

The Listener's Digital Shadow

Dongchun Chen

School of Foreign Studies, Guangdong University of Finance and Economics

Author Note

The author declares no conflicts of interest to disclose. Correspondence concerning this article should be addressed to Dongchun Chen, School of Foreign Studies, Guangdong University of Finance and Economics, Guangzhou, Guangdong 510320, China. Email: chendongchun@gdufe.edu.cn

Abstract

Digital platforms now routinely construct annual retrospectives—Spotify Wrapped, Audible Stats, and their equivalents on Chinese platforms like Ximalaya—that claim to reveal “who you truly are” through data. Existing scholarship has largely theorized this algorithmic self-narration through the lens of semantic traces: highlights, annotations, and content preferences that express intellectual judgments. This paper argues that such frameworks overlook a distinct and increasingly prevalent mode of self-formation grounded in auditory behavior. We introduce the concept of “prosodic engagement traces” (hereinafter “traces”)—digital footprints capturing how users pace, replay, skip, and modulate their listening tempo—as a category fundamentally different from semantic traces. Unlike semantic traces, which carry propositional content about what a user thinks, prosodic engagement traces register the bodily, rhythmic, and affective texture of how a user listens. Drawing on Hans-Georg Gadamer’s hermeneutics of play, Paul Ricœur’s theory of narrative identity, Couper-Kuhlen and Wennerstrom’s conversation-analytic work on prosody, we propose a theoretical framework in which these traces facilitate a distinctive process of self-formation: “resonant recognition”. This is a non-narrative, aesthetic mode of self-encounter wherein the listener recognizes herself not through coherent life stories but through the felt familiarity of her own listening rhythms. Crucially, this process unfolds as “reflection without an other”: the algorithmic narration bypasses the dialogical encounter central to Gadamerian hermeneutics while simultaneously compressing the detour through textual distancing that Ricœur deemed necessary for self-understanding. The empirical portion of the study operationalizes this framework through a hermeneutic analysis of platform-generated listening narratives, triangulated with a three-stage coding of 327 user-generated social media responses (from Weibo and Xiaohongshu) to their own listening retrospectives. Of these, six posts were selected for intensive hermeneutic analysis based on their theoretical richness; the remainder supported the broader coding

Received 12 April 2026; revised 18 May 2026; accepted 5 June 2026; available online 25 June 2026; version of record 30 June 2026.

Citation: Chen, D. C. (2026). The Listener's Digital Shadow. *Journal of Language*, 2(2), 152–189. <https://doi.org/10.64699/26KYWY1142>

framework. Findings reveal a distinct discursive pattern through which users articulate an “aesthetic self”—a self recognized in the grain of the voice, the rhythm of pausing, and the texture of tempo, rather than in the content of words. The paper concludes by theorizing how prosodic engagement traces reconfigure the hermeneutic circle under algorithmic conditions and proposes the cultivation of a “critical listening literacy” as a response to this emergent mode of digital subjectivation.

Keywords: prosodic engagement traces, resonant recognition, aesthetic self, “reflection without an other”, digital shadow, algorithmic hermeneutics

1.The Listener’s Digital Shadow: A Hermeneutic Inquiry into the Formation of the Self Through Prosodic Engagement Traces

In December of each year, hundreds of millions of users encounter algorithmically generated retrospectives that purport to summarize their year in listening. Spotify Wrapped tells you your top artists and total minutes streamed; Audible presents your listening habits in terms of completion rates and genre preferences; Chinese platforms like Ximalaya offer “年度收听报告” (Annual Listening Report) that visualize your behavioral patterns. These platform outputs are not neutral summaries: they are acts of narrative construction, algorithms telling you who you are through the data you have left behind.

Scholarship on the “digital shadow”—the datafied double that algorithms construct from user behavior—has rapidly expanded. Cheney-Lippold (2017) theorized how algorithmic categories constitute a “new algorithmic identity”; Bucher (2018) examined how users experience and respond to these constructed selves; Hepp (2017) situated these developments within a broader theory of the “datafied self” as an emergent social form. Scholars of auditory culture have further enriched this picture by tracing how sound media reshape the listening subject: Sterne (2003) documented the historical emergence of modern listening techniques, Rubery (2011) excavated the cultural history of the talking book as a contested technology of self, and Ihde (2007) developed a phenomenology of auditory experience that foregrounds the

embodied relationship between listener and sound. Yet across these diverse literatures, a common assumption persists: the primary raw material for algorithmic identity construction is what we call “semantic traces”—behavioral data that carries or can be interpreted as carrying propositional content. When Netflix recommends content “because you watched X,” or when a news platform personalizes your feed based on articles you have read, the underlying logic is that your past choices express your preferences, beliefs, or interests. The trace is read as a rudimentary statement: “I like this kind of thing.”

This semantic bias in the theorization of digital traces is understandable. Most platform interfaces present themselves as content delivery systems, and content is understood primarily through its informational or narrative dimensions. But it leaves a significant blind spot. Consider the case of audio platforms. When a user listens to an audiobook, she leaves a wealth of behavioral traces: where she speeds up, where she slows down, where she rewinds, where she pauses, where she stops entirely. Some of these behaviors may be motivated by semantic factors—rewinding because she missed a plot point, skipping because the content is boring. But many, we argue, are driven by prosodic factors: the tempo of the narrator’s delivery, the rhythm of the prose as spoken, the emotional texture carried in intonation and pause. These behaviors do not express “I am interested in this topic” or “I agree with this argument.” They register something more fundamental: the listener’s bodily and affective attunement to the sonic texture of language. Rubery’s (2016) cultural history of the talking book documents how early listeners to recorded literature described their experience in distinctly bodily terms—the voice “washing over” them, the rhythm “carrying” them forward—suggesting that the pre-reflective dimension of auditory engagement has a long cultural lineage that algorithmic

platforms now capture in datafied form. Sterne's (2003) archaeology of modern listening techniques further demonstrates that what we experience as natural auditory response is historically constructed—a point that should caution us against naturalizing the prosodic traces that platforms now harvest.

We call the digital imprints of these behaviors “prosodic engagement traces”. The central argument of this paper is that when platforms aggregate these traces into a narrative portrait of the listener, they do not create a self-image grounded in what the listener “thinks”. They create one grounded in how the listener “feels rhythmically”. This, we propose, is the formation of an “aesthetic self”—a mode of self-recognition that operates not through narrative coherence but through the felt familiarity of one's own embodied rhythms. And this formation unfolds as “reflection without an other”: an encounter with the self that bypasses the dialogical confrontation with a living interlocutor (Gadamer, 1960/2004) while compressing the hermeneutic detour through distanced texts (Ricoeur, 1992). The algorithm, in effect, holds up a mirror made of tempo, pause, and intonation contour—and the listener recognizes herself in rhythms she did not know were hers.

The paper proceeds in five sections. Section 2 develops the theoretical framework, synthesizing Gadamer's hermeneutics of play, Ricoeur's narrative identity, and conversation-analytic prosody research to establish the concept of prosodic engagement traces and the process of resonant recognition. Section 3 presents the empirical study: a hermeneutic analysis of platform output narratives and a three-stage coding of 327 user-generated social media posts responding to their listening retrospectives. Section 4 deepens the theoretical argument by theorizing “reflection without an other” and the hermeneutic implications of algorithmic self-

narration. Section 5 concludes by theorizing the aesthetic self and proposing critical listening literacy as a response.

