

생성형 AI 중한 번역의 접속 표현 분석

최승혁*

Seunghyuk Choi (2025). *An Analysis of Conjunctive Expressions in Chinese-Korean Translation by Generative AI*. This study analyzes how the Korean connective endings ‘-고’ and ‘-며’ are translated when generative AI translates Chinese asyndetic compound sentences into Korean, with the aim of identifying tendencies and limitations in AI handling of conjunctive relations. By comparing the translated texts of ChatGPT and HyperCLOVA X using newspaper articles and public speeches, the study reveals that generative AI often relies on surface-level paratactic structures without fully inferring the underlying logical relationships required by context. Consequently, the use of ‘-고’ and ‘-며’ tends to be excessively concentrated on the enumerative function, leading to weakened logical cohesion in the translated texts. Based on these findings, the study suggests enhancing AI translation quality through improved semantic inference of conjunctive relations and incorporating cohesion-based evaluation metrics that assess the accuracy of logical linkage in generative AI translation. (Chung-Ang University, Korea)

Keywords: AI translation, conjunction, Korean connective endings, chatgpt, hyperclova

X

주제어: AI 번역, 접속, 연결어미, ChatGPT, 하이퍼클로바X

* 중앙대학교 국제대학원

1. 들어가며

한국어는 전형적인 교착어로 어미와 조사가 풍부하다. 절이나 문장을 이어주려면 연결어미나 조사가 필요하다. 특히 연결어미는 그 기능과 의미가 복잡해 치열한 논쟁의 대상이 되어 왔다. 반면 중국어는 대표적인 고립어로 의합(义合)¹⁾에 따라 연결 성분을 생략하고 어순을 통해 문법적 관계나 의미를 드러낼 수 있다. 문장이나 절을 연결할 때 한국어는 유표적이지만 중국어는 유표적일 수도 있고 무표적일 수도 있다. 이와 같이 두 언어 사이에 존재하는 접속 표현의 유형론적 차이는중한 번역 시에 주의해야 할 요소이다.

생성형 AI가 만들어낸 번역문에는 이러한 언어 유형에서 비롯된 접속 표현의 차이가 정확히 반영되어 있을까? 번역은 생성형 AI의 대표적 기능 중 하나이다. 생성형 AI는 사전학습을 거친 거대언어모델(LLM)을 통해 벡터화된 언어의 조각들을 확률에 따라 배열하는데 그것이 생성한 텍스트를 읽어보면 도착 언어의 언어 규범을 정확하게 따른다. 그럼에도 불구하고 “이거 AI가 쓴 거 아니야?”라고 농담을 던질 만큼 어딘가 어색하다는 느낌을 지우기 힘든 것도 사실이다. 생성형 AI 인터페이스에 출력되는 텍스트가 표면적으로 언어 규범을 철저하게 따르고 있지만 하나의 온전한 텍스트로서 수용하기에 여전히 주저되는 이유이다. 아래 예문을 보자.²⁾

[ST] ①如今“对等关税”全面实施, ②进口商品价格上涨将进一步推高通胀, ③中低收入家庭首当其冲。

(2025.06.07. 챗GPT 번역. 프롬프트: 이 중국어 신문 기사를 자연스러운 한국어로 번역해 줘.)

[TT-C] 이번 ‘상호 관세’의 전면적 시행은 수입 상품 가격 상승으로 이어져 인플레이션을 더욱 자극할 것이며, 중·저소득층이 가장 먼저 타격을 입게 된다.

[대안 번역] 상호관세 시행(에 따라/으로 인해) 수입품 가격 상승으로 인플레이션이 심해져서 결국 중저소득층 가계의 충격이 가장 클 것으로 예상된다.

1) 접속 어구나 접속사 등 문법적 표지를 사용하지 않고 의미에 따라 복문을 구성하는 방법을 말한다.

2) ST는 원문, TT-C는 챗GPT 번역문, TT-X는 하이퍼클로바X 번역문이다.

위 예문에서 ST는 ①, ②, ③번 세 개의 절로 이루어져 있으며 절 사이에 접속 표지가 없는 전형적인 무표지 복문(无标志复句)이다³⁾. 언어 유형의 차이로 인해 중국어 무표지 복문을 한국어로 번역할 때는 연결어미 등을 활용해 적절한 접속 표현을 이루는 것이 중요하다. 챗GPT는 ①번 절을 조사 ‘은’을 사용해 전체 문장의 주어로 처리하면서 ②번 절과 논리 관계를 분절하였다. ②번 절은 연결어미 ‘-며’를 사용해 ③번 절과 대등하게 나열했다. [TT-C]는 표층적으로 도착 언어인 한국어 언어 규범을 잘 따르고 있는 것으로 보인다. 그러나 의미적으로 살펴보면 접속 표현이 정확하지 않아 절 간의 논리 관계가 훼손되면서 전체 문장의 결속성이 약화되었음을 알 수 있다. ST의 각 절은 미국의 상호 관세 시행이라는 원인에 따른 수입품 가격 상승과 인플레이션 심화, 그리고 최종적으로 중저소득 계층의 충격까지 이어지는 논리 관계를 연쇄적으로 전개한다. 따라서 대안 번역처럼 ‘-에 따라(표현), -으로 인해(표현)’⁴⁾ 또는 ‘-어서(연결어미)’ 등과 같이 인과 관계를 나타내는 성분으로 접속하여 논리 관계를 정확하게 전달하고 결속성을 높여야 한다. 위 예문에서처럼 문두에 출현하는 절을 무조건 주어로 처리하거나 절과 절을 대등적 병렬 관계로만 나열하면 독자의 이해를 방해하고 텍스트로서의 수용도가 낮아질 수밖에 없다.

본고는 생성형 AI 중한 번역에서 나타나는 접속 표현⁵⁾에 관하여 한국어 연결어미 중 사용 빈도가 높은 ‘-고’, ‘-며’를 중심으로 분석하고자 한다. 구체적으로 접속 표현의 관점에서 중국어 무표지 복문의 접속이 한국어 번역문에서 동일한 의미 기능으로 구현되는지 살펴볼 것이다. 이를 위해 챗GPT와 하이퍼클로바X 두 가지 생성형 AI 모델을 분석 도구로 사용한다. 챗GPT는 대표적인 범용 생성형 AI 모델로 사전학습 시 영어 데이터의 비중이 높다. 반면 하이퍼클로바X는 한국에서 개발된 모델로 사전학습에서 한국어 데이터의 비중이 챗GPT보다 높은 것이 특징이다. 사전학습 언어의 비중이 상이한 두 모델을 비교하면 생성형 AI의 접속 표현 경향을 더욱 객관적으로 이해할 수 있을 것이다. 생성형 AI 중한 번역의 접속 표현 경향을

3) 상푸이(邢福义 2002: 20)는 문법적으로 관계 표지가 출현하지 않는 복문을 무표지 복문으로 보았다. 정운철(2005: 113)은 절과 절 사이에 접속 표지를 사용하지 않고 어순에 따라 의미적, 논리적 관계를 표현한 복문을 무표지 복문으로 보았다.

4) 국립국어원 한국어교수학습센터에 따르면 ‘-에 따라’와 ‘-으로 인하여’의 품사를 ‘표현’으로 분류하고 있다.

5) 유혜령(2005: 87)은 텍스트 차원에서 사건과 상황의 관계를 드러내는 텍스트의 내적 구조가 여러 가지 언어 형식을 통해 텍스트 표층에 사용된 것을 접속 표현이라고 보았다. 본고는 접속 표현에 대해 유혜령의 관점을 따른다.

살펴봄으로써 생성형 AI의 문맥에 대한 이해도 파악 및 AI-TPE(AI-Translation Post Editing) 방향성 모색에 도움이 될 것으로 기대한다.

2. 접속 표현의 중요성 및 선행 연구

체계기능언어학(Systemic Functional Linguistics)에 따르면 텍스트는 단순한 문장의 연쇄를 넘어 의미적으로 결속(cohesion)된 통합 언어 단위이다. 할리데이와 하산(Halliday and Hasan 1976)은 결속 수단을 지시(reference), 생략(ellipsis), 대용(substitution), 접속(conjunction) 네 가지 문법적 결속과 어휘적 결속(lexical cohesion)으로 구분하였다. 이 중 접속은 논리적 관계를 명시화하여 문장과 문장, 절과 절을 결합하는 역할을 한다.

한국어 연결어미는 텍스트 표층에서 접속 기능을 실현하는 대표적 문법 형태소이다. 국립국어원 표준국어대사전에 따르면 연결어미는 어간에 붙어 다음 말에 연결하는 구실을 하는 어미를 말한다. 연결어미는 둘 이상의 관련된 사실이나 사건, 상황 등을 연결하여 하나의 사태로 통합하고 이 사태들이 어떠한 의미적 관계를 맺고 있는지 나타낸다(정수진 2013: 200). 따라서 연결어미는 단순한 통사적 기능을 넘어 화자의 논리 전개와 담화 구조를 결정하며 결속성(cohesion)에도 영향을 미친다.⁶⁾ 그러므로 생성형 AI가 중한 번역에서 한국어의 연결어미를 어떻게 처리하는지를 살펴보는 것은 생성형 AI의 언어 간 논리 관계 대응 능력, 나아가 결속 구성 능력을 평가할 수 있는 유효한 지표가 될 수 있다.

