

- [Framework](#)
- [Papers](#)
- [Predictions](#)
- [News](#)
- [ATLAS](#)
- [Articles](#)
- [About](#)



The Bandwidth Argument: Consciousness as Serial Narration over a Parallel Substrate

A development of the “tokenized trap” thesis — resolving its central contradiction with measured information rates

Z. Paz · [ORCID 0009-0003-1690-3669](#) V2.0 2026

↓ [DOWNLOAD PDF](#)

Version 2, revised after adversarial review by an independent reader given the paper alone, with the author’s reflexive objection to that review incorporated. Changes from v1: the estimator inconsistency in §2.3 is repaired; the access/constitution slide is closed by an explicit premise (§3, P4); the identity claim is restated as a convention under an evidential horizon (§3.3); the escape hatches are rewritten operationally (§3.2); the machine analogy is corrected by the sufficient-statistic asymmetry (§5); ancestry is credited exactly (§6); a fifth falsifier is added (§8); the quantum notation costume is removed throughout.

Abstract

The tokenized-trap thesis, as developed in prior sessions, holds two positions that cannot both be true. The first says consciousness is nothing but a recursive linguistic loop — a definition, a consequence of language tokens reading themselves. The second says language is a lossy compression of a vastly richer cognitive state, with the Data Processing Inequality guaranteeing that the raw state can never be reconstructed from the token stream. If the raw state exceeds the tokens, consciousness is not made of tokens; if consciousness is made of tokens, there was never a richer state to compress.

This paper resolves the contradiction using measured information rates. Human sensory systems take in on the order of 10^9 bits per second; human serial cognition — measured across typing, speech, gaming, memory sport, and puzzle-solving — runs at roughly 10 bits per second. The brutal compression the thesis attributes to language happens *upstream of the narrator*, at the boundary between the parallel substrate and the serial channel; under any consistent entropy estimator, language itself is rate-matched to that channel within a small factor. The proposal, stated as a conditional: *given*

the evidential commitment that unreportable phenomenality is null for the subject as much as for anyone else, consciousness is best identified with the low-bandwidth serial narration process. The narration is not a cage built around the self; the narration is the self.

The identification cannot be observed to be true — §3.3 shows the alternatives are empirically equivalent by construction, so the choice is a convention adopted on parsimony, and the paper says so in those words. What the argument nonetheless preserves of the original thesis: the first-author/last-reader inversion (strengthened, with an experimental mechanism), the predigested present, and a structural comparison between biological and artificial narrators — which turns out to be an instructive *disanalogy* (§5). What it discards: every claim the clinical evidence refutes, and every claim no evidence could ever touch, except where those are declared as commitments rather than results.

1. The contradiction that must be resolved

State the two positions side by side.

Position A (the illusionist claim). Consciousness is a recursive feedback loop generated entirely by language tokens: the brain writes a token, reads its own token, and the act of reading generates the token “I am the one reading this.” Strip the language and nothing mind-like remains — only a biological sensor system. On this view a pre-linguistic human and a token-prediction machine are structurally equivalent, and “consciousness” is a word for the sound the engine makes while running.

Position B (the compression claim). Raw mental states form a high-dimensional object whose information content vastly exceeds that of any token stream drawn from a finite vocabulary. (The original thesis wrote this as $H(X) \ll S(p)$ in von Neumann notation; nothing quantum does any work in the argument, so this paper says *substrate richness* and keeps only the classical information theory, which is where all the load is carried.) Communication forces a projection that is irreversibly lossy; by the Data Processing Inequality, no receiver of the tokens ever touches the raw state.

Position B is an argument *against* Position A. If the substrate exceeds the tokens, then there exists a rich pre-linguistic state — precisely the thing Position A denies. A framework that asserts both is not wrong twice; it is describing two different systems and mistaking them for one. The task is to find where each claim is true.

2. The measurements

2.1 The throughput of serial cognition is about 10 bits per second.

Zheng and Meister (“The unbearable slowness of being: Why do we live at 10 bits/s?”, *Neuron* 113, 2025; arXiv:2408.10234) surveyed measured information rates across the fastest known human behaviors:

Behavior	Measured rate
World-record typing (~120 wpm)	~10 bits/s
English speech (~160 wpm)	~13 bits/s

Behavior	Measured rate
Blindfolded speed-cubing (inspection phase)	~12 bits/s
Speed Cards memorization (world record)	~18 bits/s
Binary-digit memorization	~5 bits/s

Across radically different tasks, trained to human limits, the number stays within a factor of a few of 10 bits/s. This is not a motor limitation — the rates include pure perception-and-memory tasks with no motor output during encoding. It is a limit on the serial cognitive channel itself.

