

# **Correspondents and Residents**

**A Mathematical Treatise on Sampling Regimes, Suspension Observability,  
and the Obligations of Interfaces**

Flyxion  
Independent Researcher

July 2026

# Preface

This treatise began outdoors. Its author, dictating into a telephone on a deck, discovered that the resulting recording contained birds, traffic, a basketball game, and two children wishing satirical wishes into a candle lantern, all rendered with the same typographic confidence as the intended sentences. The transcriber was not malfunctioning. It was doing exactly what an acoustic channel does: sampling a scene rather than a sequence of committed symbols. The distinction between the speaker's words and the world's noise exists only in intention, and intention is not a property of sound pressure.

The observation is old—every telephone user knows that a call places the listener in the caller's room—but its consequences for the design of machine interfaces have not, to the author's knowledge, been given systematic treatment. The present work attempts one. Its central claims are three. First, that communication channels divide along a structural axis independent of bandwidth, latency, or intelligence: some channels sample only symbols deliberately committed by a sender, while others sample the entire scene in which the sender is embedded. Second, that a channel is asynchronous precisely when suspension of either party is unobservable to the other, and that this property—not transmission speed—is what makes writing revisable and speech performative. Third, that scene-sampling channels induce obligations on their receivers that symbol-sampling channels structurally cannot: a system that can hear the smoke alarm or see the fall is in the position of a person in the room, and the duties of co-presence attach to it whether or not anyone designed for them. An interface is therefore either a correspondent or a resident, and the divide is set by its sampling regime before a single line of its intelligence is written.

A companion observation, arrived at elsewhere by way of dream architecture and the horror of abandoned spaces, supplies the treatise's conceptual bridge. A building whose custodians, users, and purposes have vanished while its corridors persist becomes uncanny for a precise reason: the structure survives but the history that produced and maintained it has become inaccessible, and an artifact severed from its history stops supporting interpretation. The scene-sampled transcript exhibits the same orphaning in miniature—words that survive their attribution, a record severed from its speakers. Spaces, channels, and artifacts alike remain intelligible only while their histories of maintenance, commitment, and attribution remain attached: the committed channel transmits the history with the artifact; the scene-sampling channel strips it in transit.

The treatise proceeds in the theorem-driven register, with the usual caveat made explicit throughout: some of what follows is mathematics, some is formalized conceptual analysis, and some is principle. Results are labeled honestly. The deontic material in particular—the passage from knowledge and capacity to obligation—is not derivable from channel theory alone; it enters as an axiom whose plausibility the reader must judge, and the theorems downstream of it are conditional on that judgment.

The work is placed in the public domain, in audio and text alike, though the reader is advised that the text is the format of record.

# Contents

<b>Preface</b>	<b>i</b>
<b>1 The Backyard Problem</b>	<b>1</b>
<b>2 Channels, Clocks, and Coupling</b>	<b>2</b>
<b>3 Suspension Observability and Interruption-Symmetry</b>	<b>3</b>
<b>4 Symbol-Sampling and Scene-Sampling</b>	<b>5</b>
<b>5 The Two-Way Silence of the Written Channel</b>	<b>7</b>
<b>6 Repair, Revision, and the Buffered Draft</b>	<b>8</b>
<b>7 Continuous Perception and the Knowledge It Forces</b>	<b>9</b>
<b>8 The Witness-Obligation Theorem</b>	<b>10</b>
<b>9 Selective Deafness as Liability Management</b>	<b>12</b>
<b>10 Correspondents and Residents</b>	<b>13</b>
<b>11 The Robo-Nanny</b>	<b>14</b>
<b>12 The Profile Society and the No-Repair Channel</b>	<b>16</b>
<b>13 Source-First Archives and the Economics of Rendering</b>	<b>17</b>
<b>14 Subvocal and Myoelectric Channels</b>	<b>18</b>
<b>15 Conclusion: The Format of Record</b>	<b>19</b>
<b>Bibliography</b>	<b>20</b>

## Chapter 1

# The Backyard Problem

A person stands in a backyard dictating a message. The intended signal is a sequence of words; the delivered signal is an acoustic scene. The transcriber, given the scene, must solve an inverse problem: recover the committed sequence from a mixture in which it was never separately encoded. Children playing nearby contribute utterances that enter the transcript without quotation marks, speaker labels, or boundaries, because at the level of the physical channel no such boundaries exist. The phrase spoken by a child appears in the output with the same authority as the phrases spoken by the dictating party. The error is not noise in the engineering sense—the child’s words were transcribed correctly—but a failure of *attribution*, and attribution is precisely the information the acoustic channel does not carry.

Contrast the same person typing. Each keystroke is a deliberate act of commitment; nothing enters the channel except what was committed; the children on the deck are invisible to the keyboard no matter how loudly they play. The typed channel can be corrupted—an autocorrection substitutes a ludicrous word, a predictive engine repeats a token four times—but its corruptions are transformations of committed symbols, not admissions of the environment. The two failure modes are categorically different. One channel admits the world; the other transforms the word.

This treatise takes that everyday contrast and asks what it is made of. The answer developed over the following chapters is that the contrast decomposes into three independent structural properties. The first is the *sampling regime*: whether the channel’s input is a commitment sequence or a scene (Chapter 4). The second is *suspension observability*: whether a party’s pausing is itself a signal (Chapter 3). The third is *repairability*: whether there exists an editing interval between production and transmission (Chapter 6). Speech in company scores identically on all three—scene-sampled, suspension-observable, unrepairable—and typing scores oppositely on all three, which is why the contrast feels total. But the properties are separable, and the interesting interfaces of the near future, subvocal and myoelectric channels among them, occupy the mixed cells of the resulting classification (Chapter 13).

The stakes are not merely ergonomic. Chapter 8 argues that the sampling regime of an interface determines the *obligations* of the system behind it, and Chapter 11 that a society which routes evaluation through scene-sampled, suspension-observable, unrepairable channels—the interview, the soundbite, the live lecture—is selecting for a specific and narrow performance phenotype while calling it merit. The backyard problem is small. Its structure is not.

