EMPIRICAL COMPANION TO Â€ŒTHE ROOT OF LANGUAGE AND THE MAZE INTAKEÂ€

# Measuring the Intake: A Convergence Study of the MAZE Address Gate

J. Konstapel â€” Constable Research, Leiden Â· with Claude Code

**ABSTRACT**

The MAZE is a self-addressing net: the address of a statement is computed from its content, and the intake â€” the gate that turns language into an address â€” therefore determines the structure of the whole. A companion essay argued on historical grounds that the intake had inherited the accidental structure of natural language, and that the root of language is not the word but the discharge event with context. This paper puts that argument to measurement.

We define a *convergence test*: the same content, expressed in two languages, must arrive at the same address after intake. Measured on the running system, the current word-based intake scores **0.03** â€” the language is fully inside the address. We then descend the ladder of intakes and measure each. A concept layer built on open resolution reaches **0.40**, held back not by the idea but by resolution noise; a language-model resolver that normalises to canonical concepts reaches **0.91**; a small, closed, curated stock of act-roots â€” PÄ□á¹‡ini's method â€” reaches **0.96** with perfect discrimination. The lever is resolution quality, not the concept. We show that the fix requires *no re-indexing* of the net, because material is already anchored by its canonical name, and we report the query-side resolver now running. The pure step â€” re-addressing the geometry itself on events â€” is stated as the next programme.

## 1 The question

The intake determines the structure of the MAZE. Whatever survives the gate becomes the palace.

The MAZE stores receipted entries at addresses in one shared address space, and every address is

computed from the content of the entry rather than assigned by hand. This is the property that makes the net verifiable and lets similar things come to rest near one another. It also makes the gate decisive: if the

intake keys on the surface form of words, the net inherits the accidental morphology of whatever language walked through it. The companion essay [1] traced this surface back down a ladder â€” word, act-root, event â€” and argued that the root of language is the discharge event with context, shared across species because it operates in the medium itself, and that the word surface, where the tie to context is cut, is the youngest and most arbitrary rung [2].

That argument was historical. The question here is empirical and narrow: *does the current intake keep the language inside the address, and does descending the ladder measurably remove it?* If the claim is right, the same meaning in two languages should currently land in two different places, and each step down the ladder should pull them together.

## 2 Method

**The convergence metric.** For a pair of sentences with the same content in different languages, we pass each through an intake to obtain its address fingerprint â€” the normalised, hashed vector the net projects onto its axes to choose a route â€” and take the cosine similarity of the two fingerprints. A score of 1 means identical address (full convergence); 0 means orthogonal (the language is still inside). This is the running system's own coder, not a model of it.

**The discrimination metric.** Convergence alone can be faked by an intake that collapses everything together. We therefore also measure unrelated cross-language pairs, and report *separation* = mean(same content) â^’ mean(unrelated). A good intake scores high on same content and near zero on unrelated.

**The intakes.** Five gates were compared: (i) the current *word-hash* â€” tokens and bigrams hashed into a 4096-dimensional signed vector; (ii) a *sound fold* â€” a coarse, language-independent phonetic reduction (after PÄ□á¹‡ini's articulatory ordering [3]); (iii) an *open concept* layer â€” each content word resolved to a language-independent Wikidata identifier, optionally lifted one step up the subclass hierarchy; (iv) a *language-model resolver* â€” a model normalising the sentence to canonical English concept lemmas; (v) a *curated act-root stock* â€” a small closed dictionary mapping words in both languages to shared roots, in the manner of PÄ□á¹‡ini's *dhÄ□tupÄ□á¹ha*.

**Corpus and honesty of scale.** The test set is a small hand-built collection of paired declarative sentences (water, sun, dog, democracy, photosynthesis and a few more) with matched unrelated controls. This is a probe, not a proof: the numbers below are stable and directionally unambiguous, but a production claim requires validation on a large, natural corpus. We state this plainly rather than dress five sentences as a benchmark.

## 3 Results

The current intake keeps the language fully inside the address. Same-content pairs in Dutch and English score a mean cosine of 0.03; the only non-zero cases are accidental shared tokens (the word `water` appears in both). The failure is not even cross-lingual: a Dutch sentence and a Dutch paraphrase of the *same* content â€” different words, one language â€” reach only 0.14. The gate addresses word-forms, not content.

content.