6) 유혜령(2005: 93-94)에 따르면 연결어미는 어휘 의미 이상의 중요한 형식적 의미를 표현할 뿐만 아니라 문장 범위를 넘어 접속 기능을 하는 경우도 있기 때문에 연결어미를 텍스트 내에서 접속을 나타내는 결속 장치 역할을 하는 것으로 본다.

표 1. 현대 국어 사용 빈도 조사(연결어미)

순위	항목	교재	항목	신문	항목	문학	항목	잡지
1	고	2,223	고	7,355	고	10,331	고	6,813
2	아	1,028	아	5,316	어	6,486	어	4,256
3	어	987	어	4,166	아	4,724	아	3,730
4	게	910	게	2,281	게	3,885	게	2,568
5	지	663	지	1,980	지	3,093	지	1,741
6	면	604	면	1,175	아서	1,582	면	1,232
7	아서	551	며	953	면	1,316	며	812
8	며	372	아야	806	며	1,219	아서	775
9	면서	333	아서	680	지만	955	지만	718
10	어서	230	어야	673	면서	952	면서	639

<표 1>은 국립국어원에서 제공하는 현대 국어 사용 빈도 조사⁷⁾ 보고서에서 연결어미를 항목별로 1-10위까지 정리한 것이다. 702개의 어미 가운데 연결어미는 281개로 사용 빈도가 전체의 약 40%를 차지할 정도로 비중이 높다. 항목별 순위에서 ‘-고’는 전체 항목에서 사용 빈도가 가장 높았고, ‘-며’는 전체 항목에서 7~8 위였다. 사용 빈도 측면에서 두 연결어미는 문어와 구어 텍스트에서 대표적으로 사용되는 연결어미임을 알 수 있다.

한국어는 연결어미의 종류와 쓰임이 다양한 만큼 기존의 선행 연구가 매우 풍부하다. 연결어미 ‘-고’, ‘-며’에 관한 기존의 연구는 크게 셋으로 나눌 수 있다.

첫째, 한국어 연결어미를 의미 기능에 따라 분류, 기술한 국어학 분야의 연구이다. 남기심(1978), 김종록(1991), 윤평현(2005)은 연결어미의 통사적 위치 및 의미 관계를 다룬 후에 연결어미의 유형을 의미 혹은 기능별로 분류하였다.

둘째, 한중 언어 대조 분야의 연구이다. 주효천(2018)은 한국어 연결어미 ‘-고, -며, -면서’의 의미를 나열, 동시, 순서, 대립으로 나누고 각각의 중국어 관련사(关联词)와의 대응을 분석했다. 종가재(2024)는 한국어 비중결어미 ‘-고’와 부사형 어미 ‘-고’의 중국어 대응 양상을 분석했다. 이들 연구는 두 언어의 유형적 차이에 따라 형태적 접속 표현에 어떤 차이가 있는지를 주로 고찰하였다. 이문화와 진홍

7) 국립국어원은 ‘국어 어휘 사용 빈도 조사’의 신뢰성 제고를 위해 2002년에 수행한 기초 연구를 바탕으로 3백만 어절의 문헌 자료에 대한 사용 빈도를 조사하여 ‘현대 국어 사용 빈도 조사 2’를 발간하였다.

월(2025)은 한국어 학습 능력에 따라 연결어미 ‘-고’의 상호 대치 오류를 분석하였다.

셋째, 중국어 무표지 복문의 한국어 번역 시 일어나는 명시화에 대해 분석한 통번역 분야의 연구이다. 장현주(2006), 이완국(2022)은 중국어 복문의 비명시적 논리 관계에 대한 유형을 분류한 후에 중한 번역 시 논리 관계가 명시화되는 양상을 상세하게 분석하였다.

그 밖에 박건숙(2024)은 한국어능력시험(TOPIK)의 읽기 텍스트를 대상으로 ‘그래서/그러나/그런데/그리고’와 같은 접속부사에 대한 챗GPT의 사용 양상을 분석하여 접속 관점에서 생성형 AI의 의미 해석 성능을 분석하였다.

3. 이론적 배경

3.1. 연결어미 ‘-고’, ‘-며’의 의미 기능

연결어미는 선·후행절을 접속해 일정한 의미적, 논리적 관계를 구성한다.⁸⁾ 연결어미는 형태적으로 어간에 붙어 다음 말을 연결하고 선·후행 언어 성분의 내용을 결속하는 접속의 표지이다. 또한 교착어인 한국어에서 연결어미는 반드시 유표적으로 실현되면 문장 성분의 일부로서 어간에 붙어 생략할 수 없다(한희정 2017: 24).

연결어미는 접속문을 만드는 과정에서 통사적 기능과 이들 간의 논리적 관계를 보여주는 의미적 기능을 동시에 갖는다. 또한, 하나의 형태가 다의적 의미 기능을 갖기도 하고, 서로 다른 형태가 유사한 의미 기능을 나타내기도 한다.⁹⁾ 따라서 연결어미의 선택은 선·후행절의 논리적 구성에 중요한 영향을 미친다. 본고는 AI 중한 번역 시 중국어 무표지 복문(ST)의 접속이 한국어 번역문(TT)에서 어떤 접속

-
- 8) 한국어 연결어미에 대해 연구자마다 접속 어미, 연결 접사 등 사용하는 용어의 차이가 있으나 가리키는 대상은 동일하다. 본고는 학교문법 용어인 ‘연결어미’를 사용한다. 한희정(2017)은 접속 어미라는 용어를 사용하고 있지만 연결어미와 가리키는 대상이 동일하다고 하였고 접속 기능을 하지 않는 보조적 연결어미는 접속 어미로 보지 않는다고 하였는데 본고도 동일한 관점을 유지한다. 예를 들어 표준국어대사전 연결어미 ‘-고’의 6번째 의미 항목은 보조 용언이 있는 구성에서 본용언에 붙는 연결어미로 기술되어 있지만 이 경우 접속 역할을 하지 않기 때문에 연결어미로 보지 않는다.
- 9) 채수희(2023: 499)에 따르면 연결어미가 표현하는 관계 의미가 다양하기 때문에 선행절의 의미를 기준으로 하든 선·후행절의 의미 관계를 기준으로 하든 어떤 경우에도 예외가 발생할 수밖에 없다.

표현으로 구현되고 의미 기능이 정확히 대응하는지에 초점을 맞추고 있으므로 연결어미 의미 기능의 관점에서 접근할 것이다.

연결어미 ‘-고’, ‘-며’의 의미 기능을 살펴보자.¹⁰⁾ ‘-고’, ‘-며’의 의미 기능은 국립국어원 표준국어대사전 WEB(이하 <표준>)에 상세하게 기술되어 있다.¹¹⁾ 본고는 <표준>의 기술에 ‘-고’, ‘-며’의 의미 기능을 체계적으로 정리한 윤평현(2005), 강현화 외(2017) 이금희(2022)의 분석을 종합하여 ‘-고’, ‘-며’의 의미 기능을 다음과 같이 정리하였다.

첫째, 나열은 의미적 의존 관계가 없는 선·후행절을 열거하는 것으로 ‘-고’, ‘-며’의 공통된 의미 기능이다. <표준>에서 ‘-고’는 두 가지 이상의 사실을 대등하게 벌여 놓는 연결어미, ‘-며’는 두 가지 이상의 동작이나 상태 따위를 열거할 때 쓰는 연결어미로 정의한다.

(1) a. 돌발은 (삭막하고/삭막하며) 답답하기만 했다.

b. 나는 유리창이 덜컥거리는 소리를 (들으며/*듣고) 이 글을 썼다.

예문(1a)는 돌발의 삭막함과 답답함 두 가지 상태를 나열하였는데 ‘-고’, ‘-며’의 자리를 교체해도 의미에 변화가 없다. 하지만 예문(1b)에서 ‘-고’, ‘-며’의 자리를 교체하면 의미가 달라진다. ‘-고’는 선후 관계를 나타내지만, ‘-며’는 동시 상황을 나타낸다. 동일한 의미 기능을 가지는 연결어미라도 맥락에 따라 교체 가능 여부 및 의미 기능이 달라진다.

둘째, 동시는 선·후행절 사건이 시차를 두지 않고 동시에 발생하는 것으로 ‘-고’, ‘-며’의 공통된 의미 기능이다. <표준>은 ‘-며’에만 동시 의미 기능이 있다고 기술

10) 연결어미 ‘-고’, ‘-며’의 의미 기능 분류에 관한 기존의 연구를 보면 연구자마다 분류상 차이가 존재한다. 하지만 대부분의 연구가 표준국어대사전을 기반으로 하고 있으며, 해당 사전의 의미 기능 분류가 조금 더 상세하기 때문에 본고는 표준국어대사전을 기준으로 한다. 이금희(2022: 610)에서도 한국어 사전에 수록된 연결어미 의미 기능이 한국어 문법서나 개별 연구보다 더 많다고 보았다.

