Topology or Metaphor?
Substrate Combinatorics, Emergent Invariants, and Formal Obligations for Models of Language and Subject Development

Wanhong HUANG
huangwanhong@serendip.ngo

Abstract

Topological vocabulary appears in models of neural activity, semantic organization, linguistic structure, and development, and its inferential content depends on the object to which topology is assigned. This paper poses that question inside a background-independent substrate in which topology is not metaphorical: a group field generates two-complexes whose combinatorics constitute the theory's own data. The reformulation sharpens rather than dissolves the problem, because two questions separate. At the substrate level a new result shows that combinatorial topology is underdetermined by the subsystem's accessible boundary algebra: complexes differing by a closed component induce identical accessible states while differing in their invariants, so a system can carry literal topology that nothing available to it records. At the effective level, topological claims about language, categories, and development remain governed by a warrant architecture, and a two-level version of that architecture is supplied, together with the result that invariants do not transfer across coarse-graining without a stated preservation condition. Several separation results are retained. An observationally silent factor changes the fundamental group without changing behavior; recognition of unbounded bracket nesting and nontrivial fundamental group vary independently; a contractible carrier supports arbitrarily many category labels; homeomorphic carriers support dynamically inequivalent maps; and a bifurcation can occur on a fixed ambient topology. A positive construction is retained and rebased on accessible response laws: intervention-indexed responses induce a pseudometric on developmental histories whose metric quotient admits a Vietoris–Rips filtration, and uniform pairwise metric error at most δ yields a δ-interleaving and bottleneck error at most δ. The proposal that universal grammar is a topological invariant is preserved as an explicitly labelled conjecture with its obligations and its existing obstruction stated. The position is constructively conservative.

Keywords: persistent homology; spin-foam combinatorics; language development; topological warrant; construct validity.

Statements

Status and correspondence

This paper is a working discussion document. Its formal claims are conditional on declared assumptions, and several of its bridges to empirical and philosophical questions remain open. Objections, counterexamples, corrections, alternative formulations, and pointers to relevant literature are all welcome and may be sent to `huangwanhong@serendip.ngo`.

Generative AI use

Generative AI systems were used in the preparation of this work. Anthropic's Claude and OpenAI's ChatGPT supported exploratory discussion of the conceptual framework, development of the formal construction and its notation, identification of candidate literature for subsequent checking, and drafting and revision of the manuscript in LaTeX. The research questions, theoretical commitments, formal claims, and the epistemic status assigned to each were determined and approved by the author, who bears sole responsibility for the content of this paper, including any errors it contains. Neither system is an author of this work and neither holds authorship credit, in accordance with the position that authorship carries responsibilities a generative AI system cannot assume.

License

1 Research Scope and Methodological Position

This section establishes the paper's role, target, and sequence. It defines a graded modeling vocabulary, states the limited thesis, and separates mathematical validity from empirical and subject-level attribution.

The motivating problem concerns expressions such as *cognitive manifold*, *topological invariant of language*, *semantic loop*, and *topological phase transition in development*. Such expressions can guide inquiry. Their scientific status depends on whether the relevant carrier, topology, invariant, estimator, and target interpretation have been supplied.

An earlier version of this paper treated every topology-bearing claim as occupying a single ladder from metaphor to explanation. The present version distinguishes two levels, because the surrounding research program now works with a background-independent substrate in which combinatorial topology is constitutive rather than figurative. A group field over a declared group generates two-complexes; the amplitudes depend on their combinatorics; and a subsystem is a subcomplex with a boundary. In that setting, asking whether topology is a metaphor is the wrong question at the substrate level and the right question at every level above it.

Definition 1.1 (Model-status ladder)**.** A *topological metaphor* transfers a partial structural image while leaving the represented space and invariant open. A *topological heuristic* identifies a candidate mathematical structure and a target phenomenon while retaining unresolved construction or validation steps. A *formal topological model* specifies the represented object, topology, admissible transformations, and mathematical consequences. An *estimable topological model* additionally specifies data, construction, scale, estimator, and uncertainty. A *topological explanation* further supports the bridge to the target phenomenon and discriminates relevant rival models.

The ladder records evidential maturity. Metaphors can be intellectually productive, and formally correct invariants can remain empirically uninformative. Causal explanation requires an additional intervention or mechanism warrant.

Claim 1.2 (Level-relative topology)**.** A topological calculation licenses a conclusion about the object to which its topology is assigned and about nothing else. Substrate combinatorics, effective representations, and target-level linguistic or subject-level constructs are three distinct objects, and each transition between them requires a bridge whose validity is independent of the calculation.

Claim 1.3 (Literal topology is not accessible topology)**.** A system can carry genuine, non-metaphorical topology whose invariants are not determined by anything the system or an experimenter can record from its interface. Constitutive topology and identifiable topology are therefore separate properties, and establishing the first supplies no evidence for the second.

Claim 1.3 is the paper's principal new content and is proved as Proposition 3.2. It has an uncomfortable consequence for the surrounding research program: adopting a substrate in which topology is real does not relax the warrant obligations on topological claims, and in one respect it tightens them, because it makes the gap between what exists and what can be identified explicit.

The manuscript proceeds in five stages. Section 2 records the antecedents. Section 3 defines the substrate and proves the boundary-invisibility result. Section 4 defines the two-level warrant architecture. Section 5 derives the effective-level separation results. Section 6 constructs an estimable response topology and proves its stability bound. Sections 7–10 translate the results into developmental, linguistic, empirical, and philosophical obligations, and Section 11 states the revisable position.

2 Mathematical, Linguistic, and Developmental Antecedents

This section positions the proposal within topology, data analysis, background-independent quantum geometry, representational modeling, developmental dynamics, construct validity, and scientific analogy. Each source receives a bounded role, and the discussion distinguishes mathematical ancestry from evidence for a cognitive ontology.

Betti's work on higher-dimensional spaces and Poincaré's *Analysis situs* supply historical ancestry for connectivity invariants and systematic algebraic-topological analysis [1, 20]. Persistent homology later organized the birth and death of homology classes across a filtration [7]. Stability results show that declared geometric complexes vary controllably under declared metric perturbations [3]. Mapper supplies another constructive pipeline whose output depends on a lens, cover, clustering procedure, and nerve [22]. Manifold-homology recovery can receive high-probability guarantees under explicit smoothness, reach, sampling, and noise assumptions [16]. These results establish formal resources and display the assumptions an empirical application must earn.