Descending the ladder removes the language, and the amount removed tracks the quality of resolution, not the depth of the idea.

Cross-language convergence and discrimination by intake (0 = language inside, 1 = same address).

| INTAKE | SAME | UNRELATED | SEPARATION |
|---|---|---|---|
| Word-hash (current) | 0.03 | 0.00 | 0.03 |
| Sound fold (blunt) | 0.46 | 0.14 | 0.32 |
| Open concept (Wikidata ID) | 0.40 | 0.00 | 0.40 |
| Language-model resolver | 0.91 | 0.00 | 0.91 |
| Curated act-root stock | 0.96 | 0.00 | 0.96 |

**The sound fold is a deceptive rung.** Its raw same-content score (0.46) looks high, but it also scores unrelated pairs at 0.14: it collapses too much, so distinct things resemble one another. Dutch *hond* and English *dog* do not sound alike; the register catches cognates, not meaning. Sound is a real and ancient layer, but a lower one, and it is not the cure for same-content-different-words.

**The open concept layer is capped by resolution, not by the idea.** Where a noun resolves correctly, both languages land on the identical anchor â€” *water*/*water* â† `Q283`, *hond*/*dog* â† `Q144`, *democratie*/*democracy* â† `Q7174`, *fotosynthese*/*photosynthesis* â† `Q11982`. But a naive top-hit search mis-resolves verbs, adjectives, and even common ambiguous nouns (*sun* returned the wrong entity), and lifting a wrong identifier one step up the hierarchy inherits the error rather than repairing it â€” the score stays at 0.40 either way. The concept idea is sound; the open resolver is noisy.

**Two resolvers clear the noise.** A language-model resolver instructed to output canonical English lemmas reaches 0.91 â€ automatically, without hand-curation. (A first prompt returned the concepts in the source language, scoring 0.07; the failure was instruction-following, not capability, and an explicit â€œEnglish onlyâ€ instruction lifted it to 0.91 â€ a small caution about taking model fluency for competence [6].) A closed, curated stock of just **22 roots** reaches 0.96 with unrelated at 0.00, covering 89% of the content words in the test set. The determinism of curation, not the semantics of the anchor, is what buys the last stretch.

## 4 Interpretation

**The current gate addresses words, not content.** The 0.03 baseline, and the 0.14 for a same-language paraphrase, confirm the companion essay's diagnosis directly: structure was harvested from the surface of language where it should have been constructed from the event beneath it. The reason the running net answers named questions at all is a name bridge â€ a word-layer patch that routes a question to an entity's existing address â€ not the intake.

**The lever is disambiguation.** The jump from 0.40 (open) to 0.91 (model) to 0.96 (curated) is not a

jump in the concept idea; all three use language-independent anchors. It is a jump in how cleanly a word-in-context is mapped to its anchor. This is precisely the problem PÄ□á¹‡ini solved for Sanskrit with a closed, deterministic list of roughly two thousand roots rather than an open lexicon [4], and it is the division of labour that current work on normative and neuro-symbolic systems recommends â€” statistical extraction into the gate, deterministic inference behind it [6].

**The problem shifts from convergence to coverage.** Once a deterministic resolver is in place, convergence is effectively solved (0.96). What remains is the design problem PÄ□á¹‡ini actually faced: choosing a closed stock of roots that *covers* the content. Twenty-two roots covered a handful of sentences; two thousand covered a language; a Wikipedia-scale net needs either a larger stock or an automatic, disambiguating resolver to stand in for one.

**The two halves of the family are on different rungs.** This is the sharpest consequence, and it joins this study to the deontic account of the MAZE-Bank [5]. In the bank, the register â€” the trit â€” already attaches to a *deed*: a directed event of who did what for whom, a closure and its discharge. The bank is therefore already at the bottom rung, the event register, for its domain, and it behaves accordingly. The MAZE-net still addresses *text* and sits at the word surface. The bank works; the net gives strange answers; they are the same ladder, one half descended and one half not. Unifying the family means bringing the net's intake down to where the bank already stands: the register on the event, not the word.

## 5 What we did

**No re-indexing is required.** A natural fear is that changing the intake means re-coding every entry, since a new gate computes new addresses. It does not, for a specific reason: the material is already anchored by its canonical name. Each article carries its canonical title as an index, which is why the name bridge already reaches it without re-indexing. The strange answers are a *query-side* failure â€” the path from a question to the right place â€” not a placement failure.