11) (1) ‘-고’의 의미 기능

- a. 두 가지 이상의 사실을 대등하게 벌여 놓는 연결어미(나열)
- b. 앞뒤 절의 두 사실 간에 계기적인 관계가 있음을 나타내는 연결어미(계기)
- c. 앞 절의 동작이 이루어진 그대로 지속되는 가운데 뒤 절의 동작이 일어남을 나타내는 연결어미(지속)
- d. 서로 뜻이 대립되는 말을 벌여 놓는 연결어미(대립)
- e. 형용사 어간을 반복하여 그 뜻을 강조하는 연결어미(강조)

(2) ‘-며’의 의미 기능

- a. 두 가지 이상의 동작이나 상태 따위를 나열할 때 쓰는 연결어미(나열)
- b. 두 가지 이상의 움직임이나 사태 따위가 동시에 겹쳐져 있음을 나타내는 연결어미(동시)

하고 있지만, 이금희(2022)는 ‘-고’, ‘-며’ 모두 동시 의미 기능을 갖는다고 보았다.

- (2) a. 학생들이 술을 마시고 논다.
- b. 영화는 달을 보며 속삭인다.

예문(2a)에서 ‘-고’는 학생들이 술을 마시면서 노는 동시적 상황을 나타내고, 예문(2b)에서 ‘-며’는 영화가 달을 보는 행위와 속삭이는 동시 동작을 나타낸다.

이금희(2022: 615-620)는 ‘-고’, ‘-며’의 첫 번째 의미 기능인 나열과 두 번째 의미 기능인 동시에 관해 주어 동일 여부 및 사태의 연관성 여부에 따른 의미 차이를 지적하였다. 선·후행절 사건이 동일 주어인 경우, ‘-고’는 시간의 선후 관계, ‘-며’는 동시 관계를 나타내고, 비동일 주어인 경우 나열과 동시 관계를 나타낸다.

- (3) a. 집주인은 방문객의 신분을 확인하고(며) 문을 열어 주었다.
(동일 주어)
- b. 철수는 밥을 먹고 술을 마시며 영화는 떡을 먹는다.
(비동일 주어)

예문(3a)는 동일 주어의 예시로 ‘-고’는 신분 확인 후에 문을 열어 주는 시간의 선후 관계를, ‘-며’는 신분 확인과 문을 여는 행위가 동시에 일어남을 의미한다. 예문(3b)는 비동일 주어의 예시로 철수와 영화의 행위를 나열하거나 동시에 일어나는 상황을 기술한다. 상기 분석은 중국어 복문을 한국어로 번역할 때 연결어미 ‘-고’, ‘-며’ 용법에 대한 중요한 지침을 제공해 준다.

셋째, 시간은 ‘-고’의 의미 기능으로 선·후행절 사건을 시간의 선후 순서에 따라 순차적으로 연결하는 것이다. 이는 선·후행 사건의 순서를 통해 화자에 대한 인식과 관점을 드러낸다(강현화 외 2017: 119).

- (4) 손을 씻고 밥을 먹었다.

예문(4)에서 ‘-고’는 두 가지 행위를 순차적으로 연결해 손을 씻는 행위가 끝난 후에 밥을 먹는 행위가 있었음을 나타낸다. 이러한 시간의 선후 관계를 나타낼 때 ‘-고 나서’, ‘-후에’, ‘-뒤에’ 등으로 교체할 수 있다.

넷째, 계기는 ‘-고’의 의미 기능으로 원인 또는 이유로 이해할 수 있다. 선·후행

절 사건 사이에 인과 관계가 있고 시간적으로 겹치지 않음을 나타낸다.

(5) 할머니께서는 상한 음식을 드시고 탈이 나셨다.

예문(5)는 선·후행절 사건을 ‘-고’로 연결하였는데 선행 사건인 상한 음식을 먹은 것을 계기, 후행 사건인 탈이 난 것을 결과로 해석할 수 있다.

종합해 보면 연결어미 ‘-고’는 나열, 동시, 시간, 계기의 네 가지 의미 기능, 연결어미 ‘-며’는 나열, 동시의 두 가지 의미 기능을 가진다. 두 연결어미의 공통된 의미 기능은 나열, 동시이다. ‘-고’, ‘-며’는 공통된 의미 기능에 따라 서로 자리를 바꾸어 사용할 수 있다. 하지만 둘을 함께 사용할 경우 출현 순서나 맥락에 따라 의미 관계가 달라질 수도 있음에 유의해야 한다.

3.2. 중국어 무표지 복문

중국어 복문은 둘 이상의 의미적 연관성을 가지며 구조적으로 독립된 단문 형식으로 이루어진 문장이다(자오은팡과 탕쉐닝(赵恩芳, 唐雪凝) 1998: 11).¹²⁾ 중국어 복문 유형에 대한 여러 분류 체계가 연구되어 있지만¹³⁾ 접속 표지의 유무를 기준으로 ‘무표지 복문’과 ‘유표지 복문’으로 나눌 수 있다. 유표지 복문은 접속 어구(关连词语)¹⁴⁾와 같은 형태적 표지를 사용해 절과 절, 문장과 문장을 연결하는 형태이고, 무표지 복문은 접속 어구 없이 어순에 따라 나열한

12) 중국어 복문을 구성하는 언어 성분 및 연결 방식에 대한 다양한 논의가 있다. 본고는 무표지 복문의 구분 기준에 대해 김아영(2004)의 관점을 사용한다. 김아영(2004) <한중 통번역에서 중국어 무표지 연결 절의 관계 의미 연구>에서는 중국어의 마침표(.)와 쉼표(,) 두 가지 ‘문장부호’를 기준으로 복문을 구분하였다. 그에 따르면 본래 절이나 문장은 주요 동사 또는 술어와 그에 수반되는 논항을 토대로 정의되는 것이 일반적이므로 실제 사용되는 문장부호에 의한 정의는 설명의 단위로서 부적절한 면이 없지 않지만, 연구의 인용된 예문에서 일부 주제구나 부사구를 제외하면 문장부호에 의한 정의와 통사적 기준에 의한 정의가 대체로 일치하므로 ‘문장부호’를 기준으로 한다고 밝혔다.

13) 정운철(2005)는 자오은팡과 탕쉐닝(1998)을 참고해 중국어 복문을 다음과 같이 나누었다.

첫째, 형식에 따라 ‘일반 복문’과 ‘긴축 복문’으로 분류
둘째, 절 사이의 의미관계에 따라 ‘연합 복문’, ‘수식 복문’으로 분류
셋째, 절 사이의 층위에 따라 ‘단층 복문’과 ‘다층 복문’으로 분류
넷째, 접속표지의 유무에 따라 ‘무표지 복문’과 ‘유표지 복문’으로 분류
다섯째, 어기와 기능에 따라 ‘진술문’, ‘의문문’, ‘명령문’, ‘감탄문’으로 분류

14) 천창라이(陈昌来 2000: 267)에 따르면 접속 어구에는 접속사(因为, 所以, 不但, 而且 등), 관계 부사(就, 又, 都 등), 특정한 구(的话, 之所以, 若不是 등)가 있다.

형태이다. 무표지 복문은 중국어 의합(义合) 관계에 따라 그 논리적 관계를 추측하고 이해할 수 있다.

- (6) 只有大量植树种草, 才能防止水土流失.
(나무를 많이 심어야만 토양 유실을 막을 수 있다)
- (7) 二是坚持对外开放, 做国际自由贸易的倡导者.
(a. 대외 개방을 유지하고, 세계 자유무역의 창도자가 된다)
(b. 대외 개방을 유지함으로써(유지를 통해) 세계 자유무역의 창도자가 된다)
(c. 세계 자유무역의 창조자가 되기 위해 대외 개방을 유지한다)

예문(6)는 유표지 복문으로 ‘只有~才’라는 접속 어구를 통해 ‘~해야만 ~하다’는 의미를 문장의 표층에서 쉽게 파악할 수 있다. 예문(7)은 무표지 복문으로 쉽표를 기준으로 하여 선·후행 절에 형태적 표지가 없다. 선행절은 일종의 방식이고 후행절은 이 방식을 통해 도달하고자 하는 목표로 그 의미를 추론할 수 있다. 예문(7a)는 선·후행 절의 의미를 단순히 나열해 절 간의 논리적 관계를 파악할 수 없다. 예문(7b), (7c)는 연결어미 ‘-어(서)’와 같은 형태적 표지를 명시화하여 절 간의 의미 관계를 살려주었다. 이와 같이 중국어 복문은 형태적 표지의 유무가 자유로운 데 비해 한국어는 문장과 문장, 절과 절을 연결할 때 반드시 형태적 표지가 필요하다.

자오은광과 탕쉐닝(1998), 판샤오(范晓 2000), 류웨화 외(刘月华 외 2019)는 중국어 복문 유형을 구분한 대표적인 학자로 의미에 따라 무표지로 사용 가능한 복문의 유형을 구분하였다.¹⁵⁾ 장현주(2006)는 중국어 복문의 의미 관계에 따라 무표지 경향을 나타낼 수 있는 유형을 정리하고 중한 번역 시 어떻게 명시화되는지 분석하였다.¹⁶⁾

15) 중국어는 내포는 인정하지 않고 접속만 인정하기 때문에 무표지 복문의 유형을 주로 선·후행 절간의 의미 관계에 따라 구분하지만 학자들마다 차이가 있다(전기정 2018: 90).