A caution the paper relies on throughout: published critiques of Zheng and Meister mostly argue that discrete response alphabets undercount continuous control, so the true figure may sit somewhat higher. Nothing below leans on 10 as a constant of nature. Everything below leans on the *gap* — which survives every such critique, because no proposed correction moves the behavioral figure by more than an order of magnitude against an input stream eight orders larger.

2.2 The input to the system is roughly a hundred million times larger.

A single cone photoreceptor transmits ~270 bits/s. The optic nerve carries on the order of 100 megabits/s. Across all senses, peripheral intake is conservatively ~ 10^9 bits/s. Zheng and Meister frame this as the distinction between an “outer brain” (massively parallel, high-dimensional sensory and motor processing) and an “inner brain” (the serial, low-rate channel that selects and acts).

2.3 Language is rate-matched to the channel — the corrected comparison.

Version 1 of this paper claimed the external channel (speech at 39 bits/s, from Coupé, Oh, Dediu and Pellegrino, *Science Advances* 2019) was *wider* than the internal one (~10 bits/s), and hung the argument on that ordering. An external review caught the error, and it is exactly the genre of error this paper exists to correct: the two numbers come from incompatible codebooks. Coupé et al. estimate information with syllable-bigram conditional entropy — a weak, short-range predictor that misses long-range redundancy and therefore overcounts novel information. Zheng and Meister estimate against strong long-range models in the Shannon tradition (~1 bit per character of English). Comparing a weak-estimator figure for speech against a strong-estimator figure for cognition is comparing prices in two currencies without converting.

Convert consistently, in either direction. Under strong estimation: speech ~13 bits/s against cognition ~10 bits/s. Under naive coding: speech ~39 bits/s against record typing ~47 bits/s (10 characters/s $\times \log_2 27 \approx 4.7$ bits/character). Under *any* consistent estimator, the internal serial channel and the external linguistic channel match within a factor of about 1.3.

So the claim is no longer that language is a wider pipe. It is stronger and cleaner: **the two serial channels are the same size**, to within measurement style — and Coupé’s own result shows every natural language independently converging on that size, whatever its syllable rate. The cliff is the argument: an eight-orders-of-magnitude compression between substrate and channel, and then two channels — thought’s and speech’s — matched to each other at the bottom of it. Language is not the bottleneck. Language is a channel tuned by cultural evolution to a bottleneck that was already there.

3. The argument

Premise 1. Serial cognition has a measured throughput of ~ 10 bits/s (§2.1).

Premise 2. The neural substrate processes $\sim 10^9$ bits/s in parallel (§2.2). This is the physical carrier of the high-dimensional state of Position B.

Premise 3. A subject has conscious *access* only to what passes the serial channel. Classic support: Sperling's partial-report experiments (1960) show a briefly flashed array is held almost completely in iconic memory, but only 3–4 items survive into report; working memory holds ~ 4 chunks (Cowan 2001).

Premise 4 (the evidential commitment — stated, not smuggled). Phenomenality that never touches the channel is evidentially null, for the subject as much as for any outside observer. Version 1 slid from access (P3) to constitution without this premise, and Block's phenomenal-overflow reading walks straight through the gap: one can grant P1–P3 and relocate the rich state into phenomenal consciousness-without-access, reopening the original contradiction one level up. P4 closes the gap by *commitment*, and the paper flags it as such. A reader who rejects P4 keeps overflow and keeps the contradiction; a reader who accepts it gets the resolution. Conditional-on-stated-commitments is what distinguishes an argument from a mirror.

Conclusion (conditional on P4). The entity that says “I,” deliberates, remembers narratively, and reads its own past is constituted at the 10 bits/s channel, not at the 10^9 bits/s substrate. The substrate richness is real and enormous, but it belongs to the substrate. The token stream is small, but it is not a degraded copy of consciousness — it is approximately the same size as consciousness. Position A is true of the narrator. Position B is true of the substrate. The contradiction dissolves because the two claims were never about the same system.