The substrate contributes a different kind of ancestor. Group field theory defines a field on a group manifold whose perturbative expansion generates two-complexes [2, 10, 17], spin-foam models assign amplitudes to those complexes [19], and the reading of a Feynman diagram as a spacetime history is explicit in the literature [21]. Colored tensor models make the topological content of the generated complexes especially explicit [12]. In these theories topology is data, not imagery. This paper borrows that mathematical situation and borrows none of its empirical warrant: no result cited here provides evidence that linguistic or developmental systems are group-field systems.

Adjacent language and cognition research occupies several registers. Corpus co-occurrence can define graph-theoretic small-world structure [9]. Conceptual-spaces theory supplies a geometrical account of similarity and quality dimensions [11]. Formal-grammar research supplies an ancestor for distinguishing generative systems and finite-state restrictions [4]; it supplies no mapping from grammatical operations to a cognitive fundamental group. Representational distance matrices provide a comparison level across behavioral, computational, and neural measurements [14]. Perceptual-manifold theory derives classification capacity from explicitly defined response ensembles and their geometry [5]. Cross-linguistic embedding studies can compare local and global semantic geometry [15]. Graph structure, metric geometry, manifold geometry, and algebraic topology remain mathematically distinct even when they concern related data.

Developmental systems provide equally important comparison models. A dynamic field model of infant reaching connects variables, equations, simulations, experiments, and novel predictions within a bounded task [23]. Recurrent-network work demonstrates that a developmental schedule can alter artificial-grammar learning in a declared architecture [8]. Neither result establishes topology change in a subject.

Construct validity requires a network of relations among a theoretical construct, observables, and other claims [6]. Philosophy of scientific modeling likewise treats analogy as potentially generative while preserving positive, negative, and unresolved source–target relations [13]. Mathematical stability and construct validity therefore answer different questions.

Table 1 summarizes the formal registers used throughout the paper, with the substrate row added. The same term, especially *topology*, cannot license every row.

Table 1: Formal registers and licensed developmental inferences.

Register	Primitive structure	Licensed statement	Additional warrant
Substrate combinatorics	group field, generated two-complexes, configuration amplitudes	combinatorial and homological statements about the generated complexes	identifiability from an accessible algebra; evidence that a cognitive system realizes the substrate
Graph structure	$G = (V, E)$ with declared node and edge rules	adjacency, paths, components, degree, and graph cycles	semantic meaning of vertices, edges, thresholds, and time variation
Metric or differential geometry	metric space, or a manifold with declared tensor or connection	distance and similarity; geodesic, curvature, or volume under the relevant added structure	measurement model and invariance to nuisance transformations
Topology	(X, τ) and a transformation class	continuity, connectedness, compactness, and homeomorphism	represented carrier and empirical construction of τ
Algebraic topology	space or complex with homology or homotopy functor	H_k, π_k, and Betti numbers for that object	coefficient choice and bridge from invariant to target capacity
Persistent topology	filtration $\{K_\varepsilon\}$	birth and death of homology classes across scale	metric, filtration, estimator, uncertainty, and scale interpretation
Dynamical systems	state space with evolution law	attractors, bifurcations, stability, and conjugacy	state meaning, parameters, interventions, and relation to topology change
Scientific analogy	partial source–target correspondence	heuristic transfer and question formation	preserved relations, disanalogies, predictions, and failure conditions

3 Substrate Topology and Its Invisibility

This section defines the substrate, states what is literally topological in it, and proves that those topological facts are underdetermined by a subsystem's accessible algebra.

3.1 Generated Complexes and Accessible Algebras

Let $\Psi : G^{\times d} \to \mathbb{C}$ be a group field over a declared group G with fundamental action S_0. Perturbative expansion of the partition function generates a sum over two-complexes,

$$Z_0 = \sum_{\mathcal{C}} \frac{\lambda^{V(\mathcal{C})}}{\mathrm{sym}(\mathcal{C})}\, \mathcal{A}[\mathcal{C}], \qquad \mathcal{A}[X] = \prod_f A_f \prod_e A_e \prod_v A_v, \tag{1}$$

where a configuration X assigns representation labels to faces, intertwiners to edges, and declared data to vertices. Each $\mathcal{C}$ is a combinatorial object with genuine topological invariants: connectivity, cycle structure, the homology of its dual complex, and, for suitably restricted models, the topology of the generated pseudo-manifold [12].

A subsystem is a subcomplex $\mathcal{C}_I \subseteq \mathcal{C}$ with boundary $\partial\mathcal{C}_I$ and boundary Hilbert space $\mathcal{H}_{\partial I}$. A declared accessible algebra $\mathfrak{A}_I \subseteq \mathcal{B}(\mathcal{H}_{\partial I})$ collects the operators whose expectations the subsystem's own dynamics, or an experimenter working through its interface, can register. Summing the configuration away from the boundary gives a reduced density operator ρ_I and an accessible state $\omega_I = \rho_I \restriction_{\mathfrak{A}_I}$.

Claim 3.1 (Constitutive topology)**.** Within Equation (1), topological properties of $\mathcal{C}$ are part of the theory's data rather than an interpretive overlay. Statements about those properties are ordinary mathematical statements about the declared model, and calling them metaphorical would be a category error.

3.2 Boundary Invisibility of Bulk Topology

Proposition 3.2 (Boundary-invisible topology). *Let $\mathcal{C}_I$ be a subcomplex with boundary $\partial\mathcal{C}_I$, and let $\mathcal{B}$ be a closed complex, that is one with empty boundary, satisfying $\mathcal{A}[\mathcal{B}] \neq 0$. Set $\mathcal{C}_I' = \mathcal{C}_I \sqcup \mathcal{B}$, so that $\partial\mathcal{C}_I' = \partial\mathcal{C}_I$. Then the normalized accessible states coincide,*

$$\omega_{I'}(M) = \omega_I(M) \qquad \textit{for every } M \in \mathfrak{A}_I, \tag{2}$$

while the two subcomplexes differ in their topological invariants whenever $\mathcal{B}$ has nontrivial homology; for instance $\beta_k(\mathcal{C}_I') = \beta_k(\mathcal{C}_I) + \beta_k(\mathcal{B})$ for every k.

