Linguistic Frontiers • Ahead of Print • 2025
DOI: 10.2478/lf-2025-0021

Linguistic **Frontiers**

AI and Animal Communication: A Generative Zoosemiotics Perspective

Original Study

Nicola Zengiaro
University of Turin
Faculty of Philosophy and Science Education
nicola.zengiaro@unito.it
Orcid: 0000-0003-3268-7567

Received: 11 April 2025; Accepted: 19 August 2025

Abstract: In recent years, the application of artificial intelligence (AI) to the study of non-human communication has gained significant momentum. This article critically explores the intersection between AI and zoosemiotics by analysing a corpus of three emblematic case studies – Project CETI, ISPA, and current bioacoustic classification systems – which apply generative models and machine learning techniques to animal communication. Through close textual and conceptual analysis, the article shows how these approaches, despite their technical sophistication, tend to reduce animal semiosis to its external form, overlooking its embodied, ecological, and relational dimensions. The study exposes the epistemological, methodological, and semiotic limits of such models, demonstrating how the situated nature of animal signification is often flattened into abstract algorithmic encoding, devoid of pragmatic context. In response to these issues, the article introduces the concept of Generative Zoosemiotics (GZ), a new theoretical framework that combines tools from zoosemiotics and cultural semiotics with a critical inquiry into AI technologies. GZ proposes to analyse how machines interpret, simulate, or reconfigure animal signs, questioning the epistemic and ethical implications of these practices and asserting the irreducibility of interspecific communication to the logic of symbolic translation.

Keywords: Generative Zoosemiotics, AI, Project CETI, ISPA, NatureLM, bioacoustics, interspecific communication

INTRODUCTION

The use of artificial intelligence (AI) in the study of animal communication has recently gained unprecedented prominence, generating a new epistemological horizon in which the possibility of "translating" signals produced by non-human species into a language accessible to human interpretation is increasingly being considered. Some of the most ambitious projects in this direction, such as Project CETI (Cetacean Translation Initiative), propose applying techniques from natural language processing (NLP) and deep learning to the study of sperm whale vocalizations, imagining linguistic models capable of recognizing, and even generating, "whalish"[1] language structures.

However, from the perspective of semiotics, and in particular zoosemiotics, this scenario raises fundamental

1 *Whalish* is not a formal linguistic category but a speculative term used to refer to a hypothetical language system of sperm whales, mainly in AI-driven modeling contexts. Specifically, it is used in Goldwasser et al. (2023, 2): "This is in contrast to our analysis suggesting that the number of Whalish samples necessary for translation is on the order of what is required for ordinary (supervised) MT". However, in popular contexts, the term is commonly used to refer to the language of whales (e.g. Wired, "How to Use AI to Talk to Whales—and Save Life on Earth": https://www.wired.com/story/use-ai-talk-to-whales-save-life-on-earth/ [visited 27/07/2025]).

theoretical questions: is it legitimate to consider animal signs as translatable according to the categories of human language? Is it epistemologically sound to reduce animal semiosis to algorithmic structures? And what are the theoretical, ethical, and ontological consequences of a computational approach to animal signification?

The enthusiasm that is driving the debates, especially those of popularisation[2], do not take into account the inherent difficulties in "translating" a non-human, non-linguistic communication, whose structural characteristic depends on forms of expression that emerge from different physiologies, cognitions, ecologies. One only has to look in bookshops, newspapers, TV series, documentaries to realise that the discourses on AI and animal communication promise to give rise to an uncritical, linear revolution, banalising the communicative complexity of other species, and on the other hand to indirectly claim that the human is capable of translating through its own technology every mystery hidden within non-human life (Yovel, Rechavi 2023). In fact, the debate seems to have discovered animal communication, trying to translate it, to make animals respond with reproduced sounds or interact with them, but at the same time keeping it distant and distinguishing it strongly from human language. This debate that is now very fashionable seems to have forgotten the legacy of Sebeok and zoosemiotics, which determined different language models (Sebeok, Danesi 2000) and distinguished animal communication and human language precisely (Sebeok 1990).

The discipline of zoosemiotics, as originally conceived by Thomas A. Sebeok (in an exchange with Rulon Wells), emerged from a fundamental distinction between human and non-human forms of signification. Sebeok introduced the term "zoosemiotics" to refer to the study of animal communication, but explicitly rejected the idea that non-human species possessed language in the human sense. In his writings, he insisted that while all animals operate within species-specific semiotic systems, language – as a symbolic, recursive and syntactically complex system – is unique to humans. According to him, it is misleading to speak of "language" in reference to animal communication, except in a metaphorical sense (Sebeok 1990, 106). This distinction led Sebeok to propose a bifurcation of semiotic enquiry: on the one hand, anthroposemiotics, devoted to the verbal and symbolic capacities of human beings[3]; on the other, zoosemiotics, which deals with the codes and sign systems of animals, including aspects of communication shared with humans. This theoretical architecture reflected his growing scepticism, shared with figures such as Chomsky, about the possibility of studying human and animal communication within the same analytical framework (Sebeok 1972; 1977). Zoosemiotics was thus conceived not as a theory of animal language, but as a rigorous semiotic approach to animal semiosis, understood as the ability to produce and interpret signs.

Over time, however, the field has evolved considerably. As contemporary scholars such as Maran, Martinelli and Turovski (2011, 1–3) have pointed out, the focus of zoosemiotics has shifted from the analysis of communication alone to a broader investigation of semiosis, such as the relational and context-dependent processes through which signs acquire meaning within ecological and social systems. This evolution has broadened the scope of the discipline to include not only inter- and intra-specific communication, but also the cultural and biological entanglements that define human-animal relations. Rather than reinforcing a dualistic division between human and non-human semiosis, contemporary zoosemiotics tends towards a more integrated and systemic view that refuses to reduce animal signification to mere instinct or simplified linguistic analogy.

From this perspective, approaches that attempt to oppose human and animal communication in binary terms (as found, for example, in some recent philosophical attempts to critique language by comparing it with animal signalling) often fail to capture the semiotic complexity of both. Today, however, zoosemiotics invites us to understand sign processes as emergent and embedded in shared ecologies of life, perception and meaning.

Moving within a critical and comparative framework, this article analyses three studies that are representative of this new trend about translating animal communication through AI: "Toward Understanding the Communication in Sperm Whales" (Andreas et al. 2022), "Using Generative Artificial Intelligence to Test Hypotheses about Animal Signal Evolution" (Héjja-Brichard et al. 2023), later republished in another article in 2024 entitled "Using neural style transfer to study the evolution of animal signal design: A case study in an ornamented fish" (Héjja-Brichard et al. 2024), and "Using Machine Learning to Decode Animal Communication" (Rutz et al. 2023). This corpus is examined through conceptual tools drawn from communication semiotics, biosemiotics, and zoosemiotics, with the aim of highlighting not only the internal aporias of the algorithmic paradigm, but also

2 https://www.nature.com/immersive/d41586-024-04050-5/index.html; https://www.economist.com/podcasts/2025/02/19/could-ai-translate-animal-communication; with title like "Artificial Intelligence Could Finally Let Us Talk with Animals": https://www.scientificamerican.com/article/artificial-intelligence-could-finally-let-us-talk-with-animals/ (visited 10/04/2025). This is not simply a matter of public popularisation, as the investments raised in 2020 amounted to $33 million. This is to emphasise that this is not a classic case of media bias (cf. https://bioneers.org/can-ai-decode-whale-sounds-project-ceti-is-here-to-find-out-ztvz2504/ [visited 25/07/2025]).

3 Jamin Pelkey, during the congress "Animality and Biosemiotics" at the University of Cagliari (20/06/2025), introduced me to Frederik Stjernfelt's (2013, 96–105) engagement with Barsalou's theory of *perceptual symbols*, which rethinks abstraction not as the elimination of sensory detail, but as the selective reactivation of perceptual traces (or simulators) stored in memory. In this framework, concepts and symbols are not detached from experience but grounded in multimodal simulations, dynamically tailored to context. This idea resonates with a semiotic understanding of meaning as situated and embodied, offering a counterpoint to Sebeok's more categorical distinction between symbolic language and animal semiosis.

the possibilities opened by a qualitative, situated, and relational approach to animal sign use.

1. WHALE LANGUAGE? A SEMIOTIC ANALYSIS OF THE COMPUTATIONAL PARADIGM IN PROJECT CETI

The paper "Toward Understanding the Communication in Sperm Whales" (Andreas et al. 2022) presents the ambitious endeavour of Project CETI, a transdisciplinary initiative that aims to apply the most advanced techniques in artificial intelligence to the communication of sperm whale (*Physeter macrocephalus*), with the stated goal of understanding its structure, deciphering its functions, and ideally establishing a communicative bridge between humans and cetaceans. The article outlines a detailed methodological roadmap involving the large-scale collection of bioacoustic data, its automatic segmentation, the training of machine learning models for natural language processing (NLP), which in this context act primarily as analytical tools rather than generative systems, and finally, controlled playback experiments for semantic validation. The lexicon employed is systematically drawn from formal linguistics: the authors refer to "phonetic units", "syntax", "semantics", "pragmatics", "turn-taking", and "sentence structure". Whale communication is thus reinterpreted in the terms of a natural language, equipped with grammar, vocabulary, and potentially subject to translation.

This operation, which translates animal communication into human linguistic terms through AI-driven data acquisition, is anything but neutral. It rests on a specific semiotic and ideological assumption: that animal communication can be modelled according to the structural categories of human language. This hypothesis, although recurrent in the history of ethology and cognitive science, has long been criticised by zoosemiotics and biosemiotics as an anthropocentric projection lacking empirical foundation. In the animal world, the sign does not necessarily take the form of a symbolic, discrete, or combinatory structure; it may manifest iconically, indexically, or as a fully embodied and situated form. Moreover, communication is not always oriented toward the transmission of content, it may instead be performative, affective, identificatory, or relational. To speak of "decoding" sperm whale language thus implies the implicit assumption of a predefined code, abstracted from context and accessible to algorithmic analysis, a view that ultimately elides the radically pragmatic nature of animal semiosis.

The epistemological paradox of Project CETI[4] lies precisely here: in attempting to acknowledge the communicative alterity of sperm whales, it ultimately reabsorbs that alterity into a formalized linguistic model that presupposes, or imposes, a structural homology with human language. Cetacean vocalizations are thus treated as a set of linguistically interpretable signals, endowed with latent syntax and potentially computable semantics. However, what appears in sperm whales as *codas* (distinctive sequences of acoustic clicks produced over time) is not a form of writing to be translated, but rather a communicative praxis embedded in a web of ecological, social, and phylogenetic relations that shape its structures.

Even within human languages, as Roman Jakobson (1959; 1960) demonstrated, translation is never a neutral substitution of signs: it implies a transposition between semiotic systems; "languages differ essentially in what they *must* convey and not in what they *may* convey" (Jakobson 1959, 236). Jakobson famously identified three types of translation: intralingual (within the same language), interlingual (between languages), and intersemiotic or transmutation[5], "an interpretation of verbal signs by means of signs of nonverbal sign systems" (Jakobson 1959, 233). In this typology, intersemiotic translation proceeds from the verbal to the non-verbal: for instance, from a novel to a painting, from a poem to a musical piece, or from literature to cinema (cf. Greimas, Courtés 1979; Helbo 1997; Metz 1971; Genette 1982). These forms of transposition highlight the partial incommensurability of expressive planes[6], while maintaining some degree of content translatability across media and sensory regimes.

However, the procedure adopted in Project CETI reverses this semiotic logic: rather than interpreting verbal signs through nonverbal ones, it seeks to interpret nonverbal acoustic behavior (sperm whale codas) through the analytical and generative tools of natural language processing, re-inscribing them into a verbal-semiotic framework. This operation is not an intersemiotic translation in Jakobson's sense, but its inverse: a retro-projection of human language structure onto a non-human communicative form, whereby vocalizations are treated as translatable utterances with latent syntax and computable semantics. In doing so, Project CETI enacts not a transmutation but a formal recoding of whale communication according to the grammar of human linguistic systems.

4 Personal note: It must be said that I absolutely support the CETI project, mainly because this approach more than translating the animals and talking to them brings their existential importance closer to the social and political context. Studying the language of whales somehow protects them, so I see these studies as absolutely necessary.
5 In the context of intersemiotic translation, the term "transmutation" is used by Roman Jakobson to emphasize the transformative nature of interpreting verbal signs through non-verbal sign systems, such as adapting a novel into a film. This concept highlights a significant shift in the medium and expression, rather than a direct, word-for-word equivalence.
6 The distinction between expression plane and content plane is central to Louis Hjelmslev's glossematics, as developed in his *Prolegomena to a Theory of Language* (1961). Hjelmslev posits that every semiotic system is a correlation between these two inseparable planes. The expression plane refers to the material and formal manifestation of a sign (e.g., sounds, written characters, images), while the content plane refers to the meaning or concept conveyed. Crucially, Hjelmslev argues that both planes possess both a form (their structured organization) and a substance (their specific material realization), allowing for a rigorous analysis of how meaning is structured and expressed across diverse semiotic systems.

Such an approach, nonetheless, overlooks the fact that, even among human languages, translation is never fully isomorphic, as it involves interpretive processes that mediate between differing semantic, pragmatic, and cultural frameworks. Jakobson himself argued that poetry, for instance, is "by definition untranslatable", and that "only creative transposition is possible: either intralingual transposition – from one poetic shape into another, or interlingual transposition – from one language into another, or finally intersemiotic transposition – from one system of signs into another, e.g. from verbal art into music, dance, cinema, or painting" (Jakobson 1959, 238). This is due to the predominance of the poetic function of language, in which the form of the message becomes the message itself (Jakobson 1960, 356). In contexts where the emotive, phatic, or performative dimensions are predominant, meaning cannot be abstracted from its embodied, situated expression. If this holds true within the human domain, it becomes all the more evident across species boundaries, where meaning is not only entangled with embodiment and context, but also with species-specific perceptual and ecological constraints. Thus, the assumption that sperm whale vocalizations can be translated into a human language framework does not merely involve anthropocentric projection, it exemplifies an epistemological inversion of Jakobson's intersemiotic model: an attempt to transform the living, relational semiosis of a non-human species into a static code legible to machines.

Ethological studies have shown that these vocalizations play roles in identity recognition, clan cohesion, and cultural transmission, all within a highly cooperative and intergenerational social system (Gero et al. 2016, 9–10). The resulting semiosis is therefore relational, affective, and embodied: it is neither translatable nor formalizable as a natural language. Furthermore, semiosis in the natural world derives from a continuous structuring of the habits that the organism and the environment sediment in their communicative interaction (Bernardi della Rosa, Zengiaro 2025); a communication that demands the resolution of environmental problems from its own species-specific endowments.

In contemporary zoosemiotics, this communicative process should be understood in terms of the species-specific *Umwelt* (Cerrone, Mäekivi 2021), that is, as a situated form of signification shaped by the way sperm whales perceive, inhabit, and construct their subjective and intersubjective world. In this framework, the function of the sign is not to convey informational content, but to orient behaviour within a meaningful environment. The replacement of this perspective with an informational paradigm based on associations between acoustic patterns and syntactic structures gives rise to an epistemology of semiotic simulation, in which the sign is no longer enacted by a subject but computed by a system. In other words, it becomes a grammar without life, a language without a subject. Moreover, the data-intensive approach employed here tells us little about the animal's lifeworld, nor about the transmission of the ecosystemic, i.e., semantic content of animal communication (Farina 2014, 115).

Particularly emblematic is the tension between the vast quantity of data collected (up to four billion potential vocalizations) and the absence of any reflection on the ecological context, the animal's body, or the interpretive processes distributed across environment, social group, and individual. The interpretant, central in the Peircean and Sebeokian tradition of semiosis, is replaced here by an algorithm that does not interpret, but rather recognizes and classifies signal regularities based on statistical correlations. This automation of interpretation carries with it a deep theoretical ambiguity: the sign is no longer identified by what it does within a living relation, but by what it appears to be, as a recognisable and replicable form of data, temporally and spatially located, yet without a subject. The risk is the production of an abstract *technosemiotics*[7], one that simulates communication without ever engaging with it, and that conflates symbolic representation (linguistic) with semiotic signification (relational).

A zoosemiotic alternative would not consist in attempting to "translate" the language of sperm whales, but rather in recognising the specificity of their mode of signification, in its opacity and its irreducibility to human grammar. This was also clarified by Kull and Torop (2003) in their "biotranslation" proposal. The animal sign is not a hidden word, but a situated gesture, an act of relation that takes shape within a historical, ecological, and affective context. The true epistemic challenge is not decoding, but the construction of an objective posture toward interspecific interpretation, one that acknowledges the otherness of animal communication and its resistance to formalization. Semiotics teaches us that any process of translation must take into account a subjective world that remains inexhaustible, regardless of the potentially infinite quantity of data collected. In this sense, Project CETI, despite its technical sophistication, risks enacting a symbolic reduction of the semiotic alterity of the animal, transforming the whale's song into an alien grammar, ready for vocal synthesis, yet devoid of all subjectivity.

1.1 The Alphabet of Alterity? ISPA and the Limits of Interspecies Phonetic Analogy

As part of the broader attempt to formalize animal communication through the lens of human linguistic models, the ISPA project (Inter-Species Phonetic Alphabet for Transcribing Animal Sounds) (Hagiwara, Miron, Liu 2024, 828) represents a further extension of the grammatical paradigm applied to bioacoustics. Developed within the

7 By "technosemiotics" I refer here not to a formalized discipline, but to a process whereby signs are abstracted from ecological, embodied contexts and reformulated in algorithmic, representational systems. The term designates a semiotic condition where the material and intersubjective dimensions of communication are replaced by symbolic approximations optimized for machine processing.

Earth Species Project (ESP), ISPA proposes a system for transcribing animal sounds into text, inspired by the International Phonetic Alphabet (IPA), which is widely used to transcribe human linguistic sounds in a standardized and interpretable format. The stated goal is to construct an interspecies alphabet that is "precise, concise, and interpretable", capable of rendering animal vocalizations in textual form and thereby making them compatible with linguistic modeling technologies used in natural language processing (NLP).

More specifically, ISPA is divided into two approaches: ISPA-A (Acoustics-based), which is an acoustic transcription based on pitch, duration, spectral features, and tonal slope of the signal; and ISPA-F (Feature-based), a transcription based on latent acoustic features, obtained via vector clustering and subsequent phonetic mapping. Both systems convert the continuous audio signal into a discrete sequence of symbols, which can then be used to train NLP models such as RoBERTa: "a powerful masked language model (MLM), after tokenization, extract its pooled output, and feed it into a classification head" (Hagiwara, Miron, Liu 2024, 830). The logic is straightforward: to treat animal communication as if it were a foreign language, with its own phonetic alphabet, its own lexicon, and a structure that can be decoded and manipulated by NLP models. The results reported suggest that, at least for bioacoustic classification tasks, ISPA transcriptions, particularly those based on latent features (ISPA-F), achieve competitive, and at times superior, performance compared to models that rely on continuous audio representations.

But it is precisely here that the central question for semiotic reflection emerges, which can direct the criticism of an AI approach to animal communication in a more coherent manner. ISPA is not designed to understand animal communication, but rather to formally represent it within a system compatible with the symbolic architecture of linguistic models. The transition from animal vocalization to written text is, in itself, a semiotically significant act: it entails an ontological transformation of the sign, from a corporeal, ecologically situated sign into a computable symbol. The analogy with the human phonetic alphabet is not merely technical but conceptual: it assumes that animal sounds can be structured according to discrete, combinatory, and repeatable logics, the very premises that underpin the grammatical model of human language.

> Also, while approaches of linguistic origin tend to search distinct meaningful word-like units in animal communication, the Uexküllian approach is in contrast surprisingly close to earth, showing for instance, how being in the world or embodiment itself can be meaningful, like the body of a hermit crab fitting into the shell of sea snail, or the body of an octopus fitting into the relationship with seawater. Such questions are often not treated by anthroposemiotics, since in the intraspecific communication between humans, body structures, organs used for perceptions and expressions are just invariants. For the comparative perspective of zoosemiotics, by contrast, the embodiment of meaning has central importance. (Maran, Martinelli, Turovski 2011, 52)

This operation entails a second-order semiosis, a metarepresentation based on the extraction and re-encoding of sound, in which the sign is detached from its environment and reformatted according to principles of informational efficiency and syntactic parsimony, along with symbolic semantics. Yet from a zoosemiotic perspective, the animal sign is not simply a sound: it is an embodied event, one that occurs between bodies, within environments, and under specific ecological and historical conditions.

Current linguistic research and recent studies on cetacean communication (Beguš et al. 2025) increasingly draw analogies between human vowels and whale codas. As seen below, such parallels are not only accepted but often serve as a starting point for exploring animal communication from a linguistic perspective.

> Human vowels consist of a sequence of glottal pulses produced by vocal folds. Whale codas consist of a sequence of clicks produced by vibrating phonic lips, which play a role similar to the human vocal folds. In human languages, the frequency of glottal pulses corresponds to pitch—closely spaced glottal pulses give rise to a higher pitch, while more widely spaced pulses give rise to a lower pitch. In linguistics, *tone* refers to pitch as recruited to express linguistic meaning. Many languages use tone to distinguish between different words. For example, in Mandarin Chinese, the following four words differ only in their tonal contour, while having the same consonants and vowels: high and level tone *mā* 'mother,' rising tone *má* 'hemp,' falling-rising tone *mǎ* 'horse,' and falling tone *mà* 'scold.' The coda types are therefore analogous to human tone: "regular" coda types can be compared to level tones, codas with "increasing" inