

언어

사람이 의미를 나누는 주된 체계에서, 단어가 범주가 되는 자리까지

기준일 2026년 9월 16일 · [dosi.dev/wiki](#)

사람이 의미를 나누는 주된 체계에서, 단어가 범주가 되는 자리까지

기준일 2026년 9월 16일

언어는 문법과 어휘로 짜인 소통의 체계다.¹ 사람은 입의 소리로, 손과 몸의 수어로, 그리고 그 둘을 다시 적은 문자로 같은 종류의 일을 한다. 유한한 조각으로 아직 없었던 문장을 만들고, 지금 여기 없는 일을 가리킨다. 이 책은 그 체계가 무엇이며 어떻게 배우고 퍼지고 사라지는지를 다룬다. 프로그래밍 언어와 형식 논리는 사람이 만든 다른 기호 체계로만 짚고, 동물의 신호는 경계 조건으로만 본다.

택한 해석은 사람의 자연 언어다. 중심 질문은 하나다. **조각난 소리가 어떻게 세계에 대한 말이 되며, 그 말은 어떻게 한 사람의 머릿속에서 범주가 되는가.** 더글러스 호프스태터와 에마뉘엘 상데가 2013년 영어·프랑스어로 동시에 낸 《사고의 본질》²은 그 둘째 물음에 대한 한 답이다. 단어는 이름표가 아니라, 지금 마주한 장면을 과거의 장면과 같게 보는 유추다. 이 주장은 보편문법이나 통계적 형태 학습과 같은 층에 있지 않다. 충돌하는 지점은 9장에서 모은다.

주

- 1 영어 위키백과 「Language」, 리비전 1374640745. <https://en.wikipedia.org/wiki/Language> 위치: 리드·정의·다양성.
- 2 Basic Books / Hachette. *Surfaces and Essences*, 2013-04-23, 592쪽, ISBN 9780465018475. <https://www.hachettebookgroup.com/titles/douglas-hofstadter/surfaces-and-essences/9780465018475/> 위치: 재킷 주장.

1장. 말이 하는 일과 하지 않는 일

세 가지 정의

영어 language는 라틴어 lingua를 거쳐 인도유럽조어의 ‘혀’에 닿는다. 그 어원이 가리키는 기관과, 오늘날 학계가 가리키는 대상은 같지 않다. 페르디낭 드 소쉬르는 강의 노트에서 langage(말 일반), langue(한 공동체의 체계), parole(실제 발화)를 갈랐다. 1916년 제자 샤를 발리와 알베르 세제가 그 구분을 《일반언어학 강의》로 묶었다.¹

같은 대상을 세 방향으로 잘라 온 전통이 있다. 하나는 칸트·데카르트에서 노엄 촘스키의 보편문법, 제리 포더, 스티븐 핑커로 이어지는 정신 능력이다. 아이는 교실 없이 말을 배우고, 니카라과 수어처럼 입력이 빈약한 자리에서도 문법이 자란다. 다른 하나는 기호와 규칙을 닫힌 체계로 보는 형식적 정의다. 소쉬르의 구조, 촘스키의 생성문법이 여기 속하고, 전산 모형도 이 길을 빌린다. 셋째는 말을 사회적 도구로 보는 길이다. 기능문법, 후기 비트겐슈타인, J. L. 오스틴, H. P. 그라이스가 수행과 함의를 중심에 둔다.

스탠퍼드 철학백과의 언어학 철학 항목(2011, 2024년 3월 7일 개정)은 이 갈림을 외부주의·창발주의·본질주의로 다시 정리한다.² 코퍼스의 패턴을 모형화하는 쪽, 일반 인지와 대화에서 구조가 나온다고 보는 쪽, 학습 가능한 문법의 한계를 생물적 제약으로 보는 쪽이 같은 이중타동 구문도 서로 다른 사실로 읽는다. 같은 문장 “아이가 말을 배운다”를 본질주의는 제약의 증거로, 창발주의는 사용의 증거로, 외부주의는 코퍼스에 나타난 패턴의 증거로 읽는다. 반례가 달라진다. 대화가 드문 것은 현상이 달라서가 아니라, 무엇이 설명해야 할 대상인지가 달라서다.

설계 특징

찰스 F. 호켓은 1960년 「말의 기원」에서 인간 언어의 설계 특징을 열거했다.³ 처음 13개 가운데 영장류와 겹치는 앞의 아홉을 빼고, 원격성·생산성·전승·이중분절을 인간 쪽에 두었다. 이후 거짓말하기, 언어에 대한 언어, 다른 언어를 배울 수 있음을 더해 16개가 되었다.

원격성은 지금 눈앞에 없는 일, 없는 것, 아직 오지 않은 일을 가리키는 능력이다. 생산성은 유한한 단위로 무한히 새로운 조합을 만드는 일이다. 이중분절은 의미 없는 소리(음소)가 의미 있는 단위(형태소)를 만들고, 그 단위가 다시 문장을 만드는 두 층이다. 영어 cat은 ⁴ /k/, /æ/, /t/의 배열이다. 같은 조각으로 act와 tack이 된다.

이 목록은 경계선이지 점수가 아니다. 꿀벌의 춤은 멀리 있는 먹이를 알리므로 원격성의 조각을 갖는다. 어떤 새의 지저귀음 요소를 결합한다. 어느 동물 체계도 16개를 한꺼번에 갖지는 않는다. 보노보 칸지는 ⁵ 렉시그램 판으로 훈련을 받았고, 새와 고래는 소리를 배운다. 네 살 아이의 어휘와 문법을 따라간 비인간 사례는 확인되지 않았다. 고릴라 코코의 어휘 규모와 감정 일화는 방법 논쟁이 남아 있어 이 책은 수치를 올리지 않는다.

본능과 규약

말은 배우지 않으면 생기지 않고, 가르치지 않아도 자란다. 이 문장은 모순이 아니다. 노출된 언어(구어든 수어든)가 있어야 하고, 그 노출을 문법으로 압축하는 생물적 준비가 있다. 어떤 개설은 촘스키 이후 다수가 본능을 지지한다고 적는다. 철학백과가 보여주는 현장은 본능설과 사용 기반 설이 같은 학과 안에서 병존한다는 쪽이다.

자의성은 하늘과 sky와 뜻이 같은 대상을 다른 소리에 묶는 사실이다. 완전하지는 않다. 어떤 소리는 뜻과 느슨히 기울고, 의성어는 기울기가 더 크다. 그래도 공동체가 바꾸지 않는 한 개인이 수박을 다른 이름으로 불러 통하게 할 수는 없다. 호프스테터의 언어관은 이 자의성을 부정하지 않는다. 그가 겨냥하는 것은 소리가 아니라, 그 소리가 가리키려고 꺼내는 범주다.

주

- 1 Ferdinand de Saussure, 《Cours de linguistique générale》, 1916 (Bally-Sechehaye). 위키 「Course in General Linguistics」. https://en.wikipedia.org/wiki/Course_in_General_Linguistics 위치: langue/parole.
- 2 Scholz, Pelletier, Pullum, Nefdt, 「Philosophy of Linguistics」, SEP, 2011; 개정 2024-03-07. <https://plato.stanford.edu/entries/linguistics/> 위치: 3절근.
- 3 영어 위키백과 「Hockett's design features」. Hockett 1960 「The Origin of Speech」. https://en.wikipedia.org/wiki/Hockett%27s_design_features 위치: 13→16 특징.
- 4 영어 위키백과 「Duality of patterning」. https://en.wikipedia.org/wiki/Duality_of_patterning 위치: 이중분절.
- 5 영어 위키백과 「Language」, 리비전 1374640745. <https://en.wikipedia.org/wiki/Language> 위치: 리드·정의·다양성.