Proof. Because $\mathcal{B}$ is closed and disjoint from $\mathcal{C}_I$, no face, edge, or vertex of $\mathcal{B}$ meets $\partial\mathcal{C}_I$, and the boundary Hilbert space is unchanged. The amplitude of Equation (1) factorizes over connected components, so summing the configuration away from the boundary gives $\widetilde{\rho}_{I'} = \mathcal{A}[\mathcal{B}]\,\widetilde{\rho}_I$ for the unnormalized reduced operators. Normalization divides by the trace, and the nonzero scalar $\mathcal{A}[\mathcal{B}]$ cancels, giving Equation (2). Additivity of Betti numbers over disjoint unions gives the second statement. □

Corollary 3.3 (No identification theorem from accessible fit). *Agreement of a model with every accessible expectation supplies no constraint on the homology of the subsystem's bulk complex. In particular, an inference from observed behavior to a claim about the topology of the system generating that behavior requires a minimality, observability, or intervention assumption that excludes constructions of the kind used in Proposition 3.2.*

Proposition 3.2 is the substrate-level counterpart of the effective-level silent-factor result proved as Proposition 5.1. The two are structurally the same construction at different levels: an unobserved component carries topology that the accessible data cannot see. Their agreement is evidence that the difficulty is not an artifact of a particular modeling choice.

The construction used here is deliberately the simplest available. A disconnected closed component is easy to exclude by fiat, and a serious identification theory would need to characterize the connected modifications that are also boundary-silent. That characterization is stated as Open Problem 11.1 and is not attempted here.

4 Two-Level Topological Warrant Architecture

This section defines the inferential chain required for a topology-bearing claim, extends it to two levels, and states the non-transfer result that governs movement between them.

Let $\mathcal{O}$ denote the empirical record space generated by an observation and intervention protocol, let R map records into a represented space (X, τ), and let $\mathcal{T}$ denote the transformations under which the target structure is intended to remain invariant. When a filtration is used, let Filt denote the admitted filtration class and let the construction $\mathcal{F}$ produce $\mathbf{K} := \{K_\varepsilon\}_{\varepsilon\in\mathbb{R}}$. Let $I : \mathsf{Filt} \to \mathcal{Z}$ take the full filtration to an invariant space, let $\widehat{I}$ be a data-based estimator, let B bridge invariant values to a linguistic or developmental construct C, and let $\mathcal{N}$ be a family of comparison models.

Definition 4.1 (Two-level topological warrant specification). A complete specification is the tuple

$$\mathfrak{W} = (\mathcal{C}_I, \Lambda, \mathcal{O}, R, (X, \tau), \mathcal{T}, \mathcal{F}, I, \widehat{I}, B, \mathcal{N}), \tag{3}$$

where $\mathcal{C}_I$ is the substrate representand and Λ is the declared coarse-graining from substrate configurations to the effective representand. Its primary target chain is

$$\mathcal{C}_I \xrightarrow{\Lambda} \mathcal{O} \xrightarrow{R} (X, \tau; \mathcal{T}) \xrightarrow{\mathcal{F}} \mathbf{K} \xrightarrow{I} \mathcal{Z} \xrightarrow{B} C, \tag{4}$$

and its cross-cutting obligations are the estimator $\widehat{I} : \mathcal{O} \to \mathcal{Z}$ and the comparison family $\mathcal{N} = \{N_j\}_{j \in J}$. Each arrow, the estimator, and each comparison model carries its own assumptions and uncertainty.

Proposition 4.2 (Invariants do not transfer across coarse-graining)**.** *There is no general relation between $I(\mathcal{C}_I)$ and $I(\Lambda(\mathcal{C}_I))$. Both directions of failure occur: a coarse-graining can destroy a nontrivial invariant, and a coarse-graining can create one.*

Proof. For destruction, let $\mathcal{C}_I$ contain a cycle and let Λ be the map collapsing that cycle to a point, which is a legitimate coarse-graining of the configuration data; then β_1 drops from a positive value to zero while every retained observable defined after the collapse is unchanged by construction. For creation, let $\mathcal{C}_I$ be a contractible complex and let Λ retain two coordinates whose joint range is an annulus because the excluded coordinate was the one along which the region was filled; then β_1 rises from zero to one. Both maps are ordinary quotients or projections, so neither is excluded by the definition of a coarse-graining. □

Claim 4.3 (Arrow-local inference)**.** Correct computation of $I(\mathbf{K})$ establishes a statement about the declared filtration. A conclusion about C additionally depends on the coarse-graining Λ and its preservation properties, representation validity, scale selection, estimation, robustness, the bridge B, and comparison with $\mathcal{N}$. The strength of the conclusion follows the weakest warranted link.

This accounting rule distinguishes topological invariance from robustness. Ordinary homology is invariant under homeomorphism of the represented spaces. Vietoris–Rips persistence is derived from a metric and is stable under controlled metric change. Stability relative to one metric leaves the justification of that metric open, and Proposition 4.2 adds that stability at one level leaves the other level open as well.

4.1 Filtration Scale and Finite Samples

This subsection isolates scale dependence in the smallest useful example, showing where a topological feature enters the construction and why a Betti number requires a metric, complex, and scale.

Let four points occupy the corners of a unit square. For scale $r \geq 0$, let VR_r contain a simplex when all its pairwise distances are at most r. The side length is 1 and each diagonal has length $\sqrt{2}$. Consequently,

$$(\beta_0(r), \beta_1(r)) = \begin{cases} (4,0), & 0 \leq r < 1, \\ (1,1), & 1 \leq r < \sqrt{2}, \\ (1,0), & r \geq \sqrt{2}. \end{cases} \tag{5}$$

For $r < 1$ the complex contains four vertices. At $r = 1$ the four side edges form a cycle without filled triangles. At $r = \sqrt{2}$ every pair is connected and the full three-simplex fills the cycle. Persistent homology records this scale-indexed history; it does not remove the analyst's responsibility for the point representation, metric, and interpretive bridge.

5 Effective-Level Underdetermination

This section derives the paper's identification limits at the effective level. The argument proceeds from behaviorally silent factors to recursion, categories, dynamical equivalence, and bifurcation, with each result restricted to its declared model class.

5.1 Output-Silent Topological Factors

Let a deterministic controlled model be

$$M = (X, U, Y, F, O), \qquad x_{t+1} = F(x_t, u_t), \qquad y_t = O(x_t). \tag{6}$$

For a based topological space (K, k_0), define the extension

$$F^K((x, k), u) = (F(x, u), k), \qquad O^K(x, k) = O(x). \tag{7}$$

Proposition 5.1 (Silent-factor underdetermination)**.** *Starting from (x_0, k_0), the model on $X \times K$ defined by Equation (7) generates the same output trajectory as Equation (6) for every admissible input history. If X is path-connected and $K = S^1$, then*

$$\pi_1(X \times S^1, (x_0, k_0)) \cong \pi_1(X, x_0) \times \mathbb{Z}. \tag{8}$$

For simply connected X, the extension changes the fundamental group from the trivial group to $\mathbb{Z}$ without changing behavior.

