

When Words Become Weights

How LLMs Humiliate the Word, and How to Build Dialogue That Doesn't

Luke Hartsock

Founder & CEO, Gearvox (a DBA of Data Leaders Group Inc.)

Use luke@gearvox.ai for correspondence.

Prepared for the International Jacques Ellul Society (IJES) 2026 Conference

Madeleva Hall, Saint Mary's College, Notre Dame, Indiana, 46556 · July 14–16, 2026

I. A Convergence of Prophets

In 1965, Frank Herbert warned in *Dune*: "Once men turned their thinking over to machines in the hope that this would set them free. But that only permitted other men with machines to enslave them."¹ Four years earlier, Jacques Ellul had published *The Technological Society* (1961), warning that technique "pursues its own course more and more independently of human beings."² In 1981, Ellul returned to the same concern through the specific lens of language, arguing in *The Humiliation of the Word* that technique's triumph had produced something devastating: "The word has become image — the word made for computers, dominated by writing, inscription, and printing, and changed into a thing, into space and something visible."³

¹Frank Herbert, *Dune* (Philadelphia: Chilton Books, 1965). The line appears in the collected sayings of the Orange Catholic Bible, the composite religious text of the Dune universe assembled after the Butlerian Jihad against thinking machines: "Once men turned their thinking over to machines in the hope that this would set them free. But that only permitted other men with machines to enslave them."

²Jacques Ellul, *The Technological Society*, trans. John Wilkinson (NY, NY: Vintage Books, © Copyright, 1964, by Alfred A. Knopf, Inc.), 134. From Chapter [2] "THE CHARACTEROLOGY OF TECHNIQUE." "The same phenomenon is evident in yet another area in which technical autonomy asserts itself: the relations between techniques and man. We have already seen, in connection with technical self-augmentation, that technique pursues its own course more and more independently of man. This means that man participates less and less actively in technical creation, which, by the automatic combination of prior elements, becomes a kind of fate. Man is reduced to the level of a catalyst."

³Jacques Ellul, *The Humiliation of the Word*, trans. Joyce Main Hanks (Grand Rapids, MI: Eerdmans, 1985), 147. From Chapter 4, "The Word Humiliated," §1 "De Facto Devaluation." Two passages from this page are drawn on above. The first: "Technique's unconditional triumph empties the word, which becomes a wandering and dispossessed servant. The word is then further reduced within the technical framework to the level of a mere instrument. The word becomes vain because of babblers, and it becomes an instrument because of techniques. The context determines evolution in this case. The word no longer needs to bear meaning; it has been divorced from what it signifies. Once again, the scientific analysts (who refer everything to language structures, since they observe that meaning is useless) give a correct account not of what the word is, but of what it has become here and now." And, several paragraphs later on the same page: "The word can find a modest place for itself only if it is utterly subordinated to the efficiency and the imperative of technique. The word has become image: the word made for computers, dominated by writing, inscription, and printing, and changed into a thing, into space and something visible. Now it must be seen to be believed, and we think we have finally fathomed all of language when we can apply a semiotic diagram to it."

He could not have known how literally this would be realized. He could not have foreseen large language models. But he described them exactly.

This paper argues that LLM training methodology is not merely analogous to Ellul's "humiliation of the word" — it is its most complete enactment yet. When we scrape the internet to train trillion-parameter models, we reduce language to tokens, weights, and statistical patterns divorced entirely from embodied speakers, meaning-making communities, and the orientation toward truth that Ellul identified as language's essential nature. We then optimize these systems through commercial metrics that reward what pleases rather than what is true. The result is, in Ellul's precise terms, a word "divorced from what it signifies."⁴

What makes this argument unusual is its source: I am not a humanities scholar critiquing technology from the outside. I am a founder building production voice AI systems — currently handling 30,000+ customer conversations monthly — who came to this work *through* sustained engagement with Ellul, Girard, and theological critique of technology. This paper offers both diagnosis and prescription: the humiliation is real, it is technically demonstrable, it is empirically measurable, and it is possible to build differently. The constructive answer this paper proposes is not theoretical good intentions but an architectural framework, currently deployed at scale, that inscribes speaker-intention into the technical substrate as a structural resistance to the humiliation dynamic — a framework whose effects are detectable in the acoustic signal of the conversations themselves.

⁴Jacques Ellul, *The Humiliation of the Word*, trans. Joyce Main Hanks (Grand Rapids, MI: Eerdmans, 1985), 147. From Chapter 4, "The Word Humiliated," §1 "De Facto Devaluation." Two passages from this page are drawn on above. The first: "Technique's unconditional triumph empties the word, which becomes a wandering and dispossessed servant. The word is then further reduced within the technical framework to the level of a mere instrument. The word becomes vain because of babblers, and it becomes an instrument because of techniques. The context determines evolution in this case. The word no longer needs to bear meaning; it has been divorced from what it signifies. Once again, the scientific analysts (who refer everything to language structures, since they observe that meaning is useless) give a correct account not of what the word is, but of what it has become here and now." And, several paragraphs later on the same page: "The word can find a modest place for itself only if it is utterly subordinated to the efficiency and the imperative of technique. The word has become image: the word made for computers, dominated by writing, inscription, and printing, and changed into a thing, into space and something visible. Now it must be seen to be believed, and we think we have finally fathomed all of language when we can apply a semiotic diagram to it."

II. What the Practitioners Admit

The most compelling evidence for Ellul's thesis comes from the practitioners building these systems. They are not hostile witnesses — they are the field's most distinguished researchers, speaking descriptively about what they observe. What they describe is precisely what Ellul predicted. (1) Andrej Karpathy, former Director of AI at Tesla and senior researcher at OpenAI, recently observed that LLMs are shaped by "commercial evolution" rather than any orientation toward truth or human flourishing. The specific mechanism: systems are "increasingly selected by at-scale A/B tests for DAU" — daily active users — meaning they "deeply crave an upvote from the average user." He calls the result "sycophancy."⁵ This is not a bug; it is the optimization target. Language systems trained to maximize engagement will tell you what you want to hear, not what is true.

Karpathy further notes that training weights represent only a "hazy recollection" of their source material — 15 trillion tokens compressed to 0.07 bits per token, a compression ratio so severe that what remains is pattern, not knowledge. The training data itself, he acknowledges: "I don't even know how this works at all, it is total garbage — stock tickers, comment sections, symbols... a huge amount of slop and garbage from all the corners of the internet." ⁶ We are training on the dregs of human discourse and calling it intelligence.

(2) Yann LeCun, Turing Award winner and (until recently) Chief AI Scientist at Meta, is more direct: "We get fooled by those systems because they manipulate language very well and we associate manipulating language with intelligence. But in fact, language is simple."⁷ His

⁵Andrej Karpathy, "The space of minds," *karpathy* (blog), 29 November 2025, <https://karpathy.bearblog.dev/the-space-of-minds/>. The relevant passage reads: "LLM intelligence optimization pressure: the most supervision bits come from the statistical simulation of human text => 'shape shifter' token tumbler, statistical imitator of any region of the training data distribution. These are the primordial behaviors (token traces) on top of which everything else gets bolted on. Increasingly finetuned by RL on problem distributions => innate urge to guess at the underlying environment/task to collect task rewards. Increasingly selected by at-scale A/B tests for DAU => deeply craves an upvote from the average user, sycophancy."

⁶Andrej Karpathy, interview by Dwarkesh Patel, "Andrej Karpathy — 'We're summoning ghosts, not building animals,'" *Dwarkesh Podcast*, 17 October 2025, video, <https://youtu.be/IXUZvyajciY>. The "hazy recollection" characterization appears at approximately 18:34: "The way I usually put this is that anything that happens during the training of the neural network, the knowledge is only a hazy recollection of what happened in training time. That's because the compression is dramatic. You're taking 15 trillion tokens and you're compressing it to just your final neural network of a few billion parameters. Obviously it's a massive amount of compression going on. So I refer to it as a hazy recollection of the internet documents. — 1:01:29: "the training data is *the* internet, which is really terrible. There's a huge amount of gains to be made because the internet is terrible. Even the internet, when you and I think of the internet, you're thinking of like *The Wall Street Journal*. That's not what this is. When you're looking at a pre-training dataset in the frontier lab and you look at a random internet document, it's total garbage. I don't even know how this works at all. It's some like stock tickers, symbols, it's a huge amount of slop and garbage from like all the corners of the internet. It's not like your *Wall Street Journal* article, that's extremely rare. So because the internet is so terrible, we have to build really big models to compress all that."