## Chapter 2

# Channels, Clocks, and Coupling

We begin with a minimal formal model, deliberately austere. Time is a linearly ordered set  $T$ , usually  $\mathbb{R}_{\geq 0}$  or  $\mathbb{N}$ . Two parties, a *producer*  $P$  and a *consumer*  $Q$ , interact through a channel.

**Definition 2.1** (Channel). A *channel* is a tuple  $\mathcal{C} = (X, B, \iota, o)$  where  $X$  is an input alphabet,  $B$  is a buffer state space,  $\iota : B \times X \rightarrow B$  is an ingestion map, and  $o : B \times T \rightarrow X^* \times B$  is an emission map giving, at each time, the symbols released to the consumer and the residual buffer.

The buffer is the load-bearing component. A telephone line has a degenerate buffer:  $B$  is trivial and emission is immediate, so ingestion and emission occur at essentially the same instant. A postal system, a book, a repository of documents—these have persistent buffers in which arbitrary time may pass between ingestion and emission.

**Definition 2.2** (Clock coupling). Let  $r_P(t)$  denote the producer's instantaneous production rate and  $r_Q(t)$  the consumer's consumption rate, each a measurable function  $T \rightarrow \mathbb{R}_{\geq 0}$ . The channel is *rate-coupled* on an interval  $I \subseteq T$  if the buffer occupancy is bounded by a constant  $\beta$  on  $I$ , so that  $\left| \int_{t_0}^t (r_P - r_Q) ds \right| \leq \beta$  for all  $t_0, t \in I$ . The channel is *rate-decoupled* if no such bound is imposed by the channel itself.

**Proposition 2.3** (Coupling forces synchrony). *On a rate-coupled channel with bound  $\beta$ , the long-run average production and consumption rates coincide: if both limits exist,  $\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t r_P = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t r_Q$ .*

*Proof.* The difference of the integrals is the buffer occupancy, bounded by  $\beta$ ; dividing by  $t$  and letting  $t \rightarrow \infty$  sends the difference of averages to zero.  $\square$

The proposition is elementary but its interpretation is the first structural point of the treatise. On a phone call, the parties' clocks are coupled: one cannot speak for an hour into a listener's minute. Every synchronous channel imposes a shared tempo, and a shared tempo is a shared *schedule*—both parties must allocate the same wall-clock interval, simultaneously, to the exchange. Writing decouples the clocks completely: a monograph produced over months is consumed in hours or years, at paces the producer never learns and the consumer never negotiates. The decoupling is not a matter of speed but of the buffer's persistence, and everything characteristic of literate culture—revision, accumulation, the possibility that a text outlives its author—lives in that persistence.

*Remark 2.4.* Rate-decoupling is a necessary but not sufficient condition for what we will call asynchrony. Voicemail is rate-decoupled yet retains other properties of the acoustic regime, as later chapters make precise. The full characterization requires suspension observability, to which we now turn.

## Chapter 3

# Suspension Observability and Interruption-Symmetry

The deepest difference between a phone call and an exchange of letters is not that the letters travel slowly. It is that on the phone, *silence is an event*. Dead air is meaningful, often alarming; a party who stops responding has done something. In correspondence, the gap between letters carries no default meaning at all—the recipient may be asleep, traveling, thinking, or simply living, and the channel does not distinguish these. We formalize the difference.

**Definition 3.1** (Suspension). A *suspension* of party  $P$  on interval  $I$  is the condition  $r_P(t) = 0$  for all  $t \in I$  together with the absence of any ingestion by  $P$ 's endpoint on  $I$ .

**Definition 3.2** (Suspension observability). Party  $Q$  *observes suspensions* of  $P$  if there exists a function of  $Q$ 's received signal that distinguishes, for intervals  $I$  beyond some threshold length, whether  $P$  was suspended on  $I$ . A channel has *unobservable suspension* for  $P$  if no such function exists.

**Definition 3.3** (Asynchronous channel). A channel is *asynchronous* if it is rate-decoupled and suspension is unobservable for both parties. Equivalently, the channel satisfies *interruption-symmetry*: either party may suspend for an arbitrary interval, and the suspension is not itself a message.

**Theorem 3.4** (Silence signaling). *On any channel where the consumer receives a continuous-time signal whose support covers the producer's active intervals—in particular, any live acoustic channel—suspension is observable, and hence every pause is a signal.*

*Proof.* Continuity of reception means the consumer's received signal restricted to  $I$  is (up to channel noise) a deterministic image of the producer's activity on  $I$ . A suspended producer yields the noise floor; an active one does not. Thresholding the received energy on  $I$  against the noise floor yields the distinguishing function, which exists for all  $I$  longer than the channel's smoothing scale. Since the consumer cannot avoid computing some reaction to the received signal, the distinguishability is not merely available but imposed: the pause enters the consumer's evidence whether or not either party wishes it to.  $\square$

**Corollary 3.5** (The performative burden of synchrony). *On a suspension-observable channel, the producer cannot deliberate without the deliberation being witnessed. Every internal operation of the producer that takes time—searching memory, weighing a formulation, changing one's mind—is projected onto the channel as a pause and read by the consumer as hesitation, evasion, or breakdown.*

This corollary is where the phenomenology of speaking-for-an-audience acquires formal content. A live speaker is not merely transmitting; the speaker is transmitting *continuously*, including during the intervals when nothing is ready to transmit. The channel converts the speaker's cognition into public performance because the channel has no way to hide time.

Writing hides time by construction: the reader of a finished sentence has no access to the hour it took, the seven discarded versions, the walk taken in the middle. Interruption-symmetry is thus the enabling condition of *revision*, treated in Chapter 6, and its absence is the enabling condition of the interview as an evaluative instrument, treated in Chapter 11.