Proof. Equation (7) gives $k_t = k_0$ at every time. Induction on t yields the same X-coordinate trajectory under every input history, and the output maps agree through projection onto X. The product theorem for fundamental groups gives Equation (8). □

The proposition concerns realization classes that permit output-silent factors and interventions restricted to U. Minimal realizations, full-state observability, structural priors, or interventions on K can reduce this equivalence class, and such restrictions are substantive parts of a topological model. Proposition 3.2 shows that the same difficulty recurs at the substrate level, where the analogous restriction would be a condition excluding boundary-silent components.

Observation maps can also collapse topologically different latent spaces without an explicit product factor. An interval-valued latent variable with the arcsine density

$$p(x) = \frac{1}{\pi\sqrt{1 - x^2}}, \qquad -1 < x < 1,$$

and the observation $Y = x$ has the same observable distribution as $\Theta \sim \mathrm{Unif}[0, 2\pi)$ on S^1 with $Y = \cos\Theta$. Their fundamental groups are respectively trivial and $\mathbb{Z}$. Observable fit alone therefore supplies no latent-topology identification theorem.

5.2 Formal Nesting and Fundamental Groups

This subsection separates one narrow recursive capacity from carrier-space homotopy, using recognition of the one-bracket Dyck language as an explicit test and reserving broader claims about human linguistic recursion.

Let D_1 be the language of balanced strings over one pair of brackets. A pushdown automaton recognizes D_1 with finite control and an unbounded stack. Equip its countable configuration set with the discrete topology; every based loop is constant within its singleton component, so each based fundamental group is trivial. Conversely, equip a constant-output system with state space S^1 and input-independent dynamics; its fundamental group is $\mathbb{Z}$ while it distinguishes no string family.

Proposition 5.2 (Nesting–fundamental-group separation)**.** *Within the declared realization classes, nontrivial fundamental group is neither necessary nor sufficient for recognition of D_1. A linguistic interpretation of loops requires an explicit map from paths or homotopy classes to grammatical operations and a discriminating empirical prediction.*

The result addresses formal nesting only. Generation, compositionality, semantics, processing limits, learnability, and the organization of natural languages remain separate objects. Proposition 5.2 is also the principal existing obstruction to the conjecture stated in Section 7, and it is recorded here so that the conjecture is not read as unencumbered.

5.3 Category Inventory and Ambient Homology

For every integer $m \geq 1$, let $X = [0, 1]$ and partition it into m half-open subintervals, with the final subinterval closed at 1, to define a category map $\ell_m : X \to \{1, \ldots, m\}$. The carrier remains contractible:

$$\beta_0(X) = 1, \qquad \beta_k(X) = 0 \quad (k \geq 1), \tag{9}$$

independently of m.

Proposition 5.3 (Category–Betti separation)**.** *The same ambient Betti vector supports arbitrarily many declared categories. Output-silent factors can also change ambient homology while preserving every category output. A categorical interpretation of Betti numbers therefore requires a specified object, such as decision regions or a response complex, and a validated bridge.*

Continuity constraints, decision-boundary geometry, neural coding, and behavioral generalization can restrict admissible label maps. Those constraints belong to the model specification.

5.4 Carrier Equivalence and Developmental Dynamics

Let $f : S^1 \to S^1$ be the identity and let $g : S^1 \to S^1$ be rotation by an irrational angle. Their carriers are identical. A topological conjugacy h would satisfy

$$h \circ f = g \circ h. \tag{10}$$

The identity fixes every point, while irrational rotation fixes none.

Proposition 5.4 (Homeomorphism–dynamics separation)**.** *The systems (S^1, f) and (S^1, g) are not topologically conjugate. Carrier homeomorphism therefore supplies no equivalence of acquisition dynamics, observables, labels, or linguistic operations.*

Proof. Conjugacy maps the fixed-point set of f bijectively to the fixed-point set of g. The former is S^1 and the latter is empty, which is impossible. □

For a complementary distinction, consider

$$\dot{x} = \mu x - x^3, \qquad x \in \mathbb{R}. \tag{11}$$

The equilibrium and stability structure changes as μ crosses zero, while the ambient state space remains $\mathbb{R}$. The derived equilibrium set changes cardinality, so its topology can change when that set is explicitly selected as the representand.

Claim 5.5 (Representand-indexed transition)**.** A bifurcation establishes a change in declared dynamical organization. A topology-change claim additionally identifies the topology-bearing object, before–after construction, invariant, and observation procedure.

6 Intervention-Indexed Developmental Response Topology

This section supplies a positive construction whose representand is explicit. It defines a pseudometric from accessible response laws, forms the associated metric quotient, constructs a Vietoris–Rips filtration, and derives finite-sample stability and longitudinal error bounds.

6.1 Response Representation and Metric Quotient

Let Γ be a set of developmental histories, each realized as a family of subsystem configurations, and let ω_γ denote the accessible state that history γ induces on the declared algebra $\mathfrak{A}_I$. Let $\mathcal{U}$ be a family of admissible source insertions and $\mathcal{E}$ a family of evaluation contexts. Fix a nonempty finite preregistered probe family $\mathcal{Q} \subseteq \mathcal{U} \times \mathcal{E}$ and let $\mathcal{P}(\mathcal{Y})$ be the probability laws on a common outcome space $\mathcal{Y}$. For every $\gamma \in \Gamma$ and $c = (u, e) \in \mathcal{Q}$, suppose the response law

$$P_\gamma^c := P_{\omega_\gamma}(Y \mid \mathrm{do}(u), e) \in \mathcal{P}(\mathcal{Y}) \tag{12}$$

is defined, where the outcome distribution is generated by a measurement in $\mathfrak{A}_I$ after the declared source insertion. The do notation follows the structural intervention convention [18]; it marks a target interventional law and supplies no identification result by itself. Choose a finite-valued probability metric d_c for each condition, with total variation or a bounded Wasserstein metric as examples, and set

$$d_\mathcal{Q}(\gamma, \gamma') := \max_{c \in \mathcal{Q}} d_c\left(P_\gamma^c, P_{\gamma'}^c\right). \tag{13}$$

Proposition 6.1 (Response pseudometric)**.** *Equation* (13) *defines a pseudometric on* Γ*. The relation* $\gamma \sim_\mathcal{Q} \gamma'$ *exactly when* $d_\mathcal{Q}(\gamma, \gamma') = 0$ *is an equivalence relation, and*

$$X_\mathcal{Q} := \Gamma / {\sim_\mathcal{Q}} \tag{14}$$

inherits a metric from $d_\mathcal{Q}$*.*

Proof. Nonnegativity and symmetry follow from the condition-specific metrics. For any three histories the triangle inequality holds pointwise for every $c \in \mathcal{Q}$, and taking the finite maximum preserves it. Distinct histories can induce the same entire response family, which gives the pseudometric's zero-distance pairs. The standard metric-quotient construction then proves that the induced distance on $X_\mathcal{Q}$ is well defined and separates equivalence classes. □

The subscript in $X_\mathcal{Q}$ is constitutive. Enlarging the preregistered probe family can separate previously equivalent histories, while a restricted family can collapse meaningful differences. The quotient therefore represents functional organization relative to an experimental repertoire; it supplies neither a complete internal configuration nor a unique ontological decomposition, and by Proposition 3.2 it cannot supply the substrate topology either.

