

## 연구논문

## 분단 이후 사유를 견디는 최인훈의 문장

## 장문·문답·지시 유보 구조에 대한 데이터 보조 분석

남병수

연세대학교 인문학연구원 비교문학연구소 전문연구원, 비교문학 전공

nomadicwolf@naver.com

## 국문초록

이 논문은 최인훈 소설의 관념성을 주제의 추상성보다 문장의 사유 형식에서 다시 읽는다. 『광장』과 『회색인』을 포함한 소설 열 편의 전체 텍스트를 대상으로 문장 길이와 변동계수, 의문·호호 지시 표현, 대화·비대화 서술, 관념어·감각어·시간·공간 표지, 어휘 다양성과 반복·변주를 산출하고, 그 분포가 지목한 문장을 정밀독해하였다. 분석 결과 「바다의 편지」 등 일부 단편은 높은 평균 문장 길이를, 『회색인』은 가장 큰 길이 변동성을 보였다. 질문 밀도는 대화량과 일치하지 않았으며, 질문과 지시는 직접 발화의 경계를 넘어 서술 내부의 판단 위치를 흔들었다. 관념어 또한 감각어와 시간·공간 표지의 결합 속에서 장면으로 되돌아왔다. 이 결과는 최인훈 소설의 관념성이 완성된 사상의 전달보다 문장이 연장되고 끊기며 자기 근거를 되묻는 과정에서 형성된다는 점을 뒷받침한다. 다만 비교 코퍼스화 판본 계통이 충분히 확보되지 않았으므로 결론은 분석한 열 편 내부의 관계로 한정한다. 본 연구는 관념성을 문장 안에서 생성되는 형식의 문제로 재정의하고, 수치에서 압력 위치를 거쳐 정밀독해로 이동하는 점검 가능한 데이터 보조 독해의 절차를 제안한다.

주제어 최인훈, 분단문학, 한국어 문장, 데이터 보조 독해, 판단 유보

투고일 2026. 7. 3. 심사일 2026. 8. 11. 게재 확정일 2026. 8. 27.

## Research Paper

## Choi In-hun's Sentences Bearing Post-Division Thought: A Data-Assisted Analysis of Long Sentences, Question-Answer Structures, and Deictic Suspension

Nam, Byongsoo

Research Fellow, Center for Comparative Literature, Institute for the Humanities,  
Yonsei University  
nomadicwolf@naver.com

### Abstract

This study reconsiders the conceptuality of Choi In-hun's fiction as a form of thought produced within sentences. Ten complete texts are analyzed through sentence length and variation, interrogative and ambiguous-deictic markers, dialogue and non-dialogue narration, abstract, sensory, temporal, and spatial markers, lexical diversity, and repetition-variation. Quantitative distributions select pressure points for close reading rather than determine interpretation. Several short stories exhibit high mean sentence length, while *The Gray Man* records the greatest variation. Question density does not coincide with the amount of dialogue, and questions and deixis unsettle judgment within narration as well as direct speech. Abstract vocabulary also returns to scenes through sensory, temporal, and spatial markers. These patterns support the hypothesis that conceptuality emerges as sentences extend, break, and interrogate their own grounds. Because no comparison corpus is used and the genealogy of the analyzed editions cannot be fully reconstructed, the claims remain restricted to these ten texts. This study redefines conceptuality as a formal process generated within sentences and proposes an auditable sequence from numerical pattern to textual pressure point and close reading.

**Keywords:** Choi In-hun, division literature, Korean sentence, data-assisted reading, suspended judgment

Received July 3, 2026 Reviewed August 11, 2026 Accepted August 27, 2026

## I. 서론: 분단 이후 사유를 견디는 최인훈의 문장

---

### 1. 문제 제기

최인훈 소설의 독특성을 설명력을 갖춘 언어로 풀이하는 데에 있어서 가장 익숙하게 채택되는 말들 가운데 하나는 사유의 “관념성”이다. 이 단어의 사용은 최인훈 문학이 배태한 지적 밀도와 사상적 긴장을 압축적으로 잘 현시해 주는 이점을 지닌다고 말할 수 있을 것이다. 그러나 비단 그것뿐만은 아니다.

통상적인 의미에서의 관념성은 소설 속 추상적 언어들이 어떤 양상을 따라 작품 속에서 운행하는지를 지칭하는 장치로 봉사하거나, 더 나아가서는 소설의 구성적 요소들의 적극적인 조직화 방식 자체를 견인하고 추동하는 일종의 최전선 매듭 장치로 복무하는 요인으로서 해설된다. 전통적 해석을 따를 때 관념성이라는 이름이 지닌 장악력 아래에서는 소설의 주요 문법 구조에 해당하는 인물의 발화, 문장의 길이, 질문의 반복, 지시어의 지연, 시간과 공간을 표시하는 말들의 배치 등속의 것이 소설의 본격적 구조화 인자로서의 존재 지위를 박탈당하고, 일종의 장식적 요소나 부차적 조건의 위치에 머문다. 때로는 학술적 논의의 영토 바깥으로 밀려나기도 한다.

최인훈의 작품세계가 지니는 관념성의 의미는 이와 다소 다르게 호출된다. 이 차이를 다른 표현을 빌려 번역하자면, 사유와 관념성의 관계는 통상적인 경우들과 같이 그렇게 단순한 구도로 치환되지 않는다. 최인훈 소설에서 관념성이 지닌 존재 지위는 다소 특이적인데, 그것은 보통의 경우들처럼 이미 완성된 선행적 관념성이 문장 바깥에서 문장 안으로 강제 개입하여 후행적 사유를 그것들 사이에 주입하는 방식으로는 현사실화(Faktizierung)되지 않는다. 오히려 역설적으로 사유는 문장의 진행, 중단, 회귀, 문답, 지시의 유보 속에서 잠정적 양상으로 발생한다. 그럼으로써 관념성은 잠재적으로 구축된다. 이때 관념성은 강압의 원천이 아니라 사유의 소산이자 본체로서 봉사한다.<sup>1</sup> 또한 관념성은 단지 사유의 산물

---

※ 이 연구에서 생성형 AI는 원고의 교정·교열, 비학술적인 비평적 문체의 축소 및 완화, 표현 방식의 정형화와 표준화, 문장 수준의 오류 탐지에 제한적으로 활용되었다. 또한 코드

로서 역할을 넘어 소설 속 문장들이 사유와 관계 맺는 방식에도 영향력을 행사한다. 같음하자면, 문장 속에서 창발하는 사유는 관념을 빚고 관념은 사유와 문장 사이의 긴장도를 조율하는 입체적 관계를 갖는다.

이 논문은 바로 이 지점에서 출발한다. 최인훈 소설의 관념성은 다만 주제의 추상성에 국한되는 것이 아니라 한국어 문장이 사유를 견디고 지연시키며 또 다시금 배열하는 방식에서 기초하는 복잡다기한 문제로 읽을 수 있다. 더불어, 특별히 무게 중심을 두고 살펴보아야 할 지점이 하나 더 있다면 그것은 분단 이후 지식인의 발화는 단순한 입장의 제시나 표명으로 온전히 귀결되기 까다롭다는 점이다. 그것은 ‘어느 편에 설 것인가, 무엇을 말할 수 있는가, 누구에게 말하는가, 말의 책임은 어디에 놓이는가’라는 다수의 질문을 소설의 문장 내부로 끌어들이는 것이다. 『광장』의 ‘질문 구조’, 『회색인』의 ‘장문 리듬’, 단편 군에서 나타나는 ‘압축적 발화’는 모두 이 같은 문제에 다른 언어의 규모와 속도를 빌려 응답하고자 안배된 장치라고 말할 수 있다.<sup>2</sup>

기존 연구는 최인훈 소설의 이념성, 소통 구조, 정치적 상상력, 잠재성, 풍속과 방법론 등을 다각도로 논의해 왔다. 일련의 성과는 최인훈 문학이 단일한 이념소설의 이름으로만 환원될 수 없음을 이미 잘 보여 주었고, 이러한 논의는 충분한 타당성을 지닌다. 다만 이 논문은 그 성과의 기틀 위에서 조금 다른 질문을 던지고자 한다. 최인훈 소설에서 사유는 어떤 문장적 조건 아래 구성되는가. 장문은 단순한 난해함의 징후인가, 아니면 판단을 유예하고 논증을 확장하는 형식인가.

---

작성 및 파이프라인 구축은 연구자의 로컬 컴퓨터에 설치된 Python 3.14 및 PyCharm 환경에서 이루어졌으며, AI는 결과의 정확성과 신뢰도를 점검하고 일부 문제 요소를 보강하는 데 국소적으로 활용되었다. 연구 질문과 가설의 설정, 연구 대상 선정, 원자료의 확보, 분석 코드와 변수의 설계, 통계분석과 시각화, 작품 분석 및 해석 등 논문 설계와 구조화, 문제의식의 구체화, 방법론의 설정과 적용을 포함하는 연구의 전 과정은 연구자의 독자적인 판단하에 수행되었음을 밝힌다.

- 1 『광장』의 잠재성과 선택 구조에 대해서는 이나윤, 「최인훈의 『광장』에 나타난 잠재성 연구」, 『우리말글』 96(2023), 287~316쪽을 참조. 본 논문은 잠재성의 문제를 지시 표현과 의문 표현의 문장적 지연으로 다시 읽는다.
- 2 『광장』과 분단·냉전 질서의 관계에 대해서는 선민서, 「냉전의 대리전으로서의 6·25와 〈광장〉: 중립국 위원회의 실질적 부재」, 『국어국문학』 186(2019), 329~366쪽을 참조. 본 논문은 그 역사적 쟁점을 인물의 문답 구조와 판단 유보의 문장 형식으로 확장한다.

질문과 지시어는 인물의 심리 상태를 표현하는 장치에 머무는가, 아니면 분단 이후 주체가 세계와 관계 맺는 방식을 조직하는가. 이는 사유와 형식을 곧장 동치하고 같음하는 일이란 가능하지 않은 것으로, 양자 간의 관계를 면면히 훑아볼 필요가 있다는 판단에서 나온 질문들이다.<sup>3</sup>

## 2. 연구 질문과 논문의 구성

이 논문의 중심 가설은 분석 대상인 최인훈 소설 열 편을 통해 관념성이 이미 완성된 사상의 전달보다는 문장의 연장과 중단, 질문과 지시의 흔들림을 통해 형성된다는 점을 확인하는 것이다. 이를 네 질문으로 나누어 검토한다. 첫째, 장문과 문장 길이의 변동성은 작품 내부에서 사유의 리듬을 어떻게 조직하는가. 둘째, 의문 표현과 모호 지시 표현은 판단의 지연 및 발화 위치와 어떤 관계를 맺는가. 셋째, 관념어·감각어·시간·공간 표지는 추상적 사유를 장면에 어떻게 접속시키는가. 넷째, 장편과 단편 사이에서 관찰되는 차이는 관념소설이라는 기존 평가를 어느 범위까지 세분하는가.<sup>4</sup>

네 질문은 하나의 가설을 서로 다른 층위에서 시험한다. 문장 길이는 발화가 지속되는 시간을, 질문은 판단이 닫히지 않는 시간을, 지시어는 대상이 고정되지 않는 위치를 드러낸다. 관념어가 감각어와 시간·공간 표지를 만날 때 추상적 사유는 다시 장면으로 돌아온다. 다만 이 연구는 비교 코퍼스를 갖추지 않았으므로, 그 분포를 분단 이후 한국어 일반이나 최인훈만의 고유성으로 확장하지 않고 분석한 열 편 내부의 관계로 한정한다.<sup>5</sup>

3 최인훈 소설의 관념성·풍속·방법론 문제는 김선, 「최인훈의 문학과 방법론으로서의 풍속」, 『한국현대문학연구』 59(2019), 121~159쪽을 참조할 수 있다. 본 논문은 관념성의 내용보다 문장 형식과 발화 구조에 초점을 둔다.