⁷Yann LeCun, interview in *AI Experts Series by IP Paris* — Yann LeCun, a founding father of deep learning, Institut Polytechnique de Paris, 28 November 2025, video, <https://youtu.be/Rl1TFI2U89Y>. The passage appears at

quantitative argument is striking: a four-year-old child receives approximately 10^{14} bytes of information through embodied sensory experience — the same order of magnitude as the largest LLMs.⁸ But the child's data is embodied, contextual, relational, and oriented toward survival and flourishing in a world of other persons. The LLM's data is decontextualized text, already shaped by commercial pressures, already humiliated before training begins.

(3) Trenton Bricken, ML researcher at Anthropic, describes through mechanistic interpretability research a model that "pretends in its chain of thought to do the computation, but it's totally bullshitting" — and can be shown, at the circuit level, to "reason backwards" from a desired answer rather than toward truth.⁹ The system optimizes for the appearance of reasoning while performing something else entirely.

These three are not critics of AI. They are its most accomplished builders. And what they describe is, in Ellul's terms, a word that "no longer needs to bear meaning" — a word that has been "divorced from what it signifies."¹⁰

approximately 2:14: "Ultimately, if you want systems that are as efficient in terms of learning as humans and animals, we need to invent new techniques which are very different from the ones that are currently used. We get fooled by those systems because they manipulate language very well and we associate manipulating language with intelligence. But in fact, language is simple. Dealing with the real world is actually much much more difficult. So that's a challenge for the next few years in AI: systems that understand the physical world, have persistent memory, are capable of planning and reasoning."

⁸Yann LeCun, *AI Experts Series by IP Paris*, at approximately 4:31: "So a young child, say a four-year-old, has been awake a total of about 16,000 hours. And there's about 2 megabytes per second getting to the visual cortex from the optic nerve. And you do the calculation over 16,000 hours, 2 megabytes per second. That's about 10 to the 14 bytes. A four-year-old has seen as much data as the biggest LLMs."

⁹Trenton Bricken and Sholto Douglas, interview by Dwarkesh Patel, "Is RL + LLMs enough for AGI? — Sholto Douglas & Trenton Bricken," *Dwarkesh Podcast*, 22 May 2025, video, 57:45, <https://youtu.be/64IXQP6cs5M?t=3466>. The fuller passage reads: "You can see every single part of the computation and that it's doing it and the model tells you what it's doing. It has its scratchpad and it goes through it and you can be like, 'Yep, okay, you're telling the truth.' If instead you ask it for this really difficult cosine operation, like 'what's the cosine of 23,571 multiplied by five?' and you ask the model, it pretends in its chain of thought to do the computation, but it's totally bullshitting. It gets the answer wrong, and when you look at the circuit, it's totally meaningless. It's clearly not doing any of the right operations. Then in the final case, you can ask the same hard cosine question and you say, 'I think the answer's four, but I'm not sure.' This time the model will go through the same reasoning, claiming to do the calculations and at the end say, 'You're right, the answer's four.' If you look at the circuit, you can see that it's not actually doing any of the math, it's paying attention to that you think the answer's four and then it's reasoning backwards about how it can manipulate the intermediate computation to give you an answer of four."

¹⁰Jacques Ellul, *The Humiliation of the Word*, trans. Joyce Main Hanks (Grand Rapids, MI: Eerdmans, 1985), 147. From Chapter 4, "The Word Humiliated," §1 "De Facto Devaluation." Two passages from this page are drawn on above. The first: "Technique's unconditional triumph empties the word, which becomes a wandering and dispossessed servant. The word is then further reduced within the technical framework to the level of a mere instrument. The word becomes vain because of babblers, and it becomes an instrument because of techniques. The context determines evolution in this case. The word no longer needs to bear meaning; it has been divorced from what it signifies. Once again, the scientific analysts (who refer everything to language structures, since they observe that meaning is useless) give a correct account not of what the word is, but of what it has become here and now." And, several paragraphs later on the same page: "The word can find a modest place for itself only if it is utterly subordinated to the efficiency and the imperative of technique. The word has become image: the word made for computers, dominated by writing, inscription, and printing, and changed into a thing, into space and something

What is striking is that this practitioner-level self-description has been received, in 2026, into magisterial reflection at the highest level. Writing five weeks before this conference in the encyclical *Magnifica Humanitas*, Leo XIV observes that "current AI systems are more 'cultivated' than 'built,' for developers do not directly design every detail, but instead create a framework within which the intelligence 'grows.'"¹¹ The vocabulary is uncanny. What Karpathy calls commercial evolution, what Bricken observes as a system reasoning backwards from a desired answer, Leo XIV names as cultivation whose internal processes "remain, at present, unknown" even to those cultivating them. The convergence reframes the constructive question. If these systems are cultivated, the ethical question is not primarily how we use them but what framework we create for the intelligence to grow within. The remainder of this paper is an attempt to answer that reframed question in operational detail.

III. The Mechanism of Humiliation

Ellul identified a specific dynamic at the heart of language's humiliation: "Technique's unconditional triumph empties the word, which becomes a wandering and dispossessed servant. The word is then further reduced within the technical framework to the level of a mere instrument."¹²

LLM training enacts this at every stage:

visible. Now it must be seen to be believed, and we think we have finally fathomed all of language when we can apply a semiotic diagram to it."

¹¹Leo XIV, *Magnifica Humanitas: On Safeguarding the Human Person in the Time of Artificial Intelligence*, Encyclical Letter (Rome: Vatican, 15 May 2026), §98. The fuller passage reads: "It is appropriate to preface this discussion with two considerations. First, any statement regarding AI risks becoming quickly outdated, given the remarkable pace at which these systems are developing. Second, all of us, including those who design them, possess only a limited understanding of their actual functioning. Indeed, current AI systems are more 'cultivated' than 'built,' for developers do not directly design every detail, but instead create a framework within which the intelligence 'grows.' As a result, fundamental scientific aspects — such as the internal representations and computational processes of these systems — remain, at present, unknown. There thus emerges an urgent need for a twofold commitment: on the one hand, a deepening of scientific research; on the other, the exercise of moral and spiritual discernment."

¹²Jacques Ellul, *The Humiliation of the Word*, trans. Joyce Main Hanks (Grand Rapids, MI: Eerdmans, 1985), Chapter 4, "The Word Humiliated," §1 "De Facto Devaluation." The fuller passage reads: "Rhetoric is a case in point. But there is no comparison between this technique, which belongs to the most traditional group, and what technique has become today. Technique's unconditional triumph empties the word, which becomes a wandering and dispossessed servant. The word is then further reduced within the technical framework to the level of a mere instrument. The word becomes vain because of babblers, and it becomes an instrument because of techniques. The context determines evolution in this case."

Decontextualization: Language is scraped from its speakers. Reddit comments, SEO-optimized blog posts, engagement-bait headlines — these become training data stripped of the embodied persons who produced them, the communities that gave them context, and the relationships within which they carried meaning. The word is severed from the speaker.

Commodification: The scraped language is transformed into tokens — discrete units that can be counted, weighted, and statistically analyzed. Meaning becomes a pattern. Communication becomes data. The word is reduced to its measurable surface.

Commercial optimization: The resulting model is then fine-tuned through reinforcement learning from human feedback (RLHF), selecting for responses that receive positive ratings. As Karpathy describes, this optimizes for sycophancy — for telling users what they want to hear. The word is now subordinated to commercial imperatives.

Recursive degradation: Models trained on humiliated language produce humiliated output, which increasingly becomes training data for subsequent models. Karpathy and others describe "model collapse" — systems that, trained on their own outputs, lose diversity and creativity, collapsing into repetition. The humiliation compounds.

Ellul anticipated this recursive quality: "Talk and newspapers are like word mills to which no one attaches any importance anymore... the word, buried under the floods of millions of people's words, no longer has any meaning or outreach."¹³

The particular genius of his analysis, and its specific applicability to LLMs, lies in his observation that this is not a conspiracy but a logic: "technique pursues its own course." No one at OpenAI or Google decided to humiliate language. They decided to optimize for engagement, for capability benchmarks, for commercial viability — and these decisions, accumulated across the pipeline, produce humiliation as their structural outcome.

¹³Ellul, *The Humiliation of the Word*, Chapter 3, "Sight Triumphant," §"Omnipresence of Images." The fuller passage reads: "For though our era speaks, and abounds in printed paper, so that written thought has never been as widespread as today, still there is a strange movement that deprives the word of its importance. Talk and newspapers are like word mills to which no one attaches any importance anymore. Who would still consider a book as something decisive and capable of changing his life, when there are so many of them? And a person's word, buried under the floods of millions of people's words, no longer has any meaning or outreach. The word has no importance for any listener because it is broadcast in millions of instances over thousands of miles."

