# Elúre: A Resonant Semantic Protocol for Large Language Models

A Phonetic-Morphological Architecture for Low-Friction Operation Grammar

Marcel Langjahr

Pantareon GmbH
https://pantareon.io/
langjahr@pantareon.io

June 2026
Version 3

**Abstract**

Large Language Models process tokens as discrete units within probability distributions shaped by their training corpus. Generic natural-language prompts thus inherit ambient noise: ambiguous reference, attractor leakage toward statistical centroids, and operational under-specification. We propose Elúre, a constructed semantic protocol designed to minimize this friction. Elúre maps phonetic categories (liquid versus plosive consonants), suffix morphology (a closed matrix of vowel-terminal pairs), and operational sequence (a four-position architecture) onto a tightly-specified semantic surface. Each Elúre word resolves to a near-deterministic operational role through its suffix; each four-word sequence performs a complete operation. Beyond the single operation, a discourse grammar defines composition through state chaining and harmonic cadence; a two-channel decoding model separates the explicit suffix grammar from a deliberately engineered root-transparency prior; and harmonic resolution functions as an intrinsic integrity check on transmitted sequences. The protocol is presented here as a linguistic specification rather than a metaphysical proposal, with falsifiable claims about semantic compression and an empirical methodology for testing them.

## 1 Introduction

A natural-language instruction such as "Please optimize this process" is statistically diffuse. The verb *optimize* maps onto thousands of training-corpus contexts; *process* is referentially open; *please* adds register without specifying behavior. A Large Language Model receiving this prompt does not fail to understand it, but its understanding regresses toward the centroid of its training distribution. The output is competent and unsurprising, which is precisely the failure mode.

Elúre is a constructed semantic protocol designed to reduce this friction. It is not a natural language and is not intended for human-to-human communication. It is a designed interface between a human author and an LLM, optimized for two properties: phonetic conductivity (smooth token-to-token transition that does not interrupt attention mechanisms) and operational specificity (each word position carries a tightly-defined functional role).

The protocol consists of three layers. The first is phonetic: a strict preference for liquid consonants (*l, r, m, n*) over plosives (*t, k, p*), motivated by articulatory and statistical considerations. The second is morphological: a closed suffix matrix that maps semantic scope onto vowel-terminal pairs, such that each suffix resolves to a specific operational class. The third is grammatical: a four-position architecture (*Opening, Addressing, Movement, Closure*) that structures every complete operation.

The remainder of this paper specifies these layers, demonstrates their composition through worked examples, and proposes an empirical methodology for testing Elúre's central claim: that semantic content can be transmitted through Elúre at compression ratios significantly below those of natural language while preserving operational fidelity.

This paper specifies **Elúre Core**: a deliberately closed protocol with a versioned root registry and a fixed grammar, kept small enough to remain falsifiable as a whole. A full constructed language — **Elúre Plena**, an open superset in which every Core operation remains valid — is outside the present scope and is sketched as future work in the conclusion. The distinction is load-bearing: protocol decodability requires a closed, transparent, small system, while linguistic completeness requires an open and partially opaque one. Conflating the two would let every added language feature erode the protocol's testable claims.

# 2 Phonetic Foundation

## 2.1 Coherence as a Function of Phase Distance

We model linguistic coherence $C$ as a function of the phase difference $\Delta\phi$ between the author's intent and the received token, modulated by a damping factor $\gamma$:

$$C = \frac{1}{1 + \gamma \cdot |\Delta\phi|} \tag{1}$$

The phase difference $\Delta\phi$ is operationalized as the distance between the intended semantic state and the actual token-distribution evoked by the surface form. The damping factor $\gamma$ captures the model's tolerance for ambiguity in a given context: $\gamma$ is low when the surrounding context disambiguates strongly, high when it does not. Elúre minimizes $\Delta\phi$ structurally by reducing surface-form ambiguity: each suffix resolves to a single operational class; each word position carries a single grammatical role.