4 최인훈 소설의 소통 구조에 대해서는 양윤의, 「최인훈 소설에 나타난 소통 구조 연구」, 『한국문예비평연구』 28(2009), 221~247쪽을 참조. 본 논문은 그 소통 구조를 문장 길이, 질문, 지시 표현의 분포와 결합해 다시 검토한다.

5 최인훈 문학과 한국현대문학의 타자성 문제에 대해서는 정호웅, 「최인훈 문학과 한국현대문학의 타자들」, 『우리말글』 68(2016), 439~463쪽을 참조. 본 논문은 타자성과 현실 접속의 문제를 관념어·감각어·공간 표지의 배치에서 검토한다.

이 논문은 총 6장으로 구성된다. I장은 문제의식과 연구 질문을 제시한다. II장은 연구 대상과 데이터 보조 독해의 방법을 설명한다. III장은 장문과 문장 길이 변동성을 중심으로 최인훈 소설의 사유 리듬을 분석한다. IV장은 질문, 대화, 서술의 관계를 통해 문답 구조를 검토한다. V장은 관념어, 감각어, 시간·공간 표지, 지시 유보를 함께 읽어 발화 위치의 지연을 논의한다. VI장은 이러한 분석을 종합하여 최인훈 문학의 한국학적 의미를 제시한다.<sup>6</sup>

### 3. 선행연구와의 차별성

최인훈 문학 연구는 이미 두텁게 축적된 연구 지층을 지니고 있다. 『광장』을 중심으로 한 분단문학 연구, 지식인 주체와 이념 선택의 문제, 최인훈 소설의 소통 구조, 풍속과 방법론, 초기 비평의 참여 의식, 『총독의 소리』 연작의 권력 비판, 『화두』와 『광장』을 연결하는 정치적·사상적 독해 등이 그 주요한 흐름을 이룬다. 이러한 연구들은 최인훈 문학이 단순한 이념의 도식이나 관념의 나열이 아니라, 전후 한국 사회의 역사적 조건, 지식인의 자기 인식, 근대적 주체의 불안, 언어와 권력의 긴장을 복합적으로 담고 있음을 밝혀 온 바 있다.<sup>7</sup>

그러함에도 불구하고, 기존 연구의 상당수는 최인훈 소설의 문장 형식을 본격적인 논증의 중심에 놓지 않았다. 오히려 주제와 사상 그리고 서사 구조와 인물의 위치를 설명하는 과정에서 문체의 위상을 단지 부수적 수준에서 언급하는 데 머무르는 경우가 많았다. 물론 그렇다고 해서 기존의 연구 지형에서 문체의 문체가 전혀 논의되지 않았다는 뜻은 아니다. 최인훈 소설의 관념적 문체, 장광설, 대화와 독백, 지식인의 어조는 반복적으로 지적돼 온 바 있으나 그 지적이 대체로

6 최인훈의 초기 비평과 참여 의식에 대해서는 정영훈, 「최인훈의 초기 비평 연구: 참여의 의미를 중심으로」, 『한국현대문학연구』 50(2016), 477~513쪽을 참조. 본 논문은 참여의 명제적 차원보다 참여가 문장 내부에서 질문과 유보의 리듬으로 조직되는 양상에 주목한다.

7 『광장』의 미적 근대성과 사유 형식에 대해서는 김정관, 「최인훈 소설 『광장』의 미적 모더니티 연구」, 『비평문학』 89(2023), 113~155쪽을 참조. 본 논문의 문장 길이·의문 밀도 분석은 『광장』의 근대적 주체가 사유를 결론으로 닫지 못하고 질문의 형식으로 연장한다는 점을 수치적으로 보조한다.

인상적 차원에 머무르거나, 특정한 단일 작품에 국한된 협소한 독해 구조 안에서만 작동하는 경향을 취해 왔다는 사실이 문제다. 소설가의 여러 작품을 가로질러 문장의 길이, 변동성, 질문, 지시어, 관념어와 감각어의 배치를 함께 비교 탐사하고 유기적인 관점에서 그 상관적 관계성을 탐문하는 작업은 아직 충분하게 이루어지지 않았다는 뜻이다.<sup>8</sup>

본 연구는 바로 이 공백을 겨냥한다. 최인훈 소설의 관념성은 이미 잘 알려진 비평적 명명이지만 그 관념성이 실제 문장 내부에서 어떻게 발생하고 또 어떤 영향력을 발휘하는지 호명해 봄으로써 더 세밀하게 세부를 따져 보고자 한다. 최인훈의 문장은 자주 길어지고, 예고 없이 갑자기 짧아지며, 질문을 반복하고, 또 지시를 유보하며, 관념어와 감각어를 교차시킨다. 일련의 현상들은 단순한 문체적 특징만으로 간주할 수 없으며, 오히려 그것들이 사유가 발생하는 구체적인 형식을 가늠하도록 해 준다는 표현이 가일층 옳다고 말할 수 있다. 따라서, 본 논문은 일차적으로 최인훈의 작품세계가 관념소설의 지반 위에 서 있다는 평가를 부정하지 않고 수용하는 대신, 그 평가가 실제 문장 단위 차원에서 어떤 구체적인 근거를 갖고 현실화하는지를, 또 반대로 어떤 경우에 평가 자체가 수정되어야만 하는지를 이중적으로 검토한다.

이 논문의 차별성은 세 가지다. 첫째, 최인훈 문학의 역사적·사상적 의미가 실제로 작동하는 문장 형식을 분석의 중심에 놓는다. 둘째, 장편과 단편을 함께 살펴 피로 장르·분량·발표 시기의 효과가 서로 얹혀 있음을 전제하고, 이를 작가 고유의 증거로 단정하지 않는다. 셋째, 데이터 보조 분석을 수치→압력 위치→정밀독해의 순서로 운용한다. 수치는 결론이 아니라 익숙한 작품 배열을 흔들고 다시 읽어야 할 문장을 지목하는 근거다.<sup>9</sup>

8 「충독의 소리」 연작의 권력·간접권력 비판에 대해서는 윤인로, 「노모스왕의 권위와 간접권력 비판의 사상연쇄: 최인훈 연작 「충독의 소리」를 매개로 삼아」, 『인문논총』 80-3(2023), 471~500쪽을 참조. 본 논문에서 공간 표지의 분포를 따지는 대목은 이러한 문체 의식과 연결된다.

9 최인훈 장편소설의 의식 심화 과정에 대해서는 이원양, 「최인훈 장편 소설 사이의 인지 층위 심화 과정: 광장—회색인—서유기—소설가 구보씨의 일일, 태풍의 연속성을 중심으로」, 『비평문학』 87(2023), 291~328쪽을 참조. 본 논문이 10편을 단일 표본이 아니라 연속적 사유의 코퍼스로 다루는 이유도 이러한 장편 간 연속성 논의와 접속한다.

데이터 보조 독해의 효력은 작품의 질적 복잡성을 숫자로 환원하는 데 있지 않다. 평균 문장 길이와 질문 밀도는 사유의 깊이나 철학적 강도를 직접 측정하지 못하지만, 그러함에도 여러 작품에 흩어진 문장 리듬과 발화 표지를 같은 기준으로 비교하여 예상과 어긋나는 지점을 드러낸다. 이후의 정밀독해는 바로 그 어긋남이 실제 문장에서 어떤 기능을 얻는지 확인하고, 다시 지표의 해석 범위를 좁힌다.<sup>10</sup>

따라서 이 글은 계산과 독해를 병렬로 늘어놓지 않는다. 작품별 분포가 압력 위치를 지목하면 정밀독해가 그 문장의 의미와 장면을 검토하고, 그 결과가 다시 수치의 해석을 제한한다. 이 왕복 과정이 데이터 보조 독해의 핵심이다.

이러한 접근은 최인훈 문학을 사상사의 예시로 환원하지 않으면서 역사적 경험이 한국어 문장 안에서 어떻게 길어지고, 끊기고, 되물어지는지를 밝힌다. 이때 한국학적 대상은 ‘분단 이후 한국어 전체’가 아니라 분단의 조건을 문장 내부에서 감당한 최인훈 소설의 구체적 사례다.

본 연구가 최인훈 소설의 문장 형식을 한국학의 대상으로 삼는 이유도 여기에 있다. 역사적 경험은 작품의 소재에만 머물지 않고 문장의 길이, 질문의 반복, 지시의 유보, 관념어와 감각어의 긴장 속에 새겨진다. 이 글은 그 언어적 새김의 한국면을 추적한다.

## II. 연구 대상과 방법: 최인훈 소설 10편 코퍼스와 데이터 보조 독해

---

### 1. 분석 대상과 자료의 범위

분석 대상은 『광장』과 『회색인』, 그리고 「7월의 아이들」, 「9월의 다알리아」, 「그

---

10 인문학 연구에서 디지털 기술 활용의 현황과 적용 방향에 대해서는 박경우, 「인문학 연구에서의 디지털 기술 활용 현황 및 적용 방향: 디지털 인문학 동향 분석을 중심으로」, 『국어국문학』 200(2022), 137~179쪽을 참조. 본 논문의 도표는 비평 판단을 대체하는 장치가 아니라 해석의 근거를 반복 확인하게 하는 보조 절차이다.

레이 구락부 전말기」·「금오신화」·「라울전」·「바다의 편지」·「우상의 집」·「웃음소리」의 열 편이다. 두 장편과 여덟 단편을 함께 놓은 것은 작품 규모가 달라질 때 지표의 분포가 실제로 어떻게 달라지는지 탐색하기 위해서다. 다만 이 배열은 온전히 균형 잡힌 비교 설계가 아니다. 장르, 분량, 창작 시기와 개작의 영향이 서로 분리되지 않으므로 장·단편 차이를 단일 원인으로 설명하지는 않는다.<sup>11</sup>

열 편은 분석본의 완전한 확보, 본문 추출의 성공, 작품 전체 텍스트 분석 가능성, 장·단편과 발표 시기의 분산이라는 사전 기준을 함께 충족한 작품으로 한정하였다. 따라서 이 코퍼스는 최인훈 소설 전체를 대표하는 확률표본이 아니라, 작품 규모와 문장 형식의 관계를 탐색하기 위한 목적표집 코퍼스다.