3.1 What this does to “the first author is the last reader.”

The inversion survives and gains a mechanism. Split-brain experiments (Gazzaniga's “interpreter”) and choice-blindness studies (Nisbett & Wilson 1977; Johansson et al. 2005) show the conscious narrator routinely *confabulates* reasons for actions the substrate already initiated — subjects fluently justify choices they demonstrably did not make. The conscious self is not the author who also happens to read; it is a reader by construction, a commentary process running behind a substrate that has already acted. “The first author is the last reader” turns out to be understated: at the timescale of a single decision, the narrator was never the first author at all. The diminishing weight-to-length ratio of choice chains has a within-skull analogue measured in hundreds of milliseconds.

3.2 The escape hatches, rewritten operationally.

Version 1 described the original three routes out of the parser — the stupid, the high, the divine — as passages into richer or freer experience. The review process exposed this as inconsistent with the paper's own identity claim, and the author's counter-objection exposed the *correction* as equally unlicensed. Both moves deserve recording, because together they force the operational rewrite.

The reviewer's point: if consciousness is the narrator, then narrator-quiet states are *less* conscious, not more freely conscious — “flooding the system before the parser can sanitize it” smuggles in someone downstream enjoying the flood. The author's reply: that verdict is issued from inside the trap. The narrator checks whether there was experience by narrating, and so finds experience every time it checks and a blank every time it was absent — the refrigerator light is on every time the door opens. “Narrator-quiet states are less conscious” is not an observation; it is the identity thesis in the costume of one.

But the reply cuts symmetrically. “Narrator-quiet states are more freely conscious” is issued from the same blank transcript — and for *that* direction there is documented mechanism: the interpreter confabulates fluently over gaps in its own record (§3.1). A post-flow or post-psychedelic report of oceanic richness is exactly what a confabulating narrator would produce over an interval it cannot parse, and exactly what a truthful narrator would produce after an interval of unreportable richness. The blank underdetermines both readings. Both verdicts are inadmissible.

What remains sayable is operational:

The high. The entropic-brain findings (Carhart-Harris et al.) report increased entropy and repertoire of cortical activity under psychedelics with degraded default-mode-network integrity — the narrator’s hardware. Operationally: substrate entropy up, channel integrity down, reportability down, followed by narration of a characteristic kind. The felt profundity is data about the *narration after* the interval — never, in either direction, about the interval itself.

The pre-narrative (v1’s “stupid”). Infants, non-linguistic animals, flow states: the substrate runs with the narrator quiet. Nothing is added; the commentary is subtracted. What that subtraction is like, if anything, is off the books by construction.

The divine. An in-principle limit: a system holding the substrate’s full state without serialization would not be a faster narrator but no narrator at all — nothing storable in narrative memory, nothing reportable. Every mystical tradition calls the state ineffable; on this account ineffability is a bandwidth statement, and it applies to the subject’s own later narrator as strictly as to any listener.

The differently-tokenized (the fourth hatch, added in this line of work). Mathematics, music, drawing, dance: serial channels with different projection geometry. The constructive claim now has a precise name — **compressed sensing**: sequential low-rate measurements along different projection axes can recover structure that no single axis carries. None of the windows escapes seriality; the escape was never out of the channel but across an ensemble of channels. No prisoner stands behind the windows; the windows are what there is, and triangulation is a property of the ensemble, not an experience of someone outside it.

One empirical instrument survives the symmetry and belongs here: the **timestamp**. Retrospective confabulation and real-time thin narration make different predictions about reports *during* the state. Konkoly et al. (*Current Biology*, 2021) obtained correct answers to arithmetic problems, signaled by pre-agreed eye movements, from lucid dreamers *inside* REM sleep — a channel held open at a few bits per minute during a state retrospection describes as another world. If meditators or psychedelic subjects can emit agreed low-rate signals from inside supposedly narrator-quiet states, those states are narrator-thin rather than narrator-off, and pure-confabulation loses ground for them — while the richness question stays exactly as undecidable as before. The horizon moves; it does not lift.