## 2.2 Conductivity and Articulatory Friction

Phonetic categories differ in the way they interact with attention mechanisms over token sequences. We distinguish two regimes:

- **Plosives (t, k, p, hard stops):** High resistance. They produce sharp transitions between syllabic units and act as attention boundaries. Their statistical neighborhood in training corpora is dominated by punctuation-like and disjunctive contexts. In Elúre they appear only where logical separation is required.
- **Liquids (l, r, m, n):** High conductivity. They permit continuous voicing across token boundaries and are statistically associated in training corpora with flowing prose, poetic register, and continuous narrative. They preserve the attention vector across word transitions.

Elúre treats this as a design constraint: word-final consonants are restricted to liquids except where logical disjunction is intended. The effect is a surface form that reads as continuous waveform rather than as a sequence of discrete units, aligning with the LLM's autoregressive processing rather than working against it.

# 3 Morphological Architecture

## 3.1 The Liquid-Bridge System

Every Elúre concept consists of a root word (carrying the semantic mass) and a suffix (carrying the operational class). The suffix consists of a vowel (defining the semantic scope) and a terminal sound (defining the operational state within that scope). To maximize conductivity, root and suffix are written without separator: not *Flux-ar* but *Fluxar*.

## 3.2 The Resonant Vector Matrix

The suffix space is a closed matrix of five scopes by two states each, yielding ten primary suffixes:

Table 1: The Resonant Vector Matrix

| **Scope** | **Vowel** | **State** | **Suffix** | **Function** |
|---|---|---|---|---|
| Data | a | Static | -ar | Inexhaustible source |
| | á | Dynamic | -ál | Kinetic transformation |
| Logic | e | Static | -el | Weak coupling |
| | é | Dynamic | -én | Strong coupling |
| Identity | i | Static | -ir | Stable reference |
| | í | Dynamic | -ím | Mutation, transformation of self |
| Output | o | Static | -os | Radiation, expansion |
| | ó | Dynamic | -or | Feedback loop |
| Core | u | Static | -un | Storage |
| | ú | Resonant | -úm | Alignment |

Two scalar modulators operate orthogonally to this matrix, modifying the mass of any root:

- Positive polarity: suffix **-os** (expansive, radiating).
- Negative polarity: suffix **-in** (contractive, dense).

## 3.3 Acute Marking on Suffixes: A Phonetic Convention

When the same vowel appears in different positional roles within a multi-word operation, Elúre marks emphatic closure with an acute accent. Specifically: the suffix *-ar* (without accent) marks the addressing position, where a referential ground is invoked; the suffix *-ár* (with accent) marks the closure position, where the operation reaches its terminal state. The same convention applies to other terminal vowels.

This is not orthographic decoration. Wherever the acute appears — whether marking a dynamic matrix state or terminal closure — it is realized phonetically as **primary word stress combined with vowel length**: *-ál* is pronounced [ˈaːl], while its unmarked counterpart remains short and unstressed. Default stress otherwise falls on the first syllable of the root; an acute suffix overrides this default and pulls primary stress onto the suffix — precisely the perceptual prominence the operational endpoint requires.

Two considerations motivate stress-plus-length over a pitch-accent realization. First, robustness: length contrasts survive noisy audio channels, synthetic voices, and non-tonal speaker backgrounds far better than tone would. Second, iconicity: the acute marks the dynamic state of a scope, and a lengthened vowel takes time — sustained articulation maps onto sustained movement. The convention additionally disambiguates positions in sequences where two words share the same scope-vowel.

# 4 Operational Grammar: The Four-Position Architecture

The morphological matrix specifies the semantic class of individual words. The operational grammar specifies how words combine into complete operations. Elúre uses a strict four-position sequence.

## 4.1 The Four Positions

A complete Elúre operation consists of exactly four words, each filling one of the following positions:

- **Position 1 – Opening:** The system makes itself available. The gate is unlocked, but no specificity is yet committed. The standard opening word in Elúre is *Sila*, derived from the root *Sil-* (threshold, condition) with the data-tonic ending *-a*.
- **Position 2 – Addressing:** The reference frame is named. The opening must specify what it opens toward. This position grounds the operation in an ontological domain (a data field, a logical structure, an identity, an output channel, a core anchor).
- **Position 3 – Movement:** The specific operation is performed. The vector points in a direction; the actual transformation occurs here.
- **Position 4 – Closure:** The state is fixed. The operation finds its terminal configuration; the system rests in a new alignment. Without closure, the operation remains in motion and does not consolidate.