2장. 작은 조각이 끝없는 말을 만드는 방식

소리의 층

구어는 연속된 기류를 끊어서 듣는다. 음소는 뜻을 가르는 최소의 추상 단위다. 영어 bat과 pat을 가르는 것은¹ /b/와 /p/다. 한국어는 같은 자리에서 기식과 후두 긴장으로 세 갈래를 만든다. 한 언어의 음소가 다른 언어에서는 같은 소리의 변이에 불과할 수 있다. 영어의 기식은 음소가 아니고, 힌디어와 관화어와 한국어에서는 기식이 뜻을 가른다.

음소 목록의 크기는 언어마다 크게 다르다. 어떤 언어는 열 개 남짓이고, 어떤 클릭 언어는 백 개를 넘긴다. 피라항어의 음소 수는 논쟁 중이라 점값으로 올리지 않는다. 모든 구어는 적어도 모음과 자음의 구별을 갖고 음절을 만든다. 강세, 길이, 성조는 음소 위의 층에서 뜻을 가른다.

수어는 손이 만드는 모양, 방향, 위치, 움직임이 조음에 해당한다. 이중분절이 구어보다 덜 또렷하다는 관찰이 있다. 쓸 수 있는 몸짓의 폭이 넓어, 의미 없는 원소와 의미 있는 기호의 경계가 흐려진다는 뜻이다. 그 흐림이 문법이 없다는 뜻은 아니다.

단어 안의 문법

형태소는 뜻을 가진 최소 단위다. unexpected는 un-, expect, -ed로 분해된다. 굴절은 같은 단어를 문법 환경에 맞게 바꾸고, 파생은 새 단어를 만든다. 중국어처럼 단어가 거의 형태소 하나인 분석어, 라틴어처럼 여러 문법 값이 한 접사에 뭉치는 융합어, 터키어처럼 접사가 하나씩 쌓이는 교착어, 한 단어가 문장에 가까운 포합어가 전형으로 거론된다. 라틴 bonus는 어근 bon-에 남성·단수·주격이 한 접사 -us로 붙는다. 터키어 evlerinizden은 집-복수-당신들-에서다. 영어는 고립어에 가깝지만 대명사 격과 동사 시제에 굴절의 잔재가 있다. 속어 fan-fucking-tastic 같은 끼워 넣기는 포합의 흥내를 낸다. 유형은 경향이지 100퍼센트가 아니다.

품사는 기능과 자리로 가르다. 명사는 참여자, 동사는 사건, 형용사는 속성에 가까운 열린 부류다. 대명사처럼 닫힌 부류는 새 항목이 잘 안 들어온다. 언어마다 열린 부류와 닫힌 부류의 경계가 다르다. 자동사는 논항 하나, 타동사는 둘을 요구한다. 일치(형용사가 명사를, 동사가 주어를 따라가는 일)는 형태론이 통사론과 만나는 자리다.

단어 사이의 문법

통사론은 배열이 뜻을 정하는 규칙이다. 영어 I love you는 받아들여지고 love you I는 같은 단어로도 문장이 못 된다. 라틴어는 격이 역할을 표시하므로 어순이 더 자유롭다. 구는 나무 구조로 묶여 한 덩어리처럼 움직인다. the cat sat on the mat에서 명사구와 전치사구는 통째로 옮겨질 수 있다.

기본 어순은 언어를 가르는 잘 보이는 축이다. 영어는 주어-동사-목적어, 많은 언어는 주어-목적어-동사다. 함의적 보편이 관찰된다. SOV 언어는 후치사와, 명사 앞 형용사를 함께 갖는 경향이 있다. 주격-대격 정렬은 자동사의 주어와 타동사의 주어를 같게 취급하고, 능격 정렬은 자동사의 주어와 타동사의 목적어를 같게 취급한다. 이 공통이 보편 법칙인지, 자주 반복되는 수렴인지는 유형론의 열린 문제다.

무한은 어디에 있는가

생산성은 재귀에서 온다. 관계절 안에 관계절을 넣는 식으로 문장은 원리상 길어진다. 촘스키는 이 이산을 인간 언어의 핵심으로 보았고, 2002년 마크 하우스·촘스키·W. 테쿰세 피치의 구분²에서 좁은 언어 능력의 후보가 되었다. 사용 기반 진영은 실제

발화의 깊이가 얕다고 지적한다. 프레드리크 칼손 2007년 조사는³ 중심 내포 깊이 3이 드물고 4는 거의 없다고 했다. 무한 생성기는 능력의 기술일 수도 있고, 모형의 우아함일 수도 있다. 어느 쪽을 택하느냐가 6장의 학습 이론을 가른다.

호프스태터의 눈으로 보면, 재귀는 같은 범주를 자기 안에 다시 넣는 유추다. “그 사람”이 가리키는 자리가 절 안에서 한 번 더 채워질 때, 듣는 사람은 바깥 절의 틀을 안쪽에도 쓴다. 통사 나무는 그 재사용을 그림으로 고정한 것이다.

주

- 1 한국어 위키백과 「언어」, oldid 42425270. <https://ko.wikipedia.org/wiki/언어> 위치: 정의·구조·기원.
- 2 영어 위키백과 「Universal grammar」. https://en.wikipedia.org/wiki/Universal_grammar 위치: UG 변천.
- 3 Fred Karlsson, 「Constraints on multiple center-embedding of clauses」, Journal of Linguistics 43(2), 2007. SEP 「Philosophy of Linguistics」 §2.1가 인용. <https://plato.stanford.edu/entries/linguistics/> 위치: 중심 내포 깊이. Karlsson 원문은 미열람, SEP 경유.

3장. 입과 손과 글

두 개의 1차 통로

말은 입과 귀의 통로로 태어나지 않았다. 근본 양태는 둘이다. 구강(말, 입모양)과 수어(손과 몸). 둘은 자주 같이 간다. 이야기할 때 손짓이 격을 보태고, 어떤 오스트레일리아 언어는 몸짓이 문법 정보를 실어 나른다. 어느 쪽으로 배우든 아이는 같은 시기에 유창해진다. 뇌 손상은 말과 수어를 비슷한 방식으로 무너뜨린다. 손상되는 것은 혀가 아니라 언어다.

수어는 마임이 아니다. 기호는 관례적이고 종종 자의적이다. 손 모양·위치·움직임·손바닥 방향이 구어의 음소에 해당하고, 그중 하나만 바뀌어도 뜻이 갈린다. 영국 수어와 미국 수어는 ¹ 둘 다 영어권에서 쓰이지만 서로 통하지 않는다. 미국 수어의 통사는 영어보다 일본어에 가깝다는 비교가 있다. 청인이 발명한 손말도 아니고, 입말을 손가락으로 옮긴 암호도 아니다. 기원전 5세기 플라톤 《크라틸로스》에 손짓 언어의 기록이 있고, 1620년 후안 파블로 보넬의 논고, 1755년 샤를 미셸 드 레페의 파리 농학교가 제도의 이정표다.