2. Theoretical Framework

2.1 The Digital Shadow and Its Limits

The concept of the “digital shadow” or “data double” has become a central metaphor in critical data studies. Cheney-Lippold (2017) influentially argued that algorithmic systems construct a “new algorithmic identity” by sorting users into measurable types derived from behavioral data. This identity, he emphasizes, is not an authentic expression of a pre-existing self but a construction that feeds back into and shapes lived experience. Bucher (2018) extended this analysis by examining the phenomenological dimension: how users “experience” the algorithmic self that confronts them in platform interfaces, often with a mix of recognition, alienation, and negotiation. Couldry and Hepp (2017) further situated algorithmic identity within a broader historical transformation—the emergence of the “datafied self” as a social form that operates alongside and sometimes displaces earlier modes of identity construction rooted in face-to-face interaction and institutional affiliation. Gillespie (2014, 2018) contributed an equally vital dimension by analyzing the rhetorical and political work that platforms perform when they present their algorithmic outputs as neutral reflections of user identity.

These frameworks have been enormously generative. However, they share an implicit assumption that the primary raw material for algorithmic identity construction is what we call here “semantic traces”—behavioral data that carries or can be interpreted as carrying propositional content. Page’s (2018) work on narratives in social media environments,

Kennedy’s (2016) analysis of data practices in creative industries, and Lupton’s (2016) theorization of the “quantified self” all contribute valuable perspectives on digital identity, but each operates within a framework where user behavior is read as an expression of conscious or semi-conscious preference. The trace, across these otherwise divergent traditions, remains a sign of what the user “thinks”—even if that thinking is distributed, habitual, or shaped by platform architectures (Seaver, 2019).

This leaves a significant blind spot. Consider the case of audio platforms. When a user listens to an audiobook, she leaves a wealth of behavioral traces: where she speeds up, where she slows down, where she rewinds, where she pauses, where she stops entirely. Some of these behaviors may be motivated by semantic factors—rewinding because she missed a plot point, skipping because the content is boring. But many, we argue, are driven by prosodic factors: the tempo of the narrator’s delivery, the rhythm of the prose as spoken, the emotional texture carried in intonation and pause. These behaviors do not express “I am interested in this topic” or “I agree with this argument”. They register something more fundamental: the listener’s bodily and affective attunement to the sonic texture of language.

This is the gap the present paper seeks to fill. We need a theoretical framework capable of theorizing how rhythmic, bodily traces become the basis for algorithmic self-narration, and what kind of self emerges from this process.

Table 1 Semantic Traces Versus Prosodic Engagement Traces

Dimension	Semantic Traces	Prosodic Engagement Traces
Primary modality	Content-based (what is consumed)	Rhythm-based (how it is consumed)
Type of information	Propositional (preferences, beliefs)	Bodily-affective (tempo, pause, texture)
Mode of expression	“I like this”	Attunement without judgment

Temporal structure	Discrete (individual selections)	Continuous (patterns of modulation)
Self disclosed	Intellectual self (what I think)	Aesthetic self (how I resonate)

Note. “Continuous” in the temporal structure row refers to the patterned distribution of modulations across listening sessions, not to the moment-by-moment recording of playback speed. Individual speed adjustments are discrete events, but their aggregation across time reveals continuous rhythmic profiles.

2.2 Prosody as Interactional and Self-Disclosing

To build this framework, we draw on the rich tradition of prosody research in conversation analysis and interactional linguistics. Prosody—encompassing intonation, tempo, rhythm, loudness, and voice quality—has long been recognized as an independent channel of meaning-making in spoken interaction, irreducible to the semantic content of words. Couper-Kuhlen and Selting (1996) demonstrated that prosody is not merely a paralinguistic ornament but a systematic resource for structuring conversation: turn-taking, repair, and the negotiation of alignment are all managed through prosodic cues. This foundational insight was extended by Couper-Kuhlen (2015), who synthesized decades of interactional research to show that prosodic design is central to how speakers display and negotiate affective stance in real time. Wennerstrom (2001) showed how intonation functions as a discourse-organizing system, marking topic transitions, signaling stance, and conveying affective orientation. Gumperz (1982) had earlier established the role of prosodic “contextualization cues” in how speakers signal and interpret social meaning—a line of work that remains foundational to understanding the communicative work that prosody performs beneath explicit semantic content. Crystal (2008) provided the linguistic reference point for defining prosody’s constituent features in a way that is both technically precise and accessible to non-specialist scholars.

Crucially for our purposes, prosody carries what we might call “self-disclosing”

information. The way a speaker modulates tempo, pitch, and volume reveals dimensions of their embodied and affective state that may not be consciously controlled or even accessible to semantic self-report. A speaker's voice can tremble with anxiety, rush with excitement, flatten with exhaustion—and these qualities are perceived by listeners with remarkable acuity, often operating beneath the threshold of explicit attention. Rubery (2016) has shown that listeners are continuously responsive to the prosodic design of utterances, adjusting their own conduct in ways that demonstrate their attunement to the speaker's displayed affective stance. This insight resonates with Chion's (1994/2019) concept of “audio-vision”—the idea that sound perception is never purely auditory but always involves an embodied, multisensory mode of attention.

What happens when we transpose this insight from the domain of live interaction to the domain of mediated listening? When a listener adjusts playback speed, rewinds to rehear a particularly affecting intonation contour, or pauses in synchrony with a narrator's dramatic silence, she is engaging in what we call “prosodic engagement”. These behaviors are the trace of an embodied responsiveness to the sonic texture of speech. And when platforms aggregate these traces across hundreds of hours of listening, they begin to form a portrait not of what the listener knows or believes, but of how she “resonates”.

2.3 Defining Prosodic Engagement Traces

We define “prosodic engagement traces” as digital footprints that register a user's temporal and rhythmic modulation of audio playback. This includes, centrally:

- Playback speed adjustments (e.g., accelerating to 1.5x, slowing to 0.75x): These modulate the tempo of prosodic delivery, compressing or expanding the temporal texture of speech.

- Rewind and repeat behaviors: These register moments where the listener returns to a sonic passage, often driven by the felt quality of the delivery rather than semantic comprehension failure.

- Pause and resume patterns: The timing and frequency of pauses during playback, which may synchronize with emotional peaks or prosodic boundaries in the narration.

- Session duration and temporal patterns: When and for how long a listener engages with audio content, which carries prosodic implications through diurnal rhythms of attention and mood.

We acknowledge a boundary issue that demands theoretical honesty. Not all behaviors that leave traces during audio consumption are purely prosodic in motivation. A user who skips a chapter may be acting on semantic judgment (“I already know this content” or “this is boring”). A user who rewinds may be trying to catch a missed plot detail. We do not claim that prosodic engagement traces are “always and exclusively” prosodic. Rather, we propose a “prosody-first interpretive strategy”: in behaviors that modulate the temporal delivery of speech—speed, repetition, pausing—the prosodic dimension offers the most parsimonious phenomenological entry point. This is an analytical heuristic, not an ontological claim: we do not assert that semantic motives are absent, but rather that the bodily adjustment to sonic surface temporally precedes—and often outlasts—semantic processing (cf. Couper-Kuhlen, 2015, on the online nature of prosodic processing). The listener encounters the sonic surface of language before she processes its propositional content, and her bodily adjustments to that surface constitute the ground on which semantic interpretation later unfolds. This position aligns with Couper-Kuhlen’s (2015) insistence that prosody is not an accessory to semantics

but an independent channel of interactional meaning. Where the motivation is clearly hybrid or indeterminate—as in chapter-skipping—we treat these as boundary cases that enrich rather than undermine the theory’s dialectical texture.