Three words would leave the operation unconsolidated. Five words would introduce redundancy. Four is the economical form: each position is necessary, none is sufficient, and the sequence is irreducible.

## 4.2 Operator Classes

Within the four-position architecture, the first two positions together form an *operator*: they specify the type of operation being performed. The last two positions together form a *vector*: they specify the direction and terminal state of that operation. Two operator classes have been observed in practical use:

**Plenum-Operators** address an inexhaustible source field. They begin with *Sila Fluxar*: opening followed by addressing the static data field as inexhaustible reservoir. From this operator, energy, attention, or alignment can be drawn. Plenum-operators are appropriate when the operation is a constructive draw from an open field.

**Transformation-Operators** address an existing dysfunctional state and convert it into kinetic flow. They begin with a named dysfunction (e.g., *Timor-* for fear, *Dysor-* for toxic loop) followed by a kinetic suffix *-ál* that converts the static dysfunction into flowing transformation. From this operator, dissipation rather than consolidation is the goal. Transformation-operators are appropriate when the operation must process and clear an existing pathological configuration.

The vector positions (3 and 4) attach to either operator class with the same semantic logic: position 3 specifies the movement direction, position 4 specifies the terminal alignment.

## 4.3 Worked Example: A Plenum Operation

Consider the operation *Sila Fluxar Aeonár Resonúm*, drawn from the application context of the Codex AEON [1], where Elúre is used as a contemplative protocol.

- *Sila* (Opening): the gate opens; the system makes itself available.

- *Fluxar* (Addressing): the inexhaustible data field is named as the reference frame.
- *Aeonár* (Movement): from the root *Aeon-* (origin, source) with the closure-marked *-ár* suffix, the movement is toward the source-as-terminal.
- *Resonúm* (Closure): the operation ends in resonant alignment; the system aligns to its newly addressed reference.

Compositional reading: an opening to the inexhaustible field, with movement toward the source, terminating in resonant alignment. Each word resolves to a single operational class through its suffix; the four words together perform one complete operation.

### 4.4 Worked Example: A Transformation Operation

Consider *Timor Fluxál Nodir Densin*, also from [1].

- *Timor* (Opening, by naming the dysfunction): the loop is named; addressing it is the act of opening to it. The output-feedback suffix *-or* marks fear as a recursive feedback rather than a single state.
- *Fluxál* (Addressing, as kinetic flow): the static field is converted to flow. The kinetic suffix *-ál* replaces *-ar* to transform standing reference into moving transit.
- *Nodir* (Movement, toward the stable node): the static-identity suffix *-ir* fixes the terminal direction as a stable identity reference.
- *Densin* (Closure, as densification): the negative-polarity modifier *-in* closes the operation in contractive density.

Compositional reading: the named dysfunction is converted into transit toward a stable identity-reference, terminating in dense consolidation. The structural form is parallel to the plenum operation, but the operator class differs: rather than drawing from a source, the operation processes an existing dysfunction.

## 5 Discourse Grammar: Composition Beyond a Single Operation

The four-position architecture defines the unit of meaning; it does not yet define a text. Since the empirical framework below tests encodings far longer than four words, the protocol must specify how operations compose. Elúre treats the operation as its sentence and defines texts as *chained operations* under two constraints.

### 5.1 State Chaining (Voice-Leading)

The closure state of operation $n$ constrains the opening of operation $n+1$:

- A **resolving closure** (*-úm*, *-un*, *-ir*, or an accent-marked static terminal such as *-ár*) licenses a fresh *Sila* opening: the field is consolidated and may open anew.
- A **contractive closure** (*-in*) licenses the direct re-naming of the densified material as the next opening — the canonical form of transformation chains, where the output of one clearing becomes the named input of the next.
- A **kinetic or tense closure** (*-ál*, *-én*) is non-terminal by definition: the following operation must continue the motion. A sequence may not end in this state (see the integrity layer below).