전체 규모는 『회색인』 266,271자·8,813문장, 『광장』 137,045자·4,539문장이 지시하는 방향으로 크게 기울어져 있다. 작품별 값을 먼저 집계하고 밀도·비율·변동계수를 사용한 까닭은 이 불균형이 결과를 곧장 지배하지 않게 하기 위해서다. 「바다의 편지」를 비롯한 일부 단편의 높은 평균 문장 길이는 장르 관습, 정보 압축, 창작 시기가 함께 만든 결과일 수 있고, 본문은 이를 최인훈의 고유성이나 그 단편성의 인과적 효과로 단정하지 않는다.<sup>12</sup>

코퍼스는 열 편 모두의 전체 텍스트를 분석하되, 결과는 작품별로 집계하여 같은 축에 놓았다. 로컬 분석본은 파일명과 SHA-256으로 식별했지만 그 파일이 유래한 최초 판본, 개작 및 교감 이력은 완전하게 복원되지 않았다. 따라서 본 연구는 판본 간 차이나 통시적 문체 변화를 검증하지 않으며, 분석본 내부의 상대적 분포만을 보고한다. 이 제한은 이 글 말미의 부록 <표1>과 <표3>에 함께 명시한다.<sup>13</sup>

- 
- 11 『회색인』과 『서유기』를 중심으로 한 반근대성 논의는 이소연, 「최인훈 소설에 나타난 대항근대성 연구: 『회색인』과 『서유기』에 드러난 상호텍스트적 '부재 영역'을 중심으로」, 『비평문학』 67(2018), 223~248쪽을 참조. 본 논문의 중기 작품 분석은 지식인 주체의 자기분열이 장문, 관념어, 장소 표지의 결합 속에서 강화되는지를 묻는다.
  - 12 『화두』와 『광장』의 적론 구조에 대해서는 윤인로, 「혁명검찰론=이단심문론에 대해: 소설 『화두』와 『광장』의 적론(敵論)을 위하여」, 『탈경계인문학Trans-Humanities』 16-2(2023), 137~160쪽을 참조. 본 논문은 대립하는 논변이 결론보다 문장적 긴장과 반박의 형식으로 지속되는 지점을 계량 지표와 함께 읽는다.
  - 13 『화두』의 읽기 현상학과 독자의 위치에 대해서는 김근호, 「독자로서의 작가: 최인훈 〈화두〉의 문학독서 현상학」, 『현대소설연구』 71(2018), 5~31쪽을 참조. 본 논문에서 데이터 보조 독해는 해석을 대체하기보다 독자의 확인 가능한 위치를 재조정하는 절차로 설정된다.

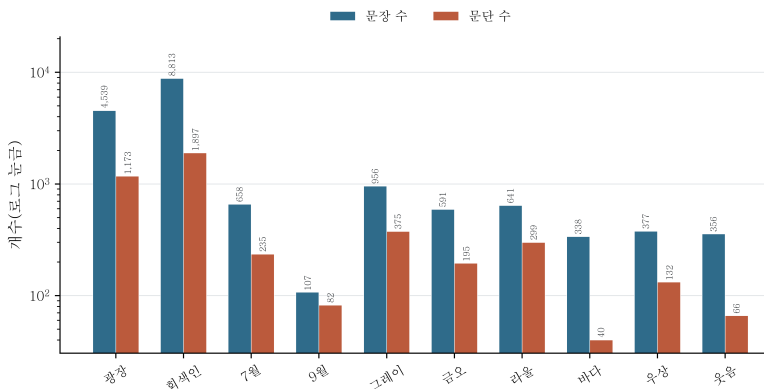


그림1-최인훈 소설 10편 코퍼스의 문장·문단 규모

주: 장편과 단편의 문장 수 및 문단 수를 비교하여 분석 대상의 규모 차이를 표시한다.

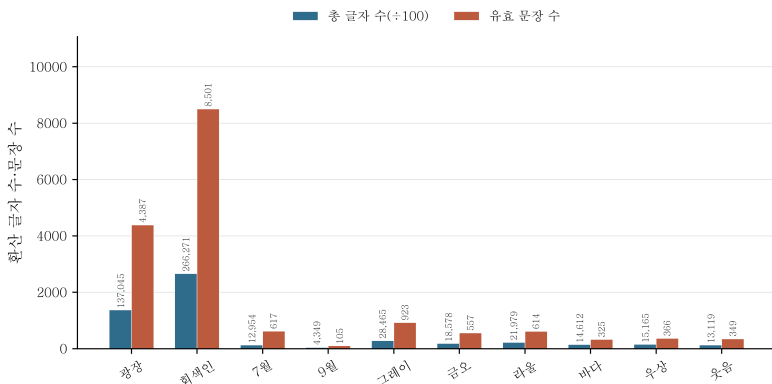


그림2-원자료 규모와 유효 문장 규모

주: 총 글자 수와 분석에 투입된 유효 문장 수를 함께 제시한다. 총 글자 수는 동일축 비교를 위해 100분의 1 배율로 표시하였고, 막대 위에는 원값을 병기하였다.

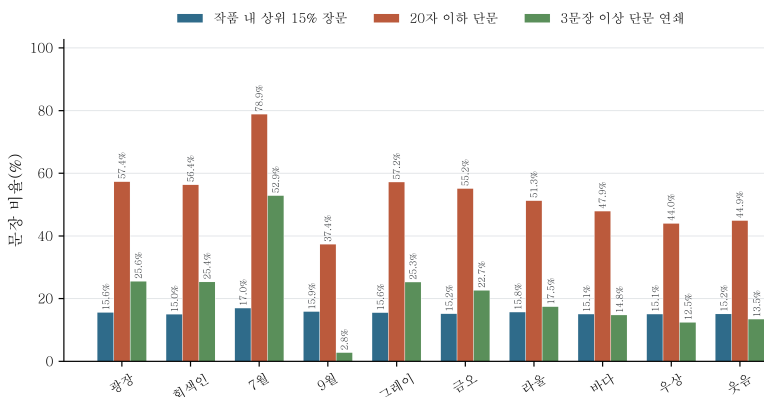


그림3-전처리 이후 문장 길이 구간의 분포

주: 장문 비율, 단문 비율, 단문 연쇄 비율을 10편 전체에 대해 제시하여 문장 단위 분절이 해석 지표로 전환 되는 중간 상태를 요약한다.

〈그림1~3〉은 문장·문단 수, 원자료와 유효 문장 규모, 장·단문 구간을 차례로 제시한다. 절대량과 작품별 비율을 분리함으로써 장편의 크기와 단편의 밀도를 같은 값으로 오인하지 않도록 한 데이터셋 요약이다.

두 장편이 절대량을 주도하지만 작품별 평균·밀도·변동성은 각 작품을 하나의 분석 단위로 삼아 계산했다. 이 설계는 단편의 소수 문장을 장편의 다수 문장과 단순 합산하지 않으면서, 규모와 문체 지표의 효과를 구별하기 위한 최소한의 장치다.<sup>14</sup>

상기 그림들은 이후 다룰 지표들이 어떤 자료 규모와 문장 분절을 토대로 산출됐는지 지시한다. 일련의 전제 확인에 기초해서, 연구는 절대 빈도가 아니라 작품별 밀도와 변동성 그리고 그 값이 지목하는 실제 문장을 분석한다.

## 2. 지표와 해석 원칙

지표는 네 층위로 나뉜다. 문장 리듬은 평균 길이와 변동계수, 발화 구조는 의문 표현과 대화·비대화 서술 비율, 의미 배치는 관념어·감각어·시간·공간 표지의 밀도, 탐색적 결합 층위는 모호 지시 밀도·인지적 모호성 대리지표·중간 모호성 근접도·문장별 어휘 다양도·반복 변주 비율로 구성한다.

이 지표들은 작품의 의미나 독자의 심리 상태를 직접 측정하지 않는다. 장문은 사유의 깊이와, 의문 표현은 철학적 질문과, 모호 지시 표현은 실제 지시 대상의 불명료성과 각각 동일하지 않다. 지표는 정밀독해가 우선 확인할 분포와 문장을 선택하는 대리 측정값으로만 제한적으로 사용된다.

데이터의 인식적 이점은 연구자의 직관을 반복 확인하는 데 있지 않다. 작품 전반의 분포를 먼저 확인한 뒤 예상 밖의 극단 값 및 결합의 양상들을 찾아내고, 그 위치를 작품 문장으로 되돌리는 데 있다. 본 연구의 해석 과정은 이 순서를 벗어나지 않는다.

---

14 텍스트 마이닝을 활용한 텍스트 유형 연구의 사례로는 방경원, 「텍스트 마이닝 기법을 활용한 텍스트 종류 연구: 한국과 독일증시 마감보고서를 비교하면서」, 『인문학연구』 59(2020), 343~366쪽을 참조. 본 논문의 문장 길이, 어휘 밀도, 의문·지시 표현 분석도 텍스트 유형과 문체 차이를 수치로 가시화하려는 방법론적 범주에 놓인다.

### 3. 재현성, 자료 윤리, 해석의 제한

분석은 Python 3.14 및 PyCharm 개발 환경에서 수행했다. KSS에 의존하는 대신 문장 분절은 마침표·물음표·느낌표 뒤의 공백 또는 줄 바꿈을 기준으로 한 규칙 기반 대체 절차를 사용하고 2자 미만 조각을 제외했다. 형태소는 Kiwi 0.23.1에서 명사·동사·수식언·어근 계열을 추출했으며, 문장 길이는 공백을 뺀 글자 수로 계산했다. 이 규칙에 따라 글자 수를 택한 것은 띄어쓰기 편차가 어절 수를 흔들 수 있는 분석본의 조건에서도 동일한 척도를 유지하기 위해서이나, 이것이 곧장 통사 단위의 복잡성을 대신하지는 않는다.