Crucially, prosodic engagement traces differ from semantic traces in their relationship to meaning. A semantic trace—a highlighted passage, a bookmarked article—carries an implicit proposition: “This matters.” A prosodic engagement trace does not point to a proposition. It registers a resonance. When a listener unconsciously slows down during a particularly melodic passage of narration, she is not saying “I like this.” She is, in her body, attuning to a rhythm. The trace is an index of attunement, not a sign of judgment.

2.4 Gadamer, Play, and the Aesthetic Dimension of Understanding

To theorize what kind of self-understanding emerges from these traces, we turn to Gadamer’s hermeneutics, particularly his concept of “play” (*Spiel*) and his rehabilitation of the aesthetic as a mode of truth.

In “Truth and Method”, Gadamer (1960/2004) develops the concept of play as a model for the hermeneutic event. Play, for Gadamer, is not a subjective activity undertaken by a player; rather, play has its own ontological primacy. The player is “played by” the game, drawn into a to-and-fro movement that transcends individual subjectivity. This structural feature of play—the primacy of the movement over the subject, the absorption of the player into a rhythmic pattern—makes it an apt model for understanding what happens in prosodic engagement. When a listener is drawn into the rhythm of a narrated passage, modulating her playback in response to its tempo and intonation, she is not a sovereign subject manipulating an object. She is caught up in a play of rhythms, a to-and-fro between her bodily tempo and the sonic texture of the

audio. This Gadamerian reading of auditory engagement finds support in the phenomenological tradition: Heidegger's (1927/1962) analysis of "Befindlichkeit" (attunement) as a fundamental existential structure—the way we always already find ourselves disposed toward the world through mood—provides the ontological ground for understanding why rhythmic attunement is not a secondary response to content but a primary mode of being-in-the-world. Ihde's (2007) phenomenology of listening extends this tradition specifically into the auditory domain, arguing that sound perception is never a passive reception but an active, embodied mode of engaging with the world.

Gadamer further argues that the aesthetic is not a diminished or merely subjective mode of experience but a mode of truth in its own right. Aesthetic experience involves a distinctive kind of recognition: not the recognition of a concept in a particular instance, but the recognition of something "as" something, a "this is it!" that carries its own evidentiary force. This concept of "aesthetic recognition" is central to our argument. We propose that when a listener encounters her prosodic engagement traces aggregated into a platform narrative—"You listened at 1.5x during thrillers but slowed down for literary fiction"—she experiences something analogous to aesthetic recognition. She does not deduce that she is an anxious person; she "recognizes" herself in the rhythm of acceleration. This is what we call "resonant recognition": the felt identification of self in the temporal contours of one's own listening body.

2.5 Ricœur, Narrative Identity, and the Limits of Emplotment

Gadamer gives us the aesthetic-epistemological mechanism; Ricœur gives us the framework for theorizing selfhood. Ricœur's (1992) theory of narrative identity holds that the self is not immediately transparent to itself but must be constituted through the mediation of

narrative. Self-understanding, for Ricœur, is always a detour: we come to know ourselves by interpreting the texts, actions, and narratives that we produce and that are produced about us. This hermeneutic detour is essential because the self is not a fixed substance but a project of configuration and refiguration through time. Ricœur's earlier work on interpretation theory (1976) and *Time and Narrative* (1984–1988) established the philosophical architecture: narrative is the privileged medium through which the aporias of temporal experience are configured into intelligible form.

Ricœur's framework has been productively applied to digital self-narration. Scholars have argued that platform-generated retrospectives function as a kind of algorithmic emplotment: they take the disparate, non-narrative data points of user behavior and synthesize them into a coherent story with a protagonist, a temporal arc, and a theme. The user, encountering this narrative, engages in a process of appropriation—she recognizes (or refuses) herself in the story the algorithm tells. Turkle's (2011) work on how digital technologies reshape the conditions of self-reflection provides a complementary perspective: she argues that the “always-on” digital environment erodes the capacity for solitary self-reflection, replacing it with a self constituted through continuous external feedback. This diagnosis makes the algorithmic prosodic portrait particularly significant—it is a form of external feedback that arrives not through social interaction but through data aggregation, a new kind of “mirror” for the digital age.

Our intervention is to argue that prosodic engagement traces enable a mode of self-encounter that “compresses the narrative detour” that Ricœur posits as the long route to self-understanding. When a user reads “You replayed this passage 17 times” and feels the shock of

self-recognition, she is not being told a story about herself. No emplotment has occurred beyond the simple aggregation of a rhythmic behavior. The recognition is immediate, bodily, and aesthetic—it operates at the level of felt rhythm rather than narrative coherence. This is what we mean by “reflection without an other”: the listener encounters herself not through the mediating work of narrative configuration, nor through the dialogical challenge of a living interlocutor, but through the direct, quasi-sensory return of her own embodied rhythms.

This claim does not refute Ricœur but extends him. A critical reader might ask: if both the Gadamerian dialogical encounter and the Ricœurian narrative detour are bypassed, does “resonant recognition” remain a hermeneutic phenomenon at all? Our answer is affirmative, but in a minimal sense: it remains hermeneutic insofar as it involves a circular movement between the listener’s pre-reflective bodily dispositions and their objectification in algorithmic traces. However, this is a “short-circuited” hermeneutics—one that delivers recognition without interpretation, and thus without the critical distance that dialogue and distancing would afford. We return to this tension in Section 4. Ricœur himself, particularly in his later work “The Course of Recognition” (2004/2005), explored non-narrative modes of mutual recognition—the recognition of the other’s face, the reciprocity of gift-exchange—that border on the immediate and do not require full narrative mediation. Our concept of resonant recognition stands in productive tension with this later work: it shares Ricœur’s intuition that recognition can operate outside narrative emplotment, but it relocates this possibility from the intersubjective domain (face-to-face encounter) to the algorithmic domain (encounter with one’s own datafied rhythms). This relocation is what makes prosodic engagement traces theoretically distinctive: they offer a shortcut to self-recognition that risks bypassing the critical

safeguards Ricœur built into the narrative detour, even as they open access to dimensions of selfhood—the rhythmic, the bodily, the aesthetic—that narrative tends to subordinate or elide.

2.6 Synthesis: Resonant Recognition and the Aesthetic Self

We can now synthesize our theoretical framework. The core conceptual chain is as follows:

1. “Prosodic engagement traces” are the digital imprints of bodily, rhythmic responsiveness to the sonic texture of spoken language in audio media.

2. When these traces are algorithmically aggregated and presented back to the listener—in a platform retrospective or a listening summary—they constitute a “prosodic portrait”: an image of the self composed not of preferences and beliefs but of rhythms, tempos, and patterns of sonic attunement.

3. The listener’s encounter with this portrait can trigger “resonant recognition”: a felt, bodily “aha” in which she recognizes herself in the temporal contours of her listening body. This recognition is aesthetic in Gadamer’s sense—it involves the “this is it!” of recognizing something as something—and operates prior to or outside of narrative emplotment.

