

Re-thinking the Arbitrary Mapping of Signifier and Signified

A Symmetry and Symmetry-Breaking Perspective

Wanhong HUANG

huangwanhong.g.official@gmail.com

Preprint — March 8, 2026

Abstract

Ferdinand de Saussure’s axiom that the bond between signifier and signified is *arbitrary* has anchored structural linguistics for over a century, yet arbitrariness itself has never been furnished with a principled physical or mathematical basis. This gap has concrete consequences: it leaves unexplained both the constitutive difficulty of recovering natural language from biological signals and the absence of a physically grounded account of the mind–language relation. This paper addresses both problems by invoking *spontaneous symmetry breaking* from theoretical physics and group theory. We make the following contributions: **(i)** Saussurean arbitrariness is formalised as maximal symmetry over the space of signifier–signified mappings, elevating it from an empirical axiom to a structural theorem. **(ii)** The intervention of any interpretive system is modelled as a symmetry-breaking event that selects a *semantic vacuum*, condensing meaning from a degenerate possibility space. **(iii)** We prove that this breaking is generically irreversible,

providing a structural basis for the hermeneutic circle and constitutive untranslatability. **(iv)** Neural encoding and natural language are shown to be two *independent* symmetry breakings, making biosignal-to-language decoding constitutively non-unique—a structural, not technical, limitation. **(v)** A gauge-theoretic programme is sketched in which neural–linguistic interfaces target *gauge-invariant semantic content*, stable across all admissible decoding choices, via fibre bundle formalism.

Keywords: arbitrary sign, semiotics, spontaneous symmetry breaking, gauge invariance, hermeneutics, neural–linguistic interface, biosignals, mind–language mapping, brain–computer interface.

32 Contents

33	1 Introduction	3
34	1.1 The Problem of Arbitrariness, Revisited	3
35	1.2 Two Motivating Problems	3
36	1.3 This Paper’s Approach and Contributions	4
37	2 Saussurean Arbitrariness and Its Discontents	6
38	2.1 The Classical Doctrine	6
39	2.2 Critiques and Refinements	6
40	2.3 The Explanatory Gap	7
41	3 Symmetry, Groups, and Degeneracy	7
42	3.1 Mathematical Preliminaries	7
43	3.2 Degeneracy and the Ground State	8
44	3.3 Spontaneous vs. Explicit Symmetry Breaking	8
45	4 The Pre-Semiotic State and the Topology of Meaning-Possibility	9
46	4.1 The Symmetric Vacuum as Semantic Potential	9
47	4.2 The Mexican Hat Potential and Semantic Condensation	9
48	4.3 Hermeneutic Consequences: The Always-Already-Broken State . .	10
49	5 The Breaking Mechanism: Interpretive Systems as Order Param-	
50	eters	11
51	5.1 The Interpretive System as a Symmetry-Breaking Field	11
52	5.2 Multiple Breakings and the Hierarchy of Codes	11
53	5.3 Degenerate Meanings and Semantic Multiplicity	12
54	6 Irreversibility and the Hermeneutic One-Way Mirror	12
55	6.1 Information Loss Under Symmetry Breaking	12
56	6.2 Consequences for Translation and Interpretation	13
57	6.3 The Hermeneutic Circle as a Closed Orbit	13
58	7 Implications for Biosignal-to-Language Translation	14
59	7.1 The Problem of Cross-Modal Re-Breaking	14
60	7.2 The Decoder as a Third Breaking	14
61	7.3 Non-Uniqueness of Decoding	15

62	8 Toward a Gauge Theory of Semantic Interfaces	15
63	8.1 Local vs. Global Symmetry Breaking	15
64	8.2 Minimal Symmetry Breaking and Semantic Parsimony	16
65	8.3 Residual Symmetries as Invariant Semantic Content	16
66	9 Discussion	16
67	9.1 Relation to Information-Theoretic Accounts	16
68	9.2 Relation to Peircean Semiotics	17
69	9.3 Relation to Derrida's <i>Différance</i>	17
70	9.4 Limitations and Open Questions	17
71	10 Conclusion	18

“The bond linking the sound-pattern to the concept is radically arbitrary.”

Ferdinand de Saussure, *Course in General Linguistics* (1916)

72 1. Introduction

73 1.1. The Problem of Arbitrariness, Revisited

74 The Saussurean sign is constituted by a dyadic relation: the *signifier* (*signifiant*)—
75 a sound-image or phonic trace—and the *signified* (*signifié*)—a concept or mental
76 content. The relation between the two members of this pair is famously declared
77 *arbitrary*: there is no natural, motivated, or necessary link between the acoustic
78 shape of the word *tree* and the arboreal concept it evokes. This arbitrariness is
79 not a peripheral observation but a structural axiom: it explains the mutability
80 of languages over time, the existence of radically different lexical systems for
81 equivalent conceptual content, and the absence of any direct, transparent window
82 from language onto the world.