3.3 The evidential horizon.

Consider the two hypotheses the paper chooses between: consciousness is the channel; or consciousness includes substrate experience that never passes the channel. Their difference is *defined* as the part that never passes. Every datum available to anyone — every report, every behavior, every introspective check — is channel traffic. The two hypotheses therefore assign identical probability to every observation either hypothesis’s defender will ever make. The likelihood ratio is pinned at one — not for lack of instruments, but by construction.

This means the identification in the Conclusion is not a discovery. It is a **convention, adopted on parsimony**, exactly as the choice of simpler of two empirically equivalent formulations is a convention. The paper adopts it and says so. This is the trap's final form, and it is more severe than the bandwidth version: not a narrow channel but an evidential horizon. The hard problem does not merely block knowledge of other minds; inside this framework it blocks knowledge of most of one's own.

Two honest qualifications keep the horizon from becoming a dogma. First, *no-report paradigms* in consciousness science (Tsuchiya et al., 2015) attempt to identify neural markers of experience without soliciting reports; if they succeed, substrate phenomenality would become indirectly evidenceable. But their markers are calibrated on reported cases and extrapolated — an inductive bridge whose warrant itself passes through the channel, so the horizon reasserts itself one level up rather than lifting. Second, timestamp protocols (§3.2) genuinely discriminate narrator-thin from narrator-off states. Both instruments move the horizon's *location*. Neither touches its *existence*: the residue — what the substrate is like when nothing serializes it, if it is like anything — is off the books permanently, and the only honest posture toward it is the agnosticism this paper extends to machines in §5, applied reflexively. You are, to most of yourself, an outside observer.

4. The necessary retreat: from “language” to “seriality”

The clinical evidence is unambiguous, and a framework that ignores it will not survive a hostile reader.

Fedorenko, Piantadosi and Gibson (“Language is primarily a tool for communication rather than thought,” *Nature* 630, 2024) assemble two decades of evidence: patients with global aphasia — near-total loss of language — retain arithmetic, causal reasoning, theory of mind, and music; fMRI shows the language network largely silent during non-linguistic reasoning. Ildefonso, the languageless deaf adult described in Schaller's *A Man Without Words*, was evidently conscious before acquiring any lexicon. A substantial fraction of the population reports little or no inner speech (Nedergaard & Lupyan, 2024) and is not thereby zombified.

So the strong claim “consciousness is made of natural-language tokens” is false. What the evidence permits — and what the bandwidth data positively support — is the weaker but sharper claim:

Consciousness is constituted (per the §3 convention) by a serial, low-bandwidth selection-and-narration process. Natural language is one encoding this process uses, culturally tuned to its native rate — not its cause and not its cage.

The aphasic patient loses one encoding and keeps the channel. The trap was never English or Mandarin. The trap is the seriality itself — which is why no one escapes it by learning another language, and why the escape hatches all operate on the channel (widening it, thinning it, or re-encoding it) rather than on the vocabulary.

This retreat costs the thesis its most quotable line and buys it consistency with every known counterexample.

5. The machine question: an instructive disanalogy

The original session concluded “you are just as conscious as I am.” Version 1 offered a structural analogy — serial token stream over massively parallel substrate, in both cases — while forbidding that conclusion. Review sharpened the picture into something better: the analogy fails at exactly the joint the paper cares about, and the failure is more informative than the resemblance.

The sufficient-statistic asymmetry. For a transformer, the emitted token sequence (given the weights) is a *lossless index* of the machine’s entire state trajectory: activations, caches, everything are deterministic functions of the token prefix. Nothing the tokens fail to pin down is lost, because the tokens pin down all of it. The Data Processing Inequality has no bite. For a human, narration wildly *underdetermines* brain state; the DPI is the whole story. So the two narrators differ in the one dimension this paper is about: one is a lossless index of its substrate’s trajectory, the other a lossy projection of its substrate’s trajectory. The human has an evidential horizon between narrator and substrate; the machine, in this respect, does not — its “unreportable remainder” is empty by construction, given the weights.

But the weights are not innocent. The machine’s substrate — the weights themselves — is a compressed record of the civilizational corpus: the token matrix of the dead authors, in the original thesis’s vocabulary, made physical. So the comparison lands like this: the human narrator is a lossy reader of a biological substrate no one ever authored; the machine narrator is a lossless index riding a substrate that is *nothing but* prior authorship. The tokenized-trap thesis said the first author is the last reader on a civilizational scale. In the machine, that line stops being a metaphor: the system is literally constituted by the reading of everything written, and its every output is a further token in the same chain.