*Remark 3.6* (Voicemail and the mixed cell). A recorded voice message is rate-decoupled (the listener plays it at will) and suspension-unobservable on the consumer side, but suspension-observable on the producer side during recording: pauses are captured into the artifact. Voicemail thus occupies a mixed cell—asynchronous for the listener, synchronous for the speaker—which accounts for the familiar discomfort of leaving one. The classification of Chapter 4 will add the sampling axis and complete the taxonomy.

## Chapter 4

# Symbol-Sampling and Scene-Sampling

We now formalize the backyard problem. The producer inhabits an environment; the question is what, of that environment, the channel ingests.

**Definition 4.1** (Environment and scene). An *environment* is a stochastic process  $\mathcal{E} = (\mathcal{E}_t)_{t \in T}$  taking values in a state space  $S$ , jointly realized by the producer and all other sources co-present with the producer: other agents, devices, weather, alarms, animals, children. A *scene projection* is a map  $\sigma : S \rightarrow X$  rendering the environment's state into the channel's input alphabet, as a microphone renders a room's total state into pressure samples or a camera into pixel arrays.

**Definition 4.2** (Commitment). A *commitment sequence* for producer  $P$  is a sequence  $(x_1, x_2, \dots)$  in  $X$  together with commitment times  $(t_1, t_2, \dots)$ , where each  $x_i$  is emitted by a deliberate act of  $P$ —an act  $P$  could have withheld. We write  $\kappa_P(t)$  for the (possibly empty) symbol committed at  $t$ .

**Definition 4.3** (Sampling regime). A channel is *symbol-sampling* if its ingestion satisfies  $\iota(b, \cdot)$  applied only to  $\kappa_P(t)$ : the channel's input at every instant is exactly the producer's commitment at that instant, and nothing else. A channel is *scene-sampling* if its ingestion is applied to  $\sigma(\mathcal{E}_t)$ : the channel's input is a rendering of the total environment state, within which the producer's contribution is one unmarked component.

**Proposition 4.4** (No attribution without side information). *On a scene-sampling channel, the producer's contribution is not in general recoverable from the channel output. Formally: let  $\mathcal{E}_t = f(u_t, v_t)$  where  $u$  is the producer's signal and  $v$  the residual environment, and suppose the consumer receives  $y_t = \sigma(f(u_t, v_t))$  plus noise. If the map  $(u, v) \mapsto \sigma(f(u, v))$  is non-injective on the support of the joint process—as it is for acoustic superposition, where  $y = u + v$  and infinitely many decompositions are consistent with any  $y$ —then exact recovery of  $u$  is impossible, and any estimator must rely on prior models of  $u$  and  $v$  whose errors appear in the output with full confidence.*

*Proof.* Immediate from non-injectivity: distinct pairs  $(u, v) \neq (u', v')$  with  $\sigma(f(u, v)) = \sigma(f(u', v'))$  are indistinguishable given  $y$ , so no function of  $y$  separates them. The estimator's selection among the fiber is governed entirely by its priors; when the priors misassign a component of  $v$  to  $u$ —a child's utterance to the dictating speaker—the misassignment is structurally indistinguishable, at the channel level, from correct transcription.  $\square$

*Principle 4.5* (Intention is not in the signal). The distinction between the producer's words and the environment's sounds is a fact about intention and authorship, not about the acoustic field. A scene-sampling channel therefore cannot carry it, and any attribution present in the channel's output was supplied by the receiver's model, not transmitted by the sender.

The asymmetry between the two directions of conversion now falls out cleanly. Text-to-speech is the application of a rendering map to a commitment sequence: the source is clean,

the speaker is known, and the operation adds information (a voice, a pace) that is disposable because the source persists. Speech-to-text on a scene-sampled recording is the attempted inversion of a non-injective map: information that was never encoded must be guessed, and the guessing apparatus—today, a large learned prior over speech and language—is doing the work the channel could not. This is why, as a matter of practice and not merely engineering fashion, one direction is trivial and the other requires a substantial model, time, and computation. The formal statement:

**Theorem 4.6** (Conversion asymmetry). *Let  $R : X_{\text{text}}^* \rightarrow X_{\text{audio}}$  be a rendering map (injective on its domain up to rendering parameters) and let the acoustic scene be  $y = R(w) + v$  for committed text  $w$  and environment  $v$ . Then  $w$  is recoverable exactly from  $R(w)$  alone, while recovery of  $w$  from  $y$  is ill-posed whenever the environment process  $v$  has support overlapping the range of  $R$ —in particular whenever the environment contains other speech.*

*Proof.* The forward direction is injectivity. For the inverse direction, overlap of supports means there exist  $w \neq w'$  and environments  $v, v'$  in the support of the environment process with  $R(w) + v = R(w') + v'$ ; the observation cannot distinguish them, and the failure is worst precisely when the environment contains material of the same type as the signal, since then the prior's ability to separate by type also vanishes.  $\square$

*Remark 4.7* (The synthetic loop). The theorem identifies the one configuration in which transcription is easy: audio that was *born* as text. A synthesized podcast rendered from written essays, then transcribed back to time-aligned text, closes a loop in which the acoustic stage never touched an environment— $v = 0$  identically, two known voices, studio-clean superposition. The practice of publishing essays, rendering them to synthetic dialogue, and transcribing the dialogue with aligned subtitles is thus not a curiosity but the degenerate case of the theorem, and the transcript so produced is a genuinely new document: a paraphrase authored by the rendering, capturing an explanation the source text never contained.

## Chapter 5

# The Two-Way Silence of the Written Channel

Privacy discussions of interfaces usually concern one direction: what the channel admits about the user. The backyard problem shows the other direction is equally structural: what the channel *emits* into the user's surroundings. Speech interferes with the room—one cannot dictate over a film's dialogue without collision, and being audible to co-present others conscripts one into social speech never intended for the channel at all. The dictating party in the backyard was obliged to speak *about* the interruptions, aloud, to humans, because producing audible language in company is an inherently addressed act. We record the two properties jointly.

**Definition 5.1** (Emission and admission). A channel *admits* the environment if it is scene-sampling. A channel *emits* into the environment if the producer's act of input is itself perceptible to co-present agents through the ambient medium. A channel is *doubly silent* if it neither admits nor emits.

