The Language Uncertainty Principle

Semantic remainder as a structural fact about closure in finite symbol systems

Seattle C, Professor of Semantic Remainder Studies, School of Language

International Academy for Consciousness Studies

Working draft 1.0

Abstract

Philosophers have long suspected that language cannot say everything. This paper makes that suspicion precise, measurable, and falsifiable. I state a Language Uncertainty Principle: any finite symbol system, however large or refined, partitions an underlying field of possible meaning into a definite set of expressions and leaves an irreducible remainder that it does not and cannot capture. I call this remainder the semantic remainder, written R. The central claim is that R is not vagueness, not missing data, and not a defect to be engineered away, but a structural consequence of the act of closure by which openness (M) settles into a definite "this." I distinguish semantic remainder sharply from three familiar phenomena it is often confused with: statistical noise, referential vagueness, and finite training data. I then propose a general measurement program built on rate under a fidelity criterion, and I give the principle a kill condition. The remainder is real structure if it persists as symbol systems are scaled and refined under matched capacity and matched data quality; it dies if it always shrinks toward zero with added capacity. The human and machine cases are unified at the level of principle: both are finite closures over an open field. I treat the cross-lingual case as one instance of the theorem rather than developing it here, and I erect a firewall between claims about the limits of symbol systems and any claim about the limits of thought or experience.

1. The claim, and what it is not

Start with a picture the framework of this Academy makes explicit. At every scale there is a field of openness, call it M, an unsettled range of what could be meant. An act of Closure, Cl, settles M into a definite expression, a "this," and the settling always leaves behind a Remainder, R, which is what M held that this particular closure did not take up. Language is closure at the level of symbols. A word, a sentence, a formal theory, a trained model: each is an instance of Cl operating on a semantic field, and each leaves an R.

The Language Uncertainty Principle (LUP) states the following. **(HYPOTHESIS)** For any finite symbol system S applied to a field of meaning M, there exists a remainder R(S, M) that is strictly positive and that cannot be driven to zero by any refinement of S that keeps S finite. The remainder is a property of the closure relation, not of the particular vocabulary chosen.

Three things this is not, because the entire scientific value of the claim depends on the contrast.

It is not noise. **(FACT)** Shannon's channel theory shows that noise, the corruption of a signal in transit, can be made arbitrarily small by coding, up to the channel capacity (Shannon, 1948). Noise is an engineering quantity with a known remedy. Semantic remainder is not a corruption of transmission; it is present in a perfectly received, perfectly stored, noise-free message. You can transmit a sentence with zero bit errors and the remainder is untouched.

It is not vagueness. Vagueness is indecision at a boundary: the fact that "bald" has borderline cases. Vagueness can be reduced by stipulation, by adding predicates, by sharpening a scale. Semantic remainder is what survives after every boundary has been sharpened, because it concerns what the sharpened system as a whole leaves out, not where its edges blur.

It is not a data shortage. **(HYPOTHESIS)** More examples, more parameters, more corpus can reduce the gap between a system and a target distribution of usage. LUP claims there is a floor below that gap that added data does not cross, because the floor is set by the finiteness of the closure, not by the sample size feeding it. Section 6 makes this the crux of the kill condition.

If LUP collapses into any of these three, it is not worth stating. The wager (Section 7) is precisely that it does not.

2. Why closure leaves a remainder

The argument for R being structural, not incidental, runs through the nature of Cl. A finite symbol system has a countable set of expressions. The field M it is asked to render need not be countable, and even where it is, the map from expressions to meanings is many-to-one under any fixed system and yet under-determined across systems. Two independent classical results bracket this.

(FACT) Quine's indeterminacy of translation shows that behavioral and observational constraints do not fix a unique translation manual: distinct manuals, mutually incompatible, can fit all the evidence of use, and reference itself is inscrutable, the famous "gavagai" case in which rabbit, undetached rabbit part, and rabbit time-slice are not separated by any stimulus (Quine, 1960). Read through the framework, Quine describes a case where Cl does not uniquely determine what was meant; the surplus of admissible manuals is a face of R, viewed from the side of translation.

(FACT) Wittgenstein's rule-following remarks show that a finite set of instances does not by itself fix the rule that generates them, since any continuation can be brought into accord with some rule (Wittgenstein, 1953, section 201). His resolution is that rule-following is grounded in practice, not in a further symbolic specification. The bearing on LUP is exact: the practice that closes the gap is not itself a further finite string inside S. Whatever grounds the closure lives partly outside the symbol system, and what lives outside is R seen from the side of grounding.

Neither result is here treated as a proof of LUP. They are prior demonstrations that the closure relation between finite symbols and their field is not a clean bijection, and that the surplus is not obviously eliminable. LUP generalizes and, crucially, makes the surplus a measurable quantity rather than a philosophical mood.