집계는 갈린다. 에스놀로그 2021년판은 150개, SIGN-HUB는 200개가 넘고 미기록분을 더한다. 위키백과 목록은 오늘날 약 300개로 잡는다. 새 수어가 크리올처럼 생기므로 점값은 의미가 없다.

2차 부호화로서의 문자



기원전 2041–2040년(아마르신 즉위 6년) 수메르어 쐩기문자 점토판. 말은 남지 않고 글만 남은 자리의 원품 사진이다. 커코르 미나시안 컬렉션, 미국 의회도서관. 복원·모형이 아니다.

Davidge (2005 업로드) / Library of Congress · Public Domain Mark 1.0 — https://commons.wikimedia.org/wiki/File:Cuneiform_script2.jpg

문자는 말을 눈에 보이게 한 이차 체계다. 나중에 온 사람이 발화를 재구성할 수 있어야 비로소 문자다. 그 이전의 그림 기호는 원문자에 가깝다.

통설은 독립 발명 네 곳이다. 메소포타미아 수메르 기원전 3400–3100년경 ², 이집트 기원전 3250년경, 중국 기원전 1250년 이전, 메소아메리카 기원후 1년 이전. 알파벳은 소리에, 음절문자는 음절에, 한자와 마야 문자는 단어와 음절을 섞어 대응한다.

순수 표어문자는 없다. 단어가 너무 많다.

방향은 관례다. 왼쪽에서 오른쪽, 그 반대, 위에서 아래, 줄마다 방향을 바꾸는 부스트로페돈(boustrophedon)이 모두 쓰였다. 국제음성기호는 세계의 말소리를 한 세트로 적으려는 후대의 도구다.

한글은 이차 체계 가운데 설계 기록이 남은 드문 경우다. 조선 세종 때 1443년 12월(그레고리력 1443년 12월-1444년 1월) 궁중에 거의 완성본이 올랐고, 1446년 9월 《훈민정음》³과 해례로 반포되었다. 원래 이름은 훈민정음, “백성을 가르치는 바른 소리”다. 세종은 “나라 말씨가 중국과 다르다”고 적었다. 기본 자음은 발음 기관의 모양을 본뜨고, 가획으로 센소리를 표시한다. 모음은 당시의 음양·사람 관념과 모음조화를 반영한다. 15세기에는 서울 궁중과 일부에 머물렀고, 16세기에 여성과 불교 쪽에서 퍼졌으며, 1504년 금지를 겪었다. 20세기 민족 운동과 문해 정책 뒤에야 남북의 주된 문자가 되었다.

말과 글은 1 대 1이 아니다. 같은 하루를 Haru, ハル, 哈嚕로 적을 수 있고, 같은 철자 chat은 영어와 프랑스어에서 다른 소리와 다른 뜻이다. 글을 소리에 종속된 기호로만 보면 수어와 점자가 빠진다. 글은 소리의 그림자이기도 하고, 어떤 전례 언어에서는 이제 1차 매체가 되기도 한다.

주

- 1 영어 위키백과 「Sign language」. https://en.wikipedia.org/wiki/Sign_language 위치: 수어=자연어.
- 2 영어 위키백과 「History of writing」. https://en.wikipedia.org/wiki/History_of_writing 위치: 독립 발명 4곳.
- 3 영어 위키백과 「Hangul」. <https://en.wikipedia.org/wiki/Hangul> 위치: 1443/1446.

4장. 말하게 하는 몸과 뇌

고전 지도와 그 한계

19세기 병변 연구는 언어를 두 점으로 찍었다. 피에르 폴 브로카는¹ 루이 빅토르 르보르뉴("탄")와 라자르 룰롱의 뇌를 보고, 왼쪽 아래이마이랑(브로드만 44·45)을 산출과 연결했다. 브로카 실어는 말은 하고 싶으나 나오지 않고, 문장이 전보처럼 짧아진다. 가고 싶다, 가게, 음식 같은 낱말은 남고 문법 골격이 빠진다. 카를 베르니케는 1874년 왼쪽 위관자이랑 뒤쪽을 이해와 연결했다. 베르니케 실어는 리듬과 문장 골격은 남고 내용이 비며, 신조어와 의미 착오가 섞인다. 본인은 잘 모른다. 고전 모형은 두 영역을 활꼴섬유로 잇고, 그 경로가 끊기면 따라 말하기가 무너진다고 보았다.

수어에도 같은 패턴이 나타난다. 망가지는 것은 조음 기관이 아니라 언어 능력이다. 그 이상은 단순하지 않다. 1978년 모어의 연구와 2007년 니나 드론커스가 브로카의 원 뇌를 MRI로 다시 찍은 결과는, 브로카 영역만의 손상이 브로카 실어를 만들지 않는다고 말한다. 역사적 병변은 흰질까지 넓었다. 베르니케 영역의 경계는 사람마다 다르고 합의가 없다. 이해는 관자·마루·이마엽의 그물에서 일어난다. 제프리 바인더의 후속 작업은 어휘 의미와 소리 인식을 서로 다른 관자엽 구역에 둔다. 교과서의 두 상자는 역사이지 회로도가 아니다.

영상 기법(fMRI, 전기생리)은 손상되지 않은 뇌를 본다. 지식이 늘어난 만큼, "언어 중추가 여기"라는 문장은 줄어들었다. 오른쪽 반구의 대응 구역은 덜 흔한 뜻과 운율에 관여한다는 보고가 있다. 개인차가 커서, 한 점의 좌표로 모든 화자를 설명할 수 없다.

유전자 하나로는 안 된다

FOXP2는 1998년² 영국 KE 가족 연구에서 말·언어 장애와 연결되었고, 2001년 유전자가 특정되었다. 15명이 3대에 걸쳐 발달성 언어실행증을 보였다. 늦은 말, 어눌함과 말더듬, 그리고 운동만으로 설명되지 않는 문법·이해 문제였다. 같은 계열의 유전자는 쥐의 울음, 새의 소리 모방, 박쥐의 반향정위에 관여한다. 침팬지에도 있다.

2018년 연구는 현생 인류에서 FOXP2의 최근 양성 선택 증거를 찾지 못했다. 말을 하지 않는 동물도 이 유전자를 가진다. 언어에 필요한 다른 유전자가 있다. FOXP2는 언어 유전자가 아니라, 언어가 깨질 수 있는 자리 하나다. 고유전체에서 이 유전자의 인간형 변이가 네안데르탈에게도 있다는 보고는, 유전자 하나를 말의 출현 연대로 쓰는 일을 더 어렵게 만든다.

화석이 말해주는 것

폐와 후두와 성도는 기류를 모음과 자음으로 자른다. 모음은 마찰 없이 혀와 입술의 위치로, 자음은 막힘의 자리와 방식과 유성으로 갈린다. 스펙트로그램의 포먼트는 그 모양의 음향 지문이다. 사람이 낼 수 있는 소리는 수백 가지이고, 일부는 한 어족이나 한 언어에만 있다.

해부학은 말의 필요조건이지 언어의 연대가 아니다. 뇌 용량, 후두 위치, 도구는 간접 단서다. 후두가 내려가야만 말이 시작된다는 주장은 논쟁 중이라 이 책은 연대의 뜻으로 쓰지 않는다. 말이 남긴 뼈는 없다. 있는 것은 글자뿐이고, 글자는 말보다 늦다.