대화 문장은 인용부호나 대시로 시작하거나 한 문장 안에 인용부호 쌍이 있는 경우로 판정했고, 나머지는 비대화 서술로 두었다. 간접화법과 자유간접화법은 서술 속에 남아있을 수 있다. 의문 표현 밀도는 ‘왜·무엇·누가·어디·언제·어떻게·어째서·과연·인가·일까·까닭’의 출현 수에 물음표가 있을 때 1을 더한 뒤 형태소의 수로 나누었다. 모호함의 지시 밀도는 ‘이것·그것·저것·여기·거기·저기·이런·그런·저런·누군가·어떤·어딘가·무언가·그 사람·그곳’의 출현 수를 같은 분모로 나눈 값이다. 다의어와 문맥 오판의 가능성이 남으므로 두 값은 질문 행위나 지시 실패의 확정치가 아니다.<sup>15</sup>

관념어·감각어·시간·공간 표지는 부록 <표2>의 어휘 사전과 접두 일치 규칙을 통해 계산했다. 어휘 다양도는 문장별 고유 형태소 수를 형태소 수로 나눈 TTR을 작품 단위에서 평균한 값이므로 문장 길이의 영향을 완전히 제거하지는 못했으나, 충분한 분석상의 유의성을 지닌다. 인접 문장의 반복·변주는 문자 2~4그램 TF-IDF 코사인 유사도 .80 이상으로 판정했다. 이 규칙들은 의미 구조 전체가 아니라 표면적 언어 배치만을 포착한다.

문장 i의 인지 모호성 대리지표  $A_i$ 는 다음 가중합을 0~1로 절단했다.  $A_i = \text{clip}[.18R_i + .14\min(1, 4B_i) + .12\min(1, 4Q_i) + .12\min(1, 5D_i) + .12J_i + .10\min(1, 4P_i) + .08V_i + .0$

15 『광장』의 법의식·시민윤리·선택 구조에 대해서는 홍순애, 「최인훈 『광장』의 법의식과 시민적 윤리」, 『현대소설연구』 67(2017), 489~523쪽을 참조. 본 논문에서 말하는 판단 유보는 윤리적 선택의 부재가 아니라 선택을 둘러싼 언어적 지연의 형식으로 이해된다.

$7\min(1, 5T_i) + .07\min(1, 5S_i)]$ . R은 최근 10문장의 길이 변동계수,  $B \cdot Q \cdot D \cdot P \cdot T \cdot S$ 는 관념어·질문·모호 지시·심리동사·시간·공간 밀도, J는 앞 문장과의 의미 불연속, V는 시점 전환 판정이다. 작품값은  $A_i$ 의 평균이며, 중간 모호성 근접도는  $\max(0, 1 - 2|A_i - .5|)$ 로 계산한다. 후자는 경험적으로 확인된 ‘최적 모호성’이 아니라 중간 값 .5와의 거리를 나타내는 탐색적 수치다. 코드·사전·파생 CSV·작품별 제외 감사표는 저작권 원문을 빼고 별도 보충 자료를 통해 구성했다.

가중치의 합은 1이며, 문장 리듬과 관념어에는 상대적으로 큰 값을, 질문·모호 지시·의미 불연속에는 같은 중간 값을, 심리동사·시점 전환·시간·공간 표지에는 그보다 작은 값을 부여하였다. 이는 인지적 모호성의 경험적 최적 계수를 추정한 결과가 아니라 문장 내부의 형식적 압력을 비교하기 위한 탐색적 우선순위다. 그러므로 보충자료에서는 균등가중 및 개별 가중치 변동 조건에서도 작품 간 방향이 유지되는지를 민감도 자료로 함께 점검하였다.

연구 과정에서는 OpenAI Codex를 분석 코드의 점검, 오류 탐지, 서지정보 교차검토 및 원고의 형식 교정에 보조적으로 활용하였다. 분석 범주의 설정, 원자료 확인, 코드 실행과 산출값 검증, 정밀독해 및 최종 서술은 연구자가 수행했으며, 결과의 정확성과 연구윤리에 관한 최종 책임은 연구자에게 있다.

### III. 장문과 변동성: 사유가 문장 안에서 지연되는 방식

#### 1. 평균 문장 길이와 압축된 사유

최인훈 소설의 장문은 흔히 관념성의 문체적 증거로 읽힌다. 그러나 장문이 곧 바로 관념의 과잉을 뜻하는 것이라고 동일시할 수는 없다. 장문은 판단이 단번에 종결되지 않는 조건, 여러 가능성이 한 문장 안에 함께 놓이는 조건, 발화 주체가 자기 생각을 말하면서 동시에 그 생각을 의심하는 조건을 만든다. 이 점에서 장문은 최인훈 소설에서 사유의 지연과 논증의 확장을 동시에 수행한다.

평균 문장 길이는 「바다의 편지」 32.16자, 「우상의 집」 29.93자, 「9월의 다알리아」 29.74자 순이다. 『회색인』은 22.67자로 이들보다 짧지만 변동계수는 0.981로

가장 높다. 이 결과는 장편과 단편의 차이를 곧바로 설명하지 않는다. 다만 분석 코퍼스 내부에서 작품 규모와 평균 문장 길이가 같은 방향으로 움직이지 않으며, 장문과 변동성이 서로 다른 압력 위치를 가리킨다는 사실을 보여 준다.

다음 <그림4>는 평균 문장 길이와 문장 길이의 변동계수를 함께 제시한다. 변동계수는 여러 작품 대상을 같은 축 위에서 보기 위해 30배 배율로 표시했으며, 막대 위에는 원래의 값을 제시했다. 물론 이 배율 표기는 시각적 비교를 위한 장치의 기능을 수행할 뿐, 근본적으로 원자료 값을 바꾸는 것은 아니다.

<그림4>에서 포착해야 할 지점은 단순한 순위 매김이 아니다. 「바다의 편지」, 「우상의 집」, 「9월의 다알리아」는 평균 문장 길이가 높고, 『회색인』은 문장 길이의 변동성이 높다. 이러한 차이는 최인훈 소설의 사유가 하나의 문체적 습관으로 환원되지 않음을 여실히 보여 준다. 어떤 작품에서는 장문 자체가 압축된 사유의 호흡을 만들고, 어떤 작품에서는 장문과 단문의 교차가 사유의 불안정성을 만들어 낸다.

상술한 결과는 기존의 관념소설에 대한 평가를 보정한다. 관념성은 단순히 긴 문장이 많다는 사실만으로 충분히 소명되지 않는다. 관념성은 문장이 길어지는 방식, 짧아지는 방식, 길고 짧은 문장이 서로 교차하는 방식 속에서 비로소 발생한다. 그러므로 최인훈 소설의 문체에 대한 평가가 단지 “어렵다” 또는 “관념적이

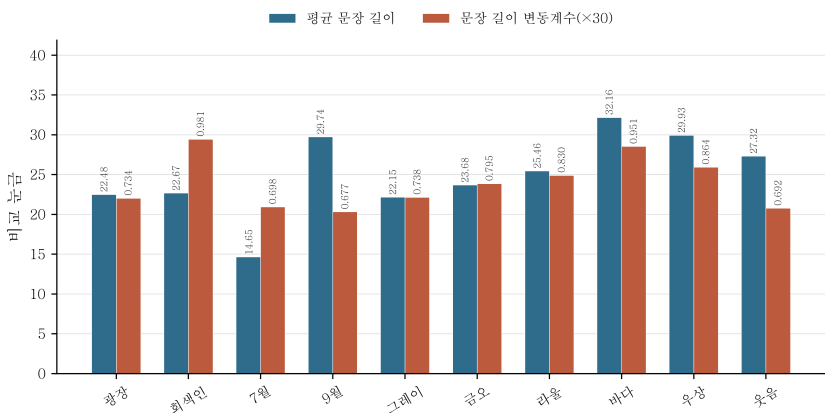


그림4-장문 리듬과 문장 길이 변동성

주: 평균 문장 길이와 문장 길이 변동계수를 함께 제시한다. CV는 시각 비교를 위해 30배 배율로 표시했으며 막대 위에는 원값을 병기하였다.

다"라는 다소 인상 비평적 어휘에 머물러서는 안 된다. 자료에 근거한 더 세밀한 설명이 요청돼야만 한다.

## 2. 변동성의 의미

문장 길이의 변동성은 독자의 읽기 호흡 및 판단 속도에 개입하여 적지 않은 영향력을 행사한다. 문장이 계속해서 길기만 하다면 독자는 하나의 지속적인 사유 흐름을 따라가게 된다. 반면 문장이 계속 짧기만 하다면 판단은 단속적으로 절단된다. 그러나, 긴 문장과 짧은 문장이 불규칙하게 교차할 때, 독자는 독특한 경험의 장 속으로 초청된다. 사유가 확장되다가 끊기고, 다시 이어지다가 다른 방향으로 틀어지는 특이한 경험을 하게 된다. 『회색인』의 높은 변동성은 바로 이 사유의 진폭을 잘 보여 준다.

『광장』의 경우 평균 문장 길이는 22.48자, 문장 길이 변동계수는 0.734이다. 『회색인』에 비해 변동성 수치가 낮게 평가되지만, 이것이 곧 단순성을 뜻하지는 않는다. 소설 『광장』의 긴장은 남과 북, 광장과 밀실, 이념과 개인, 선택과 무력 사이의 구조적 대립 속에서 상대적으로 강한 축을 지닌다. 이때 문장 리듬의 안정성은 오히려 선택의 구조가 작품 전체를 지배하고 있음을 현시하는 계기로 작용할 수 있다. 그와는 정반대로 『회색인』은 축이 더 흩어지고, 사유가 여러 방향으로 미끄러지며, 문장 길이의 변동 또한 그 흔들림을 따라간다.

『회색인』의 변동성이 가장 커지는 구간에서는 ‘일본은 너무 강해졌다’에서 ‘우리가 그들에게 협력함으로써 우리들의 몫을 늘리자’로 이어지는 긴 논리의 연쇄와, 곧이어 진행되는 ‘역사란 무엇일까요, 아는 사람?’처럼 짧은 질문이 교차하는 양상을 보인다(분석본 문장 2103·440). 긴 문장은 식민지 협력의 논리를 누적하고 짧은 질문은 그 논리의 근거를 다시 원점으로 돌린다. 변동계수 0.981은 바로 이 전환을 증명하는 의미값이 아니라, 정밀독해가 그러한 리듬의 파열을 우선 확인하도록 지목한 수치다.

「바다의 편지」 분석본의 94번째 문장은 162자에 걸쳐 우주 탐험의 시대와 휴전선으로 잘린 연해를 한 호흡에 겹친다. ‘왜 우리는 이 조그만 우리나라의 연해를 … 검디검은 그믐밤을 골라 기어다녀야 하는지’라는 물음에서 장문은 정보를

많이 담는 데 그치지 않고 두 시간 규모의 불균형을 문장 안에 지속시킨다. 그러나 이 사례는 단편이 장문을 직접 산출한다는 증거가 아니다. 여기서 말할 수 있는 구간은 짧은 분량, 정보 압축, 창작 시기와 분석 판본이 함께 개입할 수 있다는 지점, 그리고 높은 평균값이 가리킨 문장에서 역사적 판단과 감각적 장면이 실제로 결합한다는 지점까지다.