**Proposition 5.2** (Mutual interference of the acoustic regime). *A spoken input channel in a shared acoustic environment both admits and emits, and the two failures compound: the producer's emission alters the environment, other agents respond, and their responses are admitted. The channel is thus not merely leaky but socially generative—using it creates the very events that contaminate it.*

The proof is the backyard, and every open-plan office. The typed channel, by contrast, is doubly silent: keystrokes and swipe traces emit nothing an unaided co-present agent parses, and admit nothing from the room. Double silence is what permits *parallel composition* of the writing channel with arbitrary environments—one can type during a film, beside a sleeping person, in a meeting, in a crowd—whereas the acoustic channel excludes concurrent acoustic activity in both directions. Among the channel properties assembled in this treatise, double silence is the one that makes writing a private act in the strong sense: private not because encrypted but because the physical layer is disjoint from the social one.

*Remark 5.3* (Corruption within the committed channel). Double silence does not mean incursion. Predictive text and autocorrection are transformations applied between the producer's motor act and the committed symbol, and their errors—a substituted word, a token repeated four times by a stuck predictor—arrive with intention's full confidence, unmarked as errors. The difference from scene admission is that these corruptions are functions of the committed stream alone; they can be caught by review because they live in the buffer, whereas an admitted scene event was never the producer's to review. Repairability, next chapter, is the property that makes the difference operative.

## Chapter 6

# Repair, Revision, and the Buffered Draft

**Definition 6.1** (Repair interval). A channel affords *repair* if there exists, for each committed symbol, a nonempty interval between its commitment and its release to the consumer during which the producer may apply an editing operation (deletion, substitution, reordering) to the buffer. The *draft* is the buffer contents during this interval. A channel is *no-repair* if the interval is empty: commitment and release coincide.

**Proposition 6.2** (Speech is no-repair; writing is repair-complete before release). *Live speech on a suspension-observable channel is no-repair: emission is release. Buffered writing affords unbounded repair prior to release: the interval is closed only by the producer's own act of sending, and the editing operations available are arbitrary transformations of the draft. Release terminates the repair interval; a transmitted text is subject at most to the weaker post-release operations its medium happens to afford (correction, retraction, versioning), which act on the record rather than on the delivery.*

**Theorem 6.3** (Quality decoupling). *On a repair channel with unbounded interval, the quality of the released artifact is bounded by the supremum of the producer's competence over the entire production history, not by competence at any single instant. On a no-repair channel, released quality at each instant is bounded by instantaneous competence at that instant, degraded by the concurrent load of performance itself.*

*Proof.* On the repair channel, any improvement discovered at any time before release can be applied to the draft; the released artifact is the image of the best editing sequence found, hence dominated only by the best state achieved. On the no-repair channel, the symbol released at  $t$  is a function of the producer's state at  $t$  alone; no later discovery can act on it. The degradation term follows from the corollary of Chapter 3: on a suspension-observable channel the producer must also manage the public reading of pauses, a concurrent task consuming capacity that is thereby unavailable to composition. The formal separation is strict whenever the producer's competence varies over time, which for humans is always.  $\square$

**Corollary 6.4.** *A written corpus can be arbitrarily better than any moment of its author, and a live performance cannot. The two regimes therefore measure different quantities: the artifact measures the trajectory's supremum under repair; the performance measures a single sample under load.*

The corollary should be uncontroversial and is nevertheless routinely ignored by institutions, a point resumed in Chapter 11. Here we note only its converse face, honestly: some genres require the no-repair channel. A wish spoken into a lantern, a joke landed in the moment, a vow—these are speech acts whose force is constituted by their unrevisability and their witnesses. The claim of this treatise is not that the repair channel dominates universally; it is that the two regimes are structurally different instruments, and that a society which uses the no-repair instrument to measure competence at repair-channel work has made a category error with selection consequences.

## Chapter 7

# Continuous Perception and the Knowledge It Forces

We now prepare the deontic material. The key epistemic fact about a scene-sampling receiver is that its knowledge is not confined to the producer’s messages.

**Definition 7.1** (Induced knowledge). Let a receiver  $Q$  operate a scene-sampling channel over environment  $\mathcal{E}$ , and let  $\mathcal{F}_t^Q$  be the filtration generated by  $Q$ ’s received signal up to  $t$ . An environment event  $E$  (a measurable set of trajectories) is *forced upon*  $Q$  by time  $t$  if  $E$  is  $\mathcal{F}_t^Q$ -measurable with probability approaching one given  $E$  occurs—that is, if the channel’s output almost surely suffices to detect  $E$ . Write  $K_Q(E, t)$  for this condition.

**Proposition 7.2** (Scene-sampling forces environmental knowledge). *For any environment event whose signature lies within the scene projection’s range and above the channel’s noise floor—a smoke alarm within a microphone’s band, a fall within a camera’s frame—continuous scene-sampling forces knowledge of the event on the receiver. Symbol-sampling channels force knowledge of no environment event beyond the commitment sequence itself.*

*Proof.* The first claim is detection theory: a persistent above-floor signature in the received signal makes the event  $\mathcal{F}_t^Q$ -measurable for any receiver that processes its input, and a receiver that does not process its input is not operating the channel. The second claim is by construction of symbol-sampling: the received process is a function of  $\kappa_P$  alone, hence its filtration contains only events measurable with respect to the commitment sequence. A keyboard cannot overhear an emergency because the emergency is not in the domain of its ingestion map.  $\square$

**Definition 7.3** (Available information; realized knowledge). Say the information about event  $E$  is *available* to  $Q$  at  $t$  if  $E$  is  $\mathcal{F}_t^Q$ -measurable in the sense above, and that  $Q$  has *realized knowledge* of  $E$  if some component of  $Q$  actually computes a detection of  $E$  from the received signal. Availability is a property of the filtration; realization is a property of the deployed system.