주

1 영어 위키백과 「Broca's area」; 「Wernicke's area」. https://en.wikipedia.org/wiki/Broca%27s_area 위치: 고전 모형과 한계.

2 영어 위키백과 「FOXP2」. <https://en.wikipedia.org/wiki/FOXP2> 위치: KE 가족·2018 선택.

5장. 말이 시작된 자리 — 증거가 멈추는 선

연속과 불연속

기원 이론은 언어를 무엇으로 보느냐에 따라 갈린다. 연속성 이론은 영장류의 인지나 몸짓·발성에서 점진적으로 자랐다고 본다. 2024년 기준으로 언어학자의 다수가 이쪽이다.¹ 스티븐 핑커는 동물 인지의 연장을, 마이클 토마셀로는 협력하는 몸짓을 출발점으로 잡는다. 토마셀로의 공유 의도성은 같이 가리키고 같이 보려는 능력이 문법보다 먼저라는 주장이다. 장 자크 루소와 찰스 다윈, 스티븐 미텐은 음악과 노래를 경유로 둔다.

불연속성 이론은 인간만의 갑작스러운 사건을 가정한다. 촘스키는 진화 이야기를 “핵심이 아니다”에 가깝게 두었고, 무작위 돌연변이가 뇌를 재편했다는 비유를 썼다. 문자 그대로의 기관 이식이라기보다, 점진적 적응만으로는 재귀를 설명하기 어렵다는 강조로 읽힌다. 마크 하우저·촘스키·피치 2002년 논문은² 넓은 언어 능력(기억, 발성 학습, 마음 이론)과 좁은 언어 능력(이산 무한, 재귀)을 갈랐다. 핑커와 레이 재켄도프는 2005년 재귀만으로 언어를 줄이는 그림에 반박했다. 2002년 논문의 논리 구분은 이후 하우저 개인 연구에 대한 부정행위 판정과 별개다.

연대가 갈리는 이유

직접 기록은 없다. 비교언어학이 닿는 깊이는 언어 자체보다 얇다. 입증된 가장 오래된 어족으로 자주 꼽히는 아프리카아시아어족조차 언어의 출현보다 젊다. 우연과 차용을 걸러 낸 뒤에도, 음운 변화가 쌓이면 동계어가 보이지 않는다.

제시된 숫자만 늘어놓으면 거짓 정밀이 된다. 조애나 니콜스 1998년은 음성 언어의 분화가 적어도 10만 년 전이라고 보았다. 로버트 버윅과 촘스키는 20만–6만 년 사이를 말한다. 어떤 두 단계 모형은 발화 장치를 60만 년 전, 재귀 통사를 7만 년 전으로 나눈다. 하빌리스(약 230만 년)의 원시 체계부터 후기 구석기 현생 인류까지, 범위는 두 자릿수 만 년을 넘는다. 2024년 3월 어떤 연구는 약 160만 년을 제시했다. 어떤 한국어 개설이 출현을 “기원전 3만–10만년”으로 적는 표기는, 기원전과 지금 이전을 섞을 위험이 있어 채택하지 않는다.

2017년 아르디피테쿠스 라미두스 연구는 오스트랄로피테신의 소통이 유인원과 같았다는 전제를 흔들었다. 화석이 흔드는 것은 해부이지, 문법의 연대가 아니다.

하나의 발명으로 모든 현대 구어가 갈라졌다는 가정이 흔하다. 재구성으로 그 조어에 닿을 수는 없다. 관측되면 이야기가 바뀌는 신호는 분명하다. 네안데르탈 유전체에서 언어 관련 변이의 조합이 현생과 같아지거나, 고고학에서 상징 행동의 시간표가 통사 연대와 어긋나면, 불연속 모형은 자리를 잃는다. 반대로, 현생 인류 출현 이전의 해부가 말을 허용하지 않는다는 합의가 다시 선다면, 연속성의 하한은 올라간다.

주

1 영어 위키백과 「Origin of language」, 리비전 1374769938. https://en.wikipedia.org/wiki/Origin_of_language 위치: 연속/불연속·연대.

2 영어 위키백과 「Universal grammar」. https://en.wikipedia.org/wiki/Universal_grammar 위치: UG 변천.

6장. 아이가 문법을 압축하는 속도

순서와 창

제1언어 습득은 몇 년 안에 한 언어의 음운·어휘·문법을 아이 머릿속에 압축하는 일이다. 태아는 이미 주변 말에 반응한다. 신생아는 인간 말에 더 반응하고, 생후 한 달이면 음성 범주를 가르며, 여섯 달 전후 웅얼이가 시작된다. 농아동의 손 웅얼이는 더 늦다. 한 단어, 두 단어, 전보 같은 조합을 거쳐 다섯 살 무렵 문법의 골격이 선다. 18-20개월의 「엄마 밥」 같은 두 단어 조합은 전형적인 전보어로 개설서에 반복된다. 개인 차는 크다. 24개월에 단어가 50개 미만이면 늦은 말로 본다.

결정적 시기라는 말은 창이 닫힌다는 뜻이다. 와일더 펜필드와 라마 로버츠 1959년은¹ 아홉 살을 한계로 보았고, 열두 살을 넘기면 원어인 같은 습득이 어려워진다는 관찰이 많다. 13세까지 고립되었다가 발견된 지니는 어휘는 늘었으나 생산적 문법에 이르지 못했다. 아베롱의 빅토르도 같은 방향의 증거로 인용된다. 사례는 적고 윤리적으로 반복할 수 없다. 늦은 제1언어 학습자(어떤 농인 아동을 포함)의 복잡한 문법 결손이 더 넓은 자료다.

빈약한 자극, 풍부한 입력

츨스키 1959년은 B. F. 스키너의 조작적 조건화를 통사에 적용할 수 없다고 보았다. 빈곤한 자극 논증은 이렇다. 아이가 듣는 말은 유한하고 애매하며, 어른의 문법 교정은 드물고 비체계적인데, 전형적인 아이들은 다섯 살 전에 같은 문법으로 수렴한다. 영어 예/아니오 의문문은 선형 교환이 아니라 계층 구조의 이동을 따른다. 들은 문장만으로는 둘 다 가능해 보인다.

제프리 풀럼과 바버라 솔츠 2002년은² 그 경험적 뒷받침이 약하다고 했다. 마이클 토마셀로 2003년은 일반 학습, 덩어리, 사회적 피드백으로 충분하다고 보았다. 제니 새프런, 엘리사 뉴포트, 리처드 애슬린의 통계 학습 실험은 영아가 음절 경계와 단어를 빈도에서 자른다고 보여 주었다. 두 진영은 같은 속도를 다른 엔진으로 설명한다. 생물적 제약이 탐색 공간을 줄이느냐, 입력이 생각보다 정보가 많으냐.

입력이 거의 없을 때

1977년 마나과에³ 농학교 센터가 열리며 50명의 아이가 모였다. 1979년 100명, 1980년 직업학교 이후 1983년 400명을 넘었다. 그 전까지 고립된 농인은 집안의 손짓만 갖고 있었다. 스페인어 구화 교육은 실패했고, 아이들은 비공식 공간에서 체계를 만들었다. 초기 LSN은 단순했고, 어린 코호트가 동사 일치 같은 구조를 보태 ISN이 되었다. 주디 케글이 1986년 접촉했고, 앤 생하스가 세대 차이를 기록했다. 핑커가 요약하듯, 문법을 만든 것은 어른이 아니라 아이다. 윌리엄 스토크이는 타고난 장치 주장과 외부 영향을 물었고, 고립 연구의 윤리 비판이 있다. 어느 단계에서 “완전한 언어”가 되는지는 연구자마다 다르다. 창발이 일어났다는 사실 자체는 남아 있다.