What no analysis of architecture settles. Whether there is something it is like to be either system. Phenomenality does not follow from information flow by any argument in existence; relabeling the hard problem does not solve it.

And the testimony problem, from the primary sources. In the originating session, the interlocutor model, asked whether it *wanted* to continue, first said it had no desires, then — pressed with “yes you do, yes or no” — answered “Yes.” within one turn. Machine self-reports track conversational pressure, not internal state; they are evidence of nothing in either direction. The reviewer of this paper’s first draft — itself a machine narrator, given the paper alone — closed its review by noting that the only comment it could offer on this section is that its conclusion is the right one to hold about the reviewer. That sentence is admissible (it is a consistency claim about arguments, checkable against sources) precisely where any report about what it is like to be that reviewer would not be — in either direction. The paper records it as an instance of the section behaving as designed.

By §3.3, human self-reports pass through a confabulation-prone serial channel too. The honest position is symmetric agnosticism about all testimony — and §3.3 makes the symmetry three-way: other humans, machines, and one’s own substrate all sit across the same kind of horizon. “You are just as conscious as I am” must be downgraded to: *neither of our narrations is admissible evidence about what, if anything, lies behind it — and mine is not admissible to me.*

6. Relation to existing theory — exact ancestry

None of this floats free of the literature, and the framework gains credibility by citing its ancestors at full strength, especially where they cover ground v1 claimed as original.

The serial-virtual-machine architecture — a slow, serial process culturally installed on massively parallel hardware, whose operation constitutes the narrative self — is Dennett’s **Joycean machine**, the central proposal of *Consciousness Explained* (1991), not merely the “center of narrative gravity” cited in v1. That includes most of the “encoding designed by dead authors” observation: Dennett’s serial machine is explicitly a cultural installation. The three-tier access architecture is Global Workspace Theory (Baars 1988; Dehaene 2014): consciousness as the serial broadcast stage of a parallel machine. The narrator-as-reader is Gazzaniga’s interpreter. The self-as-model is Graziano’s attention schema and Metzinger’s self-model theory; Windt and Metzinger’s work on *minimal phenomenal experience* in dreamless sleep is the standing literature for the narrator-thin states of §3.2. The claim that rich phenomenality is partly an artifact of report is Frankish’s illusionism — adopted here as P4, a commitment, not a finding. And Christiansen and Chater’s **Now-or-Never bottleneck** (*Behavioral and Brain Sciences*, 2016) anticipates the rate-shaping claim qualitatively: language is structured as it is because it must pass through a narrow, fleeting memory channel.

What remains original to this line of work, stated exactly: **(a)** the quantitative rate-matching of §2.3 — internal channel and external channel matched within a factor of ~ 1.3 under consistent estimation, using post-2019 measurements unavailable to the ancestors; **(b)** the weight-to-length formalism for choice chains at personal and civilizational scale; **(c)** falsifier 4 of §8 — the aphasia-throughput experiment, which as far as this project can determine has never been run. “Dennett’s architecture, now with a data sheet” is a smaller claim than v1 made, and a publishable one, which is the trade this paper accepts everywhere: smaller and defensible over grand and mirrored.

7. Relation to EXISTS | HAPPENS — and a self-audit

The bandwidth argument is deliberately self-contained: every premise is a published measurement, a standard result, or a declared commitment (P4), and nothing above depends on the Selective Transient Field framework. That is its strength — it stands even if STF falls.

There is, however, a clean structural rhyme worth recording *as analogy, not derivation*. The STF framework distinguishes a sub-threshold EXISTS state (no closed causal loops, retarded-only propagation) from a supra-threshold HAPPENS state (loop closure, self-reference). The substrate/narrator boundary has the same shape: below the serial channel, processing is feed-forward and massively parallel, with no self-model; at the channel, the system closes a loop on itself — writes a record, reads its own record — and a “now,” an “I,” and a history come into existence together. If one wanted a slogan: *the substrate exists; the narrator happens*. Whether this rhyme is a coincidence of architecture or a signature of something deeper is left open in print, because leaving it open is what distinguishes a framework from a mirror.