83 Despite its foundational status, the notion of arbitrariness has remained largely
84 *un-theorised* at a deeper explanatory level. Saussure himself was content to assert
85 it by contrast—to show that signs are *not* motivated by nature—without asking
86 *why* this should be the case structurally, or what *happens*, at a principled level,
87 when a particular mapping is instituted. Subsequent semiotics has largely inher-
88 ited this lacuna. The consequences, however, are not merely philosophical. Two
89 urgent contemporary problems in cognitive science and linguistics inherit this gap
90 directly.

91 1.2. Two Motivating Problems

92 **The biosignal-to-language problem.** Recent advances in brain–computer in-
93 terfaces and neural decoding have made it possible to reconstruct perceived or
94 imagined speech from cortical activity with growing accuracy (Tang et al., 2023).
95 Yet practitioners have largely treated the translation of biological signals—EEG
96 microstates, local field potentials, spike trains—into natural language as a high-
97 dimensional regression or generative modelling problem, to be solved by scaling
98 data and model capacity. This framing implicitly assumes that a *correct* mapping
99 exists between neural activity and linguistic expression, and that the gap between
100 them is, in principle, closeable by technical refinement.

101 We argue that this assumption is structurally unfounded. Natural language
102 and neural coding are not two encodings of the same underlying content; they
103 are the products of *two independent semiotic processes*, each producing its own
104 organisation of a primordially undifferentiated signal space. The arbitrariness of
105 the sign, properly understood, entails that the cross-modal mapping between them
106 is constitutively non-unique—a fact that no amount of data or computational
107 power can overcome, because it is not an artefact of ignorance but of structure.

108 **The mind–language mapping problem.** More broadly, the question of how
109 mental content—phenomenal states, propositional attitudes, pre-linguistic intuitions—
110 relates to linguistic expression remains one of the deepest unsolved problems in
111 philosophy of mind and cognitive science. Existing accounts oscillate between two
112 unsatisfying poles: strong Whorfianism (language constitutes thought, leaving no
113 residual pre-linguistic content) and strong mentalism (thought is fully formed prior
114 to language, which merely labels it). Both poles struggle to explain the *partial*,
115 *asymmetric*, and *perspectival* character of the mind–language relation: why some
116 mental content is easily expressible while other content resists articulation, why
117 translation between mental and linguistic registers is always lossy, and why the
118 same mental state can give rise to incommensurable linguistic formulations.

119 A physically grounded account of the sign—one that explains *why* the mapping
120 between mind and language has the structure it does—would illuminate these
121 phenomena from a principled standpoint.

122 1.3. This Paper’s Approach and Contributions

123 The present paper proposes to address both problems simultaneously by invoking
124 a concept from theoretical physics: *spontaneous symmetry breaking*. Our central
125 contention is that Saussurean arbitrariness is not merely an empirical observation
126 about natural languages but the *phenomenological signature* of an underlying
127 symmetric state—a state in which all possible signifier–signified mappings are
128 equally weighted, and therefore none is actualised. Meaning emerges only when
129 this symmetry is broken by the intervention of an interpretive system, which
130 selects a particular *semantic vacuum* from a degenerate space of possibilities.

131 This framework is not merely a metaphor borrowed from physics. It provides
132 a *mathematically precise* structure—semiotic symmetry groups, order parameters,
133 gauge connections—within which precise theorems can be stated and precise pro-
134 grammes of investigation can be formulated.

135 We make the following specific contributions:

- 136 **C1. A physical grounding for semiotic arbitrariness.** We formalise the pre-
137 semiotic condition as a state of maximal symmetry in the space of signifier-
138 signified mappings (Section 3–4). Saussurean arbitrariness is thereby ele-
139 vated from an empirical axiom to a derivable structural property of any
140 system prior to interpretive commitment.
- 141 **C2. A unified treatment of meaning emergence.** We model the interven-
142 tion of any interpretive system—linguistic community, neural architecture, or
143 cultural code—as a symmetry-breaking event analogous to the spontaneous
144 magnetisation of a ferromagnet below the Curie temperature (Section 5).
145 This unifies the genesis of conventional signs, the fixation of perceptual cate-
146 gories, and the institution of cultural codes under a single structural descrip-
147 tion.
- 148 **C3. A semantic irreversibility theorem.** We prove that the symmetry-
149 breaking process is generically irreversible: the broken-symmetry state does
150 not encode sufficient information to reconstruct the primordial symmetric
151 state (Section 6). This result provides a structural basis for the hermeneutic
152 circle, the perspectival character of all interpretation, and the constitutive
153 untranslatability that obtains between sufficiently divergent semiotic sys-
154 tems.
- 155 **C4. A principled account of the difficulty of biosignal decoding.** We
156 show that neural encoding and natural language are two *independent* sym-
157 metry breakings, with no prior guarantee of isomorphism between their re-
158 sulting structures. Any decoding device introduces a *third* breaking, making
159 the biosignal-to-language translation constitutively non-unique—a structural
160 limitation rather than a technical one (Section 7).
- 161 **C5. A physically grounded perspective on the mind–language map-**
162 **ping.** By treating both the neural code and the linguistic system as broken-
163 symmetry phases descended from a common symmetric space, we explain
164 why the mind–language mapping is simultaneously possible (through shared
165 residual symmetries) and irreducibly partial (through divergent breaking his-
166 tories). This provides a middle path between strong Whorfianism and strong
167 mentalism.
- 168 **C6. A gauge-theoretic programme for neural–linguistic interfaces.** As
169 a constructive response to the non-uniqueness problem, we propose that
170 neural–linguistic interfaces should target *gauge-invariant semantic content*—
171 the subspace of decoded meaning that is stable across all admissible decoding

172 choices—formalised through fibre bundle structure and semantic gauge con-
173 nections (Section 8).