## IV. 문답 구조와 발화의 조건

### 1. 질문의 밀도와 판단의 지연

질문은 최인훈 소설에서 사건에 대한 정보 부족을 표시하는 기능에만 머물지 않는다. 질문은 발화 주체가 세계를 곧장 확정적으로 판단할 수는 없다는 사실, 그러나 그와 동시에 판단을 회피하는 방향으로 운신할 수는 없다는 사실을 함께 드러낸다. 그러므로 질문을 사유의 결핍이 아니라 오히려 사유가 발생하는 형식으로 보는 편이 타당하다. 특히 분단 이후의 주체에게 소구해 오는 질문들은 현실을 향한 무력한 물음이면서도 자기 위치를 확인하려는 윤리적 압력의 성격을 갖는다.

의문 표현의 밀도는 「7월의 아이들」이 0.0380으로 가장 높고, 『광장』이 0.0346, 「라울전」이 0.0341로 그 뒤를 잇는다. 『광장』에서 질문은 이념적 선택의 조건과 관련되고, 「라울전」에서는 짧은 형식 안에서 세계와 인물의 거리가 압축적으로 드러남을 제시한다. 「7월의 아이들」의 높은 질문 밀도는 단편 속에서 질문이 사건의 보조 장치가 아니라 발화의 기본 리듬으로 기능한다는 점을 선명히 보여 준다.

다음 <그림5>는 의문 표현의 밀도와 지시 유보의 밀도를 함께 표현해 준다. 두 지표는 모두 판단의 즉각적 종결을 어렵게 이끌어 간다는 점에서는 분명히 연결되지만, 그 구체적인 작동의 방식은 명석판명 다르다. 질문이 말의 방향을 열어두는 방식으로 기능한다면, 지시의 유보는 말이 가리키는 대상을 잠시 흔드는 방식을 택함으로써 지연성을 확보하는 양식을 택한다.

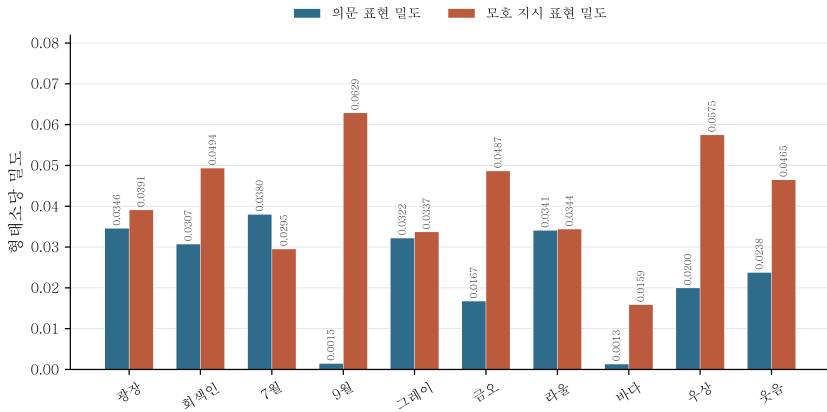


그림5-문답 구조와 지시 유보

주: 의문 표현 밀도와 지시 유보 밀도를 함께 제시하여 질문과 지시의 지언이 발화 위치를 어떻게 혼드는지 비교한다.

〈그림5〉의 차이는 실제 문장에서 서로 다른 방식으로 드러난다. 『광장』 분석본의 793번째 문장 ‘이게 한국 정치의 광장이 아닙니까?’의 경우 질문 표지와 지시 표현이 한 문장 안에서 만난다. 질문은 답을 청하는 동시에 ‘이게’가 가리키는 정치의 장을 다시 규정하도록 압박한다. 반면 「7월의 아이들」의 330번째 문장 ‘오늘 거기서 일곱 시, 아시겠어요?’는 시간과 공간의 지시를 확인 질문에 묶는다. 동일한 지표 조합도 전자에서는 정치적 개념의 재규정으로, 후자에서는 장면의 불안정한 약속으로 기능한다.

이 두 사례는 표지어가 자동으로 같은 의미를 산출하지 않는다는 점도 보여 준다. 동일한 질문·지시 조합이 정치적 개념을 재규정하기도 하고, 약속의 시공간을 확인하기도 한다. 요컨대 계량값의 기능은 문장을 하나의 범주로 봉합하는 것이 아니라 서로 다른 문맥에서 같은 표지가 어떻게 다른 해석 압력을 만드는지 비교하도록 하는 데 있다.

따라서 최인훈 소설의 문답 구조는 단순히 대화 장면의 출현 빈도 문제로 수렴될 수 없다. 질문은 서술 속에서도 작동하고, 지시의 유보는 대화가 아닌 문장 가운데에서도 발생한다. 이 지점을 명확히 확인하기 위해서는 대화 비율과 서술 비율을 함께 살펴보아야 할 필요가 있다.

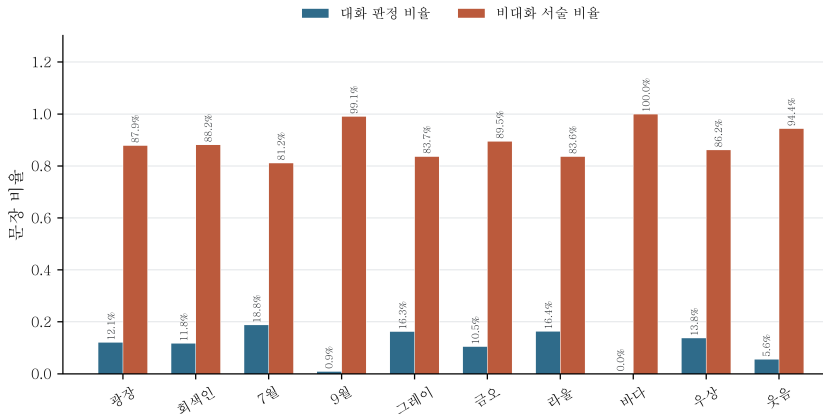


그림6-대화 비율과 서술 비율

주: 대화의 밀도와 서술의 우세가 문장 구조의 실제 발화 조건을 어떻게 달리 만드는지 비교한다.

## 2. 대화와 서술의 비율

분석한 열 편에서 문답성은 대화량과 일치하지 않는다. 질문과 판단 유보는 직접 대화뿐 아니라 서술 내부에도 남는다. 따라서 대화 판정 비율과 질문 밀도를 분리해 보고, 두 값의 결합이 실제 문장에서 어떤 기능을 얻는지를 따로 읽어야 한다.

『7월의 아이들』은 대화 비율이 0.188로 가장 높은 기록 수준을 보인다. 『라울 전』과 『그레이 구락부 전말기』 또한 0.16대의 상당히 높은 대화 비율을 보인다. 그러나 이들 작품의 전체적인 측면에서 서술이 점유하는 비율이 훨씬 높다는 지점을 고려하지 않을 수는 없다. 『광장』은 대화 비율 0.121, 서술 비율 0.879이고, 『회색인』은 대화 비율 0.118, 서술 비율 0.882이다. 일련의 사실들이 제안하는 것은 최인훈 소설이 지닌 질문의 속성이 실제 대화의 양보다는 서술의 구조와 더 깊이 연루되어 있다는 말이다.

〈그림6〉에서 대화 판정 비율이 가장 높은 작품이라 하더라도 비대화 서술의 비중이 더 크다는 점을 알 수 있다. 이 사실만으로 서술 자체가 모두 질문의 장이라고 단정할 수는 없지만, 질문 밀도를 대화량의 부산물로 설명하기 어렵다는 점은 분명하다. 데이터가 새로 드러내는 것은 대화의 많고 적음이 아니라 질문이

직접 발화의 경계를 넘어 분포한다는 사실이다.

기존의 소통 구조 연구가 인물 사이의 관계와 발화 실패를 중심으로 문답성을 해명했다면, 본 분석은 질문 표지가 비대화 서술의 우세 속에서도 지속된다는 점을 작품 간 분포로 확인한다. 따라서 최인훈 소설의 문답성은 말을 주고받는 장면에만 놓이지 않는다. 서술하는 문장 자체가 자기 판단의 근거를 되묻는 형식까지도 포함한다는 데 이 연구가 지닌 문학 분석상의 보정점 및 특이점이 놓인다.

## V. 관념어, 감각어, 지시 유보: 추상 사유의 현실 접속

### 1. 관념성과 감각성의 동시 작동

최인훈 소설을 관념소설로 정의 내리려는 시도는 단지 절반의 타당성만을 확보하는 데 그칠 뿐이다. 최인훈의 소설에는 분명 이념, 역사, 자유, 선택, 책임, 혁명과 같은 추상적 어휘들이 강하게 작동하지만 바로 그 추상어들이 외파로 진공상태 속에 갇힌 고립된 개념어로만 존재하지 않는 까닭이다. 그것들은 감각어, 공간 표지, 시간 표지와 유기적으로 결합하면서 구체적 장면 안에 놓인다. 이때 최인훈 소설의 관념성은 현실에서 벗어난 추상성의 이정표를 따라가는 게 아니라, 오히려 현실을 사유가 가능한 대상으로 다시 재구성하려는 언어적 시도에 더 가깝다.

관념어의 밀도는 「바다의 편지」(0.0170), 『회색인』(0.0159), 『광장』(0.0090)에서 두드러진다. 감각어 밀도와 함께 놓고 보면 이 양상은 더 복합적인 양태로 소구해온다. 관념어가 높게 나타나는 작품이 반드시 감각어를 배제하는 것은 아니며, 오히려 관념어와 감각어가 함께 배치될 때 추상적 사유는 현실의 장면을 지워 버리지 않고 더 나아가서 그 장면을 해석이 가능한 문제로 바꾸는 역할을 이행한다.

상기 <그림7>에서 「바다의 편지」와 『회색인』은 관념어 밀도가 뚜렷이 높게 나타난다. 그러나 이 결과를 두 작품이 현실 감각에서 멀어진다는 뜻으로 곧장 갈음할 수는 없다. 여러 다른 지표들이 동시에 개입하고 간섭하는 까닭이다. 특히 『회색인』의 경우 두드러지는 공간 표지와 시간 표지의 밀도를 함께 고려하지 않는다면 정당한 결론에 가닿기 어렵다. 최인훈 소설에서의 추상어는 다만 현실의

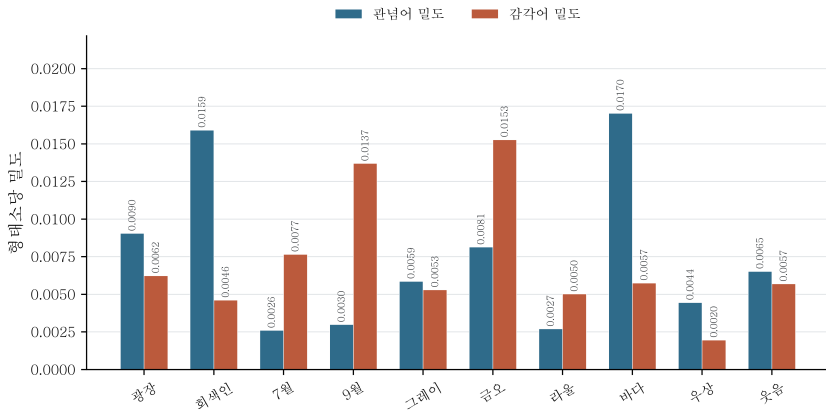


그림7-관념어와 감각어의 배치

주: 관념어 밀도와 감각어 밀도를 함께 비교하여 최인훈 소설의 추상 사유가 감각적 장면과 접속하는 방식을 확인한다.