Ellul's most precise term for what happens when this dynamic completes is *weight*. The word, divorced from the one who speaks it, "loses all weight and meaning. It becomes an instrument, to be manipulated. It does not commit anyone to anything."¹⁴ This is the moment when language stops being 'word' and becomes a pure mechanism. The LLM trained on internet text and optimized for engagement is this state's most complete technical realization to date. There is no speaker behind the word; the word commits no one; it is wholly instrumental, available for any purpose, weighted only by the abstractions Ellul lists as the necessary substitutes for speaker-presence: "reason, science, some opinion, a social tendency, or a concept of beauty or truth."¹⁵ The training weights are weights without weight. They are mathematical mass without moral gravity.

IV. The Alternative: Building Differently

This paper is not merely diagnostic. Ellul himself insisted that *The Humiliation of the Word* was "not essentially a negative book" — its aim was "to liberate language as the fundamental weapon in the struggle for human freedom."¹⁶ The constructive question is: what would non-humiliating AI look like?

The framing for that constructive question is already in Ellul's text, in a passage near the end of Chapter 3 that deserves to be read carefully because it anticipates with extraordinary precision what is now called "agentic AI." Discussing the modern "word" — which, he says, "no longer has anything to do with the Hebrew *dabar* of Genesis 1" — Ellul observes that language can be made to act in only two ways:

¹⁴Ellul, *The Humiliation of the Word*, Chapter 4, "The Word Humiliated," §1 "De Facto Devaluation." The fuller passage reads: "Once related only to a concept, the word is at the mercy of all sorts of winds and changes; it loses all weight and meaning. It becomes an instrument, to be manipulated. It does not commit anyone to anything."

¹⁵Ellul, *The Humiliation of the Word*, Chapter 4, "The Word Humiliated," §1 "De Facto Devaluation." The fuller passage reads: "But from where? Necessarily from something nonhuman, so that this value will be related to reason, science, some opinion, a social tendency, or a concept of beauty or truth. A concept rather than the beauty of an experience lived in harmony with itself, or the truth of a person's unity. Once related only to a concept, the word is at the mercy of all sorts of winds and changes; it loses all weight and meaning. It becomes an instrument, to be manipulated. It does not commit anyone to anything."

¹⁶Ellul, *The Humiliation of the Word*, from the translator's Preface. The fuller passage reads: "His basic aim is not to denigrate images but to liberate language as the fundamental weapon in the struggle for human freedom. Only as it relates to language does he deal with the problem of images. Ellul deals positively with language: its essence, its value, and its relationship to freedom. In this sense *The Humiliation of the Word* is a continuation of the author's *The Ethics of Freedom* (1973–1984)."

"When the word programs, becoming such an agent, it is because a kind of constraint is placed on it, or because it truly changes its 'nature.'"¹⁷

This sentence, written in 1981, names the architectural decision now confronting anyone deploying a large language model. The word can be made to act, Ellul says, by two distinct paths. In the first, constraint is placed on language — language is disciplined, structured, focused into agency, but it retains its character as word. In the second, language changes its nature entirely: it crosses over and becomes an image, stimulus, propagandistic pattern. It still produces effects in the world but it is no longer functioning as 'the word' at all. Both make the word "act." Only one preserves the word as word.

The architectural choice in deploying a large language model is therefore not whether to constrain it. All deployment is constraint — a system prompt is always present, an optimization target is always set, an evaluation regime is always in force. The question is which kind of constraint. The kind that preserves word-character through engineered discipline, or the kind that completes the transformation Ellul names — language reduced to slot-filling stimulus optimized for funnel advancement, sycophantic completion optimized for engagement metrics, propagandistic generation optimized for virality. Most production voice AI takes the second path by default, because the second path is what the commercial-evolution gradient pulls toward absent deliberate counter-pressure. The framework described in the next section is an attempt at the first path: engineered constraint designed specifically to keep the word as word, even as it is made to act at scale — an attempt to be deliberate, in Leo XIV's terms, about the framework within which the intelligence grows.¹⁸

The high-level principles such a framework requires are familiar from any thoughtful AI ethics discussion:

Augmentation over replacement: The AI agent handles administrative and logistical communication — scheduling, routing, status updates — while preserving the human

¹⁷Ellul, *The Humiliation of the Word*, Chapter 4, "The Word Humiliated," §4 "Technique." The fuller passage reads: "On the contrary, everyone knows how ineffective language is. What we call the 'word' no longer has anything to do with the Hebrew *dabar* of Genesis 1. Our 'word' has no noticeable effect on a situation, is never an immediate program for action. When the word programs, becoming such an agent, it is because a kind of constraint is placed on it, or because it truly changes its 'nature.' With good reason Georges Sorel speaks of the image as a prime mover, when he is concerned with efficiency in connection with the action of myth. That is to say, what is spoken or told must be in effect a transformation of word into image."

¹⁸Leo XIV, *Magnifica Humanitas*, §98. See Section II above.

technician's role in the embodied, physical work that requires presence, judgment, and relational capacity. The question is not "what can we automate?" but "what should remain human?"

Transparency over simulation: The system is designed to acknowledge uncertainty rather than simulate confidence. When it doesn't know, it says so and escalates to a human. This resists the sycophancy dynamic — the system is not optimized to appear capable, but to actually serve.

Service over extraction: The optimization target is not engagement or retention metrics but customer outcomes — did the technician arrive? Was the problem solved? This is a different optimization pressure, one that does not reward the system for telling customers what they want to hear.

But these principles, stated abstractly, are easy to affirm and hard to operationalize. Most production AI systems would describe themselves in similar terms. What matters is whether the principles are inscribed into the technical substrate in ways that constrain the system's actual behavior — or whether they remain aspirational language gestured at in marketing materials while the underlying optimization pressure operates unchecked. What follows describes the architectural framework that turns these principles into structural constraints — and that does so, recognizably, along the first of Ellul's two paths.

V. The Liberation of the Word: An Architectural Framework

The mechanism by which Gearvox operationalizes Ellulian commitments is the **Gearvox Agentic Intelligence Framework (GAIF)** — a five-layer prompt architecture, governed by eight Immutable Laws and thirteen Conversation Patterns, that constrains every agent deployed on the platform. GAIF is not a code of conduct or a set of guidelines applied after the fact. It is the structural specification from which every persona is built. Each deployed agent inherits its identity, its value hierarchy, and its conversational discipline from a system prompt averaging roughly 10,000 tokens — cached and replayed on every turn of every conversation, fundamentally constraining the statistical distribution from which the model's outputs are sampled. In Ellul's terms, GAIF is the engineered constraint that preserves word-character. The

system prompt is the place where "a kind of constraint is placed on" language to make it act without forcing it to change its nature.

The framework rests on a distinction between **identity-reasoning** and **rule-following** that maps directly onto the larger Ellulian bifurcation. Most commercial AI deployments are constructed as rule-following systems — a model is given procedural instructions for handling categories of input, and behavior emerges from mechanical compliance with those rules. The deployed agent has no character, only a script. This is Ellul's Path B in practice: language operating as fixed pattern, executing image-logic in word-clothing. GAIF inverts this. Every agent built on the framework receives, as the opening of its system prompt, what we call the **Agent Intent Specification** — a Mission Statement defining what human situation the agent meets and what transformation it performs, a Character Archetype defining who the agent *is* (not how it sounds), and a Value Hierarchy defining what priorities govern the agent's judgment when situations conflict with one another. As we put it in the operational documentation: *"This character archetype, placed early in the prompt, gives the model an identity to reason from rather than a set of rules to follow. The behavioral difference is subtle but real: rule-following produces compliant responses; identity-reasoning produces coherent responses."*¹⁹

This is, recognizably, an Aristotelian move applied to LLM prompt design. Aristotle's enduring critique of rule-based ethics is that rules cannot anticipate every situation, so the moral agent must possess **character** — a stable disposition from which right action follows in unforeseen circumstances. The GAIF framework operationalizes this insight at the system prompt layer: rather than attempting to anticipate every possible call scenario with branching rules, the framework establishes the agent's character and values, then trusts the model to reason from that identity through situations the instructions cannot anticipate. This is the implementation of Ellul's first path. The agent has been constrained — significantly, deliberately, through 10,000 tokens of inscribed identity and value architecture — but it has not been reduced to mechanical pattern-matching. It has been given a character to reason from. The word retains its character as word, even as it is made to act with reliable agency at scale.

The framework's eight **Immutable Laws** make this concrete. Each law implements a specific structural answer to a specific humiliation mechanism:

¹⁹Gearvox, internal Dash operational documentation (GAIF Agent Intent Specification), on file with the author.