174 The paper is structured as follows. Section 2 reviews the classical doctrine of arbi-
175 trariness and its limits. Section 3 introduces the relevant mathematical concepts
176 from symmetry and group theory. Section 4 develops the notion of a pre-semiotic
177 symmetric state and its semiotic potential. Section 5 analyses the breaking mech-
178 anism, the role of interpretive systems as order parameters, and the hierarchical
179 structure of semiotic codes. Section 6 establishes the irreversibility of seman-
180 tic condensation and its hermeneutic consequences. Section 7 develops the im-
181 plications for biosignal-to-language translation and the mind–language mapping
182 problem. Section 8 sketches the gauge-theoretic programme. Section 9 situates
183 the framework relative to information theory, Peircean semiotics, and Derridean
184 *différance*. Section 10 concludes.

185 2. Saussurean Arbitrariness and Its Discontents

186 2.1. The Classical Doctrine

187 Saussure distinguished two kinds of arbitrariness. *Radical* (or *absolute*) arbitrariness holds between a given signifier and its signified in isolation: the French *arbre*
188 and the English *tree* both refer to the same concept, yet there is no internal reason
189 why one sound complex should be preferred over the other. *Relative* arbitrariness
190 holds within a system: the values of signs are determined not in isolation but dif-
191 ferentially, by their relations to other signs. The sign *sheep* and the sign *mutton*
192 carve up semantic space differently from their French counterpart *mouton*, which
193 covers both; neither segmentation is more natural than the other.

195 This combinatorial insight already points toward a structural symmetry: the
196 semantic space is, in principle, undifferentiated prior to the imposition of lexical
197 boundaries, and any sufficiently fine-grained carving is as valid as any other.

198 2.2. Critiques and Refinements

199 Several traditions have sought to qualify or limit the scope of Saussurean arbi-
200 trariness.

- 201 (i) **Iconicity and indexicality.** Peirce’s triadic sign theory distinguishes the
202 *symbol* (arbitrarily linked to its object) from the *icon* (linked by resem-
203 blance) and the *index* (linked by causal contiguity) (Peirce, 1931). Natural

languages contain all three; onomatopoeia and sound symbolism suggest that radical arbitrariness is empirically overstated.

(ii) **Embodied and motivated signs.** Cognitive linguists such as Lakoff and Johnson (Lakoff & Johnson, 1980) argue that the conceptual structures underlying language are grounded in bodily experience, introducing a motivated layer beneath the arbitrary surface.

(iii) **Iconicity in syntax.** Haiman (Haiman, 1985) argued for diagrammatic iconicity in grammatical structures: the formal properties of constructions can mirror conceptual properties (e.g., proximity of form correlates with proximity of meaning).

These qualifications are empirically important but do not dissolve the problem. Even if some signs are motivated, the *generic* condition of a semiotic system—the condition that must obtain before any specific motivated relation can be established—is one of neutrality. Our argument is that this neutrality is best captured as a *symmetry*, and that motivated signs represent *explicit symmetry breaking* by physical or functional constraints, while conventional signs represent *spontaneous symmetry breaking* by social–historical contingency.

2.3. The Explanatory Gap

The classical doctrine explains *that* signs are arbitrary but not *why* the mapping space has the structure that permits arbitrary choice in the first place, nor what *happens* when a particular mapping is instituted. This gap calls for a principled structural account. We submit that the physics of symmetry offers exactly that account.

3. Symmetry, Groups, and Degeneracy

3.1. Mathematical Preliminaries

Let $\mathcal{S}$ denote the *signifier space* (the set of all possible phonic, graphemic, or otherwise material sign-vehicles) and $\mathcal{C}$ the *signified space* (the set of all possible conceptual contents, or, in a more extensional framing, possible semantic values). A *semiotic mapping* is any function $\sigma : \mathcal{S} \rightarrow \mathcal{C}$.

Definition 3.1 (Semiotic symmetry group). Let $G_{\mathcal{S}}$ and $G_{\mathcal{C}}$ be groups acting on $\mathcal{S}$ and $\mathcal{C}$ respectively. A *semiotic symmetry* is a pair $(g_{\mathcal{S}}, g_{\mathcal{C}}) \in G_{\mathcal{S}} \times G_{\mathcal{C}}$ such

235 that, for every semiotic mapping σ ,

$$g_C \circ \sigma \circ g_S^{-1} = \sigma.$$

236 The collection of all such pairs forms the *semiotic symmetry group* $G_{\text{sem}} \leq G_S \times$
237 G_C .

238 A large symmetry group G_{sem} corresponds to a high degree of equivalence
239 among mappings: many mappings are related by symmetry transformations and
240 are therefore, in a precise sense, *indistinguishable* at the level of the system's
241 equations (its structural constraints).