Proposition 6.2 (Probe-family ordering)**.** *Let nonempty finite admissible families* $\mathcal{Q}_1 \subseteq \mathcal{Q}_2$ *use the same response metrics on shared conditions. Then*

$$d_{\mathcal{Q}_1}(\gamma, \gamma') \leq d_{\mathcal{Q}_2}(\gamma, \gamma') \tag{15}$$

for all histories, and exact equivalence under $\mathcal{Q}_2$ *implies exact equivalence under* $\mathcal{Q}_1$*.*

Proof. The maximum over $\mathcal{Q}_1$ cannot exceed the maximum over $\mathcal{Q}_2$. Zero distance under the larger family therefore entails zero distance under the smaller family. □

Corollary 6.3 (No probe-free developmental topology)**.** *Because $X_{\mathcal{Q}}$ and hence its homology depend on $\mathcal{Q}$ through Equation* (15)*, the phrase "the topology of development" denotes no well-defined object until a probe family is named. Two studies reporting different developmental topologies may both be correct and simply be reporting different representands.*

6.2 Simplicial Filtration and Persistent Invariants

For a finite sample $V = \{[\gamma_1], \ldots, [\gamma_n]\} \subset X_{\mathcal{Q}}$ and scale $r \geq 0$, define

$$\mathrm{VR}_d(V; r) := \{\sigma \subseteq V : \mathrm{diam}_d(\sigma) \leq r\}\,, \tag{16}$$

where d is the quotient metric. The inclusions $\mathrm{VR}_d(V; r) \hookrightarrow \mathrm{VR}_d(V; s)$ for $r \leq s$ induce persistence modules $H_k(\mathrm{VR}_d(V; r); \mathbb{F})$ over a declared coefficient field $\mathbb{F}$.

Corollary 6.4 (Probe-family filtration ordering)**.** *For $\mathcal{Q}_1 \subseteq \mathcal{Q}_2$ and a fixed labeled history sample $V_\Gamma = \{\gamma_1, \ldots, \gamma_n\} \subset \Gamma$, apply the same diameter-threshold construction to the two response pseudometrics. Then*

$$\mathrm{VR}_{d_{\mathcal{Q}_2}}(V_\Gamma; r) \subseteq \mathrm{VR}_{d_{\mathcal{Q}_1}}(V_\Gamma; r). \tag{17}$$

Proof. Equation (15) bounds every pairwise distance under $d_{\mathcal{Q}_1}$ by its counterpart under $d_{\mathcal{Q}_2}$. Every simplex admitted under the latter metric therefore satisfies the former metric's threshold. □

Claim 6.5 (Licensed response-topology statement)**.** A persistence diagram derived from Equation (16) summarizes multiscale connectivity and holes in the sampled intervention–response representation. Its direct referent is the declared response quotient. Interpretation as linguistic organization, developmental capacity, subject continuity, or substrate combinatorics requires a separately supported bridge.

The construction differs from computing homology on raw finite records. A finite set with its subspace topology is frequently discrete, whereas Equation (16) adds a metric-dependent family of complexes. Persistence retains the scale history of this family; it preserves the analyst's representational commitments rather than eliminating them.

6.3 Metric Error and Longitudinal Contrast

Suppose d and $\widehat{d}$ are metrics on the same labeled finite set V and

$$\max_{v,w \in V} |d(v,w) - \widehat{d}(v,w)| \leq \delta. \tag{18}$$

Proposition 6.6 (Finite response-persistence stability)**.** *Under Equation* (18)*, the two Vietoris–Rips filtrations are δ-interleaved through the identity vertex map. For finite persistence modules over a field,*

$$d_{\mathrm{B}}\Big(\mathrm{Dgm}_k(d), \mathrm{Dgm}_k(\widehat{d})\Big) \leq \delta \tag{19}$$

for every homological degree k.

Proof. If $\sigma \in \mathrm{VR}_d(V; r)$, every pair of its vertices has $\widehat{d}(v,w) \leq d(v,w) + \delta \leq r + \delta$, hence

$$\mathrm{VR}_d(V; r) \subseteq \mathrm{VR}_{\widehat{d}}(V; r + \delta). \tag{20}$$

Interchanging d and $\widehat{d}$ gives the reverse shifted inclusion. The identity maps commute with filtration inclusions and therefore yield a δ-interleaving. Algebraic stability for finite persistence modules then gives Equation (19); the result is also the same-labeled-set specialization of geometric-complex stability [3]. □

For observation references χ and χ', define the true diagram displacement

$$\Delta_k(\chi, \chi') := d_{\mathrm{B}}(\mathrm{Dgm}_k(d_\chi), \mathrm{Dgm}_k(d_{\chi'})) \tag{21}$$

and define $\widehat{\Delta}_k(\chi, \chi')$ by replacing both metrics with their estimates. If the two reference-specific uniform errors are δ_χ and $\delta_{\chi'}$, the triangle inequality yields

$$\left|\widehat{\Delta}_k(\chi, \chi') - \Delta_k(\chi, \chi')\right| \le \delta_\chi + \delta_{\chi'}. \tag{22}$$

Corollary 6.7 (Error-qualified developmental displacement)**.** *A measured longitudinal displacement exceeding $\delta_\chi + \delta_{\chi'}$ cannot be explained solely by the declared uniform metric-error bounds. Attribution to learning, maturation, intervention, or subject reorganization still requires design-based comparison with those candidate processes.*

Equations (19) and (22) establish analytic robustness. They do not establish that the selected outcomes, probes, probability metric, quotient, or homological degree measures the intended construct. That obligation belongs to construct validation.

7 Universal Grammar as a Developmental Invariant

This section treats the proposal that the capacity for structured language is a topological invariant of a subject's initial organization. It states the proposal as a conjecture inside the substrate, lists the obligations that would establish it, and records the obstruction that already stands against its most natural form.