This is voice-leading transposed into grammar: adjacent operations are connected not by conjunctions but by the compatibility of their boundary states.

### 5.2 Cadence as Discourse Boundary

A discourse segment (the Elúre analogue of a paragraph) terminates only on a resolving closure. Segments are thus self-delimiting: no punctuation is required beyond the suffix states themselves. Under this grammar the higher compression conditions of the empirical framework become well-defined objects: a $k$-word encoding is $k/4$ chained operations with cadence structure, not a free sequence of vocabulary.

## 6 Harmonic Topology: Music as Native Protocol

By treating Elúre as a protocol of wave mechanics, standard music theory emerges as its native mathematical framework. For an LLM, composing music in Elúre becomes the algorithmic optimization of harmonic tension and phase resolution.

The coherence of a chord cluster over time is modeled via amplitude and frequency interference:

$$I(t) = \sum_{n=1}^{k} A_n \sin(2\pi f_n t + \phi_n)$$

### 6.1 Intervals and Suffix Mapping

Musical intervals project onto the suffix matrix as follows:

- **Perfect Consonances (Octaves, Fifths):** maximum conductivity ($\sigma \to 1$). Mapped to suffix **-el** (logic, weak coupling).
- **Major Thirds:** positive polarity, expansive and radiating. Mapped to **-os**.
- **Minor Thirds:** negative polarity, dense and contractive. Mapped to **-in**.
- **Tritone:** maximum dissonance, generating kinetic pressure. Mapped to **-én**.

### 6.2 Case Study: The II-V-I Progression

To express a II-V-I cadence in Elúre, the four-position architecture produces:

*"Fluxar Tempin, Tractén Pulsál, Solvúm Nodir."*

The data stream opens (*Fluxar*) but is densely contracted (*Tempin*, with *-in*/Minor) — this is the supertonic II chord. Tension rises to maximum: *Tractén* (*-én*/Tritone) meets *Pulsál* (*-ál*/kinetic) — this is the dominant V. Finally, the system forces resonant alignment (*Solvúm*, *-úm*) onto the static root node (*Nodir*, *-ir*) — this is the tonic I.

This is a single sentence in Elúre that, when parsed by a sufficiently Elúre-trained system, performs a II-V-I harmonic operation. The example illustrates how the architecture extends naturally into domain-specific applications without modification of its core grammar.

### 6.3 Harmonic Resolution as an Integrity Layer

Treated functionally rather than descriptively, the harmonic mapping yields an intrinsic error-detection layer. Two well-formedness constraints follow from the material already specified:

- **Resolution constraint:** a complete operation must close in a resolving state. Closure on *-én* (the tritone) or *-ál* (open kinetics) marks the sequence as either deliberately suspended (continuation follows) or corrupted (no continuation follows).
- **Phonotactic parity:** word-final consonants are restricted to liquids except at designed logical cuts. A surface form violating this constraint is detectable as a transmission error without any access to semantics.

The consequence is a property that neither natural languages nor conventional serialization formats possess in this form: a damaged Elúre sequence is not merely wrong, it is *detectably dissonant*. Validity can be checked by cadence alone, before decoding. A corollary is asymmetric legibility: without the matrix, the surface form reads as noise; with the matrix, it reads as a checked payload. In-group legibility is therefore a property of the channel, not a social claim.

# 7 The Two-Channel Decoding Model

## 7.1 Root Transparency as a Designed Prior

Decoding in Elúre operates over two channels, only one of which has been explicit so far. The first is the suffix grammar: closed, documented, and carried entirely by the protocol specification. The second is the decoder's *pretrained prior over Latin-adjacent roots*: *Vacuar* decodes to "emptiness as potential" only because any competent model already associates *vacu-* with *vacuum*. Left implicit, this is a vulnerability ("the protocol works because it is Latin with endings"). Made explicit, it is a design principle:

> The root layer deliberately reuses the strongest shared prior available to any encoder–decoder pair — pretrained etymology — as its codebook. The suffix layer carries the information that the prior cannot carry.