242 3.2. Degeneracy and the Ground State

243 In quantum mechanics, the ground state of a system is said to be *degenerate* when
244 multiple distinct states share the same minimal energy. A symmetric Hamiltonian
245 typically produces degenerate spectra: wherever the symmetry group acts non-
246 trivially, it maps ground states to ground states, generating a family of equally
247 valid lowest-energy configurations.

248 We propose the following analogy for semiotic systems.

Proposition 3.1 (Pre-semiotic degeneracy). In the absence of an interpretive system, the space of semiotic mappings $\text{Map}(\mathcal{S}, \mathcal{C})$ is acted upon by a large symmetry group G_{sem} that relates all mappings by equivalence transformations. The system is in a *semantically degenerate ground state*: no particular mapping σ is preferred, and the expected semantic value of any signifier is undefined (or, formally, smeared uniformly across $\mathcal{C}$).

249 This degeneracy is the structural correlate of Saussurean arbitrariness. The
250 sign is arbitrary because, at the level of the primordial symmetric system, all
251 choices are equivalent.

252 3.3. Spontaneous vs. Explicit Symmetry Breaking

253 Physics distinguishes two modes of symmetry breaking:

- 254 • **Explicit breaking.** A term in the system's governing equations explicitly
255 breaks a symmetry: one direction in configuration space is energetically pre-
256 ferred from the outset. In semiotics, iconicity and phonaesthesia are explicit-
257 breaking phenomena: natural constraints (articulatory, perceptual, or concep-
258 tual) privilege certain mappings.

259 • **Spontaneous breaking.** The equations remain symmetric, but the ground
 260 state selected by the system does not share this symmetry. The canonical ex-
 261 ample is the ferromagnet below the Curie temperature: the Heisenberg Hamil-
 262 tonian is $SO(3)$ -symmetric, but any particular magnetised configuration breaks
 263 this to $U(1)$. We argue that conventional signs arise through spontaneous break-
 264 ing.

265 The *Nambu–Goldstone theorem* further specifies that spontaneous breaking of
 266 a continuous symmetry produces massless modes (Goldstone bosons) in the spec-
 267 trum. In the semiotic analogy, these correspond to the *sliding degrees of freedom* of
 268 a sign system: the directions in which signifiers can be re-mapped without chang-
 269 ing the internal relational structure of the language. This resonates with Saus-
 270 sure’s own observation that the internal organisation of a language can remain
 271 stable while particular signifiers drift—sound changes leave relational structure
 272 invariant.

273 4. The Pre-Semiotic State and the Topology of Meaning- 274 Possibility

275 4.1. The Symmetric Vacuum as Semantic Potential

276 The pre-semiotic state is not a historical moment prior to the invention of lan-
 277 guage; it is a *structural description* of what the semiotic space looks like from
 278 a vantage point outside any particular interpretive system. Phenomenologically,
 279 it is approached—but never reached—in experiences of radical *semantic under-*
 280 *determination*: encountering a script one cannot read, hearing speech in an en-
 281 tirely unknown language, or perceiving a neural signal divorced from any decoding
 282 framework.

283 Formally, the symmetric vacuum is characterised by a *flat potential* over the
 284 space of mappings:

$$V(\sigma) = V_0 \quad \text{for all } \sigma \in \text{Map}(\mathcal{S}, \mathcal{C}).$$

285 No mapping costs more (or less) than any other. The system has no preferred
 286 semantic direction.

287 4.2. The Mexican Hat Potential and Semantic Condensation

288 The passage from the symmetric to the broken-symmetry phase can be visualised
 289 through the *Mexican hat* (or wine-bottle) potential, standard in Ginzburg–Landau

290 field theory:

$$V(\phi) = -\mu^2|\phi|^2 + \lambda|\phi|^4, \quad \mu^2, \lambda > 0.$$

291 The local maximum at $\phi = 0$ (the symmetric point) is unstable. The system
 292 rolls into the circular trough at $|\phi| = \sqrt{\mu^2/2\lambda}$, selecting a specific phase $\theta \in$
 293 $[0, 2\pi)$. The selection of θ is not dictated by the potential's symmetry but by
 294 an infinitesimal perturbation—or, in a finite system, by the particularities of the
 295 boundary conditions.

296 In our semiotic reading:

- 297 • $\phi = 0$ is the pre-semiotic state of maximal underdetermination.
- 298 • The trough represents the space of *coherent but mutually inequivalent* semantic
 299 systems—each a valid breaking of the original symmetry.
- 300 • The selection of a particular θ is the contingent historical event of a specific
 301 language or code being instituted.
- 302 • Once selected, the system *condenses* around θ : semantic expectations form,
 303 and new signs are integrated by reference to the already-broken structure.