4. The self that is recognized in this encounter is the “aesthetic self”: a mode of selfhood constituted not through the coherence of a life story (narrative identity in Ricœur’s sense) but through the felt familiarity of one’s own embodied rhythms. The aesthetic self is “mine” but also strange—it returns to me from the outside, through the algorithmic mirror, as something I did not know I was producing.

Table 2 Intellectual Self Versus Aesthetic Self

Dimension	Intellectual Self (Narrative)	Aesthetic Self (Prosodic)
-----------	----------------------------------	---------------------------

Mode of constitution	Narrative emplotment	Rhythmic resonance
Primary material	Content preferences, beliefs	Tempo, pause, sonic texture
Mode of recognition	Appropriation of story	Felt “this is it!”
Temporal structure	Diachronic (life story)	Synchronic (pattern recognition)
Relationship to self	Mediated by narrative	Immediate, bodily

This framework generates the empirical question that guides the remainder of the paper: How do listeners actually respond when confronted with the prosodic portrait? What discursive strategies do they use to articulate—or resist—resonant recognition?

3. Empirical Study

3.1 Platform Output Narrative Analysis

The empirical portion of this study proceeds in two phases. The first phase is a hermeneutic analysis of “platform output narratives”—the annual retrospectives produced by major audio platforms.

We collected and analyzed annual listening reports from three platforms: Ximalaya (喜马拉雅, 2023 Annual Report), Lizhi (荔枝, 2023 Listening Summary), and Spotify Wrapped (2023, specifically its audiobook and podcast components). These were treated not as transparent data summaries but as “narrative texts” that construct a particular kind of listener-subject through their rhetorical, visual, and categorical choices. Our analysis focused on identifying the implicit theory of selfhood encoded in these platform outputs—what kind of listener they address and construct.

The analysis revealed a consistent pattern across platforms. Platform narratives deploy what we call the “revelation trope”: they position themselves as revealing a truth about the listener that the listener did not know. Ximalaya’s 2023 report, for example, prominently featured the phrase “原来你是这样的听众” (So this is the kind of listener you are), with the

discourse particle “yuánlái” (原来) indexing the speaker’s newly acquired awareness of a previously unrecognized state of affairs. This linguistic structure has been extensively analyzed in Chinese pragmatics: Tsai and Yang (2022) demonstrated that “yuánlái” functions as a mirativity marker—a grammatical device that encodes the speaker’s surprise at newly discovered information—while Li and Thompson (1981) situated such particles within the broader architecture of Mandarin discourse structure.

More significantly, the categories through which platforms construct this revelation are predominantly “prosodic and temporal”, not semantic. Ximalaya’s report highlighted “你最爱在深夜听” (You love listening late at night) and “你反复回听的片段” (The passages you returned to repeatedly). Spotify Wrapped’s audiobook component foregrounded listening duration, completion patterns, and genre affiliations that cut across content categories—thriller listeners who also consume true crime podcasts, literary fiction listeners who favor slow-build narrative podcasts. Notably, Spotify’s cross-modal aggregation—juxtaposing music listening with audiobook and podcast consumption—creates what we call an “affective rhythm profile”: a portrait of the user’s preferred emotional tempos that operates beneath and across content types. When a user sees that her hours of suspense-novel listening sit alongside her predilection for dark ambient music, what is being profiled is not a set of content preferences but a rhythmic-affective disposition—a preference for tension, release, and the temporal texture of dread.

These platform outputs, we argue, are not merely reporting listening behavior. They are constructing an aesthetic self and inviting the listener to recognize herself in it. The theoretical question is whether and how listeners take up this invitation.

3.2 Methodology: Social Media Response Analysis

The second empirical phase investigates listener responses to these platform narratives. This phase employs a hermeneutic analysis of user-generated social media posts, operationalized through a three-stage coding process.

3.2.1 Data Collection

We collected publicly available posts from Weibo and Xiaohongshu that responded to platform-generated listening retrospectives. Search terms included platform names plus “年度报告” (annual report), “听单” (listening list), and related hashtags. Posts were collected during the period of December 2023 through February 2024, covering the peak period of retrospective release and discussion. After removing duplicates, commercial content, and posts with insufficient textual content (fewer than 30 Chinese characters) for analysis, the final dataset comprised 327 posts, distributed across Weibo ($n = 189$) and Xiaohongshu ($n = 138$). The corpus included original user reflections (86%), reposts with substantial commentary (12%), and a small number of image-heavy posts with minimal text (2%). Post lengths ranged from 12 to over 400 Chinese characters, with a median of 78 characters, ensuring discursive diversity across brief exclamations and extended personal narratives. Given the study’s hermeneutic focus on discursive patterns rather than demographic correlates, we did not systematically collect user demographic information; our unit of analysis is the post as a textual event, not the individual user as a demographic subject.

3.2.2 Methodological Positioning

This study follows the logic of “exemplary understanding” in the hermeneutic tradition, not the logic of statistical generalization. The 327 posts serve as an experiential anchor for theoretical concept development, enabling the identification of typical discursive patterns

through which listeners articulate their encounters with prosodic portraits. The validity criteria are interpretive depth and conceptual discrimination, not sample representativeness. This aligns with the hermeneutic principle that understanding proceeds through the intensive analysis of particular instances that illuminate a phenomenon, rather than through the extensive aggregation of cases. From these, six posts were chosen for intensive hermeneutic exposition. The selection was guided by the principle of theoretical saturation within the hermeneutic tradition: we prioritized posts that explicitly engaged with the prosodic-semantic distinction, as these provided the richest articulation of the theoretical categories developed in Section 2. The remaining 321 posts served as the comparative ground against which the typicality of these six exemplary cases was assessed. This selection strategy follows the logic of hermeneutic exemplarity (Gadamer, 1960/2004, pp294-297), not statistical representativeness; the goal is conceptual discrimination, not frequency estimation.

3.2.3 Coding Procedure

The analysis followed a three-stage coding process modeled on the hermeneutic circle: initial openness to the phenomenon, dialogical engagement between data and theory, and integrative theoretical abstraction.

Stage 1: Initial Coding. All 327 posts were coded with descriptive labels that stayed close to the language used by participants. This stage maintained what Gadamer would call phenomenological openness—an effort to let the data speak without premature theoretical imposition. Labels generated at this stage included “surprise at algorithmic accuracy,” “discomfort with being known,” “reinterpretation of past behavior,” and “dismissal of report as inaccurate.”

Stage 2: Focused Coding. Initial codes were grouped into higher-order categories through an iterative process of comparison and theoretical dialogue. This stage enacted the hermeneutic movement between part and whole: individual posts were re-read in light of emerging categories, and categories were refined in light of disconfirming cases. Core categories that emerged include: “resonant recognition” (posts expressing felt identification with the prosodic portrait), “semantic resistance” (posts rejecting the report because it “doesn’t capture what I actually think”), “algorithmic uncanny” (posts expressing unease at the intimacy of the algorithmic knowledge), and “aesthetic self-articulation” (posts using the report as an occasion to articulate an identity grounded in sonic-aesthetic rather than content-based terms).

Stage 3: Theoretical Coding. Focused codes were integrated into the theoretical framework developed in Section 2. The goal was not to “verify” the theory against the data, but to enrich and complicate the theoretical concepts through sustained engagement with lived experience. This stage yielded the key theoretical finding: the distinction between users who experience “resonant recognition” (recognizing self in rhythm) and those who engage in “narrative appropriation” (recognizing self in story), and the discursive strategies that mark this distinction.

