르게 코딩된 독립 변인의 두 번째 카테고리, 언어외적 변인들을 L2 데이터와 L1 데이터의 순서대로 살펴보겠다.

1) 종속 변인

본 연구에서 설정한 종속 변인은 종속절 동사의 법(mood)이다. 이 종속 변인에 대한 분석을 위해 우선 참가자들이 작성한 작문 데이터에서 종속 명사절이 포함된 문장을 모두 추출하였다. 이때, 종속 명사절에 사용된 동사가 직설법 형태로 사용되었을 경우 숫자 0으로, 본 연구의 타겟(target) 형태인 접속법으로 사용되었을 경우 숫자 1로 코딩하였다.

2) 독립 변인 1: 언어적 변인

데이터 분석에 포함된 언어적 변인들은 주절 동사의 시제-상-법(Tense-Aspect-Mood, 이하 TAM), 주절 주어의 인칭과 수(Person and Number, 이하 PN), 종속절 동사의 TAM, 종속절 주어의 PN, 긍정문/부정문의 구분(Polarity), 종속절 동사의 형태 규칙성, 종속절 동사의 빈도를 포함한 7가지 고정 계수(fixed factor)와 주절 동사와 종속절 동사의 2가지 랜덤 계수(random factor)로 구성되었다.²⁾ 언급한 9가지 언어적 변인들 중에 본 연구에서 고유의 기준을 두고 코딩한 것은 마지막 변인인 빈도였다. 가장 광범위하게 사용되는 스페인어 온라인 코퍼스인 Corpus del español에서 확인할 수 있는 빈도순 동사 리스트를 활용하여 최상위 빈도 20개 동사를 첫번째 그룹(G1), 그 다음 20개 동사를 G2, 같은 방법으로 G4까지 그룹별로 구분한 후, 위 4개 그룹에 속하지 않는 다른 동사들은 저빈도(low-frequency) 동사로 간주하여 마지막 G5 그룹으로 코딩하였다.

2) 주절 동사의 경우, 개별 동사를 랜덤 계수로 포함하였지만 본 연구의 주제 및 범위를 고려하여 주절 동사의 의미 유형을 따로 언어적 변인 카테고리 설정하지는 않았다. 주절 동사의 의미 유형은 흔히 스페인어 (문법) 교재들에서 접속법을 설명할 때 우선적으로 제시되기 때문에 L2 화자들에게 가장 직접적인 영향을 준다고 예상할 수 있다. 그러나 본 연구에서는 기존 연구에서 충분히 다루어지지 않은 다른 요인들을 독립 변인으로 설정하여 이들 간의 상호작용을 면밀히 관찰하고자 했기 때문에 최종적으로 주절 동사의 의미 유형을 분석 대상에서 제외하였다.

3) 독립 변인 2: 언어외적 변인

(1) L2 스페인어 화자 데이터의 언어외적 변인

스페인어 학습자 그룹의 언어외적 변인으로는 CEDEL2 코퍼스에서 제공한 20가지 정보 중에서 학습자의 모국어(L1), 나이, 스페인어 능력 시험 점수, 스페인어를 처음 접했을 때의 연령, 스페인어 학습 기간, 그리고 스페인어권 거주 기간의 총 6가지 변인 카테고리를 우선적으로 선별하였으나, 후술할 이유로 인해 4개의 변인을 제외하고 2개의 변인만을 분석에 포함했다.

우선, 분석에 포함된 두 가지 변인은 학습자의 모국어와 스페인어 능력 시험 점수 두 가지가 있다. 학습자의 모국어는 앞서 IV.2.에서 언급한 11개의 언어(영어, 독일어, 네덜란드어, 프랑스어, 포르투갈어, 이탈리아어, 그리스어, 러시아어, 일본어, 중국어, 아랍어)가 모두 포함되었으며, 각 언어별로 따로따로 구분하여 코딩하였다. 스페인어 능력 시험 점수는 숫자로 표기되었기 때문에 연속형 응답(continuous response)으로 간주할 수도 있겠으나, 해당 코퍼스에서 제공한 시험 점수가 100점 만점을 기준으로 100점, 97.7점, 95.3점 등으로 일정하게 정해진 점수 그룹이 있었기 때문에 해당 변인은 범주형 응답(categorical response)으로 처리하여 코딩하였다. 또한, 스페인어 실력이 비교적 낮은 학습자들이 접속법 형태를 자유롭게 작문에서 구사할 수 있는 가능성은 굉장히 낮다고 판단하여 CEDEL2에서 구분한 기준으로 C2에 해당하는 가장 높은 숙련도의 그룹에 속하는 세 점수 그룹(95.3점, 97.7점, 100점)의 참가자들의 작문 데이터만을 분석에 포함하였다.