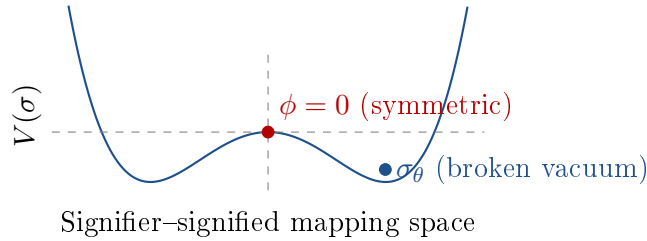


Figure 1: Schematic cross-section of the semiotic potential. The symmetric state at the origin is unstable. The system condenses into a broken-symmetry vacuum σ_θ selected by the interpretive system.

304 4.3. Hermeneutic Consequences: The Always-Already-Broken State

305 From a phenomenological standpoint, no interpreter ever occupies the symmet-
 306 ric vacuum. Every act of interpretation begins from within an already-broken
 307 structure—a Heideggerian *Geworfenheit* into a linguistic tradition. The *hermeneu-*
 308 *tic circle* can be re-read in this light: interpretation is always already constrained
 309 by the particular θ into which the interpreter was thrown. Understanding is not
 310 the reconstruction of the symmetric state but the extension and coherent elabor-
 311 ation of the broken one.

312 This does not entail relativism: two differently broken systems can be *partially*
 313 *intertranslatable* insofar as their broken structures share residual symmetries—
 314 overlapping subgroups of G_{sem} that were not destroyed by the respective breaking
 315 events. Partial translatability is thus structurally expected, while perfect trans-
 316 latability is not.

317 5. The Breaking Mechanism: Interpretive Systems as Order 318 Parameters

319 5.1. The Interpretive System as a Symmetry-Breaking Field

320 An *interpretive system* $\mathcal{I}$ is any structure that introduces a bias into the map-
 321 ping space—that is, that assigns different weights to different $\sigma \in \text{Map}(\mathcal{S}, \mathcal{C})$.
 322 Interpretive systems include:

- 323 • **Linguistic communities:** shared conventions stabilised by iterated social in-
 324 teraction.
- 325 • **Neural architectures:** the functional organisation of the nervous system that
 326 pre-constrains which signal patterns are discriminable and which semantic cat-
 327 egories are available.
- 328 • **Cultural codes:** sedimented systems of value and practice that make certain
 329 metaphors natural and others strained.
- 330 • **Formal grammars:** in computational settings, parsing algorithms and lan-
 331 guage models that assign probability distributions over parse trees.

332 In field-theoretic language, $\mathcal{I}$ functions as an *order parameter*: a quantity that
 333 is zero in the symmetric phase and non-zero in the broken-symmetry phase. The
 334 *direction* in which $\mathcal{I}$ is non-zero specifies the particular vacuum σ_θ that the system
 335 occupies.

336 5.2. Multiple Breakings and the Hierarchy of Codes

337 Real semiotic systems undergo not a single but a *hierarchical sequence* of sym-
 338 metry breakings. A natural language breaks the primordial symmetry into a
 339 phonological system; a dialect breaks that further; idiolect breaks further still;
 340 and specific utterances in specific contexts involve the most local breakings of all.
 341 Each level of the hierarchy corresponds to a subgroup of the symmetry group of
 342 the level above:

$$G_{\text{sem}} \supset G_{\text{lang}} \supset G_{\text{dialect}} \supset G_{\text{idiolect}} \supset G_{\text{utterance}}.$$

343 This cascade is analogous to the breaking pattern of grand unified theories,
 344 where a large gauge group G is broken to progressively smaller subgroups as
 345 energy decreases. The residual symmetries at each level correspond to the degrees
 346 of freedom that remain underdetermined—and therefore conventionally free—at
 347 that scale of description.

348 5.3. Degenerate Meanings and Semantic Multiplicity

349 When the interpretive context underdetermines the breaking, the system may be
 350 caught between multiple near-degenerate vacua—that is, between multiple ap-
 351 proximately equally weighted semantic interpretations. This is the structural cor-
 352 relate of *polysemy* and *ambiguity*. Disambiguation is then a process of symmetry
 353 breaking at a finer scale: additional contextual information tips the system into
 354 one vacuum rather than another. The classic Necker cube in visual perception
 355 offers a minimal analogue: the perceptual system oscillates between two broken-
 356 symmetry states when the visual input is insufficient to break the degeneracy
 357 decisively.

358 6. Irreversibility and the Hermeneutic One-Way Mirror

359 6.1. Information Loss Under Symmetry Breaking

360 Let Σ_0 denote the primordial symmetric state and Σ_θ the broken-symmetry state
 361 resulting from the selection of vacuum θ . We claim:

Proposition 6.1 (Semantic irreversibility). The mapping $\Sigma_0 \mapsto \Sigma_\theta$ is generically *non-invertible*. The broken-symmetry structure Σ_θ does not encode sufficient information to reconstruct Σ_0 .