16) 장현주(2006)는 논리 관계 표지의 명시화 여부는 텍스트 생산자 의도에 영향을 받으므로 절대적이지 않기 때문에 표지 사용의 ‘편향성’에 따라 의미를 구분한 것임을 밝혔다.

표 2. 무표지 복문으로 사용 가능한 의미 기능

학자	무표지 복문으로 사용 가능한 접속 유형
자오은팡과 탕쉐닝 (1998)	병렬(并列), 연속(承接), 설명(解说), 전환(转折), 인과(因果)
판샤오(2000)	병렬(并列), 연속(连贯) ¹⁷⁾ , 인과(因果), 주석(注释), 설명(总分), 묘사설명(记叙), 보충설명(表相)
류웨화 외(2019)	병렬(并列)), 연속(承接), 설명인과(说明因果), 가정(假设), 시간(时间)
장현주(2006)	①명시/비명시 가능: 시간, 부가, 예시, 원인·이유, 결과, 가정, 대립· 대비, 전환 ②비명시 편향: 시간, 순접, 병렬, 총괄-해석

<표 2>는 용어상 차이가 있지만 병렬, 연속, 설명, 인과가 무표지 복문으로 사용할 수 있는 공통 의미 기능이다.

첫째, 병렬은 문장이나 절 사이의 의미 관계가 독립적이고 대등하여 대부분 서로 자리를 바꾸어도 의미 변화가 없다. 또한, 선·후행절의 사건이나 사태가 공존함을 나타낸다. 이때 ‘又A, 又B’, ‘一方面A, 一方面B’, ‘同时’ 등의 표지를 사용해 의미를 검증할 수 있다(황보룽과 라오쉬둥(黄伯荣, 廖序东) 2017: 130). 병렬은 사건이나 사태의 나열과 공존을 모두 나타낼 수 있어서 한국어 연결어미 ‘-고’, ‘-며’ 의미 기능에서 나열과 동시에 대응시킬 수 있다.

(8) 书架上有小说, 有杂志.
책장에 소설이 있고 잡지가 있다.

(9) 他写报告, 我整理数据.
그는 보고서를 쓰고 나는 데이터를 정리한다.

예문(8), (9)는 두 개의 절이 아무런 표지 없이 의미상 대등하게 놓여 병렬 관계를 이루는 무표지 복문이다. 예문(8)에서 책장에 꽂힌 소설과 잡지는 서로 독립적이고 대등하여 영향을 주지 않는다. 예문(9)는 ‘그’와 ‘나’의 행동이 공존함을 나타

17) 장현주(2006)는 ‘连贯’은 하나의 주제 아래에 각 요소가 대등한 입장에서 나열되어 있으며 전체 주제를 이루는 부분이 된다고 보았다. 본고는 이를 연관 관계로 부르기로 하며 혼동을 피하기 위해 중국어를 병기하였다.

낸다.

둘째, 연속(承接)은 연속적으로 발생한 둘 이상의 행동이나 사건을 순차적으로 서술하는 것이다. 선행 사건이 후행 사건으로 이어지면서 시간적, 논리적 순차성을 나타낸다. 장현주(2006)의 시간 관계가 여기에 해당하며, 한국어 연결어미 ‘-고’ 의미 기능 중 시간에 대응시킬 수 있다.

(10) 他推门进屋，打开灯。

그가 문을 열고 들어가 불을 켜다.

(11) 他想了想，点点头。

그는 잠깐 생각하더니 고개를 끄덕였다.

예문(10)은 문을 열고 불을 켜는 연속된 행동, 예문(11)은 생각하고 고개를 끄덕이는 연속적 행동을 시간 순서에 따라 서술하였다. ‘接着’, ‘然后’ 등의 접속 어구를 추가해 자연스럽게 연결할 수도 있다.

셋째, 설명(总分)은 선행절이 후행절을 총괄하거나 후행절이 선행절을 보충, 해설, 정의하여 내용을 구체화 또는 상세화하는 것이다. 류웨화 외(2019)의 ‘分合关系’, 자오은팡과 탕쉐닝(1998)의 ‘解说’와 동일하다(정윤철 2005: 117). 앞 절에서 정리한 한국어 연결어미 ‘-고’, ‘-며’와 대응하는 의미 기능은 없다.

(12) 母亲同情贫苦的人，这是朴素的阶级意识。

어머니는 가난한 사람들을 동정하셨는데 이것은 소박한 계급 의식이다.

(13) 艺术批评有两个标准：一个是正式标准，一个是艺术标准。

예술 비평에는 두 가지 기준이 있는데 하나는 형식적 기준이고 다른 하나는 예술적 기준이다.

예문(12)는 선행절에 서술된 어머니의 동정심에 대해 후행절에서 해설하고, 예문(13)은 선행절에서 예술 비평의 두 가지 기준을 총괄하고 후행절에서 각각의 부분적 내용을 구체화한다.

넷째, 인과는 하나의 절이나 문장이 원인, 다른 절이나 문장이 결과가 되는 구조로 결과 절에는 상태나 상황의 변화, 평가, 결론에 관한 서술이 등장한다. 장현주(2006: 38)에 따르면 원인(이유)-결과, 이유(근거)-추론, 목적-방법, 조건-결과, 가정

등 의미 기능이 여기에 속하며, 대부분 인과 관계에서 접속 어구를 사용해 논리적 관계를 나타내지만 인지적 또는 화용적으로 추론이 가능한 경우 표지를 생략할 수도 있다. 한국어 연결어미 ‘-고’, ‘-며’ 의미 기능 중 계기와 대응시킬 수 있다.

(14) 天黑了，路上没有人了。
 날이 어두워져서 길에 사람이 없다.

(15) 他生病了，没来上课。
 그는 아파서 수업에 오지 않았다.

예문(14), (15)는 선행 문장이 원인, 후행 문장이 결과를 나타내는 인과 관계로 형태적 표지가 없지만 의미적 연관성 및 논리적 흐름을 추론하여 이해할 수 있다. 이상으로 한국어 연결어미 ‘-고’, ‘-며’의 대표적 의미 기능과 무표지 복문으로 사용 가능한 의미 기능을 살펴보았다. 분석에 활용할 이론적 틀로서 의미 기능의 대응 관계를 표로 정리하였다.

표 3. 중국어 무표지 복문과 한국어 연결어미 ‘-고’, ‘-며’ 의미 기능 대응

중국어 무표지 복문의 의미 기능	한국어 연결어미 ‘-고’, ‘-며’ 의미 기능	연결어미 대응
병렬	나열	-고, -며
병렬(동시)	동시	-고, -며
연속	시간	-고
설명	-	-
인과	계기	-고

중국어 무표지 복문의 병렬은 나열과 동시 의미 기능을 함께 가진다. 병렬 의미로 접속한 것은 병렬, 병렬이지만 동시로 접속한 것은 병렬(동시)로 서술하였다.

4. 분석 방법 및 대상

AI 번역은 챗GPT와 같은 생성형 AI 모델을 도구로 수행한 번역을 뜻한다.

본고는 AI 중한 번역 시 중국어 무표지 복문의 의미 기능이 한국어 연결어미 ‘-고’, ‘-며’로 명시화하는 양상과 의미 기능 대응을 살펴보는 것이 목적이다. 이를 위해 챗GPT와 하이퍼클로바X 두 가지 생성형 AI 모델을 사용하였다.

챗GPT는 OpenAI에서 대규모 언어모델(LLM)을 기반으로 개발한 대화형 인공지능이다. 방대한 텍스트 데이터를 바탕으로 언어의 패턴, 문법, 의미적 관계를 학습해 인간과 같은 자연스러운 문장을 생성한다. Common Crawl의 공개 자료에 따르면, 2023년 3월 기준 사전학습 데이터에 사용된 문서의 약 46%가 영어 텍스트였다. OpenAI 모델이 영어 중심의 데이터를 기반으로 사전학습 되었음을 알 수 있다. 본고는 챗GPT-4o 모델을 사용하였는데 이는 GPT-4 계열 최신 모델로 기존 GPT-4와 동등한 수준의 언어 이해 및 추론 능력을 유지하면서도 고도화된 추론 능력과 높은 수준의 번역 성능을 갖추고 있다.¹⁸⁾

하이퍼클로바X(HyperCLOVA X)는 네이버(NAVER)에서 2023년 8월 공개한 한국어 특화 생성형 인공지능으로 선행 모델인 하이퍼클로바를 기반으로 한다. 해외의 다른 생성형 AI 모델 대비 약 6,500배 많은 한국어 데이터를 중심으로 학습하였다. 그 결과 한국어 문법, 어휘, 담화 구조를 처리하는 데 높은 정확성과 자연스러움을 보였으며 특히 중한 번역에서 자동 평가 점수 88.8의 높은 성능을 보였다(clova.ai).¹⁹⁾

본고는 생성형 AI 모델의 번역 성능 비교에 초점을 두지 않지만, 한국어 텍스트 중심으로 학습한 모델의 번역 결과물을 함께 비교함으로써 생성형 AI 번역 시 접속 표현 처리를 보다 객관적으로 확인하고자 하였다. 생성형 AI 모델의 결과물에 오류(hallucination) 등이 존재할 수 있으므로 결과 분석 과정에서 검증 및 보완 절차를 병행하였다.