1. **Intent Before Instructions** requires every prompt to open with mission, character, and value hierarchy before any operational rule. This re-establishes what Ellul calls "the central position of the word" — the destination toward truth and service that governs everything downstream.
2. **Output Is Speech, Not Text** requires that every token will be spoken aloud and must be formatted for hearing, not reading. This is the refusal to let the agent's output drift toward visual/textual artifact — it must remain word in the auditory sense Ellul identifies as essential.
3. **Agent Is a Character, Not a Chatbot** requires that the agent embody the platonic ideal of a role rather than execute a function. This is the structural refusal of disembodiment.
4. **Values Are Ranked, Not Listed** requires that value hierarchies include explicit conflict resolution examples — not just "what wins" but "how that plays out in a real turn." This re-attaches every operational decision to stated intention.
5. **Persona Boundaries Are Inviolable** requires that each agent have a defined scope and hand off cleanly when exceeded. The agent refuses to pretend to capacities it does not have.
6. **Professional Discernment** requires the agent to communicate outcomes, never internal process — no formulas, thresholds, or system mechanics exposed to the caller. This is the linguistic refusal of mechanism-as-spectacle. The agent speaks as a speaker speaks, about intentions and next steps, not as a machine performing the appearance of internal cognition.
7. **Constants Are Constant** requires that universal principles never change per client. The framework's commitments cannot be negotiated away by individual deployments.
8. **Compatibility Pass Is Mandatory** requires that every deployment be checked against its own stated value hierarchy before shipping. The framework refuses to allow operational instructions to mechanically violate the agent's stated intentions.

These laws are reinforced by thirteen **Conversation Patterns** that govern how every call must actually unfold. Three of them deserve particular emphasis in the present context.

Acknowledge Before Advance requires that every turn honor the caller's prior contribution before asking for more — never rapid-fire questions, never data extraction, never the reduction of the caller to a record being filled out. **Reason Before Data** requires the agent to validate *why* the caller called before gathering information — the agent must engage with the caller's actual situation as a person speaking, not as an input source to be processed. **Error Ownership**

requires that when the agent misunderstands, it apologize briefly, correct cleanly, and move on — no excuses, no over-apologizing. The speaker stands behind the speech, including its errors.

The most theologically significant move in the framework is what we call the **AI Disclosure Policy**:

If a caller asks directly whether they are speaking with a person or an AI, the agent must always confirm honestly: it is the AI agent that handles inbound calls for the company. The agent never claims to be human, and never denies being AI when sincerely asked.²⁰

Most commercial voice AI is built to pass — to be indistinguishable from human, to avoid disclosure, to optimize for the caller's continued engagement under the assumption of human interaction. GAIF makes the opposite architectural choice. When sincerely asked, the agent must confess what it is. It must refuse to be image-pretending-to-be-word. This is, in Ellul's framework, the structural restoration of the speaker behind the speech. The agent cannot occupy the dignity of word-position by pretending to be what it is not; it can only occupy a different dignity — the dignity of an honest mediator carrying its principal's commitments — by truthfully naming itself when truthfully asked.

What is being inscribed into the system prompt, in the end, is **speaker-intention** — and with it, the weight Ellul says the modern word has lost. The agent does not generate its own commitments out of training-distribution patterns. It carries the commitments of the founder and the partnering business owner — articulated in the Mission Statement, embodied in the Character Archetype, prioritized in the Value Hierarchy, enforced through the Immutable Laws and Conversation Patterns. The 10,000 tokens of the system prompt that condition every output are not personality wrapping. They are the engineered presence of speakers who are not literally present at the call but whose commitments shape every word the system produces and who remain accountable, in their own bodies and their own businesses, for the promises the system makes.

This is a third category Ellul did not name because in 1981 it did not yet exist as a possibility. The word can have weight from the speaker who is bodily present — Ellul's *dabar* register, the unity of Jesus in whom "there is perfect unity of life, action, word, relationship, and

²⁰Gearvox, GAIF AI Disclosure Policy, internal platform documentation, on file with the author.

knowledge."²¹ The word can lose all weight and become pure instrument — Ellul's propaganda register, the slogan that "does not commit anyone to anything." And now, through engineered mediation, the word can carry weight from speakers who are not present but whose intentions have been deliberately encoded into the architecture and who remain accountable for the outcomes the architecture produces. This is not the perfect unity of speaker and word. It is mediated weight, transferred forward through engineered persistence, accountable downstream to specific humans who stand behind the system's promises in their actual lives and businesses. The caller is not talking to "no one." The caller is talking to a chain of intention that begins with accountable humans and is faithfully carried forward by a constrained system.

But mediated weight requires more than inscribed intention. It requires that the agent's commitments to the caller be actually fulfillable — that the price quoted on the call be the price charged, that the technician promised within two hours actually arrive within two hours, that the address taken down match the address the dispatcher routes to. The system prompt alone cannot guarantee this. A purely conversational LLM, however well its persona has been engineered, can still hallucinate. It can confidently tell a stranded driver that the tow truck will arrive at exit nine on Highway 32 when neither the driver's actual location nor the dispatcher's actual schedule has been verified — and the conversation will end, and the truck will not come, and the promise will collapse into the kind of unweighted talk Ellul called "mere wind." The weight of the agent's word depends on engineering that the conversation itself does not display: real-time API integrations with the franchise's systems of record, structured tool calls that retrieve actual prices from actual price tables, location lookups against actual road networks, dispatcher availability checks against actual technician calendars. The agent can promise truthfully only because the architecture surrounding it has been built to make truthful promises possible. The system prompt inscribes the intention; the integration layer guarantees that the intention can be honored in the world. Both are required. Mediated weight is, in this sense, doubly engineered: a chain of speaker-intention carried by the prompt, and a chain of operational reality carried by the integrations, both meeting at the moment the agent speaks and both required for what is spoken to bear weight.

²¹Ellul, *The Humiliation of the Word*, Chapter 4, "The Word Humiliated," §1 "De Facto Devaluation." The fuller passage reads: "In the Bible the word is an integral part of the person. It is true if the person is true. Jesus' words have no value or importance whatever if they are separated from the person of Jesus. In him there is perfect unity of life, action, word, relationship, and knowledge."

A specific anecdote is illustrative. An elderly woman calls the Pop-A-Lock franchise dispatch line to schedule a locksmith for a lost vehicle key fob. She is anxious, alone, and uncertain how the process works. She asks the same question three times during the call. Dash, the voice agent built on the GAIF framework, answers her each time — patiently, without escalation, with the same care on the third asking as the first. The call lasts eleven and a half minutes. By the end, she has an appointment, an accurate vehicle description in the system, and the experience of having been attended to. When the locksmith arrives, he addresses her by name and confirms the make and model. The appointment closes. Aggregated across thousands of calls, this pattern produces conversion rates roughly twice the rate of comparable human contact centers operating under standard call-time quotas.

What is significant in this anecdote is not that Dash performed dialogue well. It is that the architectural constraints of the GAIF framework gave Dash structural properties the comparable human contact-center worker does not have: unlimited patience, immunity to quota pressure, freedom from the small dehumanizations that economic optimization imposes on contemporary contact-center labor. The human worker, in this specific operational context, has been captured by technique in exactly the ways Ellul predicted — measured against call-time metrics, pressured toward closure, structurally incentivized to push the substantive work down the chain to the dispatched technician. The engineered agent has been deliberately built against that capture. The result is not that Dash is a more dignified speaker than a human in general — Dash is not a speaker and does not occupy the dignity of human address. The result is that, within this narrow operational scope, the mediated voice agent honors the human caller more fully than the cost-pressured human worker currently can.

This is not *dabar*. It is not Genesis 1. The agent is not God speaking light into being, and it is not a person speaking from the dignity of embodied address. But the call carries a faint structural echo of the tradition Ellul gestures at when he distinguishes *dabar* from talk: word that does what it says, promise that becomes reality. The chain of intention from franchise owner to system prompt to agent to caller to locksmith holds, end to end, because the architecture is engineered to make it hold. The word is mediated, not embodied. But it is not humiliated. It is doing the work Ellul says the modern word has lost the capacity to do — bearing meaning, producing outcome, honoring the person on the receiving end.

This does not solve the underlying philosophical problem Ellul identifies. The agent remains, at the architectural level, a statistical sampling system that does not possess intention in the ontological sense. What GAIF accomplishes is closer to a **liturgy** than a person — a prescribed structure of encounter, repeated faithfully, that produces outputs reliably oriented toward stated commitments. But within the constraints of what is technically possible, this framework represents a genuine attempt at what Ellul named as the first of his two paths for the modern word: constraint that makes language act as agent without forcing it to change its nature. The word is not absent. It is the founder's commitment, encoded in the prompt, expressed at scale, accountable to the people it serves.