The proposal is attractive for a specific reason. Statistical accounts of an initial linguistic capacity face a difficulty: an average over a population's linguistic behavior is not available to an individual who has encountered no language, so a shared initial capacity cannot be an inherited statistic. An inherited structural property that is preserved under the deformations constituting learning would avoid that difficulty. Topological invariants are the natural candidate class because they are exactly the properties preserved under continuous deformation.

Definition 7.1 (Developmentally invariant substrate feature)**.** Let $\mathcal{I}$ be a function on subcomplexes. Then $\mathcal{I}$ is developmentally invariant relative to a declared class $\mathcal{D}$ of admissible continuations when

$$\mathcal{I}(\mathcal{C}_I(\chi_1)) = \mathcal{I}(\mathcal{C}_I(\chi_0)) \qquad \text{for every } \mathcal{C}_I(\chi_0) \rightsquigarrow \mathcal{C}_I(\chi_1) \text{ in } \mathcal{D}. \tag{23}$$

Conjecture 7.2 (Invariant-capacity conjecture)**.** There exists a nontrivial developmentally invariant feature $\mathcal{I}$ of the initial subsystem complex, relative to the class of continuations realized in ordinary development, such that the value of $\mathcal{I}$ predicts the capacity to acquire any of a declared family of natural languages, and such that a declared alteration of $\mathcal{I}$ predicts a corresponding loss of that capacity.

Conjecture 7.2 is recorded as a conjecture and is neither assumed nor used anywhere else in this paper. Establishing it would require, at minimum: a representation of linguistic capacity as a property of $\mathcal{C}_I$ rather than of behavior alone; a proof that the declared continuation class preserves $\mathcal{I}$, since Definition 7.1 is a constraint on that class and not a consequence of topology; an estimator for $\mathcal{I}$ with uncertainty, which Proposition 3.2 shows cannot be obtained from accessible response data alone; a discriminating prediction against graph-theoretic, geometric, dynamical, and symbolic accounts of the same capacity; and a treatment of individual variation, since the conjecture asserts uniformity of $\mathcal{I}$ across a population whose bare organization differs in every other respect.

Claim 7.3 (Obstruction to the natural form)**.** The most natural instance of Conjecture 7.2, in which $\mathcal{I}$ is a fundamental group and the predicted capacity is recursion, is already obstructed. Proposition 5.2 shows that nontrivial fundamental group is neither necessary nor sufficient for recognition of the one-bracket Dyck language in the declared realization classes. A defender of the conjecture must therefore either select a different invariant, restrict the realization class in a motivated way, or supply a map from homotopy classes to grammatical operations that the separation result does not defeat.

Related developmental vocabulary inherits the same obligations. Terms such as *pruning*, *critical period*, and *topological surgery* describe changes that would move a system outside the invariance class of Definition 7.1, and they acquire topological force only when the associated representand and invariant are measured across the transition. The framework does not forbid such claims; it prices them.

8 Linguistic and Developmental Model Obligations

This section translates the formal results into target-specific obligations. It organizes recursion, categorization, cross-language comparison, developmental transition, and social interaction as distinct modeling tasks with separate objects and empirical bridges.

Table 2 records the minimum structures required for five frequent attribution patterns. The separation results limit inference; they also indicate the additional representation theorem or experiment that could turn a heuristic into a model.

Corpus networks, semantic embeddings, neural response manifolds, computational configurations, and social-interaction graphs can all support valuable structural analysis [9, 14, 5, 15]. Their vertices, neighborhoods, metrics, and nuisance transformations differ. A cross-register conclusion therefore needs a map between representations and a statement of which structure that map preserves.

The obligations permit affirmative research hypotheses. A response complex may predict transfer after controlling for conventional accuracy and complexity. A reference-local persistence change may forecast a later change in generalization. A cross-language alignment may preserve a restricted semantic neighborhood structure while acquisition dynamics remain language-specific. Each hypothesis names an observable advantage over matched graph, geometric, dynamical, and symbolic baselines. Developmental field models and schedule-sensitive learning systems show that formal developmental claims can be tested through perturbation, simulation, and behavioral prediction [23, 8]; a topological developmental model inherits those obligations and adds a declared topology-bearing construction.

Table 2: Target-specific obligations for topology-bearing claims.

Target attribution	Required object	Separation result	Evidence obligation
Recursive capacity	formal-language operation, computational state, and realization map	D_1 recognition and ambient π_1 vary independently (Prop. 5.2)	representation theorem plus processing and generalization predictions
Category organization	response regions, labels, or decision complex	a contractible interval supports any finite label count (Prop. 5.3)	category-sensitive invariant with out-of-sample behavioral or neural validation
Cross-language structure	aligned linguistic units, metric, and allowable maps	carrier homeomorphism leaves dynamics and observables open (Prop. 5.4)	alignment uncertainty, matched baselines, and predictions beyond distributional fit
Developmental transition	reference-indexed representand, correspondence, and transition rule	bifurcation can occur on fixed ambient topology (Claim 5.5)	before–after estimation, uncertainty, process model, and rival-transition comparison
Substrate organization	subsystem complex, coarse-graining, and preservation condition	bulk topology is boundary-invisible (Prop. 3.2) and invariants do not transfer (Prop. 4.2)	minimality or intervention assumption plus a preservation proof for the declared coarse-graining
Social organization	agents, relations, interaction protocol, and scale	network cycles and cognitive-state homotopy concern different carriers	multi-level bridge, intervention target, and sensitivity to edge construction

9 Subject-Level Attribution Boundary

This section locates the positive response topology within a broader theory of subject development. It defines the limited organization represented by the quotient and specifies the additional continuity, agency, embodiment, and epistemic conditions needed for subject-level interpretation.

Definition 9.1 (Probe-relative developmental organization)**.** A *probe-relative developmental organization* is an equivalence class $[\gamma]_{\mathcal{Q}} \in X_{\mathcal{Q}}$. Its identity criterion is equality of the response laws in Equation (12) over the declared probe family.

This definition supplies an operational comparison unit. It can register persistent, transferable, and intervention-sensitive differences among histories. It leaves open the internal structures that produce those differences and can merge systems that diverge under an untested context. Experimental resolution and ontological individuation therefore remain distinct. A longitudinal representation may use a reference-indexed path $\chi \mapsto [\gamma_\chi]_{\mathcal{Q}}$ only after checkpoint correspondence and probe harmonization have been defined.