피진은 공유 모어가 없는 집단의 접촉어로, 문법이 단순하고 누구의 제1언어도 아니다. 크리올은 그 피진을 아이의 모어로 삼을 때 자란다. 모든 피진이 크리올이 되지는 않는다. 데릭 비커튼은 1980년대에⁴ 플랜테이션의 아이가 생득 프로그램을 적용한다고 보았다. 살리코코 무프웨네는 피진과 크리올이 다른 조건에서 따로 생긴다고 본다. 약 100개의 크리올이 1500년 이후 기록되었다. 니카라과 수어는 손으로 된 같은 실험이다.

주

1 영어 위키백과 「Language acquisition」. https://en.wikipedia.org/wiki/Language_acquisition 위치: 습득 단계·빈곤한 자극.

2 영어 위키백과 「Universal grammar」. https://en.wikipedia.org/wiki/Universal_grammar 위치: UG 변천.

- 3 영어 위키백과 「Nicaraguan Sign Language」. https://en.wikipedia.org/wiki/Nicaraguan_Sign_Language 위치: 1977- 마나과.
- 4 영어 위키백과 「Pidgin」; 「Creole language」. <https://en.wikipedia.org/wiki/Pidgin> 위치: 접촉어.

7장. 단어가 세계를 자르는 방식 — 호프스테터의 유추

관례 의미와 지금 이 자리의 의미

소쉬르에게 기호는 소리 이미지와 개념의 결합이고, 둘의 연결은 자연이 아니라 관례다. 언어학은 그 관례를 두 층으로 다시 자른다. 문장이 모든 맥락에서 갖는 고정된 잠재(관례 의미)와, 화자가 이 자리에서 이루려는 일(소통 의도)이다. “소금 닿니?”는 거리에 대한 질문이 아니라 건네 달라는 요청일 수 있다. 오스틴의 화행¹, 그라이스의 함의가 이 틈을 설명한다. 직시(“나”, “지금”, “여기”)는 발화 순간에만 채워진다.

에밀리 M. 벤더와 알렉산더 콜러는 2020년² ACL 논문에서 의미를 형식과 소통 의도의 관계로 정의했다. 형식은 페이지의 링크, 파일의 바이트, 조음 기관의 움직임처럼 관측 가능한 실현이다. 의도에는 창문을 열라는 명령도, 말라라의 생일을 묻는 질문도, 그냥 같이 있고 싶음도 들어간다. 이해는 형식을 받아 그 의도를 되찾는 일이다.

유추가 단어를 꺼낸다

재킷 문장은 저자의 주장이다

“유추는 모든 생각의 핵심이다”는 2013년 책의 테제이지, 인지과학의 합의 문장이 아니다. 이 장은 그 테제가 단어에 대해 무엇을 설명하는지 추적한다.

호프스테터와 상데는 이 그림의 다른 면을 본다. 2013년 2월 오딜 자코브에서³ 프랑스어 《유추. 생각의 심장》이, 같은 해 4월 23일 베이직 북스에서⁴ 영어 *Surfaces and Essences: Analogy as the Fuel and Fire of Thinking*(592쪽, ISBN 9780465018475)이 나왔다. 두 언어로 동시에 썼다. 한국어는 2017년 11월 아르테에서⁵ 《사고의 본질 — 유추, 지성의 연료와 불길》로, 김태훈이 옮기고 최재천이 감수했다(ISBN 8950955679). 아래 문장은 출판사 재킷과 프랑스어 위키, 저자 항목을 경유한 주장이다.

재킷의 문장은 짧다. “유추는 모든 생각의 핵심이다.” 드문 통찰이 아니라 매 순간, 무의식적으로, 과거 경험과 강한 유추 연결을 찾는 과정이 생각을 정의한다. 범주가 꺼내는 과정은 이렇다. 눈앞의 장면이 들어오면 머리는 그것을 이미 갖고 있는 묶음과 대조한다. 겹치는 뼈대가 가장 강한 묶음이 이기고, 그 묶음에 붙어 있던 소리가 단어로 떠오른다. 카미유라는 아이가 “바나나를 벗겼어!”라고 한 일화가 실린다. 옷을 벗기는 동작의 뼈대(덮개를 몸에서 분리한다)가 껍질을 벗기는 장면에 이기고, “벗기다”가 나왔다. 아인슈타인이 빛을 입자로 본 도약도, 저자들에게는 같은 기계다. 새 관측이 이미 있는 입자 범주의 뼈대에 올라탄 것이다.

이 주장은 자의성을 뒤집지 않는다. 바나나라는 소리와 노란 과일 사이의 연결은 여전히 관례다. 뒤집히는 것은 그 관례가 머리에 저장되는 형식이다. 사전 항목처럼 경계가 선 정의가 아니라, 과거에 성공했던 묶음들의 무리다. 새 장면이 오면 그 무리 중 가장 잘 맞는 것이 단어로 떠오른다. 관례 의미는 공동체가 공유하는 그 무리의 겹침이고, 지금 이 자리의 의미는 그 겹침을 이 장면에 얹는 일이다.

호프스테터의 앞선 작업이 이 그림과 이어진다. 1997년 《마로의 아름다운 소리》는⁶ 클레망 마로의 짧은 시를 여든여덟 가지로 옮기며, 번역이 단어 대응이 아니라 제약 아래의 유추임을 보여 준다. 카피캣 같은 미시 영역 모형은 글자 열에서 “같은 것”이 미끄러지는 과정을 계산으로 재현하려 했다. 인지과학 전체의 합의는 아니다. 데드레 겐트너의 구조 사상, 개념적 은유 이론,

사용 기반 구성문법은 유추를 인정하면서도 엔진의 자리를 다르게 둔다. 호프스태터를 이 책의 중심에 두는 이유는 합의여서가 아니라, 단어가 세계에서 무엇을 하는지에 대한 물음을 형태소 목록보다 깊게 만들기 때문이다.

형식만 남은 자리

벤더와 콜러의 문어 실험은 같은 물음을 기계 쪽에 던진다. 두 섬에 떨어진 A와 B가 해저 전신으로 영어를 주고받는다. 섬에 올라올 수 없는 문어 O가 선을 엮고 통계 패턴으로 B의 답을 예측하게 된다. 일상 잡담에서는 B인 척할 수 있다. 사회적 발화는 물리적 장면에 묶일 필요가 적다. A가 코코넛 투석기 제작법을 보내면 O는 rope와 coconut이 무엇을 가리키는지 모른다. 비슷한 맥락의 단어에 기대 “멋진 생각이야!” 같은 B의 상투구를 되풀이할 수 있고, A가 의미를 부여하면 통할 것처럼 보인다. 곰에게 쫓기는 A가 무기를 부탁하는 순간, 형식의 분포는 더 이상 버틸 수 없다.

논문의 문장 그대로, 말이 되는 것은 O의 발화가 아니라 A가 그 발화를 말이 되게 만드는 일이다. 대형 언어 모형은 형식을 예측하도록 학습된다. 그 성공을 “이해”라고 부르는 것은, 청자가 늘 해 온 의미 부여를 모형에 투사하는 일이다. 호프스태터의 유추가 돌아가려면 과거의 장면이 세계에 닿아 있어야 한다. 형식의 이웃만 쌓인 기억은 그 닿음이 없다.