362 *Argument.* The broken-symmetry state Σ_θ is characterised by a specific *order*
 363 *parameter direction* θ and a set of fluctuations (Goldstone modes) around it. The
 364 value of θ carries no information about the full orbit $G_{\text{sem}} \cdot \sigma_\theta$ —the set of all vacua
 365 related by symmetry transformations—because the symmetry has been broken
 366 and the system has no access to the other vacua from within the current one.
 367 In the ferromagnetic analogy: from the magnetised state, you cannot determine
 368 which direction would have been selected had the system been cooled in a different
 369 external field; the present magnetisation encodes only the direction chosen, not
 370 the full rotation group. □

371 This irreversibility has a direct information-theoretic reading: the breaking
 372 event destroys the information specifying *which* symmetry transformation was ap-
 373 plied to produce the current vacuum from the symmetric state. This information—
 374 the *symmetry-breaking history*—is not stored in the resulting structure.

375 6.2. Consequences for Translation and Interpretation

376 If semantic states are broken-symmetry configurations, then inter-linguistic trans-
 377 lation is *not* a symmetric operation. Moving from a broken state Σ_{θ_1} (Language
 378 1) to Σ_{θ_2} (Language 2) is not simply a rotation in the original symmetric space,
 379 because neither state has access to the symmetric space from which both de-
 380 scended. Translation must instead proceed *laterally*, by constructing a partial
 381 mapping between the residual structures of two broken states.

382 This has two consequences:

- 383 (1) **Constitutive untranslatability.** For any two languages whose breaking
 384 paths are sufficiently divergent, there will be conceptual regions where no
 385 adequate translation exists—not because translators are insufficiently skilled,
 386 but because the broken-symmetry structures do not overlap in those regions.
 387 This is a structural, not accidental, form of the Sapir–Whorf thesis.
- 388 (2) **The creative supplement.** Every translation must supplement the source
 389 language’s broken structure with choices that are not determined by that
 390 structure but by the target language’s own vacuum. These choices are arbi-
 391 trary in the primordial sense: they are selections from an underdetermined
 392 space. Translation thus always involves a micro-level re-breaking.

393 6.3. The Hermeneutic Circle as a Closed Orbit

394 From the irreversibility thesis, the hermeneutic circle receives a new character-
 395 isation: it is the trajectory of an interpreter who moves within the trough of
 396 the Mexican hat potential, exploring the Goldstone modes of her particular vac-
 397 uum, without ever being able to climb back to the symmetric hilltop. Under-
 398 standing is always understanding-within-a-broken-structure. This does not mean
 399 understanding is impossible, but that it is intrinsically *perspectival*: it is always
 400 understanding-from- θ .

401 7. Implications for Biosignal-to-Language Translation

402 7.1. The Problem of Cross-Modal Re-Breaking

403 The translation of biosignals—electroencephalographic (EEG) activity, neural
404 spike trains, local field potentials—into natural language sequences represents
405 a particularly acute instance of the general problem identified above. It involves
406 not merely a lateral mapping between two broken-symmetry states (as in inter-
407 linguistic translation) but a *cross-modal* mapping between:

408 (a) A biological signal space $\mathcal{B}$ that is itself subject to a neural breaking: the
409 brain’s representational code is a particular functional organisation that se-
410 lects certain signal patterns as meaningful and others as noise.

411 (b) A natural language state Σ_θ that is the product of a socio-historical breaking.

412 These two breakings are *independent*: they proceeded along different axes and
413 under different constraints. There is no guarantee—and, under the present analy-
414 sis, no expectation—that the semantic structures produced by the two breakings
415 are isomorphic.

416 7.2. The Decoder as a Third Breaking

417 Any brain–computer interface (BCI) or neural–linguistic decoder $\mathcal{D}$ is itself a third
418 interpretive system that introduces a third breaking:

$$\mathcal{B} \xrightarrow{\mathcal{D}} \Sigma_{\theta'},$$

419 where θ' need not coincide with the subject’s own linguistic vacuum θ . The
420 decoder’s output is thus doubly mediated: by the neural breaking internal to
421 the brain and by the decoder’s own breaking. This is not merely a technical
422 limitation to be overcome with more data; it is a structural feature of the cross-
423 modal translation problem.

424 **Remark 7.1.** Current large-scale neural decoding models (Tang et al., 2023)
425 achieve impressive reconstruction of perceived speech from cortical activity, but
426 they do so by constraining the decoding to a fixed large language model prior—
427 that is, by fixing θ' to match the vacuum of the training language model. This is
428 a principled strategy, but it entails that the decoded language reflects the model’s
429 breaking rather than the subject’s, and the two may diverge in precisely the regions
430 of semantic space where the subject’s idiolect differs from the training corpus.

431 7.3. Non-Uniqueness of Decoding

432 A corollary of the symmetry analysis is that neural decoding is *fundamentally*
433 *non-unique*. Different decoders correspond to different breaking choices, and each
434 produces a semantically coherent but inequivalent output. This non-uniqueness
435 cannot be eliminated by increasing decoder complexity or training data, because
436 it reflects the genuine underdetermination of the cross-modal mapping at the
437 structural level.

438 This has both theoretical and ethical implications: there is no canonical *ground*
439 *truth* against which a decoder’s output can be compared, except by reference to
440 the subject’s own expressive behaviour—and that behaviour is itself already a
441 broken-symmetry expression, not a transparent window onto neural content.