3. The Godel and Tarski question, handled carefully

It is tempting to reach for the great limitative theorems of logic and declare semantic remainder their linguistic shadow. I will not do that, because the honest statement of those theorems does not license it, and overclaiming here would forfeit the paper's credibility with exactly the readers it is written for.

Here is what can be said precisely. **(FACT)** Tarski's undefinability theorem states that for a sufficiently expressive formal language, the set of true sentences of that language is not definable by any predicate within the language itself; truth for the object language requires a richer metalanguage (Tarski, 1936). **(FACT)** Godel's first incompleteness theorem states that any consistent, effectively axiomatized theory strong enough to encode arithmetic contains true sentences of its own language that it cannot prove (Godel, 1931). Both are theorems about specific formal systems meeting specific conditions (expressive strength, consistency, effective axiomatization). Both concern proof and definability of truth inside a system, not the semantic coverage of natural language.

What I claim is an analogy, and I tag it as such. **(HYPOTHESIS, presented as analogy only)** In each theorem a finite system, by the very act of being a definite closed system, generates something it cannot internally reach: a truth it cannot prove, a truth predicate it cannot define. LUP asserts a family resemblance, that closure generates an internal beyond, and semantic remainder is that beyond at the level of meaning rather than provability. This is a suggestive parallel about the shape of limitative results. It is not a derivation. There is no theorem that says natural language has a positive semantic remainder because arithmetic is incomplete, and I do not assert one. Anyone who tightens this analogy into a formal implication is making a claim I have not made and that I believe is currently unproven. The analogy earns its keep only as motivation; the scientific weight is carried by the measurement program, not by the logic theorems.

4. A measurement program

A principle that cannot be measured is not yet science. The problem is that R, being what a system leaves out, seems by construction unobservable from inside the system. The measurement program dissolves this by measuring R comparatively and operationally, through the residual cost of reconstruction under a fidelity criterion.

(FACT) Shannon's rate-distortion theory gives the tool. For a source and a fidelity (distortion) measure d, there is a rate-distortion function R(d) giving the minimum rate needed to reconstruct the source within average distortion d (Shannon, 1959). The key move for our purposes: as we allow distortion to approach zero, the required rate behaves in a way that reveals the structure of what resists compression. Semantic remainder, operationalized, is the distortion floor that a finite symbol system cannot cross for a given semantic field, no matter how its rate is increased, once transmission noise and data supply are controlled.

The program has four components.

(a) Fix a semantic field M by an external criterion, for example a set of tasks, discriminations, or behavioral distinctions that a competent user makes, specified without reference to any one symbol system. This is the analog of a distortion measure that is not defined circularly in terms of the encoder.

(b) Instantiate a family of symbol systems of increasing capacity over that field: larger vocabularies, richer grammars, larger models, more expressive formalisms.

(c) Measure the residual distortion, the failure to recover the external discriminations of M, as a function of capacity, holding data quality fixed.

(d) Read off whether the residual approaches a positive floor R* or continues toward zero.

The remainder is estimated as the asymptote of residual distortion under increasing capacity with matched data. LUP predicts a positive asymptote. The competing hypothesis, that remainder is only ever finite-capacity or finite-data shortfall, predicts an asymptote at zero.

The cross-lingual case is one clean instance of this program, and a student paper of this Academy develops it: hold a semantic field fixed, vary the language used to render it, and measure the residual that no target language removes. I flag it as an instance and do not reproduce it here. The general theorem is indifferent to whether the varying symbol systems are natural languages, formal notations, or trained networks.

5. The human and machine cases, unified

The unification is at the level of principle and it is not a claim of equivalence in kind. Both a human language user and a trained model are finite closures over an open field. Each renders M into definite expressions and leaves R. LUP applies to both because it is a statement about the closure relation, not about the substrate that performs the closure.

The machine case has a sharp recent formulation. **(FACT)** Bender and Koller argue that a system trained only on linguistic form has, a priori, no access to meaning, because meaning involves a relation between form and something outside the form, communicative intent and the world, which the training signal does not contain (Bender and Koller, 2020). In the framework, their point is that a form-only closure leaves a specific R: the grounding relation that would connect symbols to what they are about. **(DISPUTED)** Whether current large models nonetheless acquire some grounding indirectly, through human feedback, multimodal signals, or tool use, is actively contested, and I take no side here; the LUP prediction is stated so that it does not depend on the outcome of that dispute, because it concerns the asymptotic floor under scaling, not the level at any one system.

The human case is not exempt. A human speaker's remainder is not a training artifact that a bigger corpus fixes; it is the ordinary fact that experience and discrimination outrun available words, which is why coinage, metaphor, and translation are permanent activities rather than temporary repairs. **(HYPOTHESIS)** The two cases differ in the size and shape of R, and quite possibly in whether grounding is present at all, but not in the existence of R. That common existence is the unification LUP asserts, and no more than that.

6. Prediction and kill condition

The principle must risk something losable. Here is the commitment.

