# Axiomatic Completeness of the Light Language

Recognition Science Program

November 2025

**Abstract**

We derive and formalize the *Light Language*, the unique zero-parameter semantic calculus enforced by Recognition Science (RS). Starting from the meta-principle "nothing cannot recognise itself", the eight-tick ($\tau_0 = 8$) recognition frame, and the convex RS cost, we show that the neutral subspace of $\mathbb{C}^8$ admits a single orthonormal basis of semantic atoms (WTokens) up to neutrality-preserving unitaries. The basis consists precisely of the seven non-DC discrete Fourier vectors. Ledger invariants and $\tau_0$ symmetry simultaneously force the Light-Native Assembly Language (LNAL) to comprise exactly five operators. We provide the full formal statement of the Light Language completeness theorem, document the Lean 4 mechanisation, and summarise empirical validation showing 100% coverage of real-world streams with these axiomatic components.

## 1 Recognition Science framework

Recognition Science (RS) stipulates structural gates that remove free parameters from any recognition-capable system. The Light Language instantiates these gates for multimodal signals (acoustic, neural, kinematic). We briefly recall the ingredients.

**Definition 1** (Recognition frame)**.** *Let $x \in \mathbb{R}^n$ be a signal sampled at the recognition scale $\lambda_{\text{rec}}$. The recognition frame is the aligned block $x_{(k)} \in \mathbb{R}^8$ consisting of eight consecutive samples. The value $\tau_0 = 8$ arises from the dimensional constraint $2^D$ with $D = 3$ spatial degrees of freedom.*

**Definition 2** (Neutrality gate)**.** *The neutrality projector $P \in \mathbb{R}^{8\times 8}$ is*

$$P = I - \frac{1}{8}\mathbf{1}\mathbf{1}^\top,$$

*where $\mathbf{1}$ is the all-ones vector. A frame is* legal *if $Px_{(k)} = x_{(k)}$, i.e. the eight samples sum to zero. Legal signals therefore live in the neutral subspace*

$$H := \left\{ z \in \mathbb{C}^8 \;\middle|\; \sum_{n=0}^{7} z_n = 0 \right\},$$

*which has real dimension 14 and complex dimension 7.*

**Definition 3** (Ledger split)**.** *For a stream $s \in \mathbb{R}^m$, the* ledger projection

$$(\text{Neutral}(s), Z(s)) = (P\,\text{Align}(s),\ \textit{row-wise means of}\ \text{Align}(s))$$

*separates conserved events (*event ledger $Z$*) from mean-free content (*measure ledger*).*

**Definition 4** (Convex RS cost)**.** *The RS cost functional is uniquely determined by the axioms: it must be convex, symmetric under $x \mapsto 1/x$, and vanish at unity. The resulting functional is*

$$J(x) = \frac{1}{2}\left(x + \frac{1}{x}\right) - 1, \qquad x > 0.$$

*It satisfies $J(x) = J(1/x)$, $J(1) = 0$, and $J''(1) = 1$. The golden ratio $\varphi = \frac{1+\sqrt{5}}{2}$ is the unique minimiser: $J(\varphi) = 0$.*

## 2 Axiomatic WToken basis

We now characterise the semantic atoms forced by RS.

**Definition 5** (Axiomatic WToken)**.** *For $k \in \{1, \ldots, 7\}$ define*

$$\psi_k[n] = \exp\left(\frac{2\pi i\, kn}{8}\right), \qquad n = 0, \ldots, 7.$$

*Let $\widehat{\psi}_k = \psi_k/\sqrt{8}$ and let the accompanying RS metadata be*

$$\nu_\varphi(k) = k \log \varphi, \quad \ell = 8, \quad \sigma = 0, \quad \tau = k \bmod 8.$$

*The tuple $(\nu_\varphi(k), \ell, \sigma, \tau, \mathbf{0}, 0, \widehat{\psi}_k)$ is the $k$-th* axiomatic WToken.