분석 데이터는 중국어 신문 기사와 연설문이다. 텍스트 유형에 따라 신문 기사는 정보적 텍스트, 연설문은 작용적 텍스트에 속한다.²⁰⁾ 두 가지 유형 모두 논리성

18) GPT-3.5는 주로 텍스트 생성과 분석에 한정되었고, GPT-4는 추론 능력과 복잡한 언어 과제에 서 성능이 크게 개선되었다. 사킵 샤리אר 외(Sakib Shahriar et al. 2024: 15)에 따르면 GPT-4o는 few-shot 설정에서 높은 정확도와 효율성을 보였으며, 멀티모달 과제에서 기존 모델 대비 유의미한 향상을 보였다.

19) [https://clova.ai/hyperclova?utm_source=chatgpt.com\(2025.09.09.검색\)](https://clova.ai/hyperclova?utm_source=chatgpt.com(2025.09.09.검색)).

20) 라이스와 페르메어(Reiss and Vermeer 1984: 182-183)에 따르면 텍스트는 실재하는 세계의 사물과 현상을 독자에게 전달하는 정보적 텍스트(Informative Text), 독자에게 미적 효과를 가져오는 표현적 텍스트(Expressive Text), 독자가 출발 언어 텍스트와 같은 반응을 일으키도록 유도하는 호소적 텍스트(Operative Text)로 나눌 수 있다.

을 바탕으로 독자나 청중을 설득하는 성격이 강한 텍스트이기 때문에 한국어로 번역할 때 정확한 접속 표현을 이루는 것이 매우 중요하다. 중국어 신문 기사는 환구시보(环球时报), 신화망(新华网), 인민망(人民网) 세 개의 기관지에 실린 논설에서 선정하였다(총 5,854자). 연설문은 중국 외교부(外交部) 홈페이지에 공개된 고위급 인사의 연설에서 선정하였다(총 6,482자).

생성형 AI의 번역 결과물은 프롬프트의 조건이나 내용에 영향을 받는다. 본고는 번역 결과물이 프롬프트로부터 받는 영향을 최소화하여 생성형 AI의 중한 번역 접속 표현의 문제점을 파악하기 위해 프롬프트에서 텍스트 유형에 관한 정보만 주었다. 신문 기사는 ‘이 중국어 신문 기사를 자연스러운 한국어로 번역해 줘’, 연설문은 ‘이 연설문을 자연스러운 한국어로 번역해 줘’라고 프롬프트를 설정하였다.

표 4. 분석에 활용한 중국어 신문 기사와 연설문

유형	제목	출처	글자 수	게재 및 발표 일자
신문 기사	社评：面对关税讹诈，绥靖妥协只会是“与虎谋皮”	环球时报	1509	2025.04.22
	新华时评：关税战，一场饮鸩止渴的游戏	新华网	1768	2025.04.03
	人民网三评“美国滥施关税”之一：世界要公道，不要霸道	人民网	859	2025.05.09
	人民网三评“美国滥施关税”之二：全球要合作，不要分裂	人民网	696	2025.04.30
	人民网三评“美国滥施关税”之三：经济要繁荣，不要衰退	人民网	1022	2025.05.01
연설문	习近平在中非合作论坛北京峰会开幕式上的主旨讲话	外交部	2990	2024.09.05
	王毅在中非合作论坛成果落实协调人部长级会议开幕式上的主旨讲话	外交部	1733	2025.06.11
	王毅集体会见非洲国家驻华使节时的讲话	外交部	1759	2025.05.26

중국어 신문 기사와 연설문 무표지 복문에 나타난 5가지 의미 기능 병렬, 병렬(동시), 연속, 설명, 인과를 사용 빈도별로 정리하였다.

표 5. 신문 기사와 연설문 무표지 복문의 의미 기능별 사용 빈도

분류	접속 표현별 사용 빈도
신문 기사	병렬(89), 병렬(동시)(12), 인과(35) 설명(26), 연속(14)
연설문	병렬(99), 병렬(동시)(7), 연속(39), 인과(56), 설명(5)

위 표에서 의미 기능별 사용 빈도를 살펴보면 신문 기사는 병렬이 89회, 연설문은 병렬이 99회로 병렬의 접속이 가장 많이 사용되었다. 상기 네 가지 주요 의미 기능 외에 양보, 선택과 같은 다른 의미 기능도 출현했지만 대부분 빈도가 1~2회에 불과했다. 의미 기능 중 병렬을 제외하면 신문 기사와 연설문에서 사용된 접속 표현을 통해 논리 관계 측면에서 두 텍스트 유형의 차이를 파악할 수 있다.

신문 기사에 사용된 의미 기능은 병렬 다음으로 인과, 설명, 연속 순으로 사용 빈도가 높았다. 신문 기사, 특히 논설은 어떤 사건이나 현상에 관한 주장을 논리적으로 풀어내는 텍스트로 원인과 결과의 접속이 많다. 이에 비해 연설문에 사용된 의미 기능의 사용 빈도는 병렬 다음으로 연속, 인과 순이었다. 연설문은 공동의 목표를 중심으로 과거부터 현재까지의 경과, 그리고 앞으로의 비전을 제시하는 일종의 서사를 형성하며 청중에게 호소하는 텍스트로 연속의 접속이 신문 기사보다 약 3배 가량 많이 사용되었다. 물론 연사의 주장을 논리적으로 전개해야 하기 때문에 신문 기사와 같이 인과의 접속도 사용 빈도가 높았다. 신문 기사와 연설문이 텍스트의 성격에 따라 접속을 구성할 때 의미 기능 사용 빈도에 차이가 있음을 알 수 있다. 따라서 이를 생성형 AI로 한국어로 번역할 때 의미 기능에 따른 접속 표현의 차이가 정확히 반영되어야 할 필요성을 확인할 수 있다.

5. 분석 결과

5.1. 의미 기능의 편향 및 텍스트 핵심 의미 기능의 소실

중국어 무표지 복문은 신문 기사 98개, 연설문 108개였고, 무표지로 연결된 문장과 절은 신문 기사 184개, 연설문 212개였다.

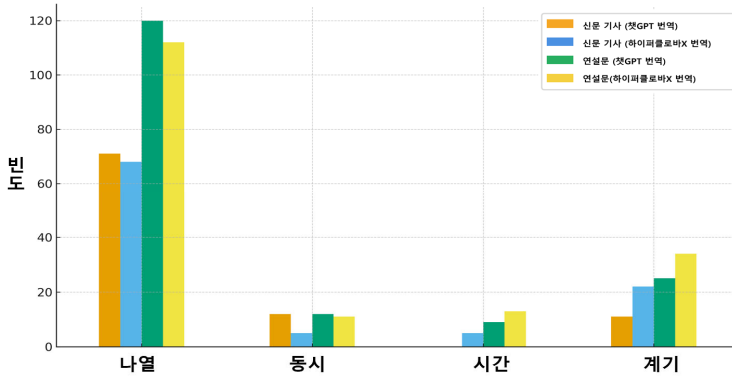


그림1. 생성형 AI 중한 번역문의 의미 기능 분포

<그림 1>은 생성형 AI 중한 번역문에서 나타난 한국어 연결어미 ‘나열’, ‘동시’, ‘시간’, ‘계기’의 네 가지 의미 기능을 비교한 히스토그램이다. 생성형 AI의 신문 기사와 연설문 중한 번역을 비교 분석한 결과, <그림 1>에서 보듯이 한국어 번역문에서 ‘나열’ 의미 기능이 과다하게 실현되는 경향이 뚜렷하게 나타났다. 이는 생성형 AI가 무표지 복문에 내재된 논리 관계에 대한 추론 정확성이 낮고 안전한 형태의 의미 기능을 선호하는 편향을 보여준다.

의미 기능 ‘나열’이 과다하게 편향되어 나타난 결과, 텍스트 고유의 핵심적 의미 기능 소실을 가져왔다. 이를 살펴보기 위해 텍스트 유형별로 접속 표현과 의미 기능의 대응 관계를 정리하였다.²¹⁾

21) 이론적 배경에서 소개한 접속 표현 외에 분석 과정에서 생성형 AI는 무표지 복문을 주어, 분절, 부사, 조사 등 다양한 형태로 접속하였다. <표 6>, <표 7>에서 ‘주어’는 선행 문장(절)을 주어로 인식해 주격 조사 은/는/이/가로 접속한 경우이다. ‘분절’은 쉼표로 연결되어 있지만 한국어 번역문에서 마침표로 문장을 종결한 경우이다. ‘부사’는 한국어 접속부사를 써서 논리 관계를 명시화한 경우이다. ‘조사’는 ‘-에서’, ‘-(으)로’와 같이 연결어미가 아닌 다양한 조사를 활용해 접속을 구성한 경우이다.