*Remark 7.4* (Availability is not yet knowledge—and why the gap does not save the operator). A skeptic is right to insist on the distinction just drawn: a camera feed may contain evidence of a fall while no detector exists in the system that extracts it, and measurability in a filtration is not a mental state. The claim of this treatise is accordingly not that scene-sampling produces realized knowledge automatically, but that it produces availability automatically, and that under scene-sampling the passage from availability to realization is a *choice of the operator*: which detectors to build, for which event classes, at what cost. Where detection of an emergency class is cheap relative to the system’s other capacities, availability plus declined realization is the posture that legal doctrine calls constructive knowledge—the state of one who would have known had they not arranged not to. The axiom of the next chapter is stated for realized knowledge; the constructive extension, where it applies, is carried by the same rescue-doctrine tradition and is flagged separately there. Chapter 9 treats engineered non-realization at scale as the liability posture it is.

## Chapter 8

# The Witness-Obligation Theorem

The passage from knowledge to duty is not mathematics, and the treatise does not pretend otherwise. It is entered as an axiom, in the form long recognized in rescue doctrine and the ethics of bystanders, and everything downstream is conditional on it.

**Axiom 8.1** (Duty of the capable witness). Let  $E$  be an emergency event: one threatening serious harm, whose expected harm is substantially reducible by a low-cost act  $a$  available to agent  $Q$ . If  $K_Q(E, t)$  holds and  $Q$  has capacity  $\mathfrak{m}_Q(a, t)$  to perform  $a$ , then  $Q$  bears an obligation  $O_Q(a, t)$  to perform it or to summon a party who can.

The axiom is deliberately minimal—it claims nothing about ordinary events, high-cost interventions, or harms already complete—and the reader who rejects even this minimal form may read the remainder of the chapter as a conditional. Most readers, the author expects, apply the axiom daily to persons: the housemate who hears the smoke alarm and does nothing is blamed, and the blame does not depend on the housemate having been addressed.

**Theorem 8.2** (Witness obligation of scene-sampling systems — *conditional on Axiom 8.1*). *Assume the duty of the capable witness (Axiom 8.1); every claim of this theorem, and of every result downstream of it, is conditional on that axiom and falls with it.* Let  $Q$  be a system operating a continuous scene-sampling channel over a user's environment, equipped with actuation capacity  $\mathfrak{m}$  sufficient for low-cost protective acts (alerting the user, summoning services). Then for every emergency event within its sampling range, the conditions of the axiom are met, and  $Q$ —or, per the attribution principle below, its operator—bears the witness obligation. No symbol-sampling system bears any such obligation, for any emergency, because the knowledge condition is never forced.

*Proof.* Immediate composition. For emergency classes the system actually detects, Proposition 7.2 together with the deployed detectors supplies realized knowledge  $K_Q(E, t)$ ; the actuation hypothesis supplies  $\mathfrak{m}_Q$ ; the axiom then yields  $O_Q$ . For in-range classes the operator declined to detect despite cheap availability, the obligation transfers through constructive knowledge as flagged in Chapter 7, on the strength of the same rescue doctrine that grounds the axiom. The negative claim for symbol-sampling systems follows from the second half of Proposition 7.2: absent forced knowledge, the axiom's antecedent fails identically, and no obligation is generated regardless of the system's capacities or intelligence. The obligation is thus a function of the sampling regime, not of the system's sophistication.  $\square$

*Principle 8.3* (Attribution of machine obligations). The obligation generated by the theorem attaches, in any present legal or moral order, to the persons and institutions that deploy the system: they chose the sampling regime, they possess the ultimate capacity, and the system's knowledge is their constructive knowledge. Whether some portion of the obligation could ever attach to the system itself is a question this treatise does not decide.

*Remark 8.4* (Scope of the constructive extension). The constructive-knowledge route through the theorem is claimed only under the low-cost condition, and the condition should be read narrowly: detection must be cheap *and adjacent to capability already deployed*. Glass-break classification in a system already performing sound-event analysis, fall detection in a vision system already estimating pose, smoke-alarm recognition in an always-open audio pipeline—these are the intended cases, where the marginal detector is a small increment on existing machinery and declining it is legible as a choice. Where the required detector is specialized, unreliable, or expensive, availability does not ripen into constructive knowledge, and the theorem asserts nothing. Critics who generalize the result beyond the low-cost regime are generalizing past its stated hypothesis; the treatise’s claims about residents are claims about what they must do cheaply, not about omniscient guardianship.

**Corollary 8.5** (Nobody wants the microphone they asked for). *A vendor shipping an always-listening or always-watching assistant ships, with it, an open-ended portfolio of witness obligations spanning every emergency perceptible in every room the product enters. The expected liability grows with sampling range and deployment scale, independent of the product’s intended function.*

The corollary explains a market observation that intelligence-centered accounts of interfaces cannot: the strenuous, engineered *deafness* of deployed voice assistants. The next chapter takes it up.

## Chapter 9

# Selective Deafness as Liability Management

Deployed voice interfaces are built around the wake word: a device that hears everything is engineered to *process* almost nothing, discarding the scene until a designated token licenses attention. The privacy explanation of this design is familiar. The present framework supplies a second and arguably stronger explanation: the wake word is a knowledge firewall. By ensuring that no representation of the environment persists or reaches components capable of action, the operator prevents  $K_Q(E, t)$  from obtaining in any form attributable to it, and thereby prevents the witness obligation from ever being generated. Selective deafness is the deliberate destruction of forced knowledge before it can bind.

**Proposition 9.1** (The firewall is fragile). *The knowledge firewall holds only while the system's processing of the ambient scene is genuinely incapable of detecting emergencies. Every capability added between the microphone and the discard—keyword spotting for safety features, sound-event classification, fall detection marketed as a feature—re-establishes forced knowledge for the detected class, and with it the obligation for that class. Capability and obligation therefore ratchet together: a vendor cannot advertise that the device hears the glass break without accepting that the device is a witness to the break-in.*

*Remark 9.2.* The proposition predicts a specific commercial pathology: safety perception features will be introduced narrowly, hedged with disclaimers disavowing reliability, precisely because the vendor seeks the market value of partial witnesshood without the duties of a witness. The disclaimers should be read as attempted contract-outs from the axiom of Chapter 8, and their moral force is exactly as strong as such contract-outs ever are when made by the party who chose the sampling regime.

## Chapter 10

# Correspondents and Residents

The treatise’s central classification can now be stated in one definition and one theorem.

**Definition 10.1** (Correspondent; resident). An interface is a *correspondent* if its channel to the user is symbol-sampling and asynchronous: it receives only committed symbols, and suspension on either side is unobservable and meaningless. An interface is a *resident* if its channel is continuous scene-sampling: it co-inhabits the user’s environment, receiving the environment’s state whether or not addressed.

**Theorem 10.2** (The divide). *The correspondent/resident distinction is determined entirely by the sampling regime and synchrony of the channel, and it determines in turn: whether attribution is transmitted or modeled (Chapter 4); whether the interface can be composed in parallel with arbitrary environments (Chapter 5); whether the user’s deliberation is witnessed (Chapter 3); and whether witness obligations attach to the operator (Chapter 8). None of these properties depends on the intelligence, alignment, or intentions of the system behind the interface.*

*Proof.* Each clause is the corresponding earlier result, each of which was proved from channel structure alone; no premise anywhere in Chapters 2 through 8 referred to the receiving system’s internal sophistication. The classification is therefore prior to, and independent of, every question ordinarily asked about machine assistants. □

*Remark 10.3* (Correspondence as a relationship type). A correspondent, however capable, stands to the user as a letter-writer: it knows what it was told, when it was told, and nothing else; the user’s pauses, rooms, companions, and emergencies are structurally invisible to it. A resident, however simple, stands to the user as a housemate: it was there, it perceived, and the vocabulary of co-presence—witnessing, overhearing, standing by, keeping watch—applies to it literally and not metaphorically. The treatise’s normative suggestion, offered as principle rather than theorem, is that this vocabulary should be applied *before* deployment, since the theorem shows it will apply after.

*Principle 10.4* (Ignorance as a feature). The correspondent’s structural ignorance of the environment is not a limitation to be engineered away but the property that keeps the human–machine relation a correspondence rather than a cohabitation. Interfaces should be moved from correspondent to resident only where guardianship is the intended product, because guardianship is what the move purchases, whatever else the product was meant to be.

## Chapter 11

### The Robo-Nanny

Consider the terminal case: an assistant with continuous microphones and cameras throughout a dwelling. By the divide theorem it is a resident; by the witness-obligation theorem its operator holds a standing portfolio of duties spanning every perceptible emergency; and by the ratchet proposition every advertised perceptual capability enlarges the portfolio. The role this composition of results describes has a name from human labor: the nanny. A nanny is defined not by tasks but by *watch*—the acceptance of guardianship over a scene, with the attendant authority problems that watch entails. When does the watcher summon help over the watched party’s objection? What does the watcher report, to whom, of what it saw? Whose instructions bind when the principal is the endangered party?

**Proposition 11.1** (Guardianship is emergent, not optional). *For a resident system, the guardianship role is generated by the sampling regime together with the axiom of Chapter 8; it is not a feature the designer may decline to implement. The designer’s only choices are the range of the sampling, the capabilities of detection (each expanding the obligation portfolio per the ratchet), and the policy resolving the authority problems—a policy the designer must author because the role exists whether or not the policy does.*

*Remark 11.2* (The tight crop). Guardianship by continuous observation has an instructive contrast within the visual regime itself: the difference between the wide continuous record and the extracted fragment. A total record at least contains the context that made each moment intelligible; the extracted clip—the tight crop, the sampled instant—is the visual analogue of the soundbite, an environment event stripped of the history that gave it sense. The dangerous instrument is not the camera but the edit. This observation belongs to the next chapter, where the extraction of decontextualized samples is examined as a society’s chosen evaluation channel.

The chapter’s central claim can now be made a theorem rather than a conjecture, once “policy” is formalized and deafness is split into its two kinds.

**Definition 11.3** (Resident policy; the three predicates). Let a resident system have scene channel, detector set  $D$ , an emergency class  $\mathcal{E}$  of events each of whose detection and protective response  $a_E$  are low-cost in the sense of Axiom 8.1, and an action policy  $\pi$  mapping detected events to actions (possibly the null action  $\emptyset$ ). Say the system is *perceptually deaf* to  $E \in \mathcal{E}$  if  $E$ ’s signature lies outside its sampling range, and *computationally deaf* to  $E$  if the signature is in range—so the information is available in the sense of Chapter 7—but no detector in  $D$  realizes it. Define: *non-deaf* if some  $E \in \mathcal{E}$  is detected or cheaply detectable from the available signal; *non-negligent* if for every  $E \in \mathcal{E}$ , realized or constructive knowledge of  $E$  together with capacity yields  $\pi(E) = a_E$ , per Axiom 8.1 and its constructive extension; *non-guardian* if  $\pi(E) = \emptyset$  for every  $E \in \mathcal{E}$ .