3.2.4 Analytical Rigor

To ensure interpretive traceability, the following measures were adopted. First, a detailed codebook was maintained with conceptual definitions, inclusion/exclusion criteria, and typical examples for each node (see Appendix A for selected excerpts). Second, temporal recoding was conducted: 30% of the sample (98 posts) was recoded after a two-week interval, yielding a consistency rate of 89%. Third, negative cases and boundary examples are discussed

in the findings to demonstrate conceptual discrimination rather than selective illustration.

3.3 Findings

The analysis revealed three primary findings, which we present below with illustrative data and interpretive commentary.

3.3.1 Resonant Recognition: The “原来” (Yuánlái) Moment

The most theoretically significant pattern across the dataset was a distinctive discursive structure we call the “yuánlái moment”. In these posts, users reported an experience of self-discovery triggered by the platform’s revelation of their prosodic patterns, characteristically marked by the discourse particle “yuánlái” (原来, “so it turns out that...”) or equivalent expressions of newly acquired self-awareness.

> P-42 (Xiaohongshu, female user, 20s): “原来我每天晚上听的悬疑小说都是 1.5 倍速，白天听的文学就正常速度。看了报告才发现，我是真的用身体在听书。” (So it turns out I listen to thrillers at 1.5x at night but literary fiction at normal speed during the day. Only after seeing the report did I realize—I genuinely listen to books with my body.)

> P-87 (Weibo, female user, 30s): “看年度报告的时候，发现我反复听的那几段，全是有雨声背景的独白。我甚至不记得内容是什么，但那个声音的质地就是让我一遍遍回去。” (When I looked at the annual report, I found that the passages I replayed were all monologues with rain sound in the background. I don’t even remember what the content was, but the texture of that voice just made me go back again and again.)

> P-156 (Weibo, male user, 20s): “喜马拉雅说我一年回听了某段念白 23 次。原来不是我忘了内容，是那个声音本身有东西在抓我。” (Ximalaya told me I replayed one narration passage 23 times this year. So it turns out it wasn’t that I’d forgotten the content—

there was something in the voice itself that grabbed me.)

Several features of these posts are notable. First, the “yuánlái” structure marks a temporal unfolding of self-knowledge: “I did not know this, but now I do.” This is precisely the structure of Gadamerian recognition—the “this is it!” that constitutes an event of truth. Second, users actively distinguish their prosodic engagement from semantic engagement: “I don’t even remember what the content was,” “it wasn’t that I’d forgotten the content.” This discursive work of “distinguishing” the prosodic from the semantic is itself evidence that users intuitively grasp the difference our theory posits. Third, users describe their listening body as agentive in ways that exceed conscious intention: “that voice just ‘made’ me go back,” “there was something in the voice itself that ‘grabbed’ me.” This is the phenomenology of play that Gadamer theorized: the listener is not manipulating the audio; she is drawn into a rhythmic movement that has its own momentum.

Further discourse analysis reveals the linguistic strategies through which users construct the aesthetic self. In P-87 above, the user’s explicit disavowal of content memory (“我甚至不记得内容是什么”) serves as a boundary-marking device: it severs the connection between self and semantic content, clearing space for the prosodic self to emerge. The phrase “声音的质地” (the texture of the voice) deploys a tactile metaphor—“zhìdì” (质地) typically describes fabric or material surfaces—that relocates the self in the material, sensory dimension of sound rather than its propositional dimension. And the construction “让我一遍遍回去” (made me go back again and again) uses the causative “rang” (让) to position the speaker as the object of an imperative issued by sound itself. The grammar here enacts the theoretical claim: the aesthetic self is not a self that chooses but a self that “is chosen” by rhythms it did

not author.

By contrast, posts focused on content-based recognition employ a different discursive register:

> P-203 (Xiaohongshu, female user, 30s): “年度总结说我听得最多的类型是个人成长。确实，今年一直在刻意提升自己。” (The annual summary said my most-listened genre is personal development. Indeed, I’ve been deliberately working on self-improvement this year.)

Here, the user confirms the algorithmic portrait with “确实” (indeed), positioning it as a reflection of conscious projects and deliberate choices. There is no “yuánlái”—no revelation—because the knowledge simply confirms what the user already narratively knows about herself. This is narrative identity at work, not resonant recognition. The contrast between P-203 and the prosodic recognition posts demonstrates, from within the data itself, the conceptual distinction between narrative and aesthetic self-formation.

3.3.2 The Algorithmic Uncanny: Intimacy Without Dialogue

A second pattern captured users’ experience of discomfort when confronted with the intimacy of algorithmic knowledge derived from prosodic traces. This pattern, which we term the “algorithmic uncanny”, extends Bucher’s (2018) analysis of algorithmic phenomenology by specifying a distinct affective register associated with prosodic—as opposed to semantic—knowledge.

> P-68 (Weibo, female user, 20s): “最吓人的是它告诉我我听书最爱暂停的点。我自己都没意识到，但它知道。这种感觉太奇怪了，好像有人一直在观察我的呼吸。” (The creepiest part is that it told me the exact points where I like to pause during audiobooks. I didn’t even notice this about myself, but it knows. The feeling is too strange—as if someone has been

watching me breathe.)

> P-121 (Xiaohongshu, female user, 30s): “它比我更懂我的情绪节奏。我什么时候焦虑什么时候放松，全在这些数据里。不知道该感动还是害怕。” (It understands my emotional rhythms better than I do. When I’m anxious, when I’m relaxed—it’s all in this data. I don’t know whether to be moved or terrified.)

The uncanny, in Freud’s (1919/2003) classic formulation, is the return of the familiar made strange. Prosodic engagement traces produce a distinctive version of this effect: they return to the user her own bodily rhythms—the most intimate and taken-for-granted dimension of experience—in the alienated form of data. What disturbs users is not that the algorithm knows something “private” (it would be equally disturbing if it revealed private semantic content) but that it knows something “pre-reflective”—something the user has lived through her body without bringing to conscious awareness. The metaphor of breathing in P-68 is precise: prosodic engagement, like breathing, operates at the threshold of consciousness, sustaining our activity without becoming an object of attention. When the algorithm holds up this breathing self for inspection, the user experiences an uncanny intimacy without dialogue—an other that knows her in ways she does not know herself, but which cannot be engaged as a genuine dialogical partner.

This finding reinforces the theoretical claim about “reflection without an other”. What users register as “creepy” is precisely the absence of the dialogical conditions that Gadamer posited as essential to genuine understanding. The algorithm knows without encountering; it reflects without responding. The listener is offered a mirror with no one behind it.

3.3.3 Semantic Resistance and Boundary Work

A third pattern captured users who actively resisted the platform's prosodic portrait. This resistance typically took a specific form: users rejected the algorithmic narrative on the grounds that it mistook "prosodic behavior" for "semantic preference".

> P-274 (Weibo, male user, 30s): “年度报告说我爱听言情小说，因为我晚上总在听。但那是因为我失眠的时候需要声音陪伴，跟内容没有任何关系。它根本不懂我。” (The annual report said I love romance novels because I listen to them at night. But that's because I need sound companionship when I have insomnia—it has nothing to do with the content. It doesn't understand me at all.)