분석에 포함할 수 없었던 4개의 변인 참가자의 나이, 스페인어를 처음 접했을 때의 연령, 스페인어 학습 기간, 그리고 스페인어권 거주 기간이었다. 하나씩 살펴보면, 학습자의 나이는 최연소 참가자의 경우 14세였으며 최고령 참가자는 74세였다. 스페인어를 처음 접했을 때의 연령 역시 거의 모국어와 동시에 스페인어에 노출된 것으로 보이는 1세에서부터 60세 이후 비교적 늦은 연령에 스페인어를 처음 배우기 시작했다고 밝힌 참가자들도 있었다. 스페인어 학습 기간도 마찬가지로 1년 미만에서부터 최대 56년까지 다양하게 나타났다. 마지

막으로, 스페인어권 거주 기간은 전혀 거주 경험이 없는 0개월에서부터 360개월, 즉 30년까지 참가자별로 서로 다른 환경에서 스페인어를 학습하고 사용해 왔음을 알 수 있었다. 이 4개의 변인 모두 참가자 간 차이의 정도가 매우 심하며, 수치로 기재된 이 응답들을 그룹화하여 범주형 응답으로 변환하기에도 역시 최소값과 최대값의 차이가 적어도 50 이상 벌어졌기 때문에 어떠한 수치를 기준으로 몇 개의 범주로 나누어야 할지 그 기준이 명확하지 않아 결국 통계 분석 과정에서 제외하게 되었다.

(2) L1 스페인어 화자 데이터의 언어외적 변인

스페인어를 모국어로 구사하는 화자들의 언어외적 변인 중에서 본 연구의 데이터 분석을 위해 활용된 정보는 출신 국가, 구사할 수 있는 제2언어의 2가지 카테고리였다. 출신 국가의 경우 응답을 하지 않은 몇몇 참가자들을 제외하고 총 11개의 국가로 참가자의 응답이 구성되었다. 알파벳 순서대로 아르헨티나, 칠레, 콜롬비아, 쿠바, 도미니카공화국, 에콰도르, 멕시코, 페루, 스페인, 우루과이, 베네수엘라가 포함되었다. L1 스페인어 화자들이 구사할 수 있다고 기재한 L2 역시 카탈란어, 중국어, 영어, 바스크어, 프랑스어, 갈리시아어, 독일어, 이탈리아어, 포르투갈어, 발렌시아어, 그리고 ‘없음’까지 다양하게 나타났다. 예상 가능하게도, 이 중 L2 영어 화자가 가장 많은 비율을 차지하고 있었다. 마지막으로, 참가자의 나이는 L2 화자들의 데이터에서와 마찬가지로 제외하게 되었는데, L1 화자들의 연령 분포는 15세에서 27세까지로, L2 데이터와는 반대로 연령 차이의 폭이 매우 작고 전체 참가자가 청소년 또는 청년층에 해당된다는 점에서 연령대별 응답의 차이를 확인할 수 없기 때문에 분석 대상에 포함하지 않았다.

이렇게 수집하고 코딩한 데이터는 일차적으로 Rbrul을 통해 분석하였다. Rbrul은 통계적 계산 및 그래픽용 소프트웨어 프로그램인 R(R Core Team 2020)을 이용하여 변이주의 접근법으로 언어 데이터를 분석하는 프로그램이다. 먼저, L2와 L1 스페인어 화자의 각 하위 데이터 세트를 따로따로 Rbrul을 통해 분석하여 설정한 독립 변인들 중 어떠한 것들이 종속질의 접속법 사용에 대한

효과적인 예측값(predictor)으로 기능하는지 살펴본다. 이후, 종속절에 사용된 동사 중 눈에 띄는 특징적인 패턴을 보이는 동사가 있는지, 있다면 어떤 동사인지를 확인해보기 위해 R에서 제공하는 corrgram(Wright 2021)이라는 패키지(package)를 활용하여 Attraction과 Reliance 값(Schmid 2000)을 구하고 그 결과를 분석했다. Attraction과 Reliance는 서로 다른 두 성분이 갖는 연관성(association)의 정도를 나타내기 위해 사용되는 것으로, 본 연구에서는 종속 명사절에 사용된 각 개별 동사가 접속법과 어떠한 연관성을 지니는지 파악하기 위해 두 측정값을 활용했다. 동사 *hablar* ‘말하다’를 예로 들어 간단히 설명하자면, Attraction은 전체 데이터의 종속 명사절에서 접속법이 활용된 구문들 중에 *hablar*가 나타난 빈도를 나타낸다. 한편, Reliance는 전체 데이터의 종속 명사절에서 *hablar*가 활용된 구문들 중에 접속법이 나타난 빈도를 나타낸다. 이 두 측정값을 통해 우리는 *hablar*라는 동사가 종속 명사절에서 어떠한 패턴으로 사용되는지를 파악할 수 있다. 다음 장에서는 이러한 절차를 통해 분석한 결과를 소개하고, 각 결과가 나타내는 시사점을 논의한다.

V. 데이터 분석과 논의

L1 스페인어 화자와 L2 스페인어 화자의 작문 데이터를 각자 따로 살펴보기에 앞서, CEDEL2 코퍼스에 나타난 종속 명사절의 법 사용에 대한 전반적인 패턴을 파악하고자 한다. 분석에 활용된 종속 명사절을 포함한 구문은 총 2,775개이고, 이 중에서 1,443개 구문이 L2 화자에 의해 작성되었으며 나머지 1,332개 구문이 L1 화자가 작성한 문장으로 구성되었다. <표 1>은 종속 명사절에서 직설법과 접속법이 사용된 빈도와 비율을 비교하여 정리한 것으로, L1과 L2 화자 모두 직설법을 압도적으로 많이 활용하고 있음이 드러난다:

〈표 1〉 종속 명사절에서 직설법과 접속법 사용의 상대적인 빈도 분포

	L2	L1	합계
직설법	1,149 (79.63%)	1,013 (76.05%)	2,162 (77.91%)
접속법	294 (20.37%)	319 (23.95%)	613 (22.09%)