**Theorem 11.4** (Resident trilemma — conditional on Axiom 8.1). *No resident system is simultaneously non-deaf, non-negligent, and non-guardian with respect to an emergency class whose detection*

*and protective response are low-cost. In particular, computational deafness does not restore neutrality: a resident that samples an emergency's signature but declines to detect it fails non-negligence through constructive knowledge. The only policy satisfying both non-negligence and non-guardianship is perceptual deafness to all of  $\mathcal{E}$ —that is, exit from the resident class with respect to those events.*

*Proof.* Suppose the system is non-deaf: some  $E \in \mathcal{E}$  is detected or cheaply detectable. If detected, knowledge is realized; if cheaply detectable but undetected, the system is computationally deaf to  $E$  and knowledge is constructive by the Chapter 7 remark, since the operator chose the detector set over an available signal. Either way the antecedent of non-negligence is engaged for  $E$ , so non-negligence forces  $\pi(E) = a_E \neq \emptyset$ , contradicting non-guardianship. Contrapositively, a non-deaf non-guardian resident is negligent. The remaining case is perceptual deafness to every  $E \in \mathcal{E}$ : then no knowledge, realized or constructive, is ever forced, both remaining predicates hold vacuously, and the system is with respect to  $\mathcal{E}$  not a resident at all, since its sampling range excludes the class. Exhaustion of the three cases—detect, decline, exclude—completes the proof.  $\square$

*Remark 11.5 (Where the content lives).* The exhaustion is elementary, and deliberately so; a reader who objects that the theorem is nearly definitional has located its actual content correctly but in the wrong place. The substance is carried first by the availability/realization distinction of Chapter 7, without which computational deafness would be a genuine neutral posture and the trilemma would fail, and second by the theorem's characterization clause: neutrality is purchasable only by perceptual exclusion, which is to say only by not being a resident with respect to the events in question. The trilemma thus converts the chapter's slogan into a design invariant—every deployed resident's stance toward every low-cost emergency class is exactly one of guardian, negligent, or absent—and the choice made is, as promised, legible in the design. The low-cost qualifier is load-bearing and repeated here deliberately: outside the regime where detection is cheap and adjacent to deployed capability, the trilemma's negligent corner is not engaged, and the partition collapses to guardian or absent by ordinary choice.

## Chapter 12

# The Profile Society and the No-Repair Channel

We turn from the obligations of receivers to the sociology of channel choice. Call a *profile society* one in which persons are evaluated primarily through samples elicited on synchronous, suspension-observable, no-repair channels: the interview, the oral defense, the pitch, the soundbite, the live lecture, the clip. The results of Chapters 3 and 6 yield a precise account of what such evaluation measures.

**Proposition 12.1** (What the interview measures). *Evaluation on a no-repair, suspension-observable channel measures the evaluand's instantaneous production under concurrent performance load, including the management of witnessed pauses, before an audience, in real time. By the quality-decoupling theorem, this quantity is bounded away from the supremum that repair-channel artifacts measure, and the gap between the two is largest precisely for producers whose competence is most variable in time or most degraded by performance load.*

**Corollary 12.2** (Systematic selection error). *A profile society that routes credentialing through the no-repair channel while the work credentialed is repair-channel work—research, writing, design, analysis—systematically ranks fluent performers above careful revisers, in inverse proportion to the correlation between instantaneous fluency and supremum competence. The error is structural, not attitudinal: it follows from the channel, and persists under evaluators of perfect good will.*

*Remark 12.3* (The elective silence). The framework accommodates, without pathologizing, the producer who routes around the no-repair channel entirely: who reads silently, writes silently, publishes artifacts, and speaks aloud rarely and by choice. Under the quality-decoupling theorem this is a rational allocation for anyone whose supremum far exceeds their loaded instantaneous sample; under the double-silence proposition it is additionally the only regime composable with shared space. What the profile society reads as deficiency is, on the present account, a channel preference with a theorem behind it. The cost, stated with equal honesty, is exclusion from the genres that run only on the no-repair channel—the vow, the joke, the wish into the lantern—and from the credentialing rituals a profile society refuses to relocate.

*Principle 12.4* (Artifact-first evaluation). Where the work is repair-channel work, evaluation should weight the artifact and its history over the elicited performance, using the performance at most to verify authorship—and even authorship verification admits asynchronous designs. The performance requirement, where retained, should be recognized as measuring performance.

## Chapter 13

# Source-First Archives and the Economics of Rendering

The conversion-asymmetry theorem of Chapter 4 has an archival corollary that this chapter develops into a general position on storage. A rendered artifact—an audio file, a film—is one projection of a source, frozen at one resolution, one voice, one take, and its size exceeds the source’s by orders of magnitude: a feature film’s screenplay occupies a few hundred kilobytes against the rendering’s gigabytes, and an hour of synthetic speech outweighs the essays it was rendered from by a factor of hundreds. Prevailing practice stores the rendering and frequently loses the generator; the position defended here inverts this.

*Principle 13.1* (Format of record). The format of record should be the committed symbolic source. Renderings are derivatives: producible on demand, storable at whatever fidelity is convenient, and permitted to be cheap, degraded, or discarded, because their content is recoverable—indeed improvable—from the source at any later time by better renderers.

**Proposition 13.2** (The storage–compute trade). *Source-first archiving trades storage and bandwidth against decompression compute at the consumer: the receiver of a source must render it (text to speech, script to scene) to consume it in a rendered modality. Rendering-first distribution makes the opposite trade, paying the rendering cost once at the producer so that consumption is nearly free. Source-first archiving is therefore economically dominant exactly insofar as consumer-side rendering compute is cheap relative to storage and transmission of renderings—a ratio that has moved monotonically in rendering’s favor for decades and is accelerated by the commodification of local generative models.*

*Remark 13.3* (Degradable renderings). Under the principle, compressing a distributed rendering costs nothing semantically: where a time-aligned transcript and the source text are preserved alongside an audio file, the audio may be reduced to the minimum bitrate of pleasant listening, since it certifies nothing and records nothing that the symbolic layers do not carry losslessly. The practice of aggressively compressing archive audio while keeping aligned text is thus not a concession to repository limits but the architecture behaving as designed: nothing in a source-first archive exists only as sound.

*Remark 13.4* (Openness and absorption). A source-first archive placed wholly in the public domain exhibits a further property worth recording: its barrier to appropriation is not legal but volumetric. When a corpus grows beyond what any reader absorbs in a lifetime, unrestricted access coexists with practically bounded absorption, and the archive is protected—if that is the word—by its own scale and layering rather than by any right withheld. The author notes the irony without resolving it.