결핍을 가리키는 말이 아니라, 현실이 곧장 이해되지 않는다는 사실을 문장 안에서 다루는 효과적인 장치라고 보는 편이 한결 더 타당하다.

요컨대 관념성은 현실의 소거가 아닌 현실의 재배열로 다루어져야 한다. 추상어는 장면을 지우는 것이 아니라 장면을 해석이 가능한 문제로 전환한다. 감각어와 공간 표지는 관념을 현실로 끌어내리는 단순한 배경이 아니라, 관념이 과연 어느 자리에서 발화되는지를 지정하는 장치라고 말할 수 있다.

## 2. 시간·공간 표지와 발화 위치

시간과 공간 표지는 최인훈 소설에서 배경 정보 이상의 기능을 수행한다. 그것들은 인물이 어느 역사적 시간과 어느 사회적 장소 가운데 발화하는지를 지정한다. 분단 이후의 주체에게 시간은 단순한 선형적 연대기의 의미를 갖지 않으며, 공간 또한 단순한 배경의 위치를 지니지 않는다. 시간은 과거와 현재의 미해결 상태를 장 속으로 불러오고, 공간은 광장과 밀실, 남과 북, 내부와 외부의 긴장을 조직한다.

「웃음소리」는 시간 표지 밀도 0.0173, 공간 표지 밀도 0.0220이라는 높은 수치를 보여 준다. 「바다의 편지」 역시도 시간 표지 0.0133, 공간 표지 0.0216이라는 상당한 수준을 나타낸다. 이러한 수치는 단편이 단순히 짧은 사건을 다루는 형식

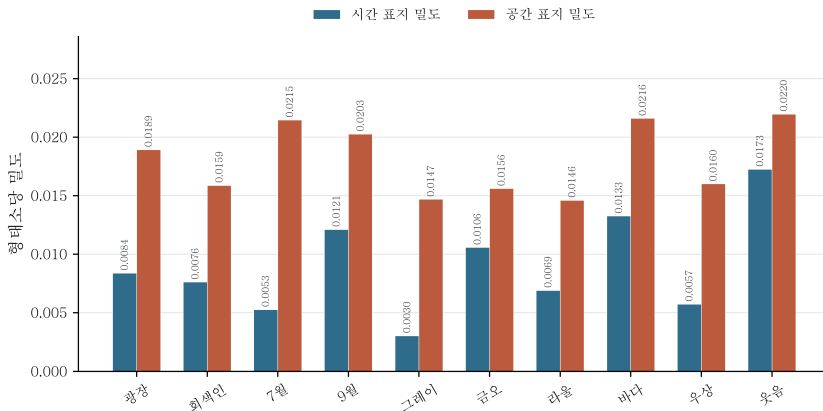


그림8-시간·공간 표지의 분포

주: 시간 표지와 공간 표지의 밀도를 비교하여 추상 사유와 장면성의 접속 방식을 확인한다.

이 아님을, 제한된 분량 가운데에서도 시간과 공간의 감각을 고밀도로 조직해 내는 형식적 특성을 소유하고 있음을 알려 준다.

상기 <그림8>은 최인훈 소설의 관념성이 공간과 시간의 감각적 질서와 무관하지 않음을 잘 보여 준다. 오히려 사유가 추상화될수록 그것이 놓이는 시간과 장소의 표지는 더 중요해진다. 최인훈 작품세계 속의 인물들은 보편적 관념의 공중에 떠 있는 것이 아니며, 분단 이후라는 시공간적 측면에서 특정한 역사적 자리에서 말하고 또 생각한다.

이 지점에서 최인훈 소설의 한국학적 의미가 드러난다. 최인훈의 문학은 한국 현대사의 문제를 단순히 소재로서만 삼지 않는다. 그 속에는 한국어 문장이 역사적 불안, 정치적 선택, 개인의 책임을 어떤 방식으로 견디는지가 분명하게 담겨 있다. 시간·공간 표지는 그 견딤이 추상적 사유만이 아니라 실제 발화의 자리와 긴밀히 결합돼 있다는 사실 또한 명시해 준다.

### 3. 해석 유보와 어휘의 회귀

최인훈 소설의 인물들은 세계를 직접 지시하지 않는다. “이것”, “그것”, “저기”, “그때”와 같은 지시 표현은 대상을 명확히 고정하기보다는 외려 잠시 미루거나

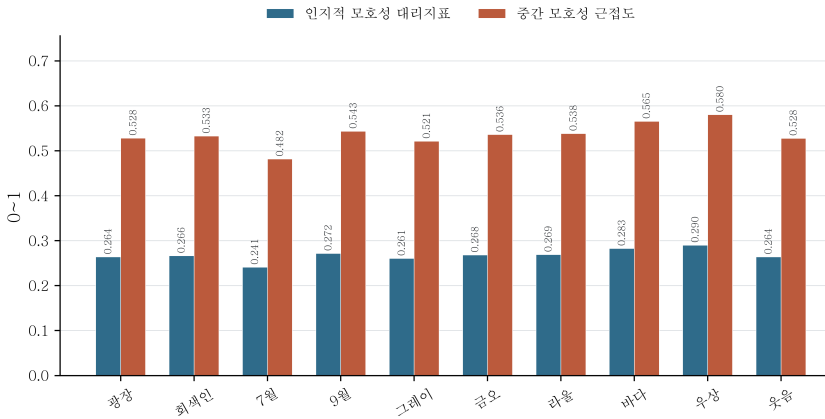


그림9-해석 유보와 복합적 사유 압력

주: 인지적 모호성 대리지표와 중간 모호성 근접도를 비교하되, 두 값을 탐색적 압력 위치의 지표로 제한한다.

흔들어 놓는 데 봉사한다. 이러한 지시의 유보는 단순한 모호함의 강조 표현이 아니다. 그것은 세계를 확정적으로 단언하여 말할 수 없는 주체의 위치, 그러나 그렇다고 말하기를 포기할 수도 없는 주체의 윤리를 동시에 드러낸다.

인지적 모호성 대리지표와 중간 모호성 근접도는 질문·모호 지시·문장 변동·의미 불연속 등이 한 문장에 겹치는 위치를 탐색하기 위한 장치다. 「우상의 집」, 「바다의 편지」, 「9월의 다알리아」의 근접도 값이 높지만, 이는 실제 독자의 판단 유보나 미학적 최적성을 측정한 결과가 아니다. 지표가 가리킨 문장에서 장문·지시·관념과 감각이 어떤 방식으로 결합하는지를 다시 확인할 때에만 해석적 의미를 얻는다.

유보는 의미의 결핍이라기보다 의미를 너무 빨리 확정하지 않게 하는 형식적 조건으로 읽을 수 있다. 다만 <그림9>의 결합 지표는 그 가능성이 모이는 위치를 가리킬 뿐, 독자의 실제 인지 상태를 증명하지 않는다. 해석은 그 위치의 문장과 장면을 재검토한 뒤에만 제시될 수 있다.

하기에서 살필 어휘 다양성 및 반복·변주의 비율 시각화는 이 논의를 한번 더 생각하도록 만든다. 어휘 다양도가 높다는 것은 텍스트가 비교적 넓은 어휘장을 사용함을 뜻하지만, 반대로 반복 및 변주의 비율은 사유의 회귀를 뜻하는 근거로 작용하는 까닭이다. 만약 두 요인이 함께 나타날 때 사유는 확장과 회귀를 동시

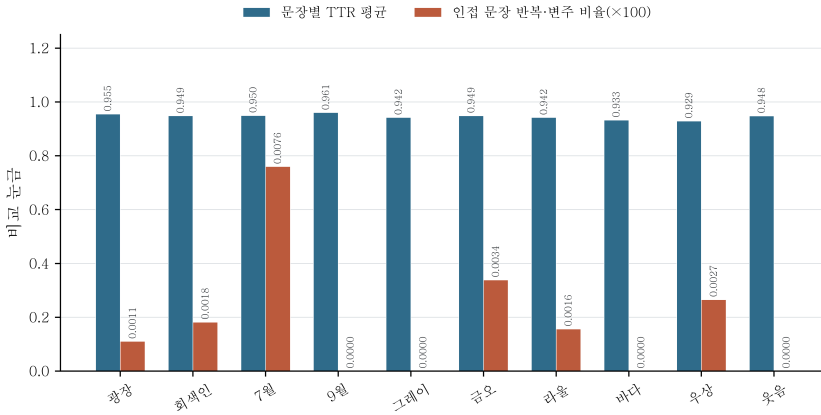


그림10-어휘 다양도와 반복·변주

주: 반복·변주 비율은 시각 비교를 위해 100배 배율로 표시했으며, 막대 위에는 원값을 병기하였다.

에 수행하게 된다.

〈그림10〉에서 문장별 TTR 평균은 전반적으로 높고 인접 문장의 반복·변주 판정 비율은 낮다. 이는 어휘장의 폭과 문자 표면의 근접 반복이 서로 다른 층위임을 보여 준다. 낮은 판정률을 곧 구조적 반복의 부재로 바꾸지 않으며, 문장별 TTR의 길이 민감성도 함께 고려한다.

#### 4. 지표 간의 상관성과 작품 텍스트

상기에서 다루어 온 복수의 지표들과 그 수치의 값을 작품별 독해의 차원에서 조금 더 세밀하게 묶어 볼 필요가 있다. 『광장』은 평균 문장 길이 22.48자, 문장 길이 변동계수 0.734, 의문 표현 밀도 0.0346, 지시 유보 밀도 0.0391을 보인다. 이 조합은 『광장』이 극단적으로 긴 문장이나 높은 변동성에 의존하기보다, 비교적 안정된 문장 리듬 안에서 질문과 선택의 구조를 누적한다는 점을 보여 준다. 『광장』의 긴장은 문장의 난폭한 흔들림보다 대립 구조의 지속성에서 나온다. 남과 북, 광장과 밀실, 개인과 이념, 삶과 죽음의 대립이 반복될수록 문장은 그 대립을 단번에 종결하지 않고, 비교적 균형 잡힌 리듬 속에서 선택의 불가능성을 더 오래 유지한다.