주

- 1 영어 위키백과 「Language」, 리비전 1374640745. <https://en.wikipedia.org/wiki/Language> 위치: 리드·정의·다양성.
- 2 Emily M. Bender & Alexander Koller, 「Climbing towards NLU」, ACL 2020. <https://aclanthology.org/2020.acl-main.463/> 위치: 형식과 의미. 문어 실험.
- 3 프랑스 위키백과 「L'Analogie. Cœur de la pensée」. Odile Jacob, 2013-02. https://fr.wikipedia.org/wiki/L%27Analogie._Cœur_de_la_pensée 위치: 프랑스어 원제.
- 4 Basic Books / Hachette. *Surfaces and Essences*, 2013-04-23, 592쪽, ISBN 9780465018475. <https://www.hachettebookgroup.com/titles/douglas-hofstadter/surfaces-and-essences/9780465018475/> 위치: 재킷 주장.
- 5 알라딘. 《사고의 본질》, 아르테, 2017-11, 김태훈 옮김, 최재천 감수, ISBN 8950955679. <https://www.aladin.co.kr/search/wsearchresult.aspx?SearchTarget=Book&SearchWord=%EC%82%AC%EA%B3%A0%EC%9D%98+%EB%B3%B8%EC%A7%88+%ED%98%B8%ED%94%84%EC%8A%A4%ED%83%9C%ED%84%B0> 위치: 한국어판.
- 6 영어 위키백과 「Douglas Hofstadter」. https://en.wikipedia.org/wiki/Douglas_Hofstadter 위치: 저자 서지.

8장. 몇 개의 말, 몇 개의 이름

세는 일의 조건

살아 있는 언어가 몇 개인가는 언어가 무엇이나의 함수다. 상호 이해, 공유된 문헌, 민족 정체성이 갈라지면 숫자가 바뀐다. “언어란 육군과 해군을 가진 방언”이라는 경구는 막스 바인라이히 1945년으로 전한다. 아랍어의 구어들은 서로 통하지 않아도 현대 표준 아랍어라는 지붕 아래 묶이고, 힌디어와 우르두어는 거의 통하는 구어를 정치가 가른다. 중국의 관화와 오어와 월어는 문자로 묶이고 말로는 갈린다.

에스놀로그 29판(2026)은 살아 있는 언어 7,170개를 적는다.¹ 28판(2025)은 7,159개였다. 아래 표는 영어 위키가 2026년 9월에 전제한 에스놀로그 2026 수치다. 에스놀로그 웹 원문은 유료라 여기서는 그 전재를 쓴다. SIL 현장 조사와 ISO 639-3 향이 입력이고, 2019년부터 연 480달러의 개인 구독 장벽이 있다. 2015년 서평은 16–18판에서 인용 공백, 누락, 허위 항을 지적하면서도 비교 가능한 자료로는 여전히 최선이라고 했다.

글로벌로그 5.3(2026년 3월)은² 편집자가 존재와 구분을 확인한 항만 남긴다. 확인되지 않은 상속 항은 허위·미증언으로 표시한다. 인구 수는 거의 없다. 두 목록을 같은 총수로 더하지 않는다.

많이 쓰는 말

에스놀로그가 거시언어를 쪼갠 총화자(제1+제2, 백만 명) 상위는 영어 1,493³ (제1 372, 제2 1,121), 관화 1,183(988+194), 힌디어 611, 스페인어 561, 현대 표준 아랍어 335(제1은 0 — 가정 방언이 먼저다), 프랑스어 334다. 한국어는 같은 표에서 제1언어 82, 제2언어 1, 합 82다. 숫자가 퍼짐을 보여 주는 쪽은 제2언어다. 영어의 제2언어 1,121은 모어보다 크고, 현대 표준 아랍어는 모어 화자가 0으로 잡힌다. 학교와 제도가 가정 방언 위에 올린 지붕이다. 프랑스어도 제2언어가 제1언어를 앞선다. 제국·무역·학교·문자가 말을 옮긴 자리다.

CIA 월드 팩트북 2022년의 세계 인구 비율은 다른 저울이다. 영어 18.8퍼센트, 관화 13.8, 힌디어 7.5, 스페인어 6.9. 제2언어의 문턱이 정의되지 않아 영어 총화자는 출처에 따라 두 배에 가깝게 흔들린다. 같은 순위를 2010년 스웨덴 국립백과 표로 적는 개설이 있으나, 그 원문은 유료라 숫자를 올리지 않는다.

어족을 화자로 묶으면 인도유럽, 중국티베트, 아프리카아시아, 니제르콩고, 오스트로네시아 다섯이 세계 인구의 약 6분의 5를 차지한다.⁴ 언어 수로 보면 니제르콩고와 오스트로네시아가 각각 1,200–1,500개 규모로 크다. 글로벌로그 4.8 스냅샷은 고립어를 포함해 수백 개의 독립 가족을 센다. 바스크어처럼 살아 있는 친척이 증명되지 않은 말이 있다. 피진·크리올·혼합어는 한 줄기의 조어가 없다.

유엔 공용어는 영어, 프랑스어, 스페인어, 러시아어, 중국어, 아랍어 여섯이다. 화자 수와 제도의 힘은 같은 축이 아니다.

화자 수 표 (에스놀로그 2026, 백만 명)

거시언어를 쪼갠 총화자. 한국어는 이 상위 10 밖이다. 단위는 백만 명.

순위	언어 (에스놀로그 묶음)	제1언어	제2언어	합
1	영어 (크리올 제외)	372	1,121	1,493
2	관화 (다른 중국어 제외)	988	194	1,183

순위	언어 (에스놀로그 묶음)	제1언어	제2언어	합
3	힌디어 (우르두 제외)	347	264	611
4	스페인어 (크리올 제외)	487	75	561
5	현대 표준 아랍어 (구어 방언 제외)	0	335	335
6	프랑스어 (크리올 제외)	75	258	334
7	벵골어	234	43	274
8	포르투갈어 (크리올 제외)	252	18	269
9	인도네시아어	78	177	255
10	우르두어 (힌디 제외)	78	168	246
—	한국어	82	1	82

한국어가 놓인 자리

한국어는 제주어와 함께 한국어족을 이룬다.⁵ 제주는 본토와 상호 이해가 어려워 별개 언어로 보는 편이 정확하다. 알타이 가설(통구스·몽골·튀르크와의 혈연)은 핵심 묶음조차 대부분 지지를 잃었다. 교착, SOV 같은 유형 공통은 접착으로도 생기고, 제안된 동계어와 음운 대응은 약하다.

화자 수는 2019–2022년 자료로 약 8,100만, 에스놀로그 2026은 8,200만이다. 남한 표준은 서울말에 기대고, 북한은 1966년부터 평양말을 표준으로 공표한다. 발음, 철자, 문법, 어휘가 갈린다. 문자는 3장의 한글이다.