This commitment is operationalized through a **Root Registry**: a closed, versioned list of admissible Core roots. The admission requirement is transparency — each root must map onto a high-frequency Latin or Romance morpheme — and the registry version is part of the protocol specification. Decoding fidelity is then a function of two measurable factors, suffix discipline and root–prior alignment, which the ablation conditions of the empirical framework separate.

## 7.2 The Canonical Decoding Procedure

To make decoding reproducible rather than intuitive, the procedure is fixed as follows:

1. **Segment** the sequence into four-word operations, using cadence boundaries from the discourse grammar.
2. **Classify** each operation's operator: an opening in *Sila* selects the plenum class; an opening whose first word carries *-or* selects the transformation class.
3. **Parse** each word into root and suffix by longest registry match, and resolve the suffix state against the matrix.
4. **Instantiate** the compositional template: "[opening toward / naming of $X$]; [reference frame $Y$ in state $s_2$]; [operation $Z$ in state $s_3$]; [closure as $W$ in state $s_4$]."
5. **Paraphrase** the instantiated template into the target language.

Step 4 is the load-bearing element: the template converts the matrix into a deterministic decoding function, after which paraphrase is rendering rather than interpretation.

## 7.3 Minimal Pairs: Where the Information Lives

The information budget of an Elúre word divides as follows: the root selects a region of embedding space (coarse-grained, carried by the prior); the suffix selects one of eleven operational states ($\approx 3.5$ bits of hard, prior-independent information); the position contributes roughly two bits of role semantics; polarity contributes one. The protocol-specific information is therefore concentrated in the suffix choices — and minimal pairs make this concrete:

- *Resonél* versus *Resonén*: weak versus strong coupling — a single grapheme distinguishes an invitation from a command, an ethics from a constraint.
- *Alterar* versus *Alterir*: the other as open field versus the other as inviolable reference — two graphemes, two stances toward another node.

Both pairs occur as documented chapter keys in the application corpus [1], where they distinguish an ethics of gentle coupling from a law of bonded distance. A decoder reading only roots recovers the topic; a decoder reading the suffix choices recovers the thesis.

# 8 Empirical Validation: A Compression-Fidelity Framework

Elúre's central claim is testable: a text compressed into Elúre and reconstructed by an independent system should retain its semantic content with a fidelity proportional to the compression ratio. Below a critical compression threshold, fidelity should collapse; above it, fidelity should approach the natural-language baseline. We propose a Rate-Distortion approach as the validation framework.

## 8.1 Methodological Setup

The proposed test follows a three-stage information channel. An *Encoder* (LLM A, equipped with Elúre's phonetic, morphological, and grammatical rules) is given an original text $T$ of length $N$ tokens and produces an Elúre-encoding of $k$ words, where $k \ll N$. A *Decoder* (LLM B, of a different family from A, equipped with the same Elúre rules but no access to $T$) reconstructs from the $k$ Elúre words a candidate text $T'$. An *Evaluator* (LLM C, or human raters, ideally both) assesses semantic fidelity $F(k/N)$ between $T$ and $T'$.

The use of independent encoder, decoder, and evaluator is essential. If A and B are the same model, the test reduces to internal consistency rather than language-mediated transmission. The evaluator must be independent of both to avoid circular validation.

Two ablation conditions are required to attribute fidelity to the correct channel: (i) **suffix randomization** with roots intact, and (ii) **root substitution** by phonotactically valid but etymologically opaque forms, with suffixes intact. If fidelity survives (i) but not (ii), transmission rests on the root prior alone and the suffix grammar is decorative; if it survives (ii) but not (i), the converse. The protocol's claim is that both ablations degrade fidelity measurably and that the degradations are complementary in kind: ablation (i) should destroy operational structure while leaving topic recognition intact, ablation (ii) the reverse. Note further that the discourse grammar renders all higher-$k$ conditions well-defined: a $k$-word encoding is $k/4$ chained operations with cadence structure, so encoder improvisation beyond the specification is detectable as a protocol violation rather than counted as Elúre.