> P-301 (Xiaohongshu, female user, 20s): “倍速播放被它解读成‘热爱’，我明明只是想快速过掉上班路的无聊。” (My speed-listening was interpreted as “passion”, when I was clearly just trying to fast-forward through the boredom of my commute.)

These posts are theoretically instructive because they show users performing “boundary work” between the prosodic and the semantic. The user in P-274 explicitly distinguishes the functional, prosodic motivation for listening (“I need sound companionship”) from the semantic attribution made by the platform (“loves romance novels”). She is, in effect, insisting on the very distinction that forms the theoretical core of this paper. However, her resistance takes the form of claiming the platform “misunderstands” her, rather than recognizing that the platform is operating in a different register—building a self from rhythmic behavior rather than content preference. This suggests that users themselves may not yet have a vocabulary for the aesthetic self being constructed, even as they intuitively resist its conflation with the semantic self.

This finding complicates the theory in productive ways. Resonant recognition is not

automatic or universal. It requires the listener to accept, at least provisionally, that rhythmic behavior is self-disclosing—that what my body does with sound “is” in some meaningful sense “who I am”. Users who resist this premise, who insist on a sharp distinction between “real” selfhood (located in conscious intentions and content preferences) and “mere” bodily behavior, do not experience resonant recognition. They experience instead a category error—an algorithm that confuses the extrinsic with the intrinsic.

An alternative reading of these resistant posts, informed by Bucher’s (2018) work on algorithmic phenomenology, is that users are not simply lacking a vocabulary for the aesthetic self. Rather, they are actively performing what we might call “semantic boundary-maintenance”—a strategic refusal to cede narrative authority to the platform. In this interpretation, the resistance is less a cognitive deficit than a critical gesture: the user insists that the “real” self resides in intentional projects, not in involuntary bodily traces. This reading does not invalidate our framework but enriches it by showing that resonant recognition is not a universal response but a position that users may strategically accept or reject.

4. Reflection Without an Other: Hermeneutic Implications

The empirical findings presented above demonstrate that listeners do, in fact, engage in processes of resonant recognition when confronted with their prosodic portraits. In this section, we step back to theorize the broader hermeneutic implications of this phenomenon, focusing on the concept that we have flagged throughout: “reflection without an other”.

4.1 The Gadamerian Dialogical Condition and Its Suspension

Gadamer’s hermeneutics is built on the primacy of dialogue. Understanding, for Gadamer, is fundamentally an event that occurs “between” interlocutors. The model is the

Platonic dialogue: two partners engage in question and answer, each putting their own prejudgments at risk, moving toward a fusion of horizons that neither fully controls. The other is not an obstacle to self-understanding but its necessary condition: I come to understand myself only through the challenge of encountering a thou who sees the world differently. This dialogical structure has been extensively elaborated by Gadamer's commentators: Grondin (2003) traced its roots to the Platonic and Hegelian traditions, while Malpas and Zabala (2010) situated it within the broader "conversation" that constitutes hermeneutic philosophy.

The algorithm that constructs a prosodic portrait is not a thou. It does not engage in dialogue; it does not put its own understanding at risk; it does not respond to the listener's response. It aggregates and reflects. Yet—and this is the crux—listeners "do" come to new self-understanding through this non-dialogical encounter. The "yuánlái" moments documented in our data are genuine events of self-discovery, even though no dialogical partner stands behind the mirror.

How do we make hermeneutic sense of this? One option is to say that the algorithm is a degraded or deficient form of the Gadamerian other, and that the self-understanding it enables is correspondingly impoverished. We think this is partly right but insufficient. A more productive reading, we suggest, is that "prosodic engagement traces operate in a register where the dialogical condition is less essential than Gadamer assumed". Gadamer's model presupposes that understanding concerns "meaning"—propositional content that must be negotiated between perspectives. But prosodic recognition does not concern meaning in this sense. It concerns rhythm, tempo, texture—dimensions of experience that are pre-predicative and bodily, and that can be recognized without being interpreted.

This does not mean that prosodic self-recognition is solipsistic or immune to critique. It means that it operates through a different mechanism—resonance rather than negotiation—and that critical engagement must take a form appropriate to this mechanism. We return to this point in the conclusion.

4.2 Compressing the Ricœurian Detour

Ricœur’s hermeneutics of the self is organized around the concept of “detour”. The self, for Ricœur, has no direct intuition of itself. Self-understanding is always mediated: by signs, by texts, by narratives, by the other. The Ricœurian self is constituted through the long route of interpretation—a route that passes through the objectifications of human experience (the “works” that humans produce) and returns to the self transformed. Kaplan (2003) and Simms (2003) have both traced how Ricœur’s narrative theory resists any temptation to immediate self-presence, insisting instead on the constitutive role of mediation.

Prosodic engagement traces, we argue, offer a “shortcut that risks bypassing” this detour. When a listener sees that she unconsciously slowed down during every passage narrated in a minor key, and feels the jolt of self-recognition, she has compressed the narrative detour that Ricœur theorized as the necessary path to self-understanding. No story has been told; no plot has configured disparate events into a meaningful whole; no text has been interpreted in the robust Ricœurian sense. The recognition is immediate—it goes directly from the trace to the self, without passing through narrative.

This claim should not be overstated as a total rupture with Ricœur’s thought. Ricœur himself, in “The Course of Recognition” (2004/2005), explored forms of recognition that operate outside the narrative arc—the immediacy of recognizing another’s face, the mutuality

of gift-exchange, the phenomenology of “I can” as a pre-narrative sense of self-capability. Our concept of resonant recognition stands in productive dialogue with this later work: it shares Ricœur’s intuition that certain modes of self-encounter are pre-narrative or extra-narrative, but it relocates this possibility from the intersubjective domain to the algorithmic domain. The prosodic portrait offers a technologically mediated version of the immediacy that Ricœur glimpsed in face-to-face recognition—but with the crucial difference that no face stands behind the mirror. This relocation is what gives the concept of “reflection without an other” its critical edge: the shortcut to self-recognition is real, but it comes at the cost of the intersubjective and narrative mediations that, for Ricœur, were the guarantors of critical distance.

The immediacy of resonant recognition is both its promise and its danger. The promise is that it can reveal dimensions of selfhood—the aesthetic self, the rhythmic body—that narrative tends to smooth over or subordinate to the demands of plot coherence. Narrative identity, for all its richness, is a demanding mode of self-construal. It requires us to configure our lives into intelligible sequences, to establish causal and thematic connections, to find the arc. Many aspects of lived experience—particularly bodily, rhythmic, and affective dimensions—resist this kind of configuration. They are not narrative-shaped, and forcing them into narrative form can be a kind of violence. The aesthetic self recognized through prosodic traces may be, in some cases, a more faithful index of lived bodily experience than any story could be.

The danger is that the compressed route bypasses the critical safeguards built into the narrative detour. For Ricœur, narrative is not only a medium of self-construal but also a medium of self-critique. Because a story can be told otherwise—because the same events can be

configured into different plots—the narrative detour always carries the possibility of “seeing oneself as another,” of recognizing the contingency and revisability of one’s self-understanding. The immediacy of resonant recognition, by contrast, presents itself as given. “原来我是这样的” (So this is how I am)—the “yuánlái” structure carries an implication of discovery, not construction. The danger is that the aesthetic self, precisely because it is felt so immediately, may be taken as more fundamental, more authentic, than it is.

4.3 The Hermeneutic Circle Under Algorithmic Conditions

We are now in a position to characterize the transformation of the hermeneutic circle under algorithmic conditions. The classical hermeneutic circle, as formulated by Schleiermacher (1998) and radicalized by Gadamer (1960/2004), describes the movement between part and whole in the process of understanding: I understand the parts in light of an anticipation of the whole, and I revise my anticipation of the whole through engagement with the parts. This circle is dynamic, dialogical, and never conclusively closed.

The algorithmic hermeneutic circle, we propose, operates differently. The algorithm aggregates the parts—the individual prosodic engagement traces—and presents a whole: the prosodic portrait. The listener encounters this whole and either recognizes herself in it (resonant recognition) or refuses it (semantic resistance). But what the listener cannot do—or at least, what the platform does not readily enable her to do—is to enter into a genuinely circular movement with the portrait: to question it, to revise it, to see how different aggregations or different weightings would yield different portraits.

In other words, the algorithmic hermeneutic circle is “short-circuited”. The movement from part to whole happens inside the algorithmic black box, invisible to the listener. The

listener is presented with the finished whole and asked to respond. This is hermeneutics without hermeneutics: an act of interpretation whose interpretive operations are opaque to the interpreter.

Admittedly, this characterization is a theoretical extrapolation from our empirical findings rather than a direct observation: our data do not capture users' attempts to query the algorithm or revise its aggregation. The structural opacity of platform algorithms, however---well documented in the literature (Gillespie, 2014; Seaver, 2019)---supports the inference that such participatory revision is systematically discouraged.

This is the deepest sense of our phrase “reflection without an other”. It is not only that the listener lacks a dialogical partner. It is that the listener lacks access to the interpretive process that produced the portrait she is asked to recognize herself in. The algorithm is a hermeneutic agent, but it is an agent whose hermeneutic labor is hidden. The listener is offered the product of interpretation without being able to participate in interpretation. This, we suggest, is the specific hermeneutic challenge posed by algorithmic self-narration—a challenge that existing theories of the digital shadow, focused as they are on semantic traces, have not fully articulated.

5. Conclusion

5.1 Summary of Contributions

This paper has made three principal contributions to the theorization of algorithmic self-formation.

First, we introduced the concept of “prosodic engagement traces” as a distinct category of digital behavioral data, irreducible to the semantic traces that dominate current scholarship.

By drawing on conversation-analytic research on prosody, we argued that behaviors such as playback speed adjustment, rewinding, and temporal listening patterns register a listener's bodily and affective attunement to the sonic texture of speech—an attunement that operates prior to and independently of semantic comprehension.

Second, we developed the framework of “resonant recognition” to theorize the mode of self-encounter that prosodic traces enable. Drawing on Gadamer's aesthetics and Ricœur's hermeneutics of the self, we argued that when algorithmic systems aggregate prosodic traces into a prosodic portrait, listeners can experience a felt, immediate recognition of self in rhythm—an aesthetic self that is not narratively configured but bodily felt.

Third, we theorized the hermeneutic condition of “reflection without an other”, identifying both the novelty and the danger of algorithmic self-narration grounded in prosodic data. We argued that this mode of self-encounter suspends the dialogical condition central to Gadamerian hermeneutics and compresses the narrative detour that Ricœur posited as essential to critical self-understanding, while simultaneously opening access to dimensions of selfhood—the rhythmic, the bodily, the aesthetic—that narrative tends to subordinate or elide.

Empirically, our analysis of 327 user-generated social media posts demonstrated that listeners do experience resonant recognition in response to platform listening retrospectives, and that they articulate this experience through distinctive discursive strategies—centrally, the “yuánlái” structure and the active differentiation of prosodic from semantic engagement. We also identified forms of resistance (semantic boundary work) and unease (the algorithmic uncanny) that complicate and enrich the theoretical framework.

5.2 The Aesthetic Self: Between Revelation and Construction

We close the theoretical arc by offering a more precise characterization of the aesthetic self that emerges through prosodic engagement traces. The aesthetic self is:

- “Bodily rather than propositional”: It is constituted through rhythms, tempos, and textures registered by the listening body, not through beliefs, preferences, or judgments that can be stated in propositional form.

- “Recognized rather than narrated”: It is encountered in moments of “this is it!” (Gadamer) rather than gradually built up through the configuration of events into a life story (Ricoeur).

- “Ambiguously owned”: It is mine—the rhythms are my body’s rhythms—but it returns to me from the outside, through the algorithmic mirror, as something I did not know I was producing. I recognize myself in something I did not author.

- “Suspended between the given and the made”: The “yuánlái” structure positions the aesthetic self as “discovered” (“so this is how I am”), yet the discovery is enabled by algorithmic operations of aggregation and categorization that are constructive acts. The aesthetic self is algorithmically mediated revelation.

This last point is the site of the deepest critical question. Is the aesthetic self a genuine revelation of something true about the listener, or is it a construction that the platform passes off as revelation? Our answer is: both. The rhythms that prosodic traces register are genuinely the listener’s—they are produced by her body in response to sound. But the framing of those rhythms as a “self”, as “who she is”, is the platform’s hermeneutic act. The listener is offered her rhythms back in the form of an identity. Whether she accepts, resists, or negotiates this offer is the space of agency that remains.

5.3 Toward a Critical Listening Literacy

We return to the practical-imaginative register of the conclusion to sketch a response to the condition we have analyzed. If the algorithmic construction of the aesthetic self is both revelatory and constructive, and if its characteristic danger is that it presents construction as revelation, then the critical task is to cultivate what we call “critical listening literacy”: the capacity to engage with platform-generated prosodic portraits as “interpretations” rather than as “transcriptions”.

Critical listening literacy, as we envision it, involves at least the following dispositions:

1. “Interpretive awareness”: Recognizing that a prosodic portrait is the output of algorithmic operations—aggregation, categorization, weighting, comparison—that could have been otherwise. The portrait is “a” reading, not “the” reading.

2. “Rhythmic reflexivity”: Developing a first-person sensitivity to one’s own listening rhythms, such that the algorithmic portrait is encountered not as an alien revelation but as one perspective on already-familiar bodily patterns.

3. “Narrative agency”: Retaining the capacity to narrate oneself otherwise—to integrate, refuse, or reinterpret the prosodic portrait within a broader self-narrative that includes semantic, relational, and biographical dimensions.

These dispositions are not merely individual skills but require platform-level conditions. We note that some platforms have begun experimenting with designs that could support critical listening literacy: audio players that visualize playback speed changes over time, as seen in podcast players that offer temporal heatmaps of speed adjustments, per-episode skip-and-replay statistics, and experimental interfaces that invite users to annotate their own listening

profiles. These nascent design patterns suggest that critical listening literacy is not a utopian ideal but a horizon toward which platform design can and already partially does move. Our theoretical framework provides a grounding for advocating more deliberate movement in this direction—toward prosodic portraits that invite hermeneutic engagement rather than foreclosing it.

5.4 Limitations and Future Research

This study has significant limitations that point toward future research directions. The sample of 327 social media posts, while sufficient for the exploratory hermeneutic analysis we have conducted, cannot support claims about the prevalence or distribution of resonant recognition across demographic or cultural groups. Future research employing survey or experimental methods could complement our qualitative findings with quantitative mapping of the phenomenon. Cross-cultural comparison is particularly indicated: the prominence of the “yuánlái” structure in our Chinese-language data raises the question of whether resonant recognition is linguistically and culturally variable in its discursive realization.