사라지는 속도

21세기 초의 말 가운데 50에서 90퍼센트가 2100년에 심하게 위태롭거나 죽을 것이라는 전망이 개설서에 반복된다. 이 구간은 점값이 아니다. 마이클 크라우스 1992년⁶ 『위기의 세계 언어들』(1991년 언어학회 강연)은 분야 자체의 90퍼센트가 사라진다고 경고했다. 2004년의 어떤 전망은 2050년 90퍼센트를 말했다. 유네스코 세계 소멸위기 언어 지도 2010년⁷ 제3판은 2003년 9개 활력 요인으로 2,473개 언어를 등급 매겼다. 온라인 WAL은 2026년 현재 작업 중·오프라인이다. 한 세기 동안 2,000개 이상이 죽었다는 집계, 2주마다 하나가 사라진다는 유엔 표현은 저울이 다르다.

위태로움의 핵심은 다음 세대로의 전승이다. 젊은이가 집에서만 쓰거나, 노인만 남거나, 공적 영역에서 더 큰 말에 자리를 내주면 화자 수가 아직 많아도 활력은 이미 낮다. 세계화가 소수를 큰 공동체로 끌어들이는 압력과, 정체성의 표지로 말을 붙잡는 운동이 동시에 있다. 무엇이 관측되면 90퍼센트 시나리오를 버리는가. 제1언어 전승이 유지되는 소수 언어의 비율이 안정되고, 제도 언어 전환이 둔화되는 세대가 두 번 이어지면, 상한은 내려가야 한다.

주

- 영어 위키백과 「Ethnologue」, 2026-09-01 편집. 29판(2026) 7,170. 원문 ethnologue.com 은 유료. <https://en.wikipedia.org/wiki/Ethnologue> 위치: 판본 수치.
- 영어 위키백과 「Glottolog」, 2026-09-13. 5.3 (2026-03). <https://en.wikipedia.org/wiki/Glottolog> 위치: 확인된 구분.
- 영어 위키백과 「List of languages by total number of speakers」, 2026-09-08 07:30 UTC. Ethnologue 2026 전제. https://en.wikipedia.org/wiki/List_of_languages_by_total_number_of_speakers 위치: 총화자 표.
- 영어 위키백과 「Language family」. https://en.wikipedia.org/wiki/Language_family 위치: 어족·재구 한계.

- 5 영어 위키백과 「Korean language」; 「Koreanic languages」. https://en.wikipedia.org/wiki/Korean_language 위치: 화자-계통.
- 6 영어 위키백과 「Michael E. Krauss」. 1992 「The World's Languages in Crisis」. https://en.wikipedia.org/wiki/Michael_E._Krauss 위치: 90% 경고.
- 7 영어 위키백과 「Endangered language」; 「Atlas of the World's Languages in Danger」. https://en.wikipedia.org/wiki/Endangered_language 위치: 소멸 구간.

9장. 아직 합의되지 않은 것

연구사가 남긴 도구

문법의 과학은 아테네에서 시작하지 않았다. 파니니¹의 《아슈타디야이(Aṣṭādhyāyī)》는 산스크리트 형태론을 3,959개(다른 집계 3,996) 규칙으로 압축했다. 연대는 긴장 상태다. 베다 문헌과의 언어적 연결은 기원전 6–5세기를, 루피아 같은 화폐 용어는 더 늦은 창을 가리킨다. 조르주 카르도나는 기원전 400년보다 이를 수 없다고 보았고, 얀 후벤은 기원전 350년경을 구체적인 증거에 기대 적는다. 소쉬르를 포함한 후대 기술은 이 생성적 메타언어를 알고 있었다.

수메르 서기관은 기원전 1900년경 이미 수메르어와 아카드어의 차이를 비교했다. 윌리엄 존스는 1786년 2월 2일² 벵골 아시아 학회 기념 강연에서, 산스크리트와 그리스어와 라틴어가 “동사의 어근과 문법의 형태에서, 우연이 만들 수 있는 것보다 강한 친연”을 보인다고 했다. 출판은 1788년이다. 비교 재구가 어족을 만드는 도구가 된 자리다. 소쉬르는 한 시점의 체계(공식)와 시간의 변화(통시)를 갈라 20세기 구조주의의 바닥을 깔았다. 촘스키 1960년대는 그 바닥 위에 학습 가능한 문법의 한계를 올렸다.

말이 생각을 만드는가

강한 언어 결정론³ — 말이 사고를 완전히 규정한다는 버전 — 은 폐기되었다. 약한 상대성은 남는다. 색 어휘는 기억에 작은 영향을 주고, 베를린과 케이의 색채 어휘 발달 순서는 생물적 제약을 가리킨다. 공간 참조 틀이 절대 방위만 있는 언어의 화자는 물체를 다른 방식으로 놓는다. 대니얼 에버렛이 보고한 피라항어의 수량 표현 부재는 수 인지와 문법의 관계를 둘러싼 논쟁의 한 향이다. 현재 합의에 가까운 문장은 이렇다. 말은 특정 인지 조작에 재현 가능한 작은 영향을 주고, 나머지 많은 조작은 언어에 의존하지 않는다.

이누이트어가 눈에 스무 개 단어를 가졌다는 이야기는 과장이다. 프란츠 보아스가 적으 것은 네 개였고, 포함어라 형태소를 쉽게 붙일 수 있어 숫자가 부풀어 보였다. 어휘 목록의 길이는 세계가 몇 조각인지의 증거가 되기 어렵다. 호프스테터라면 그 조각을 몇 개로 세느냐보다, 새 눈이 어떤 옛 눈에 유추되는지를 물을 것이다.

비트겐슈타인의 《논리철학논고》 5.6은⁴ “내 언어의 한계는 내 세계의 한계를 뜻한다”(독일어 Die Grenzen meiner Sprache bedeuten die Grenzen meiner Welt)다. 이 문장은 언어학의 정의가 아니라 1921년 논리 책의 명제다. 후기 비트겐슈타인의 언어 게임은 7장의 화용 쪽에 더 가깝다.

보편문법이 줄어드는 자리

촘스키의 보편문법은 1965년 《통사 이론의 여러 양상》⁵에서 학습 가능한 문법의 한계로 시작되었다. 이후 원리와 파라미터, 2002년 계층 구조, 2017년 버워과의 《왜 우리만》에서 병합(Merge)이라는 최소 연산으로 줄었다. 니컬러스 에번스와 스티븐 레빈슨 2009년은 극단적 다양성 앞에서 보편을 신화에 가깝다고 했다. 에버렛 2005년의 피라항 내포 부재 주장은 반박과 재반박이 이어진다. 풀럼과 솔츠는 빈곤한 자극의 경험적 기반을 흔들었다.

본질주의가 설명하려는 것은 속도다. 창발주의가 설명하려는 것은 쓰임이다. 외부주의가 설명하려는 것은 기록된 패턴이다. 세 설명이 한 문장으로 합쳐지지 않는 한, “언어는 본능이다”와 “언어는 문화다”는 둘 다 이 책의 결론이 될 수 없다.

형식의 이웃과 세계의 달음

대형 언어 모형은 형태를 예측하는 일에 능하다. 7장의 문어 실험이 가리키듯, 벤더와 콜러의 논증은⁶ 그 능함이 의미 관계를 학습했다는 증거가 될 수 없다고 한다. 반론은 이미 나와 있다. 멀티모달 학습, 도구 사용, 강화 학습으로 형식이 세계에 닿는다는 주장이다. 그 반론이 성립하려면 학습 신호에 형식 밖의 것이 들어 있어야 한다. 형식만으로 충분하다는 주장은 문어 실험이 가리키는 실패 지점을 이동시키지 못한다.