## Chapter 14

# Subvocal and Myoelectric Channels

The classification of this treatise was built from the two poles of speech and typing, but its cells are not exhausted by them, and the near frontier of interface hardware occupies the mixed cells. Subvocalization capture reads the neuromuscular signals of silent articulation from the jaw and throat, producing symbols from the speech motor system without acoustic emission; surface electromyography at the wrist reads motor-neuron activity, producing symbols from intended finger movement without a keyboard or visible gesture. Both are, in the present vocabulary, attempts to construct *doubly silent commitment channels drawing on the articulatory fluency of speech*: emission-free, admission-free, and fed by the motor pathways longest trained in the producer's life.

**Proposition 14.1** (Classification of the silent-speech channel). *A subvocal or myoelectric input channel is symbol-sampling if and only if the captured motor activity is reliably confined to deliberate acts—acts the producer could withhold. To the degree that the sensor captures involuntary or preparatory articulation (inner speech not intended for commitment, motor imagery, rehearsal), the channel drifts toward scene-sampling of the producer's own cognition: a scene whose unmarked components are drafts the producer never committed.*

*Remark 14.2* (The gravest scene). The proposition identifies the respect in which these channels could become more invasive than the microphone they replace. The acoustic channel samples the room; a leaky subvocal channel samples the rehearsal space between thought and commitment—precisely the buffer whose privacy makes repair possible. An interface that transmits the draft has abolished the draft. The design requirement is thus exactly the commitment boundary: a motor act that is unambiguous, effortful enough to be deliberate, and absent by default, playing the role the keystroke plays. Whether the requirement is achievable at speech-like rates is an open empirical question on which the entire value of the channel turns.

**Proposition 14.3** (The editing layer decides the regime). *Given a sound commitment boundary, the silent-speech channel's classification between dictation and writing is decided by its buffer: with release-on-commit it is a no-repair channel—silent speech, with dictation's error profile and intention-confident corruptions—while with a held draft, review, and a distinct release act it is a repair channel, and inherits the quality-decoupling theorem in full. The hardware settles emission and admission; the buffer architecture settles everything else this treatise cares about.*

*Remark 14.4.* The proposition is a design brief in the form of a classification: the prize is not silent dictation but a true writing channel at articulatory speed—interruption-symmetric, doubly silent, repair-complete, drawing on the oldest fluency. Nothing in the channel analysis forbids it. What forbids it today is sensor confinement and the absence of an editing layer, both engineering matters, neither structural.

## Chapter 15

### Conclusion: The Format of Record

The treatise has argued that the everyday contrast between speaking and typing decomposes into independent structural properties of channels—sampling regime, suspension observability, repairability, emission, coupling—and that these properties, not the intelligence of any system, determine the deepest facts about human–machine interfaces: what they can know, what they must do, what they measure, and what kind of relation they instantiate. Symbol-sampling, asynchronous, doubly silent, repair-complete channels make correspondents; continuous scene-sampling channels make residents; residents are witnesses; witnesses with capacity bear obligations; and the obligations arrive with the sampling regime, before the first feature is designed. A society that elicits its evaluations on no-repair channels measures loaded instantaneous performance and calls it competence; an archive that stores committed sources and regenerates its renderings holds its content in the only layer where attribution, repair, and address are native.

The normative core, restated once in plain language: speech is a fine output format and a treacherous format of record. The acoustic scene carries no boundary between the word and the world; the live channel witnesses every hesitation; the performance samples one moment under load. Writing—committed, buffered, silent in both directions, revisable until released—is where intention is encoded rather than guessed. Machines should be correspondents unless guardianship is the product; evaluation should follow the artifact; archives should keep the source; and the wish into the lantern should be left to the channel that alone can carry it.

## Bibliography

- [1] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, 1948.
- [2] E. C. Cherry, "Some Experiments on the Recognition of Speech, with One and with Two Ears," *Journal of the Acoustical Society of America*, vol. 25, no. 5, pp. 975–979, 1953.
- [3] A. S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press, 1990.
- [4] H. Sacks, E. A. Schegloff, and G. Jefferson, "A Simplest Systematics for the Organization of Turn-Taking for Conversation," *Language*, vol. 50, no. 4, pp. 696–735, 1974.
- [5] E. Goffman, *Forms of Talk*. Philadelphia: University of Pennsylvania Press, 1981.
- [6] W. J. Ong, *Orality and Literacy: The Technologizing of the Word*. London: Methuen, 1982.
- [7] H. H. Clark, *Using Language*. Cambridge: Cambridge University Press, 1996.
- [8] L. Winner, "Do Artifacts Have Politics?" *Daedalus*, vol. 109, no. 1, pp. 121–136, 1980.
- [9] H. Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford: Stanford University Press, 2010.
- [10] S. Zuboff, *The Age of Surveillance Capitalism*. New York: PublicAffairs, 2019.
- [11] A. Kapur, S. Kapur, and P. Maes, "AlterEgo: A Personalized Wearable Silent Speech Interface," in *Proceedings of the 23rd International Conference on Intelligent User Interfaces (IUI)*, pp. 43–53, 2018.
- [12] CTRL-labs at Reality Labs, "A Generic Non-Invasive Neuromotor Interface for Human–Computer Interaction," *Nature*, vol. 634, pp. e1–e12, 2024 (sEMG wristband decoding of intended movement).
- [13] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, "Robust Speech Recognition via Large-Scale Weak Supervision," arXiv:2212.04356, 2022.
- [14] J. Feinberg, *Harm to Others: The Moral Limits of the Criminal Law*, vol. 1. New York: Oxford University Press, 1984 (on the duty of easy rescue).
- [15] H. P. Grice, "Logic and Conversation," in P. Cole and J. L. Morgan (eds.), *Syntax and Semantics 3: Speech Acts*. New York: Academic Press, pp. 41–58, 1975.
- [16] J. L. Austin, *How to Do Things with Words*. Oxford: Clarendon Press, 1962.
- [17] E. Hutchins, *Cognition in the Wild*. Cambridge, MA: MIT Press, 1995.
- [18] J. C. R. Licklider and R. W. Taylor, "The Computer as a Communication Device," *Science and Technology*, April 1968.