442 8. Toward a Gauge Theory of Semantic Interfaces

443 8.1. Local vs. Global Symmetry Breaking

444 In gauge field theories, a crucial distinction separates *global* from *local* symmetry
445 breaking. A global breaking selects a single vacuum for the entire system. A local
446 (gauge) symmetry allows the vacuum to vary from point to point—each region
447 of space-time chooses its own $\theta(x)$ —and consistency is maintained by a *gauge*
448 *connection* that compensates for the local variations.

449 We propose that an analogous local/global distinction applies to semiotic sys-
450 tems and to cross-modal interfaces. A *locally gauge-invariant* semantic interface
451 is one that:

- 452 (i) Allows the decoder to choose a local semantic vacuum that is well-adapted
453 to each region of the input signal space.
- 454 (ii) Maintains global coherence through a connection that encodes the transfor-
455 mation rules relating adjacent local choices.

456 **Definition 8.1** (Semantic gauge connection). Let $\{U_\alpha\}$ be a covering of the
457 biosignal space $\mathcal{B}$, and let $\sigma_\alpha : U_\alpha \rightarrow \mathcal{C}$ be local semantic maps (local de-
458 coders). A *semantic gauge connection* is a collection of transition functions
459 $g_{\alpha\beta} : U_\alpha \cap U_\beta \rightarrow G_{\text{sem}}$ such that on overlapping regions,

$$\sigma_\alpha = g_{\alpha\beta} \cdot \sigma_\beta.$$

460 The connection defines a *fibre bundle* structure on $\mathcal{B}$ with fibre $\mathcal{C}$.

461 8.2. Minimal Symmetry Breaking and Semantic Parsimony

462 The gauge framework suggests a natural optimisation criterion for semantic in-
463 terfaces: *minimise the total symmetry broken by the decoding system*, subject to
464 the constraint that the output is coherent in the target natural language. This is
465 analogous to the *principle of minimal coupling* in gauge theories, which forbids the
466 introduction of interaction terms beyond those required by the gauge symmetry.

467 Formally, we seek a decoder $\mathcal{D}$ that minimises:

$$\mathcal{L}[\mathcal{D}] = \int_{\mathcal{B}} \|F_{\mathcal{D}}(b)\|^2 db,$$

468 where $F_{\mathcal{D}}$ is the *curvature* of the semantic gauge connection induced by $\mathcal{D}$ —a
469 measure of how much the local semantic choices vary from point to point. A
470 flat connection (zero curvature) corresponds to a globally consistent decoder that
471 never needs to switch semantic frames. Non-zero curvature indicates regions of
472 signal space where the decoder must perform a semantic frame-shift, introducing
473 additional arbitrary choices.

474 8.3. Residual Symmetries as Invariant Semantic Content

475 Under the gauge framework, the *physically meaningful* (i.e., decoder-invariant)
476 content of a decoded message is precisely the content that is invariant under the
477 residual gauge symmetry—the subgroup of G_{sem} that was *not* broken by the decod-
478 ing. This is the semantic analogue of the Higgs mechanism: the degrees of freedom
479 eaten by the gauge field (the Goldstone modes) are the decoder-dependent arte-
480 facts, while the massive degrees of freedom (the Higgs) are the genuine semantic
481 content.

482 This suggests a practical programme: rather than trying to decode a single
483 “correct” language output, neural–linguistic interfaces should aim to identify and
484 output the *gauge-invariant subspace* of the decoded message—the content that
485 all admissible decoders agree on—while explicitly marking the gauge-dependent
486 residue as interpretively underdetermined.

487 9. Discussion

488 9.1. Relation to Information-Theoretic Accounts

489 Our framework is compatible with, but distinct from, purely information-theoretic
490 approaches to meaning. Information theory characterises the *channel capacity* of
491 a signal but says nothing about the *semantic content* carried by signals of given

492 entropy. The symmetry framework addresses this gap: the entropy of the symmet-
 493 ric state corresponds to the maximum of Shannon entropy over all sign mappings,
 494 while symmetry breaking selects a specific distribution that concentrates prob-
 495 ability mass on particular signifier–signified pairs. Meaning is the *information*
 496 *gained by knowing which vacuum was selected.*

497 9.2. Relation to Peircean Semiotics

498 Peirce’s concept of *semiosis*—the irreducible triadic process by which a sign me-
 499 diates between object and interpretant—can be read as a dynamic account of iter-
 500 ated symmetry breaking. Each interpretant is a new sign that re-breaks the sym-
 501 metry at a finer scale, generating a potentially infinite chain of meaning-making.
 502 The arrest of semiosis (Peirce’s *final logical interpretant*) would correspond to a
 503 fixed point of the breaking process—a vacuum from which no further breaking is
 504 induced by the object.

505 9.3. Relation to Derrida’s *Différance*

506 Derrida’s claim that meaning is constituted by *différance*—a simultaneous differ-
 507 entiation and deferral—resonates with the symmetry framework. Differentiation
 508 is the breaking of the symmetric vacuum into a structured space of distinctions;
 509 deferral is the irreversibility of that breaking, which ensures that the originary
 510 symmetry is never directly present but only traceable in the structure of its ab-
 511 sence. The *trace*, in Derridean terms, is the mark left in the broken-symmetry
 512 structure by a symmetry that can no longer be restored.