Claim 9.2 (Subject-attribution boundary)**.** Persistent homology estimated from a declared finite sample and filtration in $X_{\mathcal{Q}}$ supports a claim about organization of probe-relative response dispositions. Attribution to one continuing subject additionally requires a defensible boundary, cross-reference correspondence, counterfactual integration across contexts, and a theory connecting the measured organization to agency or experience.

Several bridges remain available for investigation. Embodied continuity can link response histories to a persisting physical system. Causal integration can test whether candidate components jointly sustain action. Cross-context transfer can distinguish local task fit from wider organization. Self-model and report measures can inform higher-order attribution when their validity is independently assessed. Phenomenal subjecthood demands further philosophical and empirical commitments that homology alone cannot provide.

Epistemic attribution requires another layer. Initial organization can shape learning while lacking acquired content, and a durable response pattern can be accurate through artifact or environmental scaffolding. Candidate knowledge therefore requires acquisition history, retention, transfer, carrier-sensitive intervention, and content-appropriate accuracy or understanding criteria. A topological invariant can summarize the candidate organization; it does not supply truth, justification, or understanding by itself.

10 Comparative Empirical Program

This section converts the warrant architecture into a staged empirical program. It pairs each evaluation stage with a design commitment, an admissible conclusion, and a failure signal, then specifies comparative and ethical controls.

Table 3 places representation validation before target-level attribution. Advancement through the stages accumulates evidence without converting an earlier mathematical guarantee into a later semantic or causal guarantee.

A first implementation can use standardized, low-risk probes, repeated contexts, and distributional outcomes rather than point scores. Pairwise distances should be estimated with uncertainty intervals or posterior distributions, and the analysis should report sensitivity to the probability metric, filtration, coefficient field, homological degree, and scale range. Proposition 6.6 then converts a justified uniform metric-error bound into a diagram-error bound. Corollary 6.3 adds a reporting requirement that is new in this version: the probe family is part of the result and must appear in the statement of any developmental topology, not only in the methods.

Comparative evaluation should match effective flexibility and information. Relevant baselines include a response kernel without topology, graph summaries, Euclidean or manifold geometry, state-space dynamics, and a task-specific symbolic model. Negative controls can add observationally silent coordinates, permute irrelevant labels, or replace the proposed bridge while preserving marginal fit. Such controls expose representational features that the target outcome cannot identify, and they are the empirical counterpart of Propositions 3.2 and 5.1.

Longitudinal studies require a correspondence policy for persons, tasks, outcomes, and contexts. Cross-language work additionally requires alignment uncertainty and language-specific measurement checks. Hierarchical analyses can retain individual variation while estimating population regularities; averaging that erases distinct developmental routes should be reported as a model limitation.

Ethical constraints shape the admissible intervention family $\mathcal{U}$. Benign instructional variation, interface perturbation, and reversible task changes can support causal tests. Developmental deprivation, harmful social manipulation, invasive alteration, and privacy-compromising response capture remain outside an acceptable design. Observational alternatives need explicit confounding assumptions and weaker causal language.

Table 3: Evaluation stages for a developmental response-topology model.

Evaluation stage	Design commitment	Admissible description	Failure signal
Representand validation	preregister outcomes, probes, contexts, units, and nuisance transformations	declared functional response representation	conclusions reverse under equally plausible encodings
Probe-family reporting	declare $\mathcal{Q}$ and report sensitivity to its enlargement	probe-indexed developmental topology	reported topology changes qualitatively under a defensible probe extension
Response estimation	identify interventions or state assumptions for history-conditioned response laws	estimated pairwise response distance	positivity failure, unstable estimates, or unbounded confounding sensitivity
Metric robustness	propagate sampling and model uncertainty into pairwise distance bounds	error-qualified persistence diagram	topological signal falls within uncertainty or filtration-choice envelope
Predictive comparison	compare symbolic, graph, geometric, dynamical, and topological models on held-out outcomes	incremental predictive or compressive contribution	matched baselines equal or exceed performance
Causal discrimination	perturb a hypothesized carrier under an identified design; prespecify any mediated-effect estimand	identified joint response or mediated-effect estimate	target changes without predicted topology change, or topology changes without target response
Longitudinal replication	align units and protocols across references, cohorts, and sites	reproducible developmental displacement	effect depends on one alignment, cohort, reference grid, or analytic degree

11 Limitations and Open Obligations

This section records the framework's unresolved mathematical, empirical, and philosophical dependencies. The open problems identify concrete revision targets and prevent the positive construction from carrying broader attributions by rhetorical momentum.

Open Problem 11.1 (Boundary-silent modifications)**.** Characterize the connected modifications of a subcomplex that leave the accessible state on $\mathfrak{A}_I$ unchanged. Proposition 3.2 uses only a disconnected closed component; a complete characterization would determine how much substrate topology is identifiable in principle and would supply the substrate analogue of a minimal-realization theorem.

Open Problem 11.2 (Invariant-preserving coarse-graining)**.** Characterize the coarse-grainings Λ for which a declared invariant satisfies $I(\Lambda(\mathcal{C}_I)) = I(\mathcal{C}_I)$. Proposition 4.2 shows the general failure; a sufficient condition would license substrate-level conclusions from effective-level measurements for the first time.

Open Problem 11.3 (Representational minimality)**.** Characterize conditions under which an intervention–response realization is minimal or observable up to a declared equivalence. Such a result would limit output-silent topological factors and clarify which latent invariants are identifiable from the protocol.

Open Problem 11.4 (Longitudinal correspondence)**.** Develop estimators that jointly represent unit alignment, changing probe families, attrition, measurement drift, and diagram uncertainty. The desired object is a developmental comparison whose meaning survives protocol change.

Open Problem 11.5 (Linguistic bridge)**.** Specify a map from an algebraic-topological feature to a linguistic operation or capacity, together with a representation theorem, acquisition prediction, and comparison against symbolic, geometric, and dynamical alternatives. A successful construction would also have to defeat Claim 7.3.

Open Problem 11.6 (Enriched structural register)**.** Determine when topology alone is too coarse and which added structures—a measure, labels, dynamics, stratification, sheaf, category, or causal model—are required. Every enrichment introduces its own estimation and validation obligations.

Open Problem 11.7 (Subject-level bridge)**.** Relate probe-relative response organization to a justified theory of subject boundary, persistence, agency, and experience. The bridge must discriminate organizational similarity from subject identity and from epistemic status.

The countermodels use deliberately simple realization classes. The Dyck example isolates formal nesting and does not capture the full syntactic, semantic, pragmatic, or resource-bounded character of human language. The silent-factor construction establishes unrestricted input–output underdetermination; observability, minimality, mechanistic measurement, or intervention can reduce that underdetermination in a specified model class. The boundary-invisibility construction is correspondingly simple and is stated with its own open problem.