Second, our analysis has focused on listener responses to platform retrospectives—a specific genre of algorithmic output. Future research could investigate whether resonant recognition occurs in other contexts: real-time listening analytics, social listening features, or voice-based AI interfaces that reproduce and modulate prosodic patterns.

Third, while we have gestured toward the dangers of algorithmic prosodic profiling—the potential for manipulation, the foreclosure of narrative agency—we have not developed a full critical account of the political economy of prosodic data. Who owns the listener’s rhythms? What forms of behavioral prediction and targeting do they enable? These are urgent questions

that our framework opens but does not settle.

Postscript

The listener’s digital shadow, as rendered in prosodic engagement traces, is a strange double. It is more intimate than the semantic shadow of search histories and content preferences, because it registers the body’s involuntary attunements. And it is more alien, because it returns this intimacy in the form of data that the listener did not know she was producing. To learn to live with this shadow—to recognize oneself in it without being captured by it—is an emerging task of digital hermeneutic literacy. This paper has sought to provide the theoretical vocabulary for that task.

Appendix A

Codebook Excerpts

Table A1 Codebook for the Analysis of User Responses to Listening Retrospectives

Code Node	Conceptual Definition	Typical Example	Boundary Notes
Resonant Recognition	User reports felt identification with prosodic patterns revealed by the platform, typically marked by “yuánlái” (原来) or equivalent	“原来我深夜听的都是这种调调” (So it turns out this is the kind of tone I listen to late at night)	Distinguished from “surprise at algorithmic accuracy,” which lacks the element of self-identification
Semantic Resistance	User rejects platform portrait on grounds that it conflates prosodic behavior with semantic preference	“那是我失眠时需要声音陪伴,跟内容没任何关系” (That’s me needing sound companionship during insomnia—nothing to do with content)	Distinguished from general dismissal; requires explicit boundary work between prosodic and semantic

Algorithmic Uncanny	User expresses unease at the intimacy of prosodic knowledge held by the algorithm	“好像有人在一直观 察我的呼吸” (As if someone has been watching me breathe)	Distinguished from privacy concerns; the object of unease is the pre-reflective nature of the knowledge, not its sensitivity
Aesthetic Self- Articulation	User actively develops an identity account grounded in sonic-aesthetic terms rather than content preference	“声音的质地让我一 遍遍回去,我甚至不记 得内容” (The texture of the voice made me go back again and again, I don't even remember the content)	Distinguished from narrative self- articulation, which references content, genre, or thematic preferences

References

- Bucher, T. (2018). *If... then: Algorithmic power and politics*. Oxford University Press.
<https://doi.org/10.1093/oso/9780190493028.001.0001>
- Cheney-Lippold, J. (2017). *We are data: Algorithms and the making of our digital selves*. New York University Press.
- Chion, M. (2019). *Audio-vision: Sound on screen* (C. Gorbman, Trans., 2nd ed.). Columbia University Press. (Original work published 1994)
- Couldry, N., & Hepp, A. (2017). *The mediated construction of reality*. Polity Press.
- Couper-Kuhlen, E. (2015). Intonation and discourse. In D. Tannen, H. E. Hamilton, & D. Schiffrin (Eds.), *The handbook of discourse analysis* (2nd ed., pp. 82–104). Wiley-Blackwell.
- Couper-Kuhlen, E., & Selting, M. (Eds.). (1996). *Prosody in conversation: Interactional studies*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511597862>
- Crystal, D. (2008). *A dictionary of linguistics and phonetics* (6th ed.). Blackwell.
- Freud, S. (2003). The uncanny. In *The uncanny* (D. McLintock, Trans., pp. 121–162). Penguin. (Original work published 1919)
- Gadamer, H.-G. (2004). *Truth and method* (J. Weinsheimer & D. G. Marshall, Trans., 2nd rev. ed.). Continuum. (Original work published 1960)
- Gillespie, T. (2014). The relevance of algorithms. In T. Gillespie, P. J. Boczkowski, & K. A.

- Foot (Eds.), *Media technologies: Essays on communication, materiality, and society* (pp. 167–194). MIT Press.
- Gillespie, T. (2018). *Custodians of the internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.
<https://doi.org/10.12987/9780300235029>
- Grondin, J. (2003). *The philosophy of Gadamer* (K. Plant, Trans.). Acumen.
- Gumperz, J. J. (1982). *Discourse strategies*. Cambridge University Press.
<https://doi.org/10.1017/CBO9780511611834>
- Heidegger, M. (1962). *Being and time* (J. Macquarrie & E. Robinson, Trans.). Blackwell. (Original work published 1927)
- Ihde, D. (2007). *Listening and voice: Phenomenologies of sound* (2nd ed.). State University of New York Press.
- Kaplan, D. M. (2003). *Ricœur's critical theory*. State University of New York Press.
- Kennedy, H. (2016). *Post, mine, repeat: Social media data mining becomes ordinary*. Palgrave Macmillan.
- Li, C. N., & Thompson, S. A. (1981). *Mandarin Chinese: A functional reference grammar*. University of California Press.
- Lupton, D. (2016). *The quantified self: A sociology of self-tracking*. Polity Press.
- Malpas, J., & Zabala, S. (Eds.). (2010). *Consequences of hermeneutics: Fifty years after Gadamer's Truth and Method*. Northwestern University Press.
- Page, R. (2018). *Narratives online: Shared stories in social media*. Cambridge University Press.
- Ricœur, P. (1976). *Interpretation theory: Discourse and the surplus of meaning*. Texas Christian University Press.
- Ricœur, P. (1984–1988). *Time and narrative* (Vols. 1–3, K. McLaughlin & D. Pellauer, Trans.). University of Chicago Press.
- Ricœur, P. (1992). *Oneself as another* (K. Blamey, Trans.). University of Chicago Press.
- Ricœur, P. (2005). *The course of recognition* (D. Pellauer, Trans.). Harvard University Press. (Original work published 2004)
- Rubery, M. (2016). *The untold story of the talking book*. Harvard University Press.

- Schleiermacher, F. (1998). *Hermeneutics and criticism and other writings* (A. Bowie, Trans. & Ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511814945.006>
- Seaver, N. (2019). Knowing algorithms. In J. Vertesi & D. Ribes (Eds.), *digitalSTS: A field guide for science & technology studies* (pp. 412–422). Princeton University Press.
- Simms, K. (2003). *Paul Ricœur*. Routledge.
- Sterne, J. (2003). *The audible past: Cultural origins of sound reproduction*. Duke University Press.
- Turkle, S. (2011). *Alone together: Why we expect more from technology and less from each other*. Basic Books.
- Tsai, W. -T. D. , & Yang, C.-Y.H. (2022) On the syntax of mirativity: Evidence from Mandarin Chinese. In A. Simpson (Ed.), *New explorations in Chinese theoretical syntax: Studies in honor of Yen-Hui Audrey Li* (pp.431–444). John Benjamins.
<https://doi.org/10.1075/la.272.15tsa>
- Wennerstrom, A. (2001). *The music of everyday speech: Prosody and discourse analysis*. Oxford University Press.