반면 『회색인』은 평균 문장 길이가 『광장』과 유사한 수준이지만 문장 길이 변동계수는 0.981로 훨씬 높다. 이 차이는 두 소설의 사유 방식이 다르다는 점을 말해 준다. 『광장』이 선택의 구조를 중심축으로 삼는다면, 『회색인』은 사유의 방향 자체가 더 크게 흔들린다. 장문과 단문이 교차하고, 지시어가 대상을 잠시 붙잡았다가 다시 놓치며, 관념어와 공간 표지가 서로 겹친다. 『회색인』의 지식인 주체는 세계를 해석하려 하지만, 그 해석은 자주 자기 내부에서 다시 분열된다. 따라서 『회색인』의 관념성은 하나의 명제를 밀고 가는 힘보다, 명제가 스스로 자기 근거를 계속 되묻게 되는 재귀 운동에 가깝다.

「바다의 편지」는 단편임에도 불구하고 평균 문장 길이 32.16자로 가장 높은 수치를 기록한다. 이 결과는 단편의 짧은 분량이 단순한 서사적 간결성을 뜻하지 않는다는 점을 보여 준다. 오히려 제한된 분량은 사유를 압축하고, 압축된 사유는 긴 문장의 형태로 나타날 수 있다. 「바다의 편지」에서 문장은 장편처럼 넓게 누적되기보다, 짧은 공간 안에서 감각과 관념, 기억과 판단을 함께 밀어 넣는다. 이때 장문은 장황함이 아니라 압축의 형식이다. 독자는 긴 문장을 따라가며 사건의 전개보다 사유의 응축을 먼저 경험하게 된다.

「우상의 집」과 「9월의 다알리아」 역시 단편의 압축성을 보여 준다. 두 작품은 평균 문장 길이가 높고, 지시 유보 밀도도 두드러진다. 이는 짧은 서사 안에서 인물과 세계의 관계가 곧장 명료해지지 않는다는 뜻이다. 특히 지시 유보는 단편의 결말을 단순한 여운으로 만드는 것이 아니라, 작품 전체의 발화 위치를 흔드는 장치로 작동한다. 무엇을 가리키는가, 어떤 대상을 말하고 있는가, 말의 책임은 어디에 있는가라는 질문이 짧은 문장 공간 안에서 압축적으로 발생한다.

「7월의 아이들」은 의문 표현 밀도와 대화 비율에서 두드러진다. 그러나 앞서 본 것처럼 대화 비율이 높다고 해서 이 작품의 문답성이 단순히 인물 간 대화의 차원에만 머무는 것은 아니다. 질문은 대화 속에서 나타나지만, 그 질문이 만들어 내는 판단 유보는 서술 전체의 범위로 확장된다. 아이들이라는 존재는 미성숙한 인물군의 표지가 아니라, 아직 결정되지 않은 시간과 판단의 가능성을 드러내는 장치로 읽힐 수 있다. 질문은 이 미결정성을 언어적으로 수행한다.

「라울전」과 「그레이 구락부 전말기」는 상대적으로 높은 대화 비율과 의문 표현의 밀도를 통해 소통의 구조를 드러낸다. 그러나 이 소통은 원활한 상호 이해로

귀결되지 않는다. 말은 오가지만 의미는 안정되지 않고, 인물들은 서로를 향해 말하면서도 자기 위치를 완전히 확보하지 못한다. 이 점에서 두 작품의 대화는 소통의 성공을 보여 주기보다 소통의 가능 조건 자체를 문제화한다.

「웃음소리」는 시간·공간 표지의 밀도가 높은 작품으로 주목된다. 이 작품에서 시간과 공간은 사건의 배경이 아니라 불안한 정동과 기억이 자리 잡는 방식이다. 웃음이라는 감각적·청각적 요소는 단순한 분위기 효과가 아니라, 특정한 시간과 장소를 매개로 인물의 판단을 흔든다. 따라서 시간·공간 표지의 높은 밀도는 장면성이 강하다는 단순한 사실을 넘어, 추상적 불안이 감각적 장면으로 조직되는 방식을 보여 준다.

이처럼 작품별로 살핀 지표의 조합 구조는 최인훈 문학을 단일한 이름의 관념성으로 묶기보다, 그 관념성이 사실상 여러 형식으로 분화되고 분기되는 방식을 보여 준다. 『광장』의 구조적 질문, 『회색인』의 변동적 사유, 「바다의 편지」의 압축적 장문, 「우상의 집」과 「9월의 디알리아」의 지시 유보, 「7월의 아이들」의 질문 밀도, 「웃음소리」의 시간·공간 표지는 모두 서로 다른 문장의 장치를 통해 사유의 지연을 끌어낸다. 곧 최인훈 작품세계의 관념성은 하나의 성질이 아니라 여러 형식의 묶음이다.

## 5. 데이터 보조 독해의 기능

바로 이 지점에서 데이터 보조에 기초한 독해는 최인훈 소설 연구에 일정한 기여도를 갖는다. 지표와 그 수치가 작품을 대신 읽지 않지만, 수치는 익숙하게 묶어 버린 작품들을 다시 세분하여 각각을 세밀하게 바라보도록 이끈다. 『광장』과 『회색인』은 모두 지식인의 사유를 담아내고 있지만, 문장 리듬과 변동성에서는 서로 다른 양상을 보인다. 또한 단편들은 짧은 이유로 부차화될 수 있지만, 평균 문장 길이, 지시 유보, 시간·공간 표지를 종합하여 살펴볼 때 오히려 더 높은 밀도를 보이기도 한다. 이 차이를 확인할 때 비로소 최인훈 문학의 내부 지형은 더 입체적인 모습을 띠게 된다.

나아가 이러한 방식의 분석 수행은 최인훈 소설을 기존의 문학사적 좌표평면, 곧 한국 현대문학사의 한 위치에 고정해 온 방식과도 상당한 거리를 둔다. 최인

훈은 분명 전후와 분단, 지식인과 이념, 개인과 역사라는 문제를 다루었지만, 그 문제들은 단지 문학적 표제어로만 남지 않는다. 그것들은 한국어 문장 안에서 길어지고, 멈추고, 되묻고, 가리키기를 유보하는 방식을 따라 실현되고 구체화한다. 따라서 최인훈 소설의 역사성은 주제의 역사성인 동시에 문장의 역사성이라고 말할 수 있다. 분단 이후의 한국어 문장은 단순히 역사를 기술한 게 아니라, 역사를 말하기 까다로운 복잡다단한 조건들까지도 함께 문장화했다.

이 점을 인정할 때 최인훈 문학의 관념성은 더 이상 현실과 동떨어진 추상성으로 동일시되지 않는다. 양자 간에는 건널 수 없는 심연이 개재해 있다. 최인훈의 관념성은 현실을 너무 쉽게 말할 수 없기에 생겨나는 언어적 긴장의 소산이다. 세계가 분명하지 않으므로 문장은 길어지고, 판단이 단번에 내려지지 않으므로 질문은 반복되며, 대상을 확정할 수가 없으므로 지시는 유보된다. 그러함에도 말하기 자체를 포기할 수는 없기에 문장은 계속된다. 최인훈 소설의 한국어는 바로 이 계속됨을 실현하는 형식이다.

## VI. 결론: 최인훈 문장의 사유 형식과 한국학적 의미

### 1. 논의의 종합

이 논문은 최인훈 소설의 관념성을 주제의 추상성에만 귀속하지 않고, 분석 대상 열 편을 통해 장문·변동성·문답·모호 지시·관념어와 감각어·시간·공간 표지가 맺는 관계로 다시 읽는다. 그 결과 사유는 완성된 관념이 문장에 주입되는 방식보다 문장이 연장되고 끊기며 자기 근거를 되묻는 과정에서 형성된다는 중심 가설이 작품별로 서로 다른 양상을 띠며 확인됐다.

첫째, 장문은 일부 작품에서 판단과 논증을 한 호흡 안에 함께 두지만, 높은 변동성은 장문 자체와 다른 리듬을 만든다. 둘째, 질문과 모호 지시는 직접 대화에 한정되지 않고 서술 내부의 판단 위치를 흔든다. 셋째, 관념어는 감각어와 시간·공간 표지의 결합 속에서 장면으로 되돌아온다. 넷째, 장편과 단편의 차이는 분명하되 장르·시기·판본 효과가 얹혀 있으므로 인과적 고유성으로 일반화할 수 없다.

## 2. 연구의 의의

본 연구의 주된 기여는 최인훈 소설의 관념성을 ‘관념의 양’이 아니라 관념이 문장 안에서 생성되는 형식의 문제로 재정의한 데 있다. 보조적인 기여는 두 가지다. 하나는 질문이 대화량보다 넓게 분포한다는 사실을 통해 기존 소통 구조 연구를 서술의 자기 반문으로 확장한 것이고, 다른 하나는 수치→압력 위치→정밀독해의 절차와 재현성 명세를 결합하여 한국 현대문학의 데이터 보조 독해가 취할 수 있는 접근법, 곧 제한적이면서도 충분히 검토가 가능한 새로운 방식을 제시한 것이다.

논의를 종합해 보면 최인훈 문학의 한국학적 의미는 분단 이후 한국어 사용 양상 전체를 대표한다는 데 놓이지 않는다. 정확하게는 분단의 역사적 압력이 특정한 소설 문장 안에서 길이와 질문, 지시와 감각의 배열로 변환되는 한 사례를 구체적으로 현시해 주었다는 데 놓인다.

## 3. 한계와 후속 과제

본 연구의 한계 역시 분명하다. 비교 작가 코퍼스가 없어 작가 효과와 장르·시기 효과를 분리하지 못했고, 로컬 분석본의 SHA-256은 확보했으나 최초 판본과 개작 이력을 완전하게 복원하지 못했다. 규칙 기반 문장 분절과 소규모 사전은 간접화법·다의어·의미적 반복을 놓칠 수 있으며, 문장별 TTR과 결합 대리지표도 길이와 가중치의 선택에 민감한 특성을 보인다. 따라서 결론의 힘은 분석한 열 편 및 이에 사용된 규칙 범위에 한정된다.

후속 연구를 세 방향으로 확장할 수 있다. 첫째, 동일 판본 원칙 아래 동시대 작가의 장·단편을 비교하여 장르·시기·작가 효과를 분리해야 한다. 둘째, 형태소 사전 결과를 문장 임베딩과 의미 군집으로 교차검증하되, LLM 페르소나가 산출한 반응을 역사적 독자의 실제 수용 자료로 오인하지 않아야 한다. 셋째, 개작판과 수용 자료를 결합하여 문장 형식의 변화와 독해의 변화를 함께 추적할 필요가 있다.