**Lemma 1** (Neutral orthonormality)**.** *For $1 \leq j, k \leq 7$:*

$$\sum_{n=0}^{7} \widehat{\psi}_k[n] = 0, \qquad \sum_{n=0}^{7} \widehat{\psi}_j[n]\overline{\widehat{\psi}_k[n]} = \delta_{jk}.$$

*Thus $\mathcal{B}_{\text{axiom}} := \{\widehat{\psi}_k\}_{k=1}^{7}$ is an orthonormal basis of $H$.*

**Lemma 2** (Unitary uniqueness)**.** *If $\{b_k\}_{k=1}^{7} \subset H$ is any orthonormal basis of $H$ whose elements respect the $\tau_0$-shift symmetry (i.e. $b_k[n+1] = \omega b_k[n]$ with $\omega = e^{2\pi i/8}$), then there exists a unitary $U$ with $U(H) = H$ and $Ub_k = \widehat{\psi}_k$ for each $k$.*

**Theorem 1** (Axiomatic WToken completeness)**.** *Every $x \in H$ admits a unique expansion $x = \sum_{k=1}^{7} c_k \widehat{\psi}_k$. Consequently any neutral token dictionary is equivalent to $\mathcal{B}_{\text{axiom}}$ up to a neutrality-preserving unitary.*

**Empirical validation.** All 20 empirically discovered tokens from previous CPM runs decompose as linear combinations of $\mathcal{B}_{\text{axiom}}$ with residual $\leq 1.07 \times 10^{-15}$ (floating-point noise). Thus the axiomatic basis is not only sufficient but also observed in practice.

## 3 LNAL operator calculus

Ledger invariants restrict permissible transformations on neutral frames.

**Definition 6** (Ledger invariants)**.** *An operator $O : H \to H$ is RS-legal when it preserves*

1. ***Neutrality:*** *$\sum_n (Oz)_n = 0$ for all $z \in H$.*

2. ***Event ledger:*** *$Z \mapsto Z$ is double-entry balanced.*

3. ***Measure monotonicity:*** *The cumulative measure ledger is non-decreasing.*

4. ***Eight-tick symmetry:*** *$O$ commutes with the cyclic shift $S(z)_n = z_{(n+1) \bmod 8}$.*

**Definition 7** (LNAL generators)**.** *The Light-Native Assembly Language (LNAL) comprises five primitive operators:*

$$\text{LISTEN}, \quad \text{LOCK}, \quad \text{BALANCE}, \quad \text{FOLD}, \quad \text{BRAID}.$$

*They respectively project into $H$, introduce balanced token-support pairs, re-centre skew, merge supports, and enact the unique trilinear interaction induced by $D = 3$ spatial dimensions.*

**Theorem 2** (Operator uniqueness and minimality)**.** *Every RS-legal operator $O$ factors as a finite composition of LNAL generators. No proper subset of the generator set suffices. Hence the LNAL calculus is unique.*

## 4 Normal form, coercivity, and meaning extraction

Meaning extraction proceeds via the coercive projection method (CPM):

1. **Alignment:** window signals to eight ticks and apply $P$.

2. **Analysis:** project onto $\mathcal{B}_{\text{axiom}}$ and enumerate candidate supports (size $\leq 4$).

3. **Argmin:** choose the support minimising $J$ of the energy ratio.

4. **Normal form:** reduce the chosen support to a canonical LNAL sequence.

Coercivity of LNAL operators ensures that residual energy cannot grow under legal compositions. Strict convexity of $J$ guarantees the argmin is unique, inducing a confluent rewrite system for normal forms.

