

The dream

The quest for the *lingua adamica* begins as theology. Genesis gives language a single origin and a single catastrophe: Adam names, Babel scatters. Everything afterward is understood as debris. So the search for the original tongue was never merely historical curiosity — it was soteriological. Recover the first names and you recover the fit between word and thing that the Fall broke.

The classic episodes:

- **Hebrew as the *Ursprache*** — the default Christian and Jewish assumption into the 17th century, defended by Kircher, contested by nationalist rivals (Goropius Becanus's notorious claim for Dutch, Swedish for Rudbeck).
- **Magical-combinatorial recovery** — Kabbalistic letter permutation, Lull's *Ars*, Dee and Kelley's "Enochic," Bruno's mnemonic wheels: if the first language was *motivated* rather than arbitrary, it could be reconstructed by generative combination.
- **Philosophical languages** — Wilkins, Dalgarno, and above all Leibniz's *characteristica universalis*: abandon retrieval, build the perfect language instead. The lost Adamic tongue becomes a design spec.
- **Comparative philology** — Jones's 1786 Sanskrit lecture, then Grimm, Bopp, Rask, Schleicher. Here the theological project mutates into a science.

Eco's *The Search for the Perfect Language* is still the best map of this whole terrain; Foucault's *Les Mots et les choses* tells the companion story of how the resemblance-based episteme that made such a search intelligible collapsed.

How it failed

Three distinct failures, worth keeping apart.

1. **Empirical.** No candidate survived. Hebrew is one Semitic branch among many; every nationalist claim was special pleading. Reconstruction has a horizon — comparative method degrades past roughly 6,000–10,000 years because sound correspondences are swamped by chance resemblance. Proto-Indo-European is reachable; "Proto-World" is not. Nostratic, Altaic, Greenberg's Amerind, Ruhlen's global etymologies: all foundered on that horizon.
2. **Conceptual.** The premise was wrong. Saussure's arbitrariness of the sign means there is no privileged mapping to recover. Worse, a "first language" may be a category error: language plausibly emerged gradually, in populations, possibly polygenetically, without a datable *Ursprache* to be the ancestor of.
3. **Political.** The project was repeatedly annexed by racial and national mythology — Aryanism being the catastrophic case. The dream of the one origin doubles as a dream of legitimate descent.

Popper's point applies neatly: the program's early forms were unfalsifiable, and its scientific form succeeded precisely by accepting constraints that made the original goal unreachable.

How it succeeded

Spectacularly, by failing at the wrong target and hitting a better one.

- **Method.** The comparative method is one of the genuine achievements of 19th-century human science — regular sound change, Grimm's and Verner's laws, the neogrammarian principle of exceptionlessness. A predictive science of the past out of nothing but word lists.
- **Vindicated predictions.** Saussure's *coefficients sonantiques* postulated on internal grounds, then found as actual Hittite letters decades later. That is as clean a test as historical linguistics gets.
- **Displaced discovery.** The universal turned out not to be a *language* but a *faculty*. Chomsky relocated the search from lexicon to competence; Greenberg pursued typological universals; Berlin & Kay found structure in color terms. The "original language" dissolved into universal grammar, universal semantic constraints, and the neural substrate of language.
- **Disciplinary yield.** Egyptology, Assyriology, Indo-European studies, and the entire archaeology-linguistics-genetics triangulation of prehistory are all children of a failed theological errand.

What AI adds

Not the answer — the *scope*. This is where large models do real epistemic work rather than rhetorical work.

- **Distributional semantics as an argument.** Word embeddings, and then transformer representations, demonstrate that enormous amounts of meaning are recoverable from co-occurrence structure alone. That is a substantive philosophical result about how much of semantics is relational rather than referential — closer to Wittgenstein's later view, and to Firth, than to any Adamic naming.
- **Cross-lingual alignment.** Unsupervised bilingual lexicon induction (Conneau et al., Artetxe) showed embedding spaces of unrelated languages can be aligned with little or no parallel data — evidence of shared conceptual geometry. Massively multilingual models exhibit interlingual representations that support zero-shot transfer. This is the nearest thing we have to an empirical *interlingua*: not a first language, but a common latent space, which is arguably what Leibniz actually wanted.
- **Scaled comparative work.** Automated cognate detection, phylogenetic Bayesian methods on lexical databases (Gray & Atkinson and successors), computational sound-law induction. AI does not extend the reconstruction horizon much, but it makes the horizon measurable and lets us quantify how much resemblance chance alone produces.
- **Decipherment.** Statistical and neural attacks on Linear B, Ugaritic, and ongoing attempts at Linear A and the Indus script; the animal-communication projects (whale and elephant vocalization) push at the boundary of what counts as language at all.
- **The honest limit.** Models are trained on attested text. They can characterize the *space* of possible human languages and the structure common to them; they cannot recover an

unattested origin. What they give is a picture of the *invariants* — and invariants, not ancestors, were the real prize.

The reversal worth naming

The tradition sought a lost object and built a science. AI, oddly, inverts the Babel story: instead of recovering the one tongue, it constructs a shared latent space through which the scattered ones translate. Not restoration — commensuration. Leibniz would have recognized it; Kircher would not.

If you want to push on any strand — the Foucauldian reading of why the search became unthinkable, or the philosophical status of embedding spaces as interlinguae — that's where the interesting argument lives.