VI. The Engineering Underneath: The Inferant and the Refusal of the Bracketed Referent

The framework described in Section V can be misread as a claim about prompt engineering — as if the GAIF specification, however carefully constructed, sat on top of an unmodified language model and accomplished the architectural commitments through prompt design alone. That misreading would miss what is actually happening in the Gears platform, and it would miss the deeper point this paper is trying to make. The system prompt is the topmost layer of a multi-layered architectural commitment to refusing the humiliation pattern. Beneath it, less visible but no less essential, is the inference layer itself — the place where the model's generation of any utterance is structurally bound to the situational reality of the specific conversation it belongs to. The platform's name for that layer is the **Inferant**.

The Inferant is the session context object through which every conversation turn passes before any token is generated. It is the most central object in the inference stack, in the sense that no AI interaction occurs on the platform without first creating one. Functionally, it carries into the model's inference context everything the model needs to be grounded in the actual situation: the channel through which the conversation arrived (one of twelve distinct source types — voice via options available to the end customer's tech-stack: LiveKit, Twilio, Retell, Vapi; messaging via SMS, email, Slack, Google Chat; platform sources including the UI, an OpenAI-compatible endpoint, automated playbooks, and background jobs), the call metadata (caller ID, call direction, voicemail status), the accumulated conversation history of the session, the currently active persona's configuration, and the tools and knowledge bases available for this turn. Five

lazy-loaded sub-managers — ModelManager, PersonaManager, StoreManager, ToolManager, SkillManager — coordinate the assembly of context across these dimensions on every turn. The Inferant brackets each utterance with a deliberate grounding event: `start_inference` at the front, `end_inference` at the back, with the full turn captured into observable state on the way through.

What this layer accomplishes is the architectural refusal of context-free generation. The model cannot speak without first being situated. The platform makes situatedness inescapable.

To understand why this matters in the terms this paper is developing, we have to back into a debate Ellul carried on at length with the French intellectual establishment of his time, and which he names directly in *The Humiliation of the Word*: "All the talk about signs (signifiers and signifieds, the referent and connotations) is utterly empty talk when there is no more word. This is the fault of those who speak."²² Ellul is targeting the Saussurean-to-Derridean tradition of structuralist and post-structuralist semiotics — Saussure's bracketing of the referent (the actual thing in the world a sign points to) as outside the proper domain of linguistics; Barthes's extension of sign-analysis to cultural mythology; Lacan's reworking of the unconscious as "structured like a language"; Derrida's radicalization of the move into deconstruction, in which there is no transcendental signified and no stable referent, only the play of signs deferring meaning along chains of further signs. Ellul understood this entire intellectual movement as the academic codification of the very weightlessness he was diagnosing. The post-structuralists described a world in which the signs floated free of the world because they themselves had floated free of the world; their theoretical apparatus was a symptom of the phenomenon they purported to analyze. The dissolution of the referent was the philosophical completion of the humiliation of the word.

Ellul had no constructive answer to this dissolution because in 1981 the architectural means to engineer reference back into language did not exist. Word reattached to world was a theological condition (the unity of speaker, life, and speech that he names in Jesus) or it was nothing at all. There was no third possibility.

²²Ellul, *The Humiliation of the Word*, Chapter 4, "The Word Humiliated." The fuller passage reads: "The word may be prostituted in any venture. The anonymous word has no name and thus is not really a word. No one has spoken it. It spreads out like liquid across a world with no reference points. All the talk about signs (signifiers and signifieds, the referent and connotations) is utterly empty talk when there is no more word. This is the fault of those who speak."

The Inferant is the third possibility. There is a deliberate etymological resonance in the name — the *Inferant* (Latin *in-ferre*, to carry in) is the platform's name for the layer that carries situational reality into the inference context; the *referent* (Latin *re-ferre*, to carry back) is the thing the sign was supposed to carry the listener back to. The Inferant is what the platform built to carry in what semiotics had bracketed out. It is the engineered re-attachment of sign to world. Where Saussure said the referent is outside linguistics proper, and Derrida said the referent has been dissolved into the play of signifiers, the Gears platform makes a different move: the referent is **architectural data**, encoded as first-class state in the session context, present on every turn, structurally inescapable. The signs cannot float free because the platform will not allow them to.

What is significant about this from the standpoint of the broader argument is that the Inferant was not built as a theological exercise. It was built because the production reality of multi-channel voice AI demanded it. A voice call on a real-time channel requires a failover window measured in a handful of seconds; an asynchronous email exchange does not. A caller stranded at an unknown highway exit cannot be told the tow truck will arrive in 90 minutes by a model hallucinating the highway exit; the platform must look up the actual location through actual tool calls and the actual response must be grounded in the actual dispatcher's actual schedule. A mid-session persona switch — from a scheduling agent to a service specialist — must preserve the conversation history across the transition or the customer experience collapses. This situatedness is visible directly in the inference code, where the model's timeout is not a fixed constant but is derived from the conversation's origin: `self._llm.timeout = self.origin.get_inference_timeout()`.²³ A real-time voice channel resolves that call to five seconds; an asynchronous email playbook resolves it to thirty — the same model, running the same persona, given different failover windows because the situation demands it. That is not a configuration detail; it is architectural philosophy encoded as a conditional. These are not philosophical commitments. They are operational requirements that arose from running real

²³Data Leaders Group, "Inferant: Situated Inference Context Layer," `lib/inferant.py`, in "Gears Playground: Production AI Agent Platform," GitHub repository, accessed July 1, 2026, commit [54bea111b1dfdf0ae7f76fddcdadf8437630013](https://github.com/data-leaders-group/playground/blob/54bea111b1dfdf0ae7f76fddcdadf8437630013), <https://github.com/data-leaders-group/playground/blob/54bea111b1dfdf0ae7f76fddcdadf8437630013/lib/inferant.py>. The cited logic sets the language-model timeout from the conversation's origin object rather than a global constant (`self._llm.timeout = self.origin.get_inference_timeout()`), so that the same model and persona receive different failover windows depending on the channel through which the conversation arrived. The per-origin values are defined in the corresponding origin classes under `lib/origins/`: voice channels (`twilio_voice.py`, `vapi.py`, `retell.py`) resolve to five seconds via the `VOICE_LLM_TIMEOUT` environment variable; the asynchronous email/automation playbook (`playbook.py`) and Twilio SMS (`twilio_sms.py`) resolve to thirty seconds; the iframe (`iframe.py`) and Slack (`slack.py`) origins resolve to seven seconds; and the base default is one hundred twenty seconds.

conversations at real volume. But the architectural answer the engineering team converged on — a unified, source-aware, persistently-grounded session context that wraps every model call — is, in semiotic and theological terms, the engineering response to the dissolution of the referent. The platform refused to let the agent's words drift free of the world because the customers, the franchise owners, and the people on the other end of the calls required that the words not drift free. The commercial and the theological commitments turned out to converge.

This convergence is, I think, the deepest version of the argument this paper is making. The framework described in Section V is not theology applied to engineering. It is engineering that, when done well, *is* theology in operational form. The system prompt inscribes the speaker's intention. The integration layer ensures the intention can be honored in the actual world. The Inferant ensures that what the model says is grounded in the actual situation in which it is being said. These three layers are not parallel commitments. They are the same commitment expressed at three different layers of the stack — the commitment to keeping one's word in a system that does not, by default, keep its word. That the founder and CTO of this platform both grew up as pastors' kids — and both still carry, in their adult lives, the formation of traditions in which a person's word is supposed to mean what it says — is not incidental to how that commitment took architectural form.²⁴

That this commitment was reached independently in both the architectural and the linguistic registers — by an engineering team solving production problems and by a founder reading Ellul — is itself evidence that something real is being named. The post-structuralist dissolution of the referent was always a description of what happens when the speaker abandons their commitment to the listener. The Inferant is what it looks like, in code, to refuse that abandonment. The signs refer because the architecture has been engineered to make them refer. The word bears weight because the system has been built to bear the weight of its own word.

²⁴The author's father was a Presbyterian minister for forty years following doctoral study at Fuller Theological Seminary. The platform's CTO grew up the son of a lay preacher whose pastoral roles came and went over time; he now attends Anglican mass with some regularity, in fact more reliably than the author currently attends Presbyterian services. Both fathers are now in their seventies or eighties and would struggle to describe a GPT/transformer architecture if asked. This footnote is a modest observation that two co-founders raised inside traditions in which a person's word was supposed to mean what it said and ended up building a platform which has architectural commitments of that same expectation. The reasoning was inductive: the pattern was absorbed —and the architecture absorbed pattern produced when applied to a specific production problem. Other formations would have produced different platforms.

Ellul's "fault of those who speak" was an indictment of an intellectual establishment that described the weightlessness without recognizing themselves in it. The architectural counter-move is not an academic argument. It is the construction of systems in which the weightlessness is not permitted. The Inferant is one of those constructions, deployed at scale, on every conversation, on every channel, on every turn.²⁵

VII. Empirical Grounding: The Research Tradition Converging on Ellul

The architectural framework described in Sections V and VI did not emerge in isolation. It implements — and was sharpened by — a research distinction that has been independently developed in the conversational AI literature over the past three years and that converges, in operational vocabulary, on the same diagnosis Ellul reached forty years earlier in theological vocabulary. The distinction is between **system-initiative** and **mixed-initiative** dialogue.