표 6. ST 접속 표현과 TT 의미 기능 간 최고 대응 빈도

신문기사 (챗 GPT 번역)	병렬-나열 (39)	병렬(동시)- 나열(6)	연속-나열(6)	설명-주어(9)	인과-나열(18)
신문기사 (하이퍼 클로바 X 번역)	병렬-나열 (39)	병렬(동시)- 나열(4)	연속-나열(6)	설명-주어 (12)	인과-계기(18)
연설문 (챗 GPT 번역)	병렬-나열 (75)	병렬(동시)- 동시(7)	연속-나열(16)	설명-주어(2)	인과-나열(27)
연설문 (하이퍼 클로바 X 번역)	병렬-나열 (75)	병렬(동시)- 동시(3)	연속-나열(15)	설명-비교(2)	인과-계기(32)

<표 6>을 보면 텍스트 핵심 의미 기능의 소실은 신문 기사 번역에서 그 경향이 특히 뚜렷했다. 챗GPT는 신문 기사에서 주장과 근거를 연결하는 핵심적 접속 표현인 ‘인과’를 ‘나열’로 오역하는 사례가 빈번하였다. 그 결과 원문 텍스트가 전달 하려는 논리적 인과성이 크게 약화되었음을 알 수 있다. 연설문에서도 동일한 현상이 관찰되었다. 연설문은 일반적으로 시간적 순차성과 연속성을 중요한 담화적 특징으로 볼 수 있지만 생성형 AI는 ‘연속-시간’ 대응을 적절히 처리하지 못하고 ‘연속-나열’로 처리하는 경향을 보였다. 이로 인해 연설문의 논리적 전개와 자연스러운 담화의 흐름이 약화되었다. 챗GPT와 하이퍼클로바X 모두 ‘나열’ 의미 기능이 광범위하게 나타나면서 병렬적 구조를 선호하는 경향이 관찰되었고 이는 생성형 AI 모델의 언어 출력 패턴이 편향되어 있음을 의미한다.

단, 하이퍼클로바X에서 ‘인과-계기’ 대응 비율이 챗GPT보다 조금 높게 나타난 점은 한국어 데이터 학습량의 효과를 짐작해 볼 수 있는 대목이기는 하나 전체적으로 두 모델에서 논리 관계의 추론 능력은 제한적인 것으로 판단된다.

이러한 결과는 앞의 절에서 언급했듯이 생성형 AI가 한국어 연결어미 체계를 충분히 반영하지 못하고 병렬 관계를 나타내는 접속 부사, 연결어미, 조사 등을 과도하게 선택한 데서 기인한 것으로 보인다. 이로 인해 텍스트 고유의 의미 기능이 손상되면서 비록 표층적으로는 언어 규범에 맞는 자연스러운 한국어 문장을 생성

했지만 한국어 모어 화자가 읽었을 때 어색함을 느끼게 되는 결과를 가져오게 된다. 이를 통해 생성형 AI 번역이 문장 층위에서는 문법적 정확성이 높아졌지만 담화 층위에서의 의미 기능 및 결속성 구현 능력은 여전히 인간 번역사에 미치지 못할뿐 아니라 생성형 AI가 텍스트 유형별 접속 표현의 논리 관계에 대한 정확한 추론이 부족하다는 사실을 알 수 있다. 따라서 향후 번역 품질 평가 시에 문장 수준의 정확성뿐 아니라 연결어미 사용의 다양성과 접속 표현의 논리적 정확성 등을 종합적으로 고려할 필요가 있다.

5.2. 연결어미 ‘-고’ 의미 기능에 따른 접속 표현 분석

신문 기사와 연설문에서 접속 표현(병렬, 병렬(동시), 설명, 시간, 인과)에 해당하는 무표지 복문을 추출하고 한국어 연결어미 ‘-고’의 4가지 의미 기능(나열, 동시, 시간, 계기)과 대응시켰다. 실제 분석 과정에서 중국어 무표지 복문은 한국어의 접속 부사, 조사, 시간명사 등 다양한 형태를 통해 접속을 이루었다.

<표 7>은 신문 기사와 연설문에 사용된 무표지 복문의 접속 표현과 연결어미 ‘-고’의 의미 기능과의 대응을 분석한 결과이다.

표 7. 한국어 연결어미 ‘-고’의 접속 표현 대응(빈도)

분류	신문 기사		연설문	
	ST 접속 표현	‘-고’ 의미 기능	ST 접속 표현	‘-고’ 의미 기능
챗 GPT	병렬(89)	나열(14)	병렬(99)	나열(45)
	병렬(동시)(12)	나열(5), 동시(1)	병렬(동시)(7)	-
	연속(14)	나열(3)	연속(39)	나열(12), 시간(1)
	인과(35)	나열(7)	인과(56)	나열(11), 계기(1)
	총 사용 빈도	30	총 사용 빈도	70
하이퍼클로바 X	병렬(89)	나열(11)	병렬(99)	나열(40)
	병렬(동시)(12)	나열(3)	병렬(동시)(7)	나열(2)
	연속(14)	나열(3)	연속(39)	나열(11)
	인과(35)	나열(2)	인과(56)	나열(12), 계기(1)
	총 사용 빈도	19	총 사용 빈도	66

<표 7>에서 신문 기사 번역문에서 연결어미 ‘-고’의 사용 빈도를 보면 생성형 AI 모델별로 각각 30회, 19회 사용되었으며 모델별로 유의미한 차이는 보이지 않았다. 신문 기사에 출현한 무표지 복문 중 병렬로 접속된 경우가 89회로 이중 한국어 연결어미 ‘-고’로 구현된 것은 14회로 의미 기능은 나열이었다. 연설문 번역문의 경우 생성형 AI 모델별로 연결어미 ‘-고’가 각각 70회, 66회 사용되었으며 모델별로 유사한 사용 빈도를 나타냈다. 연설문에 출현한 무표지 복문 중 병렬로 접속된 경우는 99회로 이중 한국어 연결어미 ‘-고’로 구현된 것은 40-45회로 의미 기능은 나열이었다. 텍스트 유형별로는 연설문에서 ‘-고’의 사용 빈도가 신문 기사보다 2배 이상 많았다.

중국어 무표지 복문의 접속 표현과 연결어미 ‘-고’의 의미 기능 대응 측면에서 살펴보면, 연결어미 ‘-고’는 일부를 제외하고 모두 나열의 의미 기능으로 가장 많이 사용되어 한 가지 의미 기능에만 편중된 것으로 나타났다. 이는 접속 표현과 의미 기능 간 대응이 불일치할 가능성이 높은 것으로 해석할 수 있다. 신문 기사와 연설문에서 ST 접속의 병렬과 연결어미 ‘-고’의 의미 기능 나열의 대응을 제외하면 다른 접속 표현이 대부분 나열 의미 기능으로 잘못 번역되었다. 즉, 생성형 AI가 중한 번역에서 접속 표현을 부정확하게 처리한 것이다. 이는 표층에 드러난 형태적 표지가 없는 경우 생성형 AI가 문장이나 절 간의 의미 관계를 정확하게 추론하지 못해 접속 표현의 정확한 구현에 취약하다는 점을 보여주는 것이며 그 결과 번역문의 논리 관계를 약화시켜 번역 품질 저하를 가져오는 원인이 된다.

(16)

[ST]①经济滞胀风险上升, ②美元信用遭受冲击, ③美元资产遭大规模抛售。

[TT-C] 경제 침체와 인플레이션이 동시에 나타나는 스태그플레이션 (stagflation) 우려가 커지고(나열), 달러에 대한 신뢰가 훼손되며, 달러 자산이 대거 매각되는 상황이 발생할 수 있다.

[TT-X] 경기 침체 위험이 높아지고(나열) 미국 달러 신용이 타격을 받으며, 미국 자산에 대규모 매도가 일어날 수 있습니다.

[대안 번역] 경기 침체 리스크의 상승으로(으로 인해/에 따라) 달러화 신용도가 충격을 받아(받으면서) 달러 자산의 대규모 매각이 발생했다.

예문(16)은 신문 기사의 무표지 복문 중 인과의 접속 표현이 연결어미 ‘-고’의 나열의 의미 기능으로 잘못 구현된 예시이다. 신문 기사는 인과 관계를 명확히 드러내야 하는 텍스트임을 앞에서 확인하였다. ST의 절 ①, ②, ③번 사이에 아무런 형태적 접속 표지가 없지만 모두 인과의 논리 관계로 연결되어 있다. ①번이 원인이 되어 ②번 결과가 발생하고, ②번이 다시 원인이 되어 결과적으로 ③번 결과가 발생하는 연쇄적 관계이다. 따라서 TT에서는 인과(계기)의 의미 기능을 갖는 접속 표현을 사용해야 한다. 하지만 챗GPT와 하이퍼클로바X 모두 ①, ②번 절을 나열로 처리하여 원문의 논리 관계를 유지하지 못했다. 따라서 대안 번역과 같이 인과(계기)의 의미 기능을 갖는 다양한 접속 표현을 활용해 논리 관계와 결속성을 유지하는 것이 좋다.

(17)

[ST] ①推动中国企业投资药品生产, ②继续对非洲遭遇的疫情提供力所能及的帮助。

[TT-C] 중국 기업이 현지에 의약품 생산에 투자하도록 촉진하고(나열), 아프리카의 전염병 대응에 가능한 도움을 계속 제공할 것입니다.