Finite-sample stability also has a narrow scope. It protects persistence diagrams against a declared metric perturbation on a common labeled set. It does not control representand choice, sample selection, missingness, cross-person correspondence, metric misspecification, or the semantic bridge. Manifold and homology recovery theorems can strengthen inference when their sampling and regularity assumptions are independently defensible [16].

Topology may contribute descriptive compression, prediction, mechanism, or causal explanation in different applications. These roles should be stated separately. A robust descriptor can remain causally downstream, and a causal process can act on variables whose topology is descriptively uninformative.

12 Revisable Position

This section consolidates the paper's limited commitments and states the conditions under which they should be revised. It separates established mathematical consequences, proposed methodology, and open substantive claims.

Adopting a substrate in which topology is constitutive changes the question without answering it. Within the generated complexes, topological statements are ordinary mathematics and no longer candidates for the charge of metaphor. What the substrate supplies in literalness it withholds in identifiability: complexes differing by a closed component induce identical accessible states while differing in every Betti number, so a system can carry real topology that nothing available to it or to an experimenter working through its interface records. Invariants also fail to transfer across coarse-graining in both directions. The consequence is a two-level warrant architecture in which the substrate representand, the coarse-graining, and its preservation properties join the representation map, scale, transformation class, invariant, estimator, uncertainty, target bridge, and comparison family as declared parts of any topological model.

The retained effective-level results are conditional and representation-relative. A declared intervention–response family induces a pseudometric; its metric quotient supports a Vietoris–Rips filtration; uniform pairwise metric error yields the stated persistence bound; and the quotient depends constitutively on the probe family, so a developmental topology is not well defined until that family is named. Conversely, the silent-factor model changes latent fundamental group while preserving admitted behavior; nontrivial fundamental group is neither necessary nor sufficient for the declared D_1 recognition criterion; ambient Betti numbers leave category count open; carrier homeomorphism leaves dynamical conjugacy open; and a bifurcation need not change the topology of the declared ambient state space.

The proposal that a shared linguistic capacity is a topological invariant of initial organization is preserved as an explicitly labelled conjecture together with its obligations and its existing obstruction. It is neither adopted nor discarded, and the framework's contribution to it is to state exactly what a defense would have to establish.

The substantive possibilities remain open. Topological organization may prove useful for some neural, semantic, interactional, developmental, or substrate-level representations. Stronger evidence may justify enriched or causal claims, and future minimality or preservation results may narrow the underdetermination established here. The framework should be revised when a better representation chain, a sharper identification theorem, or discriminating empirical results become available.

References

[1] Enrico Betti. Sopra gli spazi di un numero qualunque di dimensioni. *Annali di Matematica Pura ed Applicata*, 4:140–158, 1870. Springer issue metadata gives 1870; some historical catalogues list 1871.

[2] D. V. Boulatov. A model of three-dimensional lattice gravity. *Modern Physics Letters A*, 7(18):1629–1646, 1992.

[3] Frédéric Chazal, Vin de Silva, and Steve Oudot. Persistence stability for geometric complexes. *Geometriae Dedicata*, 173:193–214, 2014.

[4] Noam Chomsky. On certain formal properties of grammars. *Information and Control*, 2(2):137–167, 1959.

[5] SueYeon Chung, Daniel D. Lee, and Haim Sompolinsky. Classification and geometry of general perceptual manifolds. *Physical Review X*, 8(3):031003, 2018.

[6] Lee J. Cronbach and Paul E. Meehl. Construct validity in psychological tests. *Psychological Bulletin*, 52(4):281–302, 1955.

[7] Herbert Edelsbrunner, David Letscher, and Afra Zomorodian. Topological persistence and simplification. *Discrete & Computational Geometry*, 28:511–533, 2002.

[8] Jeffrey L. Elman. Learning and development in neural networks: The importance of starting small. *Cognition*, 48(1):71–99, 1993.

[9] Ramon Ferrer i Cancho and Ricard V. Solé. The small world of human language. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1482):2261–2265, 2001.

[10] Laurent Freidel. Group field theory: An overview. *International Journal of Theoretical Physics*, 44(10):1769–1783, 2005.

[11] Peter Gärdenfors. *Conceptual Spaces: The Geometry of Thought*. MIT Press, Cambridge, MA, 2000.

[12] Razvan Gurau and James P. Ryan. Colored tensor models — a review. *Symmetry, Integrability and Geometry: Methods and Applications*, 8:020, 2012.

[13] Mary B. Hesse. *Models and Analogies in Science*. University of Notre Dame Press, Notre Dame, IN, revised edition, 1966.

[14] Nikolaus Kriegeskorte and Rogier A. Kievit. Representational geometry: Integrating cognition, computation, and the brain. *Trends in Cognitive Sciences*, 17(8):401–412, 2013.

[15] Molly Lewis, Aoife Cahill, Nitin Madnani, and James Evans. Local similarity and global variability characterize the semantic space of human languages. *Proceedings of the National Academy of Sciences*, 120(51):e2300986120, 2023.

[16] Partha Niyogi, Stephen Smale, and Shmuel Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete & Computational Geometry*, 39(1–3):419–441, 2008.

[17] Daniele Oriti. Group field theory as the second quantization of loop quantum gravity. *Classical and Quantum Gravity*, 33(8):085005, 2016.

[18] Judea Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge, 2 edition, 2009.

[19] Alejandro Perez. The spin-foam approach to quantum gravity. *Living Reviews in Relativity*, 16:3, 2013.

[20] Henri Poincaré. Analysis situs. *Journal de l'École polytechnique*, 1:1–121, 1895.

[21] Michael P. Reisenberger and Carlo Rovelli. Spacetime as a Feynman diagram: The connection formulation. *Classical and Quantum Gravity*, 18(1):121–140, 2001.

[22] Gurjeet Singh, Facundo Mémoli, and Gunnar Carlsson. Topological methods for the analysis of high dimensional data sets and 3d object recognition. In Mario Botsch, Renato Pajarola, Baoquan Chen, and Matthias Zwicker, editors, *Eurographics Symposium on Point-Based Graphics*, pages 91–100. The Eurographics Association, 2007.

[23] Esther Thelen, Gregor Schöner, Christian Scheier, and Linda B. Smith. The dynamics of embodiment: A field theory of infant perseverative reaching. *Behavioral and Brain Sciences*, 24(1):1–34, 2001. Target article; peer commentary and authors' response continue on pages 34–86.