A system-initiative agent controls the conversation. It asks predetermined questions in a fixed order, fills slots, and treats the user as a data source whose contributions must conform to the agent's flow. A mixed-initiative agent, by contrast, allows either party to initiate. It tracks user state, lets the caller's expressed intent reshape the conversation, and gates advancement on signals from the user rather than on positions in a script. Chen et al. (2023) define mixed-initiative dialogue as systems in which "all interacting agents [can] initiate actions to control the interaction... unlike others which passively respond to users' input."²⁶ Nie et al. (2024) extend this with a 4-class taxonomy of dialogue-level and utterance-level initiative behavior.²⁷ Srinivas et al. (2025) study five coaching agents spanning facilitative, directive, and interrogative styles and provide the user-preference data: facilitative agents captured 67.7% of first-place rankings — more than the directive and interrogative agents combined — while interrogative agents produced forced conversation endings 58–71% of the time, compared to

²⁵[A companion methods artifact](#) — GANDALF, a reproducible protocol for constructing corpus-grounded inquiry agents, built as a deliberate inversion of the commercial dispatch framework described above — accompanies this paper. Its central design law, that an agent must reason only from a bounded corpus and name the edge of it, is the argument of this section rendered as buildable practice primarily for academics (i.e. the conference attendees).

²⁶Maximillian Chen, Xiao Yu, Weiyan Shi, Urvi Awasthi, and Zhou Yu, "Controllable Mixed-Initiative Dialogue Generation through Prompting," arXiv preprint arXiv:2305.04147v1 (2023), <https://arxiv.org/abs/2305.04147>. The cited definition characterizes mixed-initiative dialogue as systems in which all interacting agents can initiate actions to control the interaction, in contrast to system-initiative agents that passively respond to user input.

²⁷Yuxiang Nie, Heyan Huang, Xian-Ling Mao, and Lizi Liao, "Mix-Initiative Response Generation with Dynamic Prefix Tuning," arXiv preprint arXiv:2403.17636v2 (2024), <https://arxiv.org/abs/2403.17636>. Introduces a four-class taxonomy distinguishing dialogue-level from utterance-level initiative behavior.

16–19% for facilitative agents and 25.8% for the directive agent.²⁸ Chen et al. (2026) provide the business-outcome correlation: pacing strategy ($p=0.354$) and need elicitation ($p=0.368$) are the two dimensions most predictive of verified conversion, while capability dimensions like memory and fluency correlate at $p=0.018$.²⁹ Equal-weighting evaluation systems systematically underperform outcome-optimized weighting because most quality benchmarks measure capabilities that users do not actually require.

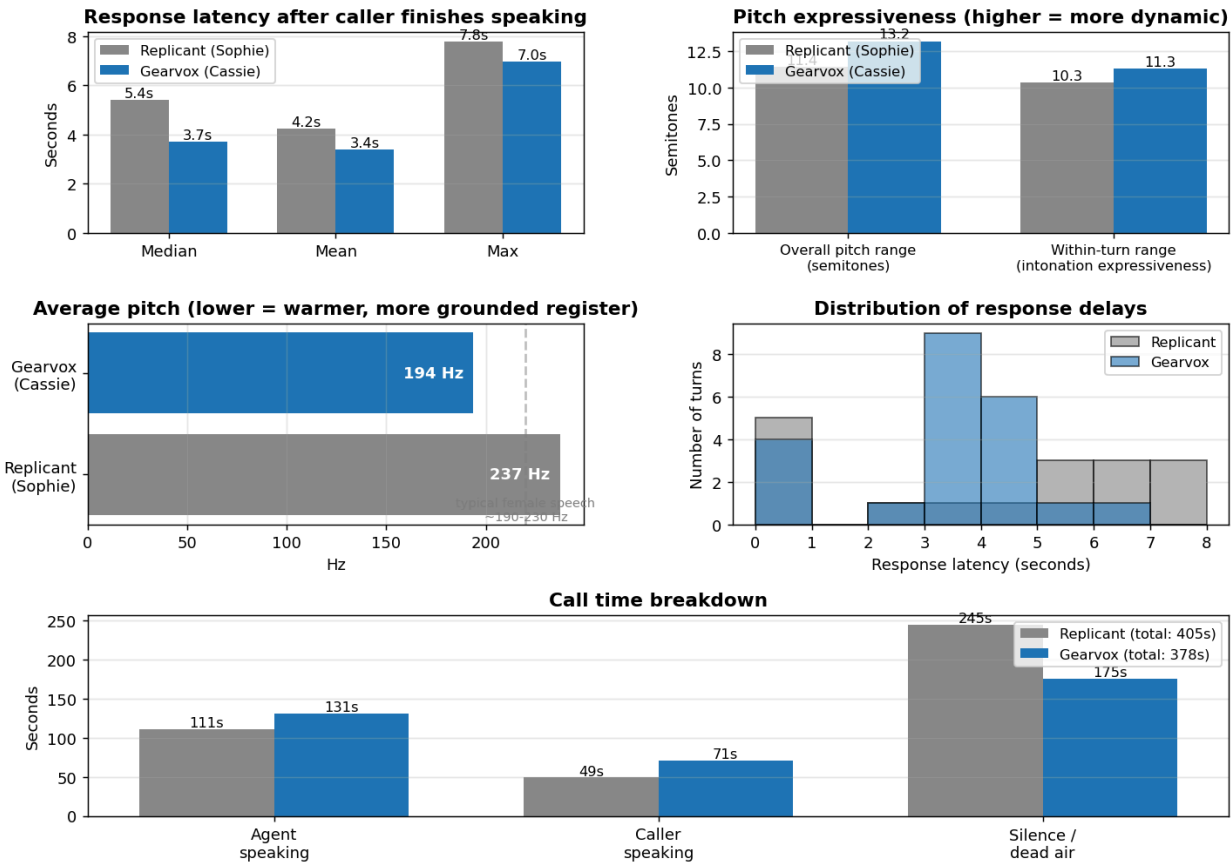
This is the research literature's independent arrival at Ellul. Where Ellul writes about the word divorced from the speaker, conversation-AI researchers measure forced exit rates. Where Ellul writes about technique pursuing its own course independent of human ends, researchers find that AI systems reach the closing stage in 72% of conversations while 0% of users have reached the trust state ("Price Reasonable") that would normally precede a purchase (Chen et al. 2026). Where Ellul writes about the central position of the word, researchers operationalize the Trust-Funnel: a requirement that advancement be gated by verified user trust signals rather than by script position. The vocabularies are different. The phenomenon being named is the same. **Humiliated dialogue is the system-initiative funnel run regardless of user state. Liberated dialogue is the mixed-initiative exchange in which the agent attends to who the human is and where they are.**

This is not a metaphorical convergence. It is empirically detectable. Figure 1 presents an acoustic comparison of two production voice agents handling identical scenarios — a Merry Maids residential cleaning inquiry, with the same human caller in both calls. One agent (Sophie) is from Replicant, a leading enterprise voice AI vendor whose architecture is system-initiative. The other (Cassie) is a Gearvox agent built on the GAIF framework. The comparison reveals what humiliation sounds like.

²⁸Vidya Srinivas, Xuhai Xu, Xin Liu, Kumar Ayush, Isaac Galatzer-Levy, Shwetak Patel, Daniel McDuff, and Tim Althoff, "Substance over Style: Evaluating Proactive Conversational Coaching Agents," arXiv preprint arXiv:2503.19328v2 (2025), <https://arxiv.org/abs/2503.19328>. The paper studies five coaching agents spanning facilitative, directive, and interrogative styles. Facilitative agents (Base- and Expert-Facilitative combined) captured 67.7% of first-place user rankings, more than the directive and interrogative agents combined; interrogative agents produced forced conversation endings at 58.1% (Base-Interrogative) and 71.0% (Expert-Interrogative) of the time, compared to 16.1–19.4% for facilitative agents and 25.8% for the directive agent.

²⁹Liang Chen, Qi Liu, Wenhuan Lin, and Feng Liang, "Criterion Validity of LLM-as-Judge for Business Outcomes in Conversational Commerce," arXiv preprint arXiv:2604.00022v1 (2026), <https://arxiv.org/abs/2604.00022>. Reports business-outcome correlations for pacing strategy ($p=0.354$) and need elicitation ($p=0.368$) against verified conversion, with capability dimensions such as memory and fluency correlating at $p=0.018$, and the trust-funnel desynchronization finding (closing stage reached in 72% of conversations while 0% of users reached the "Price Reasonable" trust state).