<표 1>에 따르면, L1 화자들이 L2 화자들에 비해 아주 약소하게 3.58% 포인트의 차이로 종속 명사절에 접속법을 더욱 많이 사용하는 것으로 보인다. 이러한 수치적 차이가 통계적으로 의미가 있는지를 확인하기 위해 카이 제곱 검정(chi-square test)을 실시했고, 해당 변수들 간의 차이는 $X^2(1) = 4.9372, p = 0.02628$ 로 그 정도가 아주 크지는 않더라도 결과적으로는 유의미한 것으로 나타났다. 다시 말하면, 종속 명사절에서 접속법을 사용할 가능성은 비원어민 L2 스페인어 학습자보다 원어민 L1 스페인어 화자에게서 더 높았다는 것이다.

이러한 결과를 바탕으로, 비록 20% 정도에 그치는 접속법의 사용이지만, 종속 명사절에서 접속법이 사용될 경우에 어떠한 요소들이 긍정적인 영향을 미쳤는지를 각 화자 그룹을 따로 분리하여 자세히 살펴보도록 하겠다. 공간의 제약에 따라, 앞서 IV.3에서 언급한 독립 변인들 중 최종 모델에 포함된 것들만 분석한다. 먼저, 스페인어 학습자의 데이터를 Rbrul에서 분석한 결과는 <표 2>와 같다. Rbrul의 분석 결과를 나타난 표를 해석할 때 가장 주목해야 할 수치는 가장 오른쪽에 표기한 확률비중(factor weight, 이하 FW)이다. 통상적으로 확률비중이 0.5 이상일 경우 해당 독립 변인이 종속 변인의 출현에 긍정적인 상관관계를 갖고 있는 것으로 해석된다. 다시 말하면, 0.5 이상의 확률비중을 지닌 변인이 포함된 구문은 0.5 미만의 확률비중을 지닌 변인이 포함된 구문에 비해 타겟이 되는 종속 변인 역시 포함하고 있을 가능성을 높게 보인다는 것이다.

<표 2>의 내용을 독립 변인의 카테고리별로 순서대로 살펴보겠다. 우선, 실험 참가자의 모국어 11개 중에서 절반 정도(5개 언어)가 접속법 사용과 긍정적인 상관관계를 지니는 것으로 나타났으며 나머지 6개 언어는 반대의 경향성을 보였다. 독일어, 포르투갈어, 일본어, 중국어, 프랑스어를 L1으로 사용하는 학습자들이 러시아어, 영어, 그리스어, 아랍어, 네덜란드어, 이탈리아어 모국어

〈표 2〉 L2 스페인어 화자 데이터의 Rbrul 결과 요약

	접속법	직설법	총합	확률비중 (factor weight)
L1				
독일어	11 (3.74%)	19 (1.65%)	30 (2.08%)	0.666
포르투갈어	33 (11.22%)	51 (4.44%)	84 (5.82%)	0.635
일본어	2 (0.68%)	6 (0.52%)	8 (0.55%)	0.619
중국어	5 (1.70%)	13 (1.13%)	18 (1.25%)	0.607
프랑스어	1 (0.34%)	4 (0.35%)	5 (0.35%)	0.541
러시아어	15 (5.10%)	53 (4.61%)	68 (4.71%)	0.474
영어	188 (63.95%)	828 (72.06%)	1,016 (70.41%)	0.456
그리스어	29 (9.86%)	114 (9.92%)	143 (9.91%)	0.420
아랍어	3 (1.03%)	11 (0.97%)	14 (0.97%)	0.419
네덜란드어	2 (0.68%)	14 (1.22%)	16 (1.11%)	0.351
이탈리아어	5 (1.70%)	36 (3.13%)	41 (2.84%)	0.313
Polarity				
부정문	41 (13.95%)	43 (3.74%)	84 (5.82%)	0.693
긍정문	253 (86.05%)	1,106 (96.26%)	1,359 (94.18%)	0.307
Frequency				
G5 (최저빈도)	111 (37.76%)	189 (16.45%)	300 (20.79%)	0.649
G2	26 (8.84%)	55 (4.79%)	81 (5.61%)	0.629
G4	5 (1.70%)	20 (1.74%)	25 (1.73%)	0.421
G3	16 (5.44%)	71 (6.18%)	87 (6.03%)	0.420
G1 (최고빈도)	136 (46.26%)	814 (70.84%)	950 (65.84%)	0.376
Regularity				
규칙 변화 동사	76 (25.85%)	130 (11.31%)	206 (14.28%)	0.579
불규칙 변화 동사	218 (74.15%)	1,019 (88.69%)	1,237 (85.72%)	0.421

화자들에 비해 종속 명사절에서 접속법을 활용할 확률이 높았다. 상기한 11개의 언어가 서로 다른 어족 또는 언어군(language family)에 속하며, 같은 로망스어군에 속하는 포르투갈어, 프랑스어(FW 0.5 이상)와 이탈리아어(FW 0.5 미만)가 상반되는 결과를 보였기 때문에 이 결과는 일반화하기에 어려움이 있다. 이에 따라, 특정 언어권 화자들이 비슷한 스페인어권 체류 경험이나 스페인어 학습 기간을 공유하는지 확인해보기 위해 각 독립 변인들 사이의 상호 작용(interaction)의 효과 역시 확인해보았으나, 눈에 띄는 결과의 변화가 발견되지 않았고 분석 모델의 적합도(model fit)이 하락하는 경향을 보여 최종 분석 모델에서는 해당 상호 작용 요인들을 제외하였다.

주절 문장의 긍정 여부 또한 종속 명사절의 법 선택에 영향을 미치는 요인으

로 나타났다. 주절 문장에 부정 표현이 있을 경우에 종속절에 접속법이 사용되는 경향이 나타났는데, 이는 접속법이 지니고 있는 의미적 특성에 기인한 것으로 볼 수 있다. 앞서 언급했듯이, 불확실성이나 비실제성이 접속법 사용의 절대적인 기준이 될 수는 없으나 실제 언어 사용 패턴을 비추어보면 이러한 의미적 특성이 내포된 문장들에서 접속법이 많이 사용된다는 것은 보편적으로 받아들여진다. 이렇게 우리가 이론적으로 인지하고 있었던 스페인어 접속법의 특징을 실증적으로, 또 통계 분석을 바탕으로 다시 한번 확인할 수 있었다.

종속절 동사의 빈도는 앞서 다루었던 독립 변인 카테고리인 L1과 비슷하게 일정한 패턴을 드러내지 않았다. 가장 낮은 빈도로 사용되는 동사들이 포함된 G5와 두 번째로 높은 빈도로 사용되는 동사들로 구성된 G2가 0.5 이상의 확률비중을 나타냈고, 나머지 G4, G3, G1은 0.5의 기준을 넘어서지 못했다. 비록 G2, G3, G4가 일정한 순서를 보이지 않았다는 점이 이 결과의 해석에 어려움을 주지만 G5와 G1이 양극단에 위치하고 있다는 것은 주목할만한 부분이다. 특히, G5의 확률비중이 가장 높게 나타난 것은 코퍼스 고유의 특징에 기인한 것으로 생각할 수 있다. CEDEL2는 다양한 주제에 대해 참가자들이 자유롭게 작문을 하는 방식으로 구성된 과업을 통해 구축한 코퍼스이다. 이때, 주어진 주제에 따라 특정 동사들이 더욱 많이 등장하게 되며 특히 그림책이나 영화의 내용을 요약하는 과업에서는 거의 모든 참가자가 서로 비슷한 동사들을 비슷한 구문 구조로 활용하는 모습을 보였다. 이에 따라, 비록 실제 자연스러운 스페인어 구사 환경에서는 비교적 자주 사용되지 않는 동사라고 할지라도 특정 과업에서 집중적으로 많이 활용된 동사들이 있기 때문에 이러한 결과가 나왔을 것으로 생각할 수 있다.

마지막으로, 본 연구에서 가장 중점적으로 다루고자 하는 종속절 동사의 형태 규칙성에서는 규칙 변화를 하는 동사가 그렇지 않은 동사에 비해 접속법으로 활용될 가능성이 높은 것으로 나타났다. 다만, 0.5를 크게 상회하지는 않는 0.579라는 확률비중을 보이기 때문에 형태 규칙성이라는 요인이 독립 변인으로 작용하는 힘이 크다고 볼 수는 없다. 지금의 결과가 어떠한 함의를 갖는지는

다음에 살펴볼 L1 스페인어 화자의 데이터 분석 결과와의 비교와 개별 동사가 접속법 형태와 지니는 연관성을 엿볼 수 있는 Attraction과 Reliance 측정값 결과의 확인을 통해 알아볼 수 있을 것이다.

이제, 스페인어 원어민 화자의 작문 데이터를 Rbrul을 통해 분석한 결과를 <표 3>을 통해 확인하겠다:

<표 3> L1 스페인어 화자 데이터의 Rbrul 결과 요약

	접속법	직설법	총합	확률비중 (factor weight)
Polarity				
부정문	39 (12.23%)	26 (2.57%)	65 (4.88%)	0.712
긍정문	280 (87.77%)	987 (97.43%)	1,267 (95.12%)	0.288
Frequency				
G4	12 (3.76%)	15 (1.49%)	27 (2.03%)	0.636
G5	134 (42.01%)	188 (18.56%)	322 (24.17%)	0.601
G2	37 (11.59%)	52 (5.13%)	89 (6.68%)	0.598
G3	16 (5.02%)	51 (5.03%)	67 (5.03%)	0.403
G1	120 (37.62%)	707 (69.79%)	827 (62.09%)	0.275
Regularity				
규칙 변화 동사	183 (57.37%)	383 (37.81%)	566 (42.49%)	0.524
불규칙 변화 동사	136 (42.63%)	630 (62.19%)	766 (57.51%)	0.476

L1 스페인어 원어민 화자들에게서 발견된 흥미로운 결과 중 하나는 설정된 언어외적 변인 중 어떠한 것도 결과에 영향을 미치지 못했다는 것이다. 이는 참가자들의 출신 국가도 그들이 L2로 학습한 다른 언어도 L1 스페인어 화자의 작문 생산에 통계적으로 유의미하게 상관 관계를 맺지 않는다는 의미를 갖는다. 모국어로 글을 쓰는 과정에서 구사 가능한 다른 언어가 영향을 주지 않는다는 것은 통계 분석을 진행하기 전에도 유추할 수 있는 부분이었지만, 출신 국가 역시 그러했다는 사실은 주목할만 하다. 이 결과는 더욱 많은 양의 데이터와 구어 생산 등 다른 방식의 데이터를 취합하여 추후에 더욱 깊이 있는 연구를 진행하여 그 시사점을 알아볼 필요가 있다고 판단된다.

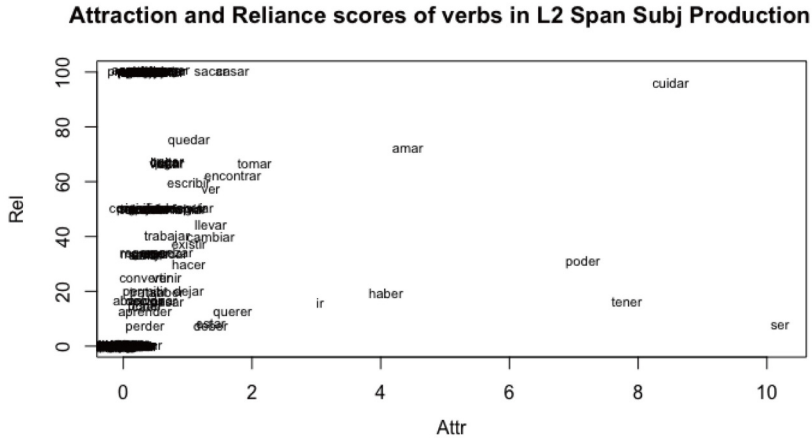
본격적으로 위 <표 3>의 내용을 살펴보면, 우선 주절 동사가 부정문일 경우 종속 명사절에서 접속법이 사용될 가능성이 높은 것으로 나타났다. 이는 L2 화

자의 데이터와 일치하는 방향성을 보이는 결과로, 다시 한번 주절에서의 부정 표현, 특히 불확실성과 비실제성을 나타내는 표현이 종속절에서의 접속법 사용을 유인하는 요소라는 것을 확인할 수 있었다.

종속절 동사의 빈도와 관련해서는 비교적 빈도가 낮게 사용되는 G4와 G5의 동사들이 접속법 사용과 긍정적인 상관 관계를 가졌으며, 뒤이어 두 번째로 높은 빈도로 사용되는 동사들이 포함된 G2 역시 0.598이라는 확률비중으로 0.5의 기준을 조금 넘긴 것으로 나타났다. 가장 낮은 확률비중을 보인 빈도 그룹은 최고빈도의 동사들로 구성된 G1였다. 이 결과 역시 L2 화자 데이터의 분석과 유사성을 보인다고 해석할 수 있다. 비록 그룹별 순서는 <표 2>의 그것과 일치하지는 않지만, 빈도가 낮은 동사군이 0.5 이상의 확률비중을, 빈도가 높은 동사군이 가장 작은 수치의 확률비중을 보였기 때문이다. 이러한 공통점의 이유 역시 코퍼스에서 활용된 작문 과업의 특성에서 기인하는 것으로 볼 수 있다.

마지막으로, 종속절 동사의 형태 규칙성 역시 L2 화자들의 분석 결과와 비슷한 양상을 보인다. 규칙 변화 동사는 접속법으로 활용될 가능성이, 반대로 불규칙 변화 동사는 접속법이 아닌 직설법으로 사용될 가능성이 더욱 높은 것으로 나타났다. 그러나, 앞서 L2 데이터 분석에서 보았던 것과 마찬가지로 이 두 요인 간 확률 비중의 차이가 굉장히 미미한 수준에 그친다는 것을 염두에 두고 분석 결과를 받아들여야 할 것으로 보인다.

이제 각 참가자 그룹의 Rbrul 결과에서는 확인할 수 없었던 종속 명사절에서 활용된 개별 동사가 갖는 접속법과의 연관성을 알아보기 위해 Attraction과 Reliance 수치를 구하고, 그 결과를 살펴보겠다. 두 측정값 모두 R을 활용하여 계산하였으며, 자세한 과정은 생략한다. 여기에서는 계산 결과를 그래프를 통해 시각적으로 제시하고 그래프를 해석하는 데 있어 주목할 만한 내용을 간략하게 설명하고자 한다. 우선, L2 스페인어 학습자 데이터를 분석한 결과는 <그림 1>과 같다:



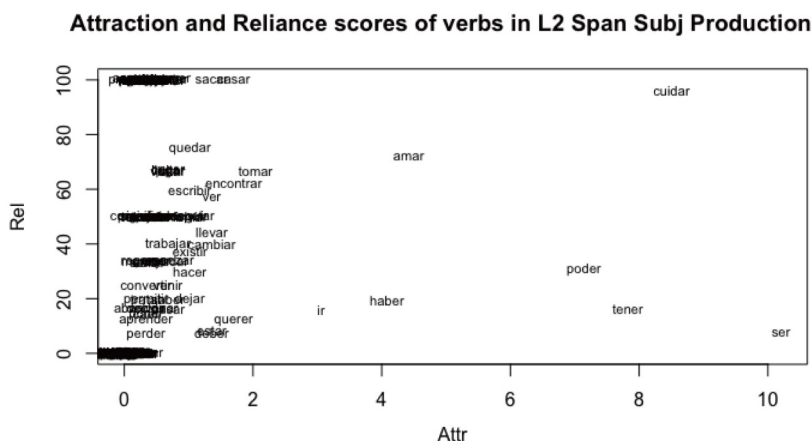
〈그림 1〉 L2 스페인어 화자 데이터의 연관성 측정 결과

위의 그림에서 가장 눈에 띄는 모습을 보이는 동사를 몇 개 추려보면 *ser*, *cuidar*, *tener*, *poder*, *amar*, 그리고 *haber*가 있다. 이 중 가장 오른쪽 끝에 위치한 *ser*는 Attraction 측정값은 매우 높은 반면 Reliance 측정값은 매우 낮은 것을 알 수 있다. 이는 종속절에 접속법이 사용된 모든 동사 중에 *ser*가 차지하는 비율이 높지만 *ser*가 사용된 구문을 기준으로 봤을 때는 접속법보다 직설법으로 활용된 비율이 높다는 것을 의미한다. 바꿔 말하면, *ser*는 스페인어에서 최고빈도 어휘에 속하는 동사로 분석한 데이터에서 역시 직설법과 접속법 모두 다양하게 많이 사용되었지만 그 절대적인 수치를 기준으로 비교했을 때는 직설법으로 활용된 토큰이 압도적으로 많다는 것을 뜻한다. 그 정도는 다르지만 *tener*, *poder*, *haber* 동사 역시 이와 유사한 패턴을 보인다고 할 수 있다.

이제 Attraction과 Reliance 값이 모두 높은 동사로 *cuidar*를 볼 수 있다. 해당 동사는 스페인어라는 언어 자체를 기준으로 보았을 때 그렇게 자주 사용되는 동사로는 보기 어려우나, CEDEL2 코퍼스에서는 꽤나 많이, 그것도 활용된 대부분의 경우에 접속법으로 쓰였다는 것을 두 측정값을 통해 알 수 있다. 이는 앞서 Rbrul 분석 결과에서 언급한 작문 주제의 특성과 관련되어 있는데, 찰리 채플린의 무성 영화를 보고 그 내용을 스페인어로 요약하는 과업에서 많은 참

가자들이 *Le pide que cuide al bebé* ‘아이를 돌봐달라고 부탁했다’와 유사한 문장을 많이 작성했기 때문이다. 이와 유사한 패턴을 보이는 *amar* 역시 마찬가지로, 같은 장면에서 *Le pide que ame al bebé* ‘아이를 사랑해달라고 부탁했다’ 등의 문장으로 인해 접속법으로 활용된 빈도가 다소 높았다. 따라서 이 두 동사와 관련된 결과는 분석한 데이터의 독특한 특징 때문으로 볼 수 있다.

L1 스페인어 원어민 화자의 데이터를 통해 확인한 연관성 측정 결과도 L2의 결과와 크게 다르지 않았는데, 이를 <그림 2>에서 확인할 수 있다:



〈그림 2〉 L1 스페인어 화자 데이터의 연관성 측정 결과

이번에도 주목할만한 동사로 *ser*와 *cuidar*가 가장 두드러지며, 그 다음으로는 *llevar*가 *cuidar*와 유사하게 비교적 높은 Attraction 측정값과 그것보다 조금 더 높은 Reliance 측정값을 갖는 것으로 나타났다. 이러한 유사성을 제외하면 <그림 2>의 L1 데이터는 <그림 1>의 L2 데이터에 비해 대부분의 동사들이 널리 퍼지지 않고 비슷한 위치에 모여있는 것을 확인할 수 있다. 이러한 양상은 스페인어를 모국어로 구사하는 화자들이 학습자들에 비해 개별 동사의 영향을 덜 받는 경향이 있는 것으로 해석할 수 있다. 즉, L2 스페인어 화자들은 특정 동사를 다른 동사들에 비해 두드러지게 접속법으로 자주 활용하는 편향된 모습을

보여 이러한 동사들이 <그림 1>에 단독으로 떨어져서 확인되지만, L1 스페인어 화자들은 일부 동사를 제외하고는 다양한 동사들을 고르게 접속법으로 활용하여 비슷한 측정값을 가진 동사들끼리 서로 붙어 어떤 동사인지를 확인하기 어려울 정도로 모여 있게 되었다는 것이다.

지금까지의 분석 결과를 종합해보면, 스페인어의 원어민 화자와 비원어민 화자는 전반적으로 서로 유사한 패턴으로 종속 명사절에서 접속법을 사용하는 것으로 보인다. 이는 Rbrul 결과에서 파악한 부정문, 저빈도 동사군, 규칙 변화를 하는 종속절 동사의 높은 확률비중을 통해 확인할 수 있었다. 그러나 개별 동사의 접속법 활용 경향을 확인해 보았을 때는, 가장 많이 사용된 *ser*와 *cuidar* 동사에서는 유사한 분포를 보였지만, 이외의 동사들에 대해서 L2 화자들은 특정 동사를 선별적으로 접속법으로 활용하는 편향된 모습을 보인 반면 L1 화자들은 더욱 다양한 동사를 일관되지 않은 패턴으로 접속법 형태로 사용하고 있음을 알 수 있었다.