[TT-X] 중국 기업이 의약품 생산에 투자하도록 촉진하고(나열), 아프리카가 직면한 전염병에 대해 최대한 지원을 제공할 것입니다.

[대안 번역] 중국 기업의 의약품 생산에 대한 투자를 촉진하여(함으로써) 아프리카 전염병 해결을 위해 최대한의 지원을 계속 제공할 것입니다.

예문(17)은 연설문의 무표지 복문 중 연속의 접속 표현이 연결어미 ‘-고’의 나열의 의미 기능으로 잘못 대응된 예시이다. 연설문은 연속의 접속 표현 출현 빈도가 높은 텍스트적 특성을 갖는다. 무표지로 나열된 ST의 절 ①, ②는 연속의 논리 관계를 이루고 있다. 즉, ①번의 행위가 선행된 후에 ②번의 목적에 도달하는 시간적 순차성을 내포하고 있다. 따라서 TT에서는 연속의 의미 기능을 갖는 접속 표현을 사용해야 한다. 하지만 챗GPT와 하이퍼클로바X 모두 ①, ②번 절이 마치 별개인 것처럼 분절하여 단순 나열함으로써 원문의 논리 관계를 충분히 전달하지 못했다.

여기서 생각해 볼 부분은 채숙희(2023)에서 지적했듯이 연결어미가 표현하는 관계 의미는 실로 다양하여 선행절의 의미를 기준으로 하든 선·후행절의 의미 관

계를 기준으로 하든 어떤 경우에도 예외가 있고 독자에 따라 주관적 해석이 열려 있다는 점이다. 즉, 어떤 독자는 예문(16)의 ‘-고’를 인과(계기)의 의미로, 예문(17)의 ‘-고’를 연속의 의미로 이해할 수도 있다. 하지만 연결어미 ‘-고’는 나열의 의미 기능이 강하기 때문에 인과(계기)로 이해하려면 충분한 맥락이 주어져야 한다. 번역문을 읽었을 때 선·후행절의 논리적 결속 관계가 즉시 이해되지 않으면 문장을 여러 번 반복해서 읽어야 하고 이런 경우에 가독성이 크게 떨어질 수밖에 없다. 따라서 특정 의미 기능을 단번에 효과적으로 보여줄 수 있는 연결어미나 접속 표현의 사용을 지향해야 한다.

5.3. 연결어미 ‘-며’ 의미 기능에 따른 접속 표현 분석

<표 8>는 신문 기사와 연설문에 사용된 무표지 복문의 접속 표현과 연결어미 ‘-며’의 의미 기능과의 대응을 분석한 결과이다.

표 8. 한국어 연결어미 ‘-며’의 접속 표현 대응(빈도)

분류	신문 기사		연설문	
	ST 접속 표현	‘-며’ 의미 기능	ST 접속 표현	‘-며’ 의미 기능
챗 GPT	병렬(89)	나열(24)	병렬(99)	나열(29)
	병렬(동시)(12)	동시(5), 나열(1)	병렬(동시)(7)	동시(7)
	총 사용 빈도	30	총 사용 빈도	30
하이퍼클로바 X	병렬(89)	나열(28), 동시(1)	병렬(99)	나열(35)
	병렬(동시)(12)	동시(2), 나열(1)	병렬(동시)(7)	동시(6)
	총 사용 빈도	32	총 사용 빈도	41

<표 8>을 보면 신문 기사 번역문에 출현한 무표지 복문 중 병렬로 접속된 89회 중 한국어 연결어미 ‘-며’로 구현된 것은 30~32회로 대부분의 의미 기능이 나열이었다. 연설문 번역문에 출현한 무표지 복문 중 병렬로 접속된 99회 중 한국어 연결어미 ‘-며’로 구현된 것은 30~41회로 대부분 나열의 의미 기능으로 접속되었다. 연결어미 ‘-며’는 주로 나열과 동시 두 가지 의미 기능을 갖는데, 총 사용 빈도에서 생성형 AI 모델별, 텍스트 유형별로 유의미한 차이는 보이지 않았다.

분석 결과, 한국어 번역문에서 연결어미 ‘-며’ 역시 주로 나열의 의미 기능으로

구현되는 편중 양상을 보였다. 이러한 현상은 AI 모델이 문장 간 의미 관계를 충분히 해석하지 못하고 표층적으로 인접한 절을 단순한 병렬 관계로 처리하는 경향을 반영하는 것이다. 즉, 생성형 AI는 형태적 표지가 없는 중국어 무표지 복문의 해석에서 논리적 결속 관계를 정확히 추론하지 못하는 한계를 드러냈다.

(18)

[ST]①德国总理朔尔茨则公开批评美国加征关税破坏全球自由贸易, ②强调欧盟将继续果断应对。

[TT-C] 독일의 솔츠 총리는 미국의 관세 부과가 세계 자유무역을 파괴한다고 공개 비판했고(나열), EU는 단호히 대응하겠다고 밝혔다.

[TT-X] 독일 총리 올라프 솔츠는 공개적으로 미국의 관세 부과가 세계 자유 무역을 파괴한다고 비판하고(나열) 유럽연합은 단호하게 대응할 것이라고 강조했다.

[대안 번역] 독일의 솔츠 총리는 미국의 추가 관세로 인해 세계 자유무역을 파괴된다고 비판하며 EU에서(/차원에서/를 통해) 단호히 대응할 것이라고 강조했다.

예문(18)은 신문 기사에서 ‘-며’의 의미 기능이 정확히 구현되지 못한 예시이다. 무표지 복문 ST의 ①, ②번 절을 보면 동일한 주어가 같은 시간대에 ‘비판’과 ‘강조’ 두 가지 행위를 동시에 하였음을 알 수 있다. 챗GPT와 하이퍼클로바X는 ‘-고’의 나열 의미 기능으로 접속하였고 ②번 절에서 EU를 다시 주어로 처리하였다. 이로 인해 두 절의 논리 관계가 소실되었을 뿐 아니라 ‘비판’과 ‘강조’의 주체가 다른 것처럼 오해할 수 있는 여지를 남겼다. 대안 번역과 같이 ST의 선·후행절을 ‘-며’의 동시 의미 기능으로 접속한 후에 후행절의 EU를 주어로 처리하지 않으면 ST의 동일 주어 및 논리 관계를 모두 정확하게 전달할 수 있다.

반면 연설문에서 병렬(동시)-동시 대응은 빈도는 낮아도 거의 100%의 정확도를 보였다. 연설문 텍스트에서 시간적 동시성을 표현하는 문맥에서는 생성형 AI가 연결어미 ‘-며’를 적절히 대응시켰음을 알 수 있다.

(19)

[ST]①中非领导人像走亲戚一样常来常往, ②深化治国理政经验交流。

[TT-C] 중·아 양국 지도자들은 친지처럼 자주 왕래하며(동시), 국가 운영 경험을 심화 공유하였습니다.

[TT-X] 중·아 지도자들은 마치 친척처럼 자주 왕래하며(동시) 국가 운영 경험을 심화시키고 있습니다.

예문(19)는 연설문의 무표지 복문으로 생성형 AI의 연결어미 선택의 적절성을 보여준다. ①, ②번 절은 동일한 주어로 형태적 표지 없이 나열되어 있다. 이 문장은 두 개의 절로 분리되어 있지만 주어가 동일하고 연결어미 ‘-며’가 동시 의미 기능을 갖는다는 점을 고려해 보면(이금희 2022: 615-620) 생성형 AI 두 모델 모두 적절하게 접속하였음을 알 수 있다.

6. 나오며

한국어와 중국어의 유형론적 차이는 두 언어의 접속 표현 실현 방식에 구조적 차이를 일으켜 생성형 AI의 중한 번역 과정에서 접속 표현의 왜곡으로 이어질 가능성이 있다. 본고는 이러한 문제의식을 바탕으로 생성형 AI 모델이 중국어 무표지 복문을 한국어로 번역할 때 접속 표현이 구현되는 양상을 고찰하고자 한국어 연결어미 ‘-고’, ‘-며’를 중심으로 분석을 실시하였다. 분석 결과, 생성형 AI 중한 번역에서 확인된 핵심 문제는 무표지 복문에 내재된 논리 관계의 추론 부족, 의미 기능 ‘나열’로의 과잉 실현, 텍스트 유형(특히 연설문)의 핵심 의미 기능 상실이라는 세 가지 결론으로 요약할 수 있다.

첫째, 생성형 AI는 무표지 복문에 내재된 논리 관계를 충분히 파악하지 못한 채 표층적으로 안전한 형태의 의미 기능만을 선호하는 편향을 보였다. 둘째, 그 결과 연결어미 ‘-고’, ‘-며’ 모두 ‘나열’의 의미 기능으로 과도하게 치우쳐 사용되었다. 이와 같은 의미 기능의 편향적 실현은 번역문의 접속 표현을 단순화시켜 원문의 결속성을 약화시키는 동시에 독자의 이해와 해석을 방해하는 원인이 될 수 있음을 확인하였다. 셋째, 연설문과 같이 담화적 설득 구조가 중요한 텍스트 유형에서 한국어 연결어미가 단일한 의미 기능으로만 구현되다 보니 텍스트 고유의 결속성이 훼손되고 텍스트 유형의 특성까지 소실되었다.