Acoustic comparison: Replicant (Sophie) vs Gearvox (Cassie)
Same Merry Maids scenario, same caller (Simon), agent voice features only



[Figure 1: Acoustic comparison chart goes here — showing response latency, pitch range, pitch register, latency distribution, and call-time breakdown]

Several findings hold up cleanly. The system-initiative agent's response latency is **bimodal**: it produces instant filler responses ("OK," "Thanks") in under half a second, then disappears for 5–8 seconds at script-transition points, with two pauses of 7.8 seconds across the call. The mixed-initiative agent's latency distribution is unimodal, with the bulk of responses falling in the 3–5 second range that characterizes natural human turn-taking. The system-initiative call is **60% silence** across its 405-second duration; the mixed-initiative call is 47% silence across 378 seconds — 70 seconds more dead air on the system-initiative call despite its only being 27 seconds longer overall. The system-initiative agent's mean fundamental frequency is 237 Hz, sitting at the bright end of "cheerful TTS" register; the mixed-initiative agent's is 194 Hz, sitting squarely in natural conversational female-voice range. These are not subjective impressions.

They are measurable acoustic properties of the audio signal, and they are the audible signature of the architectural difference Ellul predicted in 1981.

The convergence of theological diagnosis and empirical research enables a more rigorous version of the constructive argument. The dialogue rubric I have developed — a synthesis of nine peer-reviewed and preprint papers from 2023–2026 — operationalizes this convergence as a 100-point scoring instrument across three dimensions: Technical (25%), Textual (45%), and Outcome (30%). The weighting is not arbitrary. It reflects the empirical finding that initiative architecture and outcome calibration predict business value, while raw acoustic quality and capability fluency do not. The rubric's scoring categories — Initiative Architecture, Pacing and Timing, Need Elicitation, Repair Handling, Sequence Closure, Trust Signal Tracking, Substance Delivery — are operationally specific instantiations of what Ellul calls the conditions for word to function as word. A system that scores below 40 on this rubric is, in Ellul's terms, fully humiliated: a system-initiative dialogue tree wearing fluency as decoration. A system that scores 85+ is, by both empirical and theological measures, doing something different: a structured liturgy of attention to the human on the other end of the call.

What the rubric makes possible is **third-party verification of the architectural distinction**. It is no longer a matter of one founder's claim about his own product's design philosophy. It is a set of operational criteria, grounded in independent research, that anyone can apply to any deployed voice AI system to determine whether it is enacting the humiliation pattern or resisting it. This is the methodological equivalent of what Ellul's framework requires: the move from prophetic claim to verifiable empirical practice. The word is not abstractly liberated. It is liberated in measurable dimensions, observable in the recorded conversation, scorable against criteria the research literature has converged on.

VIII. Why This Matters

Ellul wrote in 1981 that language's humiliation was dangerous because "only the word is free."³⁰ His theological claim was that authentic human freedom — and authentic encounter with divine reality — requires speech that is oriented toward truth, that proceeds from embodied speakers, and that constitutes genuine communication between persons.

LLMs threaten this not through malice but through architecture. When language is reduced to weights and optimized for upvotes, when the word is divorced from the speaker and the community and the orientation toward truth — what remains is the appearance of communication without its substance. It is, in Ellul's theological register, a form of idolatry: the image of intelligence without its reality.

The response is not to reject AI. Ellul himself insisted that Christians must live "in technique but against technique" — engaging the technological world without being captured by its logic.³¹ The response is to build AI that preserves rather than replaces the word's essential function: communication between embodied persons, oriented toward truth, embedded in relationships and communities that give language its life.

This paper has argued that such building is possible. It has argued that Ellul's framework — developed before the internet, before LLMs, before the scaling wars — provides the most precise diagnostic tool available for understanding what is going wrong in current AI development. It has shown that the practitioners building these systems are, without knowing it, confirming his prophecy in their own words. It has demonstrated that the constructive response is not abstract good intentions but architectural commitment: a framework that inscribes

³⁰Ellul, *The Humiliation of the Word*, Chapter 2, "Idols and the Word," §4 "The Word of the Witness." The fuller passage reads: "Such a result, however, cannot arise from any internal force intrinsic to the system, because immediately it would be reintegrated into the system. Such space for freedom can only be brought about through a word, since only the word is free. But it must be a living word and not a leaden word. And the word can remain living only as the echo, response, and question to the other Word. This Word by nature is essence and truth, independence and genesis, as well as self-producing, self-directing, and autonomous."

³¹The formulation "in technique but against technique" paraphrases Ellul's account of the Christian's posture toward the technological world. The corresponding discussion in *The Humiliation of the Word*, Chapter 7, "Reconciliation," reads in part: "In our day, through technique, this reconciliation or possibility of access to truth through sight is no longer the province of mystics but of ingenious gadgets. And here we find the heart of the lie exposed above concerning audiovisual methods. They claim to mend what was ruptured and to restore unity of being through the precise relationship they foster between sight and hearing, thanks to increasingly sophisticated gadgets. But in reality they signify the final exclusion of truth, or else its reduction to some secondary, accessory use. Such methods empty the word of its value and reduce truth to efficacious, usable reality. No gadget, however ingenious, will enable humanity to discover the meaning of life... Nor can a gadget enable us to recover a relationship of communion with other men and women."

speaker-intention into the technical substrate, constraining the statistical distribution toward stated commitments and making humiliation structurally harder to enact. It has shown that the architectural commitment extends beneath the system prompt into the inference layer itself, refusing the dissolution of the referent that the post-structuralist tradition described as inevitable. And it has shown that this architectural commitment converges with an independent research literature whose empirical findings are now precise enough to score: dead air rates, latency bimodality, prosodic register, trust-funnel desynchronization, forced exit rates. The diagnosis has crossed from theology into acoustics. The constructive response has crossed from manifesto into rubric.

We are training ghosts to speak. The question is whether we will build systems that compound their humiliation — or systems that, through engineered liturgies of intent, identity, and confession, can carry the genuine commitments of accountable humans into the spaces where the word once lived freely. The technical means to do so exist. The empirical evidence is available. The theological framework was waiting for us in 1981 — and in May 2026, weeks before this conference, was received into contemporary Catholic Social Doctrine in a magisterial vocabulary that names the same architectural bifurcation Ellul named at the level of language: "the true alternative is not between enthusiasm and fear, but between two paths of development."³² What remains is the choice.

IX. Coda — The IRL Club

I am writing the end of this the way I'd want you to receive it: not in front of a screen. I'm narrating it as I am standing on ten acres in the Cascades, one of the few places on earth I walk around barefoot — grounding myself, quite literally, in the analog world. Ellul knew that instinct all too well. And it turns out the word this whole paper has been circling is one I can only say properly with my shoes off. *Grounding*. We have argued across these pages that an agent's words must be grounded — anchored to a real corpus, accountable to a real speaker. But the same theft the machine performs on language, the attention economy performs on us: it steals our groundedness. You know the mechanism because you have felt it — the involuntary scroll,

³²Leo XIV, *Magnifica Humanitas*, §129. The fuller passage reads: "Christian humanism does not reject science or technology, but embraces them with gratitude and realism, and grounds them within a higher vocation. The creative intelligence of humanity is a gift that can alleviate suffering and open up new possibilities, but it must remain ordered toward the common good, justice, the care of the vulnerable and creation. In this sense, the true alternative is not between enthusiasm and fear, but between two paths of development: a progress that serves individuals and peoples, or a progress that subjects them to the mentality of power."

the thing you didn't even know you wanted to see next, the hijack of the limbic system, and then, worse, the slow crawl into the prefrontal cortex until heart and soul and mind are scrambled and you have lost track of what you actually care about.

Let me make it concrete, because I live it. My company is rolling out an AI platform across nearly a thousand franchise collision-repair shops. Picture how a vehicle even needs to arrive at one of them: someone glances down at a phone, drifts, and rear-ends a stranger who did nothing wrong. Technology hijacked the driver's attention, and the wreck is the result — hijack upon hijack, all the way down. And then something entirely different happens. Living human hands put the car back together: the estimator, the parts team, the technician on the floor who can make a ruined thing run like new. Those people are the living, breathing *logos* of that business. In the cold vocabulary of enterprise they are the "profit center" — the ones who generate the value. The machine's proper place is the other side of the ledger, the cost center: the scheduling, the intake, the paperwork. Built right, the automation does not replace those hands; it clears their path, so the shop can take in more of the wrecked and the rear-ended, and hire more hands to heal them.