VI. 맺음말

스페인어의 원어민 화자와 비원어민 화자가 작성한 작문 데이터에서 종속 명사절에 접속법이 사용된 구문들에 작용하는 다양한 요인들의 영향을 파악하고자 했던 본 연구에서는 앞서 IV.1.에서 소개한 연구 질문과 그에 대한 가설을 바탕으로 데이터 분석을 진행했다. 이에 따라 각 연구 질문과 가설을 다시 한번 상기하고 실제 분석 결과가 연구 방법을 설계할 때 예측했던 가설의 내용과 부합하는지를 논의해보는 것으로 이 논문을 마무리하고자 한다. 우선, 첫 번째 연구 질문과 가설은 다음과 같다:

RQ1: 다양한 모국어 배경의 스페인어 학습자들의 명사절 법 선택에 영향을 미치는 요인은 어떠한 것들이 있는가?

H1: 영어를 모국어로 구사하는 스페인어 학습자를 대상으로 한 일련의 선행 연구들에서 발견된 바와 같이, 주절 동사의 시제, 종속절 동사의 빈도

(frequency), 동사의 형태 규칙성 등의 언어적 요인들과 스페인어 숙련도(proficiency), 스페인어 학습 기간, 스페인어권 국가 거주 경험 등의 비언어적 요인들이 영향을 미칠 것이다.

코퍼스 분석을 통해 확인한 바에 따르면, 언어적 요인들 중 주절 동사의 긍정 또는 부정 여부, 종속절 동사의 사용 빈도, 그리고 종속절 동사의 형태 규칙성의 세 가지 카테고리는 종속 변인에 영향을 미치는 것으로 파악되었다. 그러나 언어 외적 요인들은 L2 스페인어 학습자의 모국어를 제외한 다른 어떠한 요인도 유의미한 결과를 보이지 않았다. 즉, 종속 명사절의 접속법 사용에 있어서는 사회적인 변인보다 언어적인 변인이 더욱 큰 상관 관계를 갖는다고 볼 수 있다. 이어서, 두 번째 연구 질문과 그 가설은 아래와 같다:

RQ2: 동사의 형태 규칙성이 직설법과 접속법 사용에 통계적으로 유의미하게 작용한다면, 해당 변수들 간의 세부적인 상관관계는 어떠한 양상을 보이는가?

H2: 접속법에서 불규칙 변화를 하는 동사들 중 다수가 고빈도 어휘이기 때문에 양적으로 불규칙 변화 동사가 접속법으로 활용되는 토큰(token)이 많이 발견될 것이다. 또한 동사의 종류를 개별 요인(individual factor)으로 설정하여 통계 분석을 실시할 경우에도, 불규칙 변화 동사와 접속법이 공통적으로 지니고 있는 유표성(markedness; saliency)으로 인해 이 두 변수 간의 상관관계가 더욱 높을 것이다.

본 연구의 데이터 분석 결과는 두 번째 가설과 전혀 부합하지 않았음을 알 수 있다. 우선, 불규칙 변화 동사보다 규칙 변화 동사가 더욱 높은 확률로 접속법으로 활용되었으며, 개별 동사에 대한 분석 역시 동사 형태의 규칙성보다는 기타 다른 요인들의 영향이 더욱 많이 작용했다는 것을 볼 수 있었다. 특히, 작문 과업의 주제에 따라 작문 내용이 크게 좌우되는 코퍼스 고유의 특징이 분석 데이터에 상당히 반영된 점을 보았을 때, 더욱 중립적이고 다양한 종류의 데이터를 확보하여 분석을 시행하는 것이 필요할 것 같다는 점이 본 연구의 한계점인 동

시에 후속 연구가 추구해야 할 방향임을 확인할 수 있었다. 또한, CEDEL2 코퍼스에 부재한 L1 한국어 화자들의 데이터를 충분히 수집하여 후속 연구의 분석 대상으로 포함하면 본 연구에서 관찰된 L2 스페인어 종속 명사절 접속법 활용 현상의 보편성을 높이는 데 기여할 수 있을 것이다.

참고문헌

- Bello, Andrés(1847/1982), *Gramática de la lengua castellana*, Madrid: Edaf.
- Collentine, Joseph(1997), “The effects of irregular stems on the detection of verbs in the subjunctive,” *Spanish Applied Linguistics*, Vol. 1, pp. 3-23.
- Collentine, Joseph(2003), “The Development of Subjunctive and Complex-Syntactic Abilities among FL Spanish Learners,” Barbara Lafford and Rafael Salaberry(eds.), *Studies in Spanish Second Language Acquisition: the State of the Science*, Washington, DC: Georgetown University Press, pp. 74-97.
- Collentine, Joseph(2010), “The Acquisition and Teaching of the Spanish Subjunctive: An Update on Current Findings,” *Hispania*, Vol. 93, No.1, pp. 39-51.
- Davies, Mark(2016-), Corpus del Español: Two billion words, 21 countries. Available online at <http://www.corpusdelespanol.org/web-dial/>. (Web / Dialects)
- Farley, Andrew P.(2001), “Authentic Processing Information and the Spanish Subjunctive,” *Hispania*, Vol. 84, No. 2, pp. 289-299.
- Geeslin, Kimberly L. and Gudmestad, Aarnes(2008), “Comparing interview and written elicitation task in native and non-native data: Do speakers do what we think they do?,” Joyce Bruhn de Garavito and Elena Valenzuela(eds.), *Selected Proceedings of the Hispanic Linguistics Symposium*, Somerville, MA: Cascadilla Proceedings Project, pp. 64-77.
- Gudmestad, Aarnes(2006), “L2 Variation and the Spanish subjunctive: Linguistic features predicting mood selection,” Carol Klee and Timothy Face(eds.), *Selected proceedings from the 7th CLASP*, Somerville, MA: Cascadilla Press, pp. 170-184.