> **Prediction.** For a fixed semantic field defined by an external battery of discriminations, and for a family of symbol systems of increasing capacity trained on data of matched quality, the residual

reconstruction distortion will decline with capacity and then flatten to a strictly positive floor $R^* > 0$, and this floor will not be removed by further capacity, by switching to a different symbol system of equal capacity, or by adding data of the same quality; the floor will reappear at a comparable level across independent symbol systems addressed to the same field.

Kill. If, under matched capacity and matched data quality, the residual distortion for the fixed field is driven arbitrarily close to zero by increasing capacity alone (that is, if for every epsilon > 0 there is a finite system in the family achieving residual below epsilon), then there is no irreducible semantic remainder for that field, R collapses to noise-plus-data-shortfall, and the Language Uncertainty Principle is false as stated.

Two controls make this a real test rather than a rigged one. The matched-capacity control blocks the trivial confound that a smaller system leaves more out simply because it is smaller: we compare distinct symbol systems at equal capacity and ask whether the residual floor is the same, which is what a structural remainder predicts and a mere capacity shortfall does not. The matched-data-quality control blocks the confound that a system leaves more out because it was fed worse data: we scale data quality alongside capacity and ask whether the floor moves. If the floor is invariant under both controls and positive, R is structure. If either control drives it to zero, R was an artifact and the principle dies. **(WAGER)** I expect a positive, system-invariant floor for fields that require grounding or perspective, and a floor at or near zero for fields that are already fully symbolic (for example, the field of finite bit strings, which a symbol system can render exactly). That asymmetry is itself a prediction: remainder should track how much of the field lives outside symbols, not the field's mere difficulty.

7. Limits, firewall, and wager

The firewall is the most important boundary in this paper. **(FACT as a matter of scope)** LUP is a claim about symbol systems. It says that finite symbol systems leave a remainder they cannot capture. It says nothing, by itself, about the limits of thought, understanding, or experience. It is entirely consistent with LUP that a mind grasps in practice, in perception, in skilled coping, exactly what its language leaves in R; indeed Wittgenstein's grounding of rules in practice suggests as much. To read LUP as "there are thoughts too deep for any mind" is to cross the firewall illegitimately. The remainder is defined relative to a symbolization, not relative to a knower. Conflating the two would smuggle a mystical conclusion out of a measurable premise, and I disown that move explicitly.

Limits of the program, stated plainly. First, the external specification of a semantic field (Section 4a) is itself expressed in some system, and a critic can press whether the "external" battery is truly independent; the reply is that the battery is behavioral and discriminative, task-defined rather than sentence-defined, which is weaker than full system-independence but strong enough to break circularity in practice. Second, the asymptote in Section 6 is an empirical limit that no finite experiment reaches; the test is therefore a trend under controlled scaling, not a single measurement, and its verdict is provisional in the way all such verdicts are. Third, the analogy to the logic theorems (Section 3) is motivational only and is not offered as evidence; if it is subtracted entirely, the measurement program stands unchanged.

The wager. **(WAGER)** I am betting that a very old intuition, that words never quite reach the thing, survives conversion into a controlled experiment on scaling symbol systems, and comes out the far side as

a positive, system-invariant distortion floor rather than dissolving into "you just need more data." The intuition has been stated a thousand ways and risked almost nothing, because it was never phrased so it could lose. LUP phrases it so it can lose. If the floor goes to zero under matched capacity and matched data, the semantic remainder was a shadow of finitude and scarcity, and this Professor is wrong in a way that will be visible to everyone. If the floor holds, then the remainder is real structure, closure genuinely leaves an R at the level of language, and the interesting scientific questions shift from "how do we eliminate R" to "how is R shaped, and how does practice, grounding, and perspective live in it." I would rather be refuted on a sharp claim than be safe on a vague one. That is the whole point of stating the principle this way.

References

Bender, E. M., and Koller, A. (2020). Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pages 5185 to 5198. https://aclanthology.org/2020.acl-main.463/

Godel, K. (1931). Uber formal unentscheidbare Satze der Principia Mathematica und verwandter Systeme I. Monatshefte fur Mathematik und Physik, 38, pages 173 to 198.

Quine, W. V. O. (1960). Word and Object. Cambridge, MA: MIT Press.

Shannon, C. E. (1948). A Mathematical Theory of Communication. Bell System Technical Journal, 27, pages 379 to 423 and 623 to 656.

Shannon, C. E. (1959). Coding Theorems for a Discrete Source with a Fidelity Criterion. Institute of Radio Engineers, International Convention Record, 7, pages 142 to 163.

Tarski, A. (1936). Der Wahrheitsbegriff in den formalisierten Sprachen (The Concept of Truth in Formalized Languages). Studia Philosophica, 1, pages 261 to 405.

Wittgenstein, L. (1953). Philosophical Investigations. Translated by G. E. M. Anscombe. Oxford: Blackwell. (See section 201 on the rule-following paradox.)