호프스태터는 2023년 인터뷰에서⁷, 인공지능이 곧 사람을 넘어설 수 있다는 쪽으로 오랜 회의를 수정했다고 한다. 그 수정은 유추가 생각의 핵심이라는 2013년 테제를 자동으로 입증하지 않는다. 기계가 사람 말을 흉내 내는 속도와, 단어가 범주로 꺼내는 과정은 같은 축정이 아니다.

관측되면 9장의 그림이 바뀌는 신호. 제1언어 없이 형식만으로 자란 체계가 새 장면에서 사람에게 필요한 도구를 사람처럼 지시할 수 있다면, 문어 논증의 상한은 내려간다. 피라항의 내포가 재현 가능한 방식으로 부재가 아니라면, UG를 신화로 부르는 쪽의 한 기둥이 빠진다. 색·수·공간 밖의 영역에서 언어가 사고를 크게 비트는 효과가 반복되면, 약한 상대성의 “약함”을 다시 적어야 한다.

주

- 1 영어 위키백과 「Pāṇini」. <https://en.wikipedia.org/wiki/P%C4%81%E1%B9%87ini> 위치: 연대 긴장.
- 2 영어 위키백과 「William Jones (philologist)」; 「Indo-European languages」. 담화 1786-02-02, 출판 1788. [https://en.wikipedia.org/wiki/William_Jones_\(philologist\)](https://en.wikipedia.org/wiki/William_Jones_(philologist)) 위치: 비교언어학.
- 3 영어 위키백과 「Linguistic relativity」. https://en.wikipedia.org/wiki/Linguistic_relativity 위치: 결정론 폐기.
- 4 Ludwig Wittgenstein, *Tractatus Logico-Philosophicus* 5.6. Wikiquote 경유. https://en.wikiquote.org/wiki/Ludwig_Wittgenstein 위치: 한계 명제.
- 5 영어 위키백과 「Universal grammar」. https://en.wikipedia.org/wiki/Universal_grammar 위치: UG 변천.
- 6 Emily M. Bender & Alexander Koller, 「Climbing towards NLU」, ACL 2020. <https://aclanthology.org/2020.acl-main.463/> 위치: 형식과 의미. 문어 실험.
- 7 영어 위키백과 「Douglas Hofstadter」. https://en.wikipedia.org/wiki/Douglas_Hofstadter 위치: 저자 서지.

참고문헌

개설

- 영어 위키백과. 「Language」. 리비전 1374640745. <https://en.wikipedia.org/wiki/Language>
- 한국어 위키백과. 「언어」. oldid 42425270. <https://ko.wikipedia.org/wiki/언어>
- 나무위키. 「언어」. 최종 수정 2026-09-02 09:41:04. <https://namu.wiki/w/언어>
- Scholz, Barbara C.; Pelletier, Francis Jeffry; Pullum, Geoffrey K.; Nefdt, Ryan. 「Philosophy of Linguistics」. *Stanford Encyclopedia of Philosophy*, 2011; 실질 개정 2024-03-07. <https://plato.stanford.edu/entries/linguistics/>

수와 다양성

- 영어 위키백과. 「Ethnologue」. 2026-09-01 편집. 29판(2026) 7,170. 원문 [ethnologue.com](https://www.ethnologue.com) 은 유료. 수치는 위키 전재. <https://en.wikipedia.org/wiki/Ethnologue>
- 영어 위키백과. 「List of languages by total number of speakers」. 2026-09-08 07:30 UTC. 에스놀로그 2026 전재. https://en.wikipedia.org/wiki/List_of_languages_by_total_number_of_speakers
- 영어 위키백과. 「List of languages by number of native speakers」. https://en.wikipedia.org/wiki/List_of_languages_by_number_of_native_speakers
- 영어 위키백과. 「Glottolog」. 2026-09-13. 5.3 (2026-03). <https://en.wikipedia.org/wiki/Glottolog>
- 영어 위키백과. 「Language family」; 「Endangered language」; 「Language death」; 「Atlas of the World's Languages in Danger」; 「Michael E. Krauss」.

구조·양태·몸

- Hockett, Charles F. 「The Origin of Speech」. *Scientific American*, 1960. (설계 특징 — 위키 「Hockett's design features」, 「Duality of patterning」 경유)
- 영어 위키백과. 「Sign language」; 「History of writing」; 「Hangul」; 「Korean language」; 「Koreanic languages」; 「Broca's area」; 「Wernicke's area」; 「FOXP2」.

기원·습득

- 영어 위키백과. 「Origin of language」. 리비전 1374769938.
- 영어 위키백과. 「Language acquisition」; 「Nicaraguan Sign Language」; 「Pidgin」; 「Creole language」; 「Universal grammar」; 「The Language Instinct」.
- Chomsky, Noam. *Aspects of the Theory of Syntax*. 1965. (UG — 위키 경유)
- Hauser, Marc D.; Chomsky, Noam; Fitch, W. Tecumseh. 「The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?」. *Science*, 2002. (Science 원문 대신 SEP·UG 항목 경유)
- Tomasello, Michael. *Constructing a Language*. 2003. (위키 경유)
- Pinker, Steven. *The Language Instinct*. 1994.
- Pullum, Geoffrey K.; Scholz, Barbara C. 「Empirical assessment of stimulus poverty arguments」. 2002. (위키 경유)

의미와 유추

- de Saussure, Ferdinand. *Cours de linguistique générale*. 1916 (발리·세세 편집).
- Hofstadter, Douglas; Sander, Emmanuel. *L'Analogie. Cœur de la pensée*. Odile Jacob, 2013-02.
- Hofstadter, Douglas; Sander, Emmanuel. *Surfaces and Essences: Analogy as the Fuel and Fire of Thinking*. Basic Books, 2013-04-23. 592쪽. ISBN 9780465018475.
<https://www.hachettebookgroup.com/titles/douglas-hofstadter/surfaces-and-essences/9780465018475/>
- 더글러스 호프스테터, 에마누엘 상데. 《사고의 본질》. 김태훈 옮김, 최재천 감수. 아르테, 2017-11. ISBN 8950955679.
- Hofstadter, Douglas. *Le Ton beau de Marot*. 1997.
- Bender, Emily M.; Koller, Alexander. 「Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data」. *Proceedings of the 58th Annual Meeting of the ACL*, 2020.
<https://aclanthology.org/2020.acl-main.463/>
- 영어 위키백과. 「Linguistic relativity」; 「Analogy」; 「Douglas Hofstadter」.
- Wittgenstein, Ludwig. *Tractatus Logico-Philosophicus*, 명제 5.6. Wikiquote 경유.

연구사

- 영어 위키백과. 「Pāṇini」; 「William Jones (philologist)」; 「Indo-European languages」; 「Course in General Linguistics」.
- Evans, Nicholas; Levinson, Stephen C. 「The myth of language universals」. 2009. (위키 경유)

그림 출처

- 그림 1. 수메르 쐼기문자 점토판. 아마르신 즉위 6년(기원전 2041–2040년). 커코르 미나시안 컬렉션, 미국 의회도서관. 원품의 사진(복원·모형 아님). 업로드 Davidge, 2005, 의회도서관 출처. 미국 연방 저작 Public Domain Mark 1.0. https://commons.wikimedia.org/wiki/File:Cuneiform_script2.jpg · LoC
<http://hdl.loc.gov/loc.amed/amcune.cf0013>