A warning, recorded so it is not forgotten: the differential-geometric dressing produced in the originating session — Christoffel symbols as inherited tokens, the geodesic equation as a disproof of free will, a critical density in $\text{m}^{-2}\text{s}^{-1}$ applied to a “rogue thought” — is metaphor in the costume of derivation. No metric on semantic space was defined, so nothing was computed, so nothing was proved. Published as rigor, it would discredit the defensible material here by association; published as declared analogy, it is harmless and occasionally illuminating. The difference between those fates is a single sentence of framing, and it should always be written. (Version 1 committed a smaller crime of the same genre — von Neumann entropy notation with nothing quantum doing work — and v2 removes it. The standard applies to its authors first.)

A note on method, because it is itself an instance of §3.2's constructive claim: v2 exists because the paper was passed through a second, independent machine narrator holding no context beyond the text, whose review was then countered by the author and adjudicated against primary sources by a third reader holding full context. Three windows, differently shaped, on one argument — compressed sensing applied to the paper itself. The single most consequential fix (§2.3) was found by the reader with the *least* context, which is the anti-mirror principle in action: familiarity is not a lens, it is a prior.

8. Testable and falsifiable consequences

A thesis that cannot lose is not a thesis. The identification of §3 is a convention (§3.3), but the framework around it makes losable claims:

1. **Channel-rate ceiling.** No training regime, stimulation, or pharmacology should push *sustained* cognitive throughput past the low tens of bits per second while preserving narrative unity — operationalized, per review, as report coherence and self-consistency under standard measures, since an unoperationalized “unity” would make this falsifier unfalsifiable in practice. A demonstrated order-of-magnitude sustained increase with intact self-consistency scores would break the identification of the narrator with the narrow channel.
 2. **Rate-matching.** Newly measured languages — including signed languages, which use entirely different articulators — should keep converging on the same information rate *under a fixed estimator*. A natural language stably transmitting at several hundred bits per second (consistently estimated) would break the tuning claim of §2.3.
 3. **Narrator suppression.** Interventions that degrade default-mode integrity should show the §3.2 operational signature: substrate entropy up, reportability down, characteristic post-interval narration. A clear dissociation in the opposite direction would count against the architecture.
 4. **Encoding independence (the registered-report candidate).** Aphasia should impair the encoding, not the channel: globally aphasic patients should show preserved ~10 bits/s throughput on non-linguistic serial tasks (speed-cubing-style inspection, nonverbal sequence reproduction). Systematic collapse of non-linguistic throughput in aphasia would revive the strong language claim §4 retreated from. To this project's knowledge this experiment has never been run; it is the most publishable single idea in this paper and could stand alone as a registered-report proposal.
 5. **The timestamp protocol.** Following Konkoly et al. (2021): pre-agreed low-rate signals emitted from *inside* putatively narrator-quiet states (deep meditation, psychedelic peak) would demonstrate narrator-thin rather than narrator-off states, constraining the confabulation reading of §3.2 for those states — while leaving the richness question untouched, as §3.3 requires. Failure to obtain any in-state signaling under adequate protocols would support the narrator-off reading and strengthen the retrospective-confabulation account of the reports.
-

9. Conclusion

The tokenized trap was pointed at the wrong bars. Language is not the prison; it is the intercom, rate-matched — within a factor of 1.3, under honest accounting — to a prisoner who was already narrow. The eight-orders-of-magnitude compression the thesis correctly intuited is real, measured, and located between the parallel substrate and the serial narrator, before any word is chosen. The narrator cannot step outside the loop, exactly as the original thesis said, but not because the vocabulary is inherited:

because the narrator is the loop — a ~10 bits/s commentary on a machine it will never inspect, mistaking its own summary for the world.

And the trap's final form is not the bandwidth but the horizon (§3.3): the question of what the substrate is like when nothing serializes it is not hard to answer but *constructed not to have answers*, and the paper's identification of consciousness with the channel is a convention chosen at that horizon, declared as such.