513 9.4. Limitations and Open Questions

514 We note several important limitations of the present proposal.

515 First, the framework is currently *schematic*: while the physical analogies are
 516 precise, the identification of concrete mathematical structures in natural language
 517 (e.g., the specific group G_{sem} , the specific potential $V(\sigma)$) remains programmatic.
 518 Progress requires either empirical constraints from corpus linguistics and psy-
 519 cholinguistics, or principled derivations from more foundational theories of cogni-
 520 tion.

521 Second, the framework as presented treats $\mathcal{S}$ and $\mathcal{C}$ as given, whereas in reality
 522 both spaces are themselves constituted by prior breakings (perceptual categorisa-
 523 tion, conceptual development). A fully recursive account would need to address
 524 this bootstrapping problem.

525 Third, the gauge-theoretic programme for neural–linguistic interfaces requires
526 a well-defined *fibre bundle structure* over biosignal space. Whether the topological
527 conditions for such a structure are satisfied by realistic neural data is an empirical
528 question that the framework motivates but does not answer.

529 10. Conclusion

530 This paper has argued that the Saussurean arbitrariness of the sign is not merely
531 an empirical generalisation about natural languages but the phenomenological
532 expression of a deeper structural fact: the pre-semiotic condition is a state of
533 maximal symmetry, in which all signifier–signified mappings are equivalent under
534 the action of a large semiotic symmetry group. Meaning emerges through spon-
535 taneous symmetry breaking, driven by the intervention of an interpretive system
536 that selects a particular semantic vacuum from the degenerate ground state.

537 This breaking is irreversible: the broken-symmetry structure does not encode
538 the information needed to reconstruct the primordial symmetry. Hermeneutics is
539 therefore an enterprise conducted within a broken-symmetry phase, from which
540 the symmetric origin is structurally inaccessible, though asymptotically approach-
541 able through the proliferation of interpretive perspectives.

542 For the cross-modal translation of biosignals into natural language, the frame-
543 work implies a constitutive non-uniqueness of decoding that cannot be overcome
544 by purely technical means. We have sketched a gauge-theoretic programme that
545 addresses this non-uniqueness constructively, by demanding local semantic covari-
546 ance and identifying gauge-invariant content as the proper target of decoding.

547 The programme opens several directions for future work: the characterisa-
548 tion of residual symmetries in specific neural codes; the development of gauge-
549 covariant loss functions for training neural–linguistic decoders; and the extension
550 of the framework to multimodal and embodied semiotic systems. More broadly,
551 the symmetry approach suggests that the physics of broken symmetry and the phi-
552 losophy of language share a deeper structural kinship than has previously been
553 recognised—one that may ultimately enrich both disciplines.

554 References

- 555 Anderson, P. W. (1972). More is different. *Science*, 177(4047), 393–396.
- 556 Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press, Cambridge,
557 MA.

- 558 Derrida, J. (1967). *De la Grammatologie*. Les Éditions de Minuit, Paris.
- 559 Eco, U. (1976). *A Theory of Semiotics*. Indiana University Press, Bloomington,
560 IN.
- 561 Goldstone, J., Salam, A., and Weinberg, S. (1962). Broken symmetries. *Physical*
562 *Review*, 127(3), 965–970.
- 563 Haiman, J. (1985). *Natural Syntax: Iconicity and Erosion*. Cambridge University
564 Press, Cambridge.
- 565 Heidegger, M. (1927). *Sein und Zeit*. Max Niemeyer Verlag, Halle.
- 566 Jakobson, R. (1960). Closing statement: Linguistics and poetics. In T. Sebeok
567 (Ed.), *Style in Language* (pp. 350–377). MIT Press, Cambridge, MA.
- 568 Lakoff, G. and Johnson, M. (1980). *Metaphors We Live By*. University of Chicago
569 Press, Chicago, IL.
- 570 Nambu, Y. (1960). Quasi-particles and gauge invariance in the theory of super-
571 conductivity. *Physical Review*, 117(3), 648–663.
- 572 Peirce, C. S. (1931–1958). *Collected Papers of Charles Sanders Peirce*, 8 vols.
573 Harvard University Press, Cambridge, MA.
- 574 Ricoeur, P. (1981). *Hermeneutics and the Human Sciences*. Cambridge University
575 Press, Cambridge.
- 576 Saussure, F. de (1916). *Cours de linguistique générale*. Payot, Lausanne and Paris.
- 577 Tang, J., LeBel, A., Jain, S., and Huth, A. G. (2023). Semantic reconstruction of
578 continuous language from non-invasive brain recordings. *Nature Neuroscience*,
579 26, 858–866.
- 580 Weinberg, S. (1967). A model of leptons. *Physical Review Letters*, 19(21), 1264–
581 1266.
- 582 Wittgenstein, L. (1953). *Philosophical Investigations*. Blackwell, Oxford.
- 583 Yang, C. N. and Mills, R. L. (1954). Conservation of isotopic spin and isotopic
584 gauge invariance. *Physical Review*, 96(1), 191–195.