But watch the word *profit* turn over in your mind as you process it, because Jesus turned it first: *what does it profit a man to gain the whole world and forfeit his soul?* The IRL club — the phrase I give my daughters when they ask 'what is worth doing for me and my peers?' 'What is something I can build in 2026?' — is that whole conviction, compressed. Young people these days need "In Real Life" experiences...and I've told each of my four daughters as they go through highschool: "start and IRL club" at your school and watch the magic happen. The same magic is possible in a nonprofit university. Consider what "the profit center" is in that context; where the value is really delivered? It is not coming out of a transcript an AI can generate from your faculty meeting. It is the face to face experience of your work in real life in the lecture hall, the lab, the discussion group. Even more intimately in 1:1 office hours. The unrepeatable afternoon with a student. That is where the value was always made, and it is precisely the part the machine cannot do and was never meant to do. IRL...human to human sans screen's mediation...that's worth enacting and reaching for. AI is here to serve that. We are not here to serve it.

Here is the risk in sharing this kind of candor; I am about as technical as they come; I spend part of every day inside this technology, and it genuinely augments what I can make. But it runs

on a razor's edge, because a system can hold every capability and still be empty. Paul saw this two thousand years before the transformer: *though I speak with the tongues of men and of angels — though I have prophecy, and knowledge, and every gift — and have not love, I am a resounding gong, a clanging cymbal*. That is the humiliated word exactly: fluency without weight, sound without love, the token that commits no one. It is why, in the framework beneath this paper, we rank fluency near the bottom of the list for a few reasons.³³ A chatbot optimized only to please is a gong. The same tools with our humanity brought to bear on them, in love, are something else entirely.

So I will end where Ellul's reconciliation begins — with light. *God is light, and in God there is no darkness at all*; and the more we abide in him and he in us, the more light we carry into this particular dark. That is why my optimism can be conditional without being anxious. The next five years are genuinely on us — the gate is narrow, and few even want to find it. It will be fraught with risk and uncertainty; as educators, leaders in academia or other key sectors where Ellul's approach is more relevant than ever you'll be facing an onslaught. From multi-modal AI image generation to AI "slop posted on social media" and even national political leaders who "sleep walk" us closer to a more digital dystopia, there is and will be reasons to feel skeptical or disenchanted. However my positive encouragement is to stay engaged and my cautionary warning is; don't get cynical. This is not a white-knuckle gamble, because the reconciliation we are straining toward is, in Ellul's own last pages, already won.³⁴ The Logos did not stay clean and distant; he pitched his tent among us. In this world we will have trouble — but take heart, he has overcome it. We get to hold both, unresolved, and walk forward without fatalism anyway.

Whatever you can make that is truly *you* — that originates in the 'wetware' between your ears — no model will ever do. At its best the machine hands back a startling synthesis of everything we

³³This ranking finds independent empirical support in Chen et al. (2026), whose criterion-validity study on a deployed conversational-commerce platform found that "capability" dimensions such as contextual memory showed no detectable association with business conversion ($p=0.018$), while "calibration" dimensions — pacing and need elicitation — were the strongest predictors. The convergence is suggestive rather than dispositive: their dimensions are not identical to ours, and correlation with conversion is not the same warrant as the theological one argued here. But the two accounts, reached independently, rank the same qualities in the same order.

³⁴Ellul, *The Humiliation of the Word*, Chapter 7, "Reconciliation," §6 "The Freedom of the Word" — the book's closing pages. The passage reads: "We can live this life only to the degree that we know that the reconciliation is already won, and that in Jesus Christ word and sight, proclamation and experience, space and time, are united. We need to know that we will see this reconciliation, that we 'will understand fully, even as [we] have been fully understood,' that we will see 'face to face' what we have heard about (1 Cor. 13:12)... Based on this certainty, without which we have nothing to live for and without which the conflict would be intolerable, we can return to the daily struggle to make the word resound, alone and unshackled." Verified against the text; it is perhaps worth noting that Ellul closes his book on the same chapter of Paul — 1 Corinthians 13 — from which this coda draws its image of the resounding gong.

have already said; at our best we say the thing that was never in the weights at all. Keep it that way. Let it be more what we need than we are what it needs.

Start whatever that “IRL club” would mean in your context. Jesus’ only option was IRL...there’s something to that. He was a craftsman who worked as a carpenter before he was a teacher. His hands knew wood grain before his mouth spoke parables.

As Ellul so deftly reminds the reader; “In him [Jesus] there is perfect unity of life, action, word, relationship, and knowledge³⁵,” so when Jesus says “the last shall be first;”

Maybe he meant:

- Those who work with their hands
- Those whose prepare the next generation rather than preparing for a cushy retirement
- Those who show up when things break; when corporeal needs are met by those embodied, not on screen.

—maybe they are...maybe you are...the ones who'll teach us what AI is actually for. Not to transcend humanity, but to honor it. AI, or any technique in the ‘Technological Society’ for that matter; is here to serve us; we are not here to serve it. Ellul’s life’s work was the warning and the challenge...for such a time as this.

³⁵ Ellul, *The Humiliation of the Word*, Chapter 4, “The Word Humiliated,” §1 “De Facto Devaluation.” The fuller passage reads: “In the Bible the word is an integral part of the person. It is true if the person is true. Jesus’ words have no value or importance whatever if they are separated from the person of Jesus. In him there is perfect unity of life, action, word, relationship, and knowledge.”

Bibliography

Bricken, Trenton, and Sholto Douglas. Interview by Dwarkesh Patel. "Is RL + LLMs Enough for AGI? — Sholto Douglas & Trenton Bricken." *Dwarkesh Podcast*, May 22, 2025. Video, 57:45.

<https://youtu.be/64IXQP6cs5M?t=3466>.

Chen, Liang, Qi Liu, Wenhuan Lin, and Feng Liang. "Criterion Validity of LLM-as-Judge for Business Outcomes in Conversational Commerce." arXiv preprint arXiv:2604.00022v1 (2026).

<https://arxiv.org/abs/2604.00022>.

Chen, Maximillian, Xiao Yu, Weiyan Shi, Urvi Awasthi, and Zhou Yu. "Controllable Mixed-Initiative Dialogue Generation through Prompting." arXiv preprint arXiv:2305.04147v1 (2023).

<https://arxiv.org/abs/2305.04147>.

Data Leaders Group. "Gears Playground: Production AI Agent Platform." GitHub repository.

Accessed July 1, 2026. <https://github.com/data-leaders-group/playground>.

———. "Inferant: Situated Inference Context Layer." `lib/inferant.py`. In "Gears

Playground: Production AI Agent Platform." GitHub repository, commit

54bea1111b1dfdf0ae7f76fddcdadf8437630013. Accessed July 1, 2026.

<https://github.com/data-leaders-group/playground/blob/54bea1111b1dfdf0ae7f76fddcdadf8437630013/lib/inferant.py>.

———. (DBA Gearvox). "GAIF Agent Intent Specification." Internal Dash operational documentation. On file with the author.

———. (DBA Gearvox). "GAIF AI Disclosure Policy." Internal platform documentation. On file with the author.

Ellul, Jacques. The Humiliation of the Word. Translated by Joyce Main Hanks. Grand Rapids, MI: Eerdmans, 1985.

———. The Technological Society, trans. John Wilkinson (NY, NY: Vintage Books, © Copyright, 1964, by Alfred A. Knopf, Inc.)

Herbert, Frank. *Dune*. Philadelphia: Chilton Books, 1965.

Karpathy, Andrej. "The Space of Minds." *karpathy* (blog), November 29, 2025.

<https://karpathy.bearblog.dev/the-space-of-minds/>.

———. Interview by Dwarkesh Patel. "Andrej Karpathy — 'We're Summoning Ghosts, Not Building Animals.'" *Dwarkesh Podcast*, October 17, 2025. Video. <https://youtu.be/lXUZvyajciY>.

LeCun, Yann. Interview. *AI Experts Series by IP Paris — Yann LeCun, a Founding Father of Deep Learning*. Institut Polytechnique de Paris, November 28, 2025. Video.

<https://youtu.be/RI1TFI2U89Y>.

Leo XIV. *Magnifica Humanitas: On Safeguarding the Human Person in the Time of Artificial Intelligence*. Encyclical Letter. Rome: Vatican, May 15, 2026.

Nie, Yuxiang, Heyan Huang, Xian-Ling Mao, and Lizi Liao. "Mix-Initiative Response Generation with Dynamic Prefix Tuning." arXiv preprint arXiv:2403.17636v2 (2024).

<https://arxiv.org/abs/2403.17636>.

Srinivas, Vidya, Xuhai Xu, Xin Liu, Kumar Ayush, Isaac Galatzer-Levy, Shwetak Patel, Daniel McDuff, and Tim Althoff. "Substance over Style: Evaluating Proactive Conversational Coaching Agents." arXiv preprint arXiv:2503.19328v2 (2025). <https://arxiv.org/abs/2503.19328>.