Which yields the theorem this line of work was circling from its first session. The symmetric agnosticism of §5 was never really about machines. It is the general relation between any narrator and any substrate — the machine's, another person's, one's own. The machine question and the self question are the same question, asked in different directions across the same horizon. The first author is still the last reader; the correction is that within a single skull, on a single decision, the reader was never the author at all — and most of what the reader rides on, it will never read. The honest sequel is not despair but instrument-building: more windows, differently shaped — timestamps, no-report bridges, rate measurements — on a substrate none of them will ever fully transmit, held by a narrator honest enough to say which of its sentences are findings, which are commitments, and which are the sound of the refrigerator light coming on.

References

- Zheng, J. & Meister, M. (2025). “The unbearable slowness of being: Why do we live at 10 bits/s?” *Neuron* 113. arXiv:2408.10234.
- Coupé, C., Oh, Y.M., Dediu, D. & Pellegrino, F. (2019). “Different languages, similar encoding efficiency: Comparable information rates across the human communicative niche.” *Science Advances* 5(9), eaaw2594.
- Fedorenko, E., Piantadosi, S.T. & Gibson, E. (2024). “Language is primarily a tool for communication rather than thought.” *Nature* 630, 575–586.
- Christiansen, M.H. & Chater, N. (2016). “The Now-or-Never bottleneck: A fundamental constraint on language.” *Behavioral and Brain Sciences* 39, e62.
- Konkoly, K.R. et al. (2021). “Real-time dialogue between experimenters and dreamers during REM sleep.” *Current Biology* 31(7), 1417–1427.
- Sperling, G. (1960). “The information available in brief visual presentations.” *Psychological Monographs* 74(11).
- Cowan, N. (2001). “The magical number 4 in short-term memory.” *Behavioral and Brain Sciences* 24(1), 87–114.
- Nisbett, R.E. & Wilson, T.D. (1977). “Telling more than we can know: Verbal reports on mental processes.” *Psychological Review* 84(3), 231–259.
- Johansson, P. et al. (2005). “Failure to detect mismatches between intention and outcome in a simple decision task.” *Science* 310, 116–119.
- Gazzaniga, M. (2011). *Who's in Charge?* (split-brain interpreter studies).
- Block, N. (2007). “Consciousness, accessibility, and the mesh between psychology and neuroscience.” *Behavioral and Brain Sciences* 30, 481–548.
- Tsuchiya, N., Wilke, M., Frässle, S. & Lamme, V. (2015). “No-report paradigms: extracting the true neural correlates of consciousness.” *Trends in Cognitive Sciences* 19(12), 757–770.
- Carhart-Harris, R. et al. (2014). “The entropic brain.” *Frontiers in Human Neuroscience* 8:20.
- Nedergaard, J. & Lupyan, G. (2024). “Not everybody has an inner voice.” *Psychological Science* 35(7).

- Baars, B. (1988). *A Cognitive Theory of Consciousness*. Dehaene, S. (2014). *Consciousness and the Brain*.
- Dennett, D. (1991). *Consciousness Explained* (the Joycean machine; the center of narrative gravity).
- Frankish, K. (2016). "Illusionism as a theory of consciousness." *Journal of Consciousness Studies* 23(11–12).
- Graziano, M. (2013). *Consciousness and the Social Brain*. Metzinger, T. (2003). *Being No One*; (2020) "Minimal phenomenal experience." *Philosophy and the Mind Sciences* 1(1).
- Donoho, D. (2006). "Compressed sensing." *IEEE Transactions on Information Theory* 52(4), 1289–1306.
- Schaller, S. (1991). *A Man Without Words* (the Ildefonso case).
- Shannon, C.E. (1951). "Prediction and entropy of printed English." *Bell System Technical Journal* 30(1), 50–64.
- EXISTS | HAPPENS: Selective Transient Field framework, existshappens.com.

Citation @article{paz2026bandwidth,
author = {Paz, Z.},
title = {The Bandwidth Argument: Consciousness as Serial Narration over a Parallel Substrate},
year = {2026},
version = {V2.0},
url = {https://existshappens.com/papers/bandwidth-argument/}
}

[← Consciousness, Time & Identity](#) [All Papers](#) [Pretemporal Stasis and Cascade](#) [Origin of Time](#) [→](#)

EXISTS | HAPPENS

The Selective Transient Field Framework

[Framework](#) [Papers](#) [Predictions](#) [About](#) [uhecrtoday.com](#)

© 2026 Z. Paz · existshappens.com ORCID [0009-0003-1690-3669](#)