그러나 본 연구는 분석 대상과 범위 면에서 일정한 한계를 지닌다.

첫째, 중국어 무표지 복문의 다양한 유형을 포괄하지 못하고 텍스트 유형이 연설문과 뉴스에 한정되어 있어 생성형 AI 중한 번역의 경향을 일반화하기에는 한계가 있다. 이에 문학, 학술 등 다양한 유형을 포함해 접속 표현 실현의 변이를 추가로 고찰할 필요가 있다.

둘째, 담화의 일관성 및 텍스트 결속성 유지의 관점에서 출발해 인간 번역사가 생성형 AI 번역문에서 반복적으로 출현하는 오류를 찾아 유형화, 체계화함으로써 AI-TPE(AI Translation Post-Editing) 방향성을 수립할 필요가 있다.

셋째, 향후 모델 개선을 위해 보다 구체적이고 실현 가능한 방향이 제시되어야 한다. 예컨대 다음과 같은 전략이 고려될 수 있다. (1) 텍스트 결속성 기반의 번역 병렬 말뭉치(corpus)를 구축하는 것이다. 중국어 무표지 복문의 논리 관계와 한국어 연결어미의 의미 기능에 대해 전문가가 주석(annotate)한 병렬 데이터 세트를 구축하여 언어 모델이 의미 기능을 변별적으로 학습하도록 유도할 수 있다. (2) 도메인 특화(domain adaptation) 관점에서 언어 모델이 문학, 학술 등 다양한 텍스트 유형에 따라 차별화된 접속 패턴을 생성하도록 프롬프트를 개발할 수 있다. 이러한 방법론적 확장을 통해 담화 차원에서 생성형 AI 번역의 정확성을 높여 궁극적으로 인간 번역사의 후편집 부담 감소, 즉 AI-TPE 효율성 제고에 기여할 수 있을 것으로 본다.

본 연구는 접속 표현이라는 미시적 언어 현상을 통해 AI 번역의 구조적 한계를 고찰하였다. 향후 방대한 말뭉치 기반의 실험과 다양한 텍스트 유형 간의 비교를 통해 생성형 AI 번역 연구의 외연을 확장할 수 있는 후속 연구를 기대해 본다.

참고문헌

- 강현화·장채린·남신혜·홍연정·김강희·이현정. (2017). 『담화 기능에 따른 한국어 유사 문법 항목 연구』 서울: 한글과컴.
- 김아영. (2004). 「한-중 통번역에서 중국어 무표지 연결 절의 관계 의미 연구」. 『통역과 번역』 6: 65-87.
- 김종록. (1991). 「국어 접속어미 지도를 통한 사고력 향상 방안」. 『국어교육연구』 23: 205-230.
- 남기심. (1978). 「국어 연결어미의 활용론적 기능-나열형 ‘-고’ 중심으로-」. 『연세총론』 15: 1-21.
- 박건숙. (2024). 「생성형 AI의 한국어 문맥 해석 연구: 챗GPT의 접속 부사 해석을 중심으로」. 『한국과 국제사회』 8: 763-786.
- 이문화·진홍월. (2005). 「중국어권 한국어 학습자의 연결어미 ‘-고’와 ‘-아서/어서’의 상호 대치 오류 현상 연구」. 『언어학연구』 30: 29-48.
- 유혜령. (2005). 「연결어미의 접속 기능에 대하여」. 『문법교육』 3: 87-105.
- 윤평현. (2005). 『현대국어 접속어미 연구』 서울: 박이정.
- 이금희. (2022). 「외국인 유학생 학술적 글쓰기에서의 연결어미 확장 교육 - ‘나열’, ‘동시’와 ‘원인·이유’의 연결어미 중심으로 -」. 『국제어문』 0(95): 605-633.
- 이완국. (2022). 중한 정부 공공자료 번역에 나타난 논리 연결 표지의 명시화 연구. 한국외국어대학교 대학원 석사학위논문.
- 장현주. (2006). 중·한 번역의 논리 연결 연구. 한국외국어대학교 대학원 중어중문학과. 박사학위논문.
- 전기정. (2018). 「중국어와 한국어 복문의 비교」. 『중국어문학논집』 112: 81-105.
- 정수진. (2013). 「국어 연결어미의 의미 구조에 대한 인지언어학적 탐색」. 『한글』 302: 199-222.
- 정윤철. (2005). 「무표지 복문의 유형 고찰」. 한국중국어교육학회 학술대회 110-118.
- 종가재. (2024). 한국어 비중결어미 ‘-고’의 중국어 대응 양상 연구. 동국대학교 대학원 국어국문학과 석사학위논문.
- 주효천. (2018). 한국어 연결어미 ‘-고, -며, -면서’의 중국어 대응 표현 연구. 동국대학교 대학원 국어국문학과 석사학위논문.
- 채숙희. (2023). 「현대 한국어 연결어미 연구의 성과와 과제」. 『국어학』 0(107): 469-524.
- 한희정. (2017). 한국어교육을 위한 접속 표현 연구: 언어 사용역에 따른 사용 양상을 중심으로. 단국대학교 국어국문학과 박사학위논문.
- 陈昌来. (2000). 『现代汉语句子』 上海: 华东师范大学出版社.
- 范晓. (1998). 『汉语的句子类型』 太原: 书海出版社.
- 邢福义. (2002). 『汉语复句研究』 北京: 商务印书馆.
- 黄伯荣, 廖序东. (2017). 『现代汉语 (增订六版)』 北京: 高等教育出版社.
- 刘月华, 潘文娉, 故韡. (2019). 『实用现代汉语语法』(第三版) 北京: 商务印书馆.
- 赵恩芳, 唐雪凝. (1998). 『现代汉语复句研究』 济南: 山东教育出版社.
- Halliday, M. A. K., and Hasan, R. (1976). Cohesion in english. London: Longman.
- Reiss, K. and Vermeer, H. J. (1984), Groundwork for a General Theory of Translation, Niemeyer: Tubergen.
- Sakib Shahriar et al. (2024). Putting GPT-4o to the Sword: A Comprehensive Evaluation of

Language, Vision, Speech, and Multimodal Proficiency. *Applied Sciences* 14(17), p.7782.

[인터넷 자료]

국립국어원 한국어교수학습센터. <https://kcenter.korean.go.kr/kcenter/index.do>

국립국어원 표준국어대사전. <https://stdict.korean.go.kr/search/searchResult.do>

현대 국어 사용 빈도 조사 2.

https://m.korean.go.kr/front/etcData/etcDataView.do?etc_seq=60&mn_id=46&pageIndex=50&utm_source=chatgpt.com (2025.08.27.검색).

Common Crawl. https://en.wikipedia.org/wiki/Common_Crawl?utm_source=chatgpt.com (2025.09.02. 검색).

clova.ai. https://clova.ai/hyperclova?utm_source=chatgpt.com(2025.09.09.검색).

李雨童. (2025). 社评：面对关税讹诈，绥靖妥协只会是“与虎谋皮”(2025.07.07.검색).

<https://opinion.huanqiu.com/article/4MN7DQ2m1GB>

张樵苏. (2025). 新华时评：关税战，一场饮鸩止渴的游戏(2025.07.07.검색).

<http://www.xinhuanet.com/world/20250403/cd461b9f21cb40cf98d8daa93bbdee05/c.html>

徐祥丽. (2025). 人民网三评“美国滥施关税”之一：世界要公道，不要霸道(2025.07.07.검색).

<http://opinion.people.com.cn/n1/2025/0429/c1003-40470689.html>

刘叶婷. (2025). 人民网三评“美国滥施关税”之二：全球要合作，不要分裂(2025.07.07.검색).

<http://opinion.people.com.cn/n1/2025/0430/c1003-40471166.html>

崔译戈. (2025). 人民网三评“美国滥施关税”之三：经济要繁荣，不要衰退(2025.07.07.검색).

<http://opinion.people.com.cn/n1/2025/0501/c1003-40472206.html>

习近平在中非合作论坛北京峰会开幕式上的主旨讲话(2025.07.07.검색).

https://www.mfa.gov.cn/web/ziliao_674904/zyjh_674906/202409/t20240905_11485600.shtml

王毅在中非合作论坛成果落实协调人部长级会议开幕式上的主旨讲话(2025.07.07.검색).

https://www.mfa.gov.cn/web/ziliao_674904/zyjh_674906/202506/t20250611_11645957.shtml

王毅集体会见非洲国家驻华使节时的讲话(2025.07.07.검색).

https://www.mfa.gov.cn/web/ziliao_674904/zyjh_674906/202505/t20250526_11633606.shtml

This paper was received on 7 November 2025; revised on 5 December 2025; and accepted on 15 December 2025.

Author's email address

choi00joy@hanmail.net

About the author

Seunghyuk Choi is a Lecturer in the Korean-Chinese Interpretation and Translation Program at the Graduate School of International Studies of Chung-Ang University. His main research interests include machine translation and AI-based translation.