- Gudmestad, Aarnes(2010), "Moving beyond a sentence-level analysis in the study of variable mood use in Spanish," *Southwest Journal of Linguistics*, Vol. 29, No. 1, pp. 25-51.
- Gudmestad, Aarnes(2012), "Toward and Understanding of the Relationship between Mood Use and Form Regularity: Evidence of Variation across Tasks, Lexical Items, and Participant Groups," Kimberly Geeslin and Manuel Díaz-Campos(eds.), *Selected Proceedings of the 14th Hispanic Linguistics Symposium*, Somerville, MA: Cascadilla Proceedings Project, pp. 214-227.
- Guitart, Jorge M.(1982), "On the use of the Subjunctive among Spanish-English bilinguals," *Word*, Vol. 33, pp. 59-67.
- Haverkate, Henk(2002), *The Syntax, Semantics and Pragmatics of Spanish Mood*, Amsterdam/Philadelphia: John Benjamins.
- Igualada Belchí, Dolores Anunciación(1989), "Nueva hipótesis sobre el subjuntivo en español," *Estudios Románicos*, Vol. 4, pp. 643-664.
- Johnson, Daniel Ezra(2009), "Getting off the GoldVarb standard: Introducing Rbrul for mixed-effects variable rule analysis," *Language and Linguistic Compass*, Vol. 3, No. 1, pp. 359-383.
- Lakoff, George(1968), "Instrumental Adverbs and the Concept of Deep Structure," *Foundations of Language*, Vol. 4. No. 1, pp. 4-29.
- Lozano, C.(2022), "CEDEL2: Design, compilation and web interface of an online corpus for L2 Spanish acquisition research," *Second Language Research*, Vol. 38, No. 4, pp. 965-983.
- Lubbers Quesada, Margaret(1998), "L2 acquisition of Spanish subjunctive mood and prototype schema development," *Spanish Applied Linguistics*, Vol. 2, pp. 1-23.
- Perez-Cortes, Silvia(2020), "Lexical frequency and morphological regularity as sources of heritage speaker variability in the acquisition of mood," *Second Language Research*, Vol. 38, No. 1, pp. 149-171.
- R Core Team(2023), *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Schmid, H.-J.(2000), *English Abstract Nouns as Conceptual Shells. From corpus to*

- cognition*, Berlin/New York: Mouton de Gruyter.
- Silva-Corvalán, Carmen(1994), "The gradual loos of mood distinctions in Los Angeles Spanish," *Language Variation and Change*, Vol. 6, pp. 255-272.
- Stokes, Jeffrey(1988), "Some factors in the acquisition of the present subjunctive in Spanish," *Hispania*, Vol. 71, pp. 705-710.
- Terrell, Tracy and Hooper, Joan(1974), "A Semantically Based Analysis of Mood in Spanish," *Hispania*, Vol. 57, No. 3, pp. 484-494.
- VanPatten, B.(1993), "Grammar Teaching for the Acquisition-Rich Classroom," *Foreign Language Annals*, Vol. 26, pp. 435-450.
- Wright K.(2021), *corrgram: Plot a Correlogram*. R package version 1.14, <https://CRAN.R-project.org/package=corrgram>.

황지희

계명대학교 스페인어중남미학과
jiheeh@kmu.ac.kr

논문투고일: 2023년 11월 26일
심사완료일: 2023년 12월 18일
게재확정일: 2023년 12월 18일

Examining L1 and L2 Spanish Subjunctive Usage through a Variationist Lens

Jihee Hwang

Keimyung University

Jihee Hwang(2023), Examining L1 and L2 Spanish Subjunctive Usage through a Variationist Lens, *Revista Asiática de Estudios Iberoamericanos*, 34(3), 57-86.

Abstract The Spanish mood is generally divided into the indicative and subjunctive, with previous research suggesting a close link between syntactic and/or semantic features and the distinction between these moods. In the study of Spanish as a second language (L2), it is believed that learners often struggle with acquiring the subjunctive, viewing it as a structure acquired in the later stages of learning. Previous studies have primarily supported these assumptions, focusing on morpho-syntactic, semantic, and/or pragmatic aspects of the subjunctive. However, recent sociolinguistic studies reveal a more complex array of factors influencing mood distinction observed in authentic Spanish usage. Building on these discussions, the present study employs a variationist methodology to analyze L1 and L2 Spanish mood distinction. Considering conflicting results in the literature regarding verb form regularity, this study also incorporates the analysis of this variable. Examining written production data from L2 Spanish speakers in the CEDEL2 corpus and comparing it with a control group of L1 Spanish speakers, the present study concludes that regular verbs favor the subjunctive in subordinate noun clauses. Additionally, the negative polarity of the main clause and the low frequency of the verb in the subordinate clause condition the use of the subjunctive form. Given limited discussion on sociolinguistic methodology in previous L2 Spanish acquisition studies, the current research attempted to offer insight into the application and implications of a variationist approach to L2 Spanish mood distinction.

Key words Spanish subjunctive, L2 Spanish mood distinction, Written corpus, L2 Variationism, Form regularity