Accordingly we built the query half now, and only the query half. A resolver normalises a question to canonical concepts and looks them up in the existing name index. It is strictly additive: the fast word bridge runs first and catches named entities most sharply; the concept resolver fires only when the words find nothing, so it can add coverage but never regress what already works, and it costs a model call only on a miss. An early version that ran the resolver first was measured to regress entity questions â€” it placed a generic category ahead of the named thing (*â€œwho painted the Mona Lisaâ€□* â†’ *paint*) â€” and was demoted to the fallback position on that evidence. No entry was moved; the invariant that an address is never rewritten holds untouched.

## 6 What we will do

The query-side resolver is the inexpensive half. The pure step is to bring the geometry of the net itself down the ladder, so that the address *is* the meaning and neighbourhood becomes conceptual neighbourhood â€” walking from one address to a nearby one would then move between related events, not similar spellings. Concretely, a constructed intake would: strip the word surface; map each content word or act, in context, to a concept through a disambiguating resolver; snap that concept to a closed canonical stock of act-roots (which removes the residual synonym variance that separates 0.91 from 0.96); and attach the trit to the resulting event, carrying a source position â€” perceived, dreamt, felt, reported â€” so that both universes

resulting event, carrying a source position â€” perceived, dreamt, felt, reported â€” so that both universes of testimony share one address space [1].

Three tasks stand between here and that net. Build or automate the closed act-root stock, and measure how its size trades against coverage on a natural corpus. Validate convergence at scale, beyond the seven pairs of this probe. And run the two tests this paper did not: the *contour* test â€” whether a single-frequency whistled rendering of a sentence reaches the same address as its written form [7] â€” and the *depth* test against the admissible closure depths. Until then the honest position is the one measured here: the ladder is real, its rungs are numbered, and the gate has, for the first time, a floor older than writing to stand on.

## 7 Limitations

The corpus is small and hand-built; the convergence figure is a cosine over a hashed fingerprint, a faithful proxy for the address but not the address walk itself; the language-model resolver was measured on one run of one model; the contour and depth tests are unrun; and a content-addressing net, however deep its intake, still does not perform relational inference â€” *â€œthe capital of Franceâ€□* resolves to *capital* or *France*, never to *Paris*, and the system says so rather than inventing it. None of these weakens the direction; each marks where the next measurement goes.

## 8 References


[1] **Konstapel, J. (2026). The Root of Language and the MAZE Intake.** The companion essay: the intake determines the structure; the ladder from word to act-root to event; source-marking across the two universes. This paper is its measurement.

[2] **Saussure, F. de (1916). Cours de linguistique gÃ©nÃ©rale.** The arbitrariness of the sign. Read here as a claim true only at the top of the ladder, where abstraction has cut the tie between form and context; at the event layer the sign is not arbitrary.

[3] **Kiparsky, P. â€” on the ÅšivasÅ«tras.** The fourteen-sÅ«tra ordering of sounds is compression-optimal: a constructed intake alphabet in which every natural class is addressable with two symbols.

The model for a gate whose ordering is physical, not accidental.

[4] **PÄ□á¹‡ini, Aá¹£á¹Ä□dhyÄ□yÄ« (ca. 400 BCE), with the dhÄ□tupÄ□á¹ha.** A closed stock of ~2,000 verbal roots and ~4,000 deterministic rules. The existence proof that a curated, closed resolver covers a language â€” and the reason the curated stock, not the open lexicon, clears the resolution noise.

[5] **Konstapel, J. (2026). The Use of Deontic Logic in the MAZE-Bank.** The bank as a sanction-free deontic system in which the trit attaches to a deed. The evidence that one half of the family already addresses events, while the net measured here does not yet.

[6] **â€œAI and Deontic Logic: From Language Models to Normative Intelligenceâ€□**

[6] **â€œAI and Deontic Logic: From Language Models to Normative Intelligenceâ€□ (report, 2026).** The hybrid architecture â€” language model as extractor and resolver, formal system as memory and inference â€” and the caution that surface fluency is not competence, illustrated here by the source-language prompt failure.

[7] **Meyer, J. (2015). Whistled Languages: A Worldwide Inquiry on Human Whistled Speech.** The corpus for the unrun contour test: language reduced to one modulated frequency line and still intelligible â€” a written and a whistled rendering of one sentence should reach one address.