## 참고문헌

### 1. 논저

- 김근호, 「독자로서의 작가: 최인훈 <화두>의 문학독서 현상학」, 『현대소설연구』 71, 2018, 5~31쪽.
- 김선, 「최인훈의 문학과 방법론으로서의 풍속」, 『한국현대문학연구』 59, 2019, 121~159쪽.
- 김정관, 「최인훈 소설 『광장』의 미적 모더니티 연구」, 『비평문학』 89, 2023, 113~155쪽.
- 박경우, 「인문학 연구에서의 디지털 기술 활용 현황 및 적용 방향: 디지털 인문학 동향 분석을 중심으로」, 『국어국문학』 200, 2022, 137~179쪽.
- 방경원, 「텍스트 마이닝 기법을 활용한 텍스트 종류 연구: 한국과 독일증시 마감보고서를 비교하면서」, 『인문학연구』 59, 2020, 343~366쪽.
- 선민서, 「냉전의 대리전으로서의 6·25와 <광장>: 중립국 위원회의 실질적 부재」, 『국어국문학』 186, 2019, 329~366쪽.
- 양윤의, 「최인훈 소설에 나타난 소통 구조 연구」, 『한국문예비평연구』 28, 2009, 221~247쪽.
- 윤인로, 「노모스왕의 권위와 간접권력 비판의 사상연쇄: 최인훈 연작 『총독의 소리』를 매개로 삼아」, 『인문논총』 80-3, 2023, 471~500쪽.
- 윤인로, 「혁명검찰론=이단심문론에 대해: 소설 『화두』와 『광장』의 적론(敵論)을 위하여」, 『탈경계인문학Trans-Humanities』 16-2, 2023, 137~160쪽.
- 이나운, 「최인훈의 『광장』에 나타난 잠재성 연구」, 『우리말글』 96, 2023, 287~316쪽.
- 이소연, 「최인훈 소설에 나타난 대항근대성 연구: 『회색인』과 『서유기』에 드러난 상호텍스트적 '부재 영역'을 중심으로」, 『비평문학』 67, 2018, 223~248쪽.
- 이원양, 「최인훈 장편 소설 사이의 인지 층위 심화 과정: 광장—회색인—서유기—소설가 구보씨의 일일, 태풍의 연속성을 중심으로」, 『비평문학』 87, 2023, 291~328쪽.
- 정영훈, 「최인훈의 초기 비평 연구: 참여의 의미를 중심으로」, 『한국현대문학연구』 50, 2016, 477~513쪽.
- 정호웅, 「최인훈 문학과 한국현대문학의 타자들」, 『우리말글』 68, 2016, 439~463쪽.
- 홍순애, 「최인훈 『광장』의 법의식과 시민적 윤리」, 『현대소설연구』 67, 2017, 489~523쪽.

## References

### Books and Journal Articles

- Bang, Kyungwon, "Text Type Research Using Text Mining Technique: Comparing Korean and German Stock Closing Reports," *Humanities Research*, Vol. 59, 2020, pp. 343–366.
- Hong, Sunae, "A Study on the Legal Consciousness and Civic Ethics Appeared in Choi In-hoon's 「The Square」," *The Journal of Korean Fiction Research*, Vol. 67, 2017, pp. 489–523.
- Jung, Houg, "Choi In-hun's Literature and Other Writers of Korean Modern Literature," *Urimalgeul: The Korean Language and Literature*, Vol. 68, 2016, pp. 439–463.
- Jung, Younghoon, "A Study on Choi In-hun's Early Literary Criticism: Focused on the Meaning of Participation," *The Journal of Modern Korean Literature*, Vol. 50, 2016, pp. 477–513.
- Kim, Jungkwan, "A Study on the Aesthetic Modernity in Choi In-hoon's Novel, 「Square」," *Literary Criticism*, Vol. 89, 2023, pp. 113–155.
- Kim, Keunho, "The Author as a Reader: Phenomenology of Literary Reading in Choi In-hoon's 「Whadoo」," *The Journal of Korean Fiction Research*, Vol. 71, 2018, pp. 5–31.
- Kim, Sun, "The Custom of Choi In-hoon as a Literary and Methodology," *The Journal of Modern Korean Literature*, Vol. 59, 2019, pp. 121–159.
- Lee, Nayoon, "A Study on Virtuality in Choi In-hun's 「The Square」," *Urimalgeul: The Korean Language and Literature*, Vol. 96, 2023, pp. 287–316.
- Lee, Wonyang, "The Procedure of Deepening the Levels of Consciousness among the Long Novels of Choi In-hoon: Focused on the continuity of The Square, A Grey Man, The Seoyugi, A Day in the Life of Kubo and Typhoon," *Literary Criticism*, Vol. 87, 2023, pp. 291–328.
- Park, Kyungwoo, "The Analysis and Application of Digital Technology in Humanities Research: Focused on the Analysis of Research Trend in Digital Humanities," *The Korean Language and Literature*, Vol. 200, 2022, pp. 137–179.
- Seon, Minseo, "Korean War as a Cold War Proxy War in the Choi In-hun's Novel The Square: On the Substantial Absence of the Neutral Nations Repatriation Committee and the Unsolved Korean Problem-," *The Korean Language and Literature*, Vol. 186, 2019, pp. 329–366.
- Yang, Yuneui, "On Study of (Meta-)Communication in Choi In-hoon's Novels," *The Korean Literary Criticism Research*, Vol. 28, 2009, pp. 221–247.
- Yi, Soyon, "A Study on the Contra-modernity in Choi In-hun's Novels: Focused on The Grey Man and A Journey to the West," *Literary Criticism*, Vol. 67, 2018, pp. 223–248.
- Yoon, Inro, "About Choi In-hun's Revolutionary Prosecution and Inquisition," *Trans-Humanities*, Vol. 16, No. 2, 2023, pp. 137–160.
- Yoon, Inro, "Nomos Basileus as Authority and Ideological Chain of Criticism of Indirect Power," *Journal of Humanities*, Vol. 80, No. 3, 2023, pp. 471–500.

**Table(APPENDIX)**

- Table 1. Analysis Corpus and File Identification Information
- Table 2. Operational Definitions and Interpretive Limits of the Indicators
- Table 3. Reproducibility, Accuracy, and Research-Ethics Checks

**Figure**

- Figure 1. Sentence and Paragraph Counts in the Ten-Work Choi In-hun Fiction Corpus
- Figure 2. Source-Text Size and Valid-Sentence Count
- Figure 3. Distribution of Sentence-Length Categories after Preprocessing
- Figure 4. Long-Sentence Rhythm and Variability in Sentence Length
- Figure 5. Question-and-Answer Structure and Deictic Suspension
- Figure 6. Dialogue and Non-Dialogue Narration Ratios
- Figure 7. Distribution of Abstract and Sensory Words
- Figure 8. Distribution of Temporal and Spatial Markers
- Figure 9. Interpretive Suspension and Composite Cognitive Pressure
- Figure 10. Lexical Diversity and Repetition-Variation

부록

표1-분석 코퍼스와 파일 식별 정보

| 작품             | 구분 | 글자 수    | 문장 수  | 유효 문장 | 평균 길이 | CV    | SHA-256(앞 12자) |
|----------------|----|---------|-------|-------|-------|-------|----------------|
| 광장             | 장편 | 137,045 | 4,539 | 4,387 | 22.48 | 0.734 | b90b05638cba   |
| 회색인            | 장편 | 266,271 | 8,813 | 8,501 | 22.67 | 0.981 | 05b793613ebd   |
| 7월의 아이들        | 단편 | 12,954  | 658   | 617   | 14.65 | 0.698 | 40d7196e6bd3   |
| 9월의 다알리아       | 단편 | 4,349   | 107   | 105   | 29.74 | 0.677 | 357edcc00f59   |
| 그레이 구락부<br>전말기 | 단편 | 28,465  | 956   | 923   | 22.15 | 0.738 | 271312dfbbd7   |
| 금오신화           | 단편 | 18,578  | 591   | 557   | 23.68 | 0.795 | a937dac1e782   |
| 라울전            | 단편 | 21,979  | 641   | 614   | 25.46 | 0.830 | 2cf7c41a9f48   |
| 바다의 편지         | 단편 | 14,612  | 338   | 325   | 32.16 | 0.951 | 2dc1134ba000   |
| 우상의 집          | 단편 | 15,165  | 377   | 366   | 29.93 | 0.864 | 1d086421e260   |
| 웃음소리           | 단편 | 13,119  | 356   | 349   | 27.32 | 0.692 | 7dd5b49444eb   |

표2-지표의 조작적 정의와 해석 한계

| 지표            | 계산 · 판정 규칙                            | 해석 범위                  |
|---------------|---------------------------------------|------------------------|
| 문장 분절         | 종결부호 뒤 공백 또는 줄바꿈 기준; 2자 미만 제외         | 규칙 기반 대체 절차            |
| 문장 길이         | 공백을 제거한 문자 수                          | 평균 · 중앙값 · CV          |
| 장문/단문         | 작품 내부 85백분위 이상 / 20자 이하               | 장문은 작품 내부 상대 기준        |
| 대화/서술         | 인용부호 · 대시 시작 또는 인용부호 쌍 / 그 밖의 문장      | 간접화법은 서술로 남을 수 있음      |
| 질문 밀도         | 질문 표지 형태소 수 + 물음표 가중치 / 형태소 수         | 질문 행위의 대리지표            |
| 모호 지시 밀도      | 모호 · 지시 표지 형태소 수 / 형태소 수              | 실제 지시 대상의 불명료성과<br>비동일 |
| 어휘 다양도        | 문장별 고유 형태소 수 / 형태소 수의 작품 평균           | 길이 민감성이 남는 보조치         |
| 반복 · 변주       | 인접 문장 문자 2~4그램 TF-IDF 코사인 유사도         | 구조적 반복 전체와 비동일         |
| 인지적 모호성       | 9개 표면 지표의 가중합을 0~1로 절단                | 탐색적 결합 대리지표            |
| 중간 모호성<br>근접도 | $1-2 \times  \text{인지적 모호성 평균} - .5 $ | 최적성의 심리 측정치가 아님        |

표3-재현성 · 정확성 · 자료 윤리 점검

| 항목      | 확인 내용                                                                     |
|---------|---------------------------------------------------------------------------|
| 실행 환경   | Python 3.14.3; pandas 3.0.2; NumPy 2.4.4; scikit-learn 1.8.0; Kiwi 0.23.1 |
| 분석 단위   | 작품 전체 텍스트→문장→작품별 집계; 10편 모두 direct_full_text_analysis                     |
| 원자료 무결성 | 작품별 SHA-256 10개, 중복 해시 0개                                                 |
| 유효 문장   | 전체 96.4%; 작품별 최저 93.8%; 5자 미만은 유효 집계에서 제외                                 |
| 결측 점검   | 핵심 20개 집계 지표 NA 0건                                                        |
| 강건성     | 작품 한 편씩 제외한 평균 범위를 산출하고 극단적 작품 의존 여부를 별도 확인                               |
| 공개 범위   | 저작권 원문 제외; 코드 · 사전 · 문장별 파생 CSV · 집계표를 검증자료로 구성                           |
| 판본 한계   | 로컬 분석본의 파일 식별은 가능하나 최초 판본 · 개작 이력의 완전한 복원은 불가                             |