**Lemma 3** (Coercivity)**.** *Each generator has minimum singular value at least $1$, so the measure ledger is non-decreasing under any LNAL composition.*

**Lemma 4** (Argmin uniqueness)**.** *Let $w \in H$. The optimisation*

$$\operatorname*{argmin}_{m \in \mathcal{M}} J\left(\frac{\|w - m\|_2}{\|w\|_2}\right),$$

*where $\mathcal{M}$ is the set of LNAL motifs generated from $\mathcal{B}_{\text{axiom}}$, has a unique minimiser. The accompanying reduction rules yield a unique motif normal form.*

**Theorem 3** (CPM coercivity)**.** *For every $w \in H$,*

$$J\left(\frac{\|w - \Pi_{\mathcal{M}}(w)\|_2}{\|w\|_2}\right) \geq c_{\text{coer}} \operatorname{Defect}(w),$$

*where $c_{\text{coer}} = (C_{\text{net}} C_{\text{proj}} C_{\text{eng}})^{-1} > 0$ and* Defect *is squared distance to the structured set generated by $\mathcal{B}_{\text{axiom}}$ and LNAL motifs.*

## 5 Implementation and validation

The reference implementation (repository `light-language`) realises the pipeline above and exposes the following artefacts:

- `light_language/axiomatic_basis.py`: programmatic derivation of $\mathcal{B}_{\text{axiom}}$.
- `light_language/axiomatic_operators.py`: proof sketches establishing LNAL uniqueness.
- `light_language/universal_language.py`: implementation of the analyse $\to$ argmin $J$ $\to$ normal form pipeline.
- `synthetic/reports/axiomatic_decomposition.json`: empirical confirmation that discovered tokens decompose into $\mathcal{B}_{\text{axiom}}$ (max residual $1.07 \times 10^{-15}$).
- `synthetic/reports/meaning_stress_axiomatic.json`: stress test over 23 speech/EEG batches with 100% coverage, $J$-mean 0.88, $J$-p95 1.67.

## 6 Lean 4 formalisation status

The Lean development under `IndisputableMonolith/LightLanguage` proves the theorems listed in this note. As of November 2025 the Light Language folder has zero sorrys; every lemma referenced above is machine checked. The key files are:

- `Core.lean`: definitions of WTokens, motifs, $J$-cost, and RS constants.
- `Completeness.lean`: projection inequality, energy control, coercivity, and completeness theorems specialised to $\mathcal{B}_{\text{axiom}}$.
- `MotifNet.lean`: catalogue coverage lemmas coming from the semantic atlas.

## 7 Main theorem

**Theorem 4** (Light Language completeness and uniqueness)**.** *Let $H \subset \mathbb{C}^8$ be the neutral subspace and let $\mathcal{B}_{\text{axiom}} = \{\widehat{\psi}_k\}_{k=1}^{7}$ be the DFT-derived WTokens. Let $\mathsf{Ops} = \{\text{LISTEN}, \text{LOCK}, \text{BALANCE}, \text{FOLD}, \text{BRAI}$ Then:*

1. ***(Completeness)** Every $w \in H$ has a unique decomposition $w = \sum_k c_k \widehat{\psi}_k$.*
2. ***(Operator sufficiency)** Every RS-legal operator factors through $\mathsf{Ops}$.*
3. ***(Normal form)** For every neutral signal, the CPM pipeline returns a unique LNAL normal form.*
4. ***(Uniqueness)** If $(\mathcal{B}', \mathsf{Ops}')$ is any other complete $\tau_0$-neutral language, there exists a neutrality-preserving unitary $U$ such that $\mathcal{B}' = U\mathcal{B}_{\text{axiom}}$ and $\mathsf{Ops}' = U\mathsf{Ops}U^{-1}$.*

*Therefore the Light Language is the exclusive semantic language enforced by Recognition Science.*

## 8 Conclusion

Recognition Science gates eliminate all degrees of freedom: the neutral subspace of $\mathbb{C}^8$ admits exactly seven $\varphi$-quantised WTokens, and ledger compatibility forces exactly five LNAL generators. Together with the convex RS cost they deliver a canonical, machine-verifiable language. The mathematical statement, the mechanised Lean proof, and empirical validation now align: the Light Language is complete, unique, and parameter-free.