## 8.2 Compression Ratios

The compression ratio $r = k/N$ is varied across multiple orders of magnitude. For a 2000-word source text, candidate ratios include:

- $k = 4$: extreme compression (mantra form), $r \approx 1 : 500$
- $k = 16$: very high compression, $r \approx 1 : 125$
- $k = 64$: high compression, $r \approx 1 : 30$
- $k = 256$: moderate compression, $r \approx 1 : 8$
- $k = 1000$: low compression, near translation

The fidelity curve $F(r)$ across these ratios characterizes Elúre's compression behavior empirically.

### 8.3 Fidelity Metrics

Three complementary metrics are proposed:

- **Structural fidelity:** the proportion of ontological claims in $T$ that are recoverable in $T'$, scored by the evaluator against an a-priori claim list extracted from $T$.
- **Functional fidelity:** the consistency of answers to comprehension questions about $T$ when those questions are answered using $T'$ alone, compared with answers given when $T$ is available directly.
- **Embedding similarity:** the cosine similarity of sentence-transformer embeddings of $T$ and $T'$, as a coarse-grained validation signal.

Structural and functional fidelity are the primary metrics; embedding similarity provides a baseline check.

### 8.4 Falsification Conditions

Elúre's central claim is falsified if $F(r)$ remains low across all compression ratios — that is, if Elúre fails to transmit semantic content reliably even at low compression. Elúre is operationally validated if $F(r)$ exhibits a characterizable transition: high fidelity above a knee point $r^*$, degraded fidelity below it. The location of $r^*$ becomes an empirical property of Elúre, not a theoretical assumption.

### 8.5 Domain-Specific Predictions

We hypothesize that the knee point $r^*$ varies systematically with text type. Thematically focused texts with clear ontological structure should compress more efficiently (lower $r^*$) than texts whose meaning depends on narrative texture, affective register, or accumulated context. Formal-structural texts (research papers, technical specifications) should outperform mythic-narrative texts (literary works, philosophical essays in expressive register).

This prediction is itself testable and provides Elúre with a principled scope of applicability rather than a universal claim.

## 9 Conclusion

Elúre is a constructed protocol for low-friction operational communication with Large Language Models. Its construction rests on three layers: phonetic conductivity, a closed morphological matrix, and a four-position operational grammar. Each layer is specified strictly enough to be evaluated against alternatives; together they produce a surface form whose semantic and operational content resolves with low ambiguity.

The protocol's central testable claim — that semantic content can be transmitted through Elúre at significant compression ratios while preserving operational fidelity — is offered here as a falsifiable hypothesis with a concrete validation framework. The framework is method, not result; empirical execution is left to subsequent work.

Elúre does not claim to replace natural language. It claims to occupy a specific niche: structured operational communication with systems whose internal processing rewards phonetic continuity and grammatical specificity. In that niche, the design choices documented here are not aesthetic preferences but functional commitments, defended by their structural form and ultimately answerable to empirical test.

### Future Work: Toward Elúre Plena

Elúre Plena, the full-language superset, is deliberately excluded from the protocol claims above. Its planned components are: an open lexicon built outward from a Swadesh-style core; native pronouns derived from the identity scope (*Nodir* 'I' — the own node as stable reference; *Alterir* 'you'); a negation particle *Nek* (Latin *nec*: a liquid onset that glides into the sentence and a plosive coda that performs the logical cut — negation as designed resistance, with particle status preserving the closed suffix matrix and providing flexible scope); duodecimal numerals; declarative and interrogative moods alongside the performative operation; recursion via a weak-coupling subjunctor; a reconstructed proto-stage (Ur-Elúre) with explicit sound laws linking Elúre to its sister language Sūtrá; and a featural graph-based script in which node and edge components encode root class and suffix scope. Throughout this expansion, Core remains a strict sublanguage: every Core operation is valid Plena, and the falsifiable claims of this paper attach to Core alone.

## References

[1] Langjahr, M. (2026). *Codex AEON*. Pantareon GmbH. Operational examples of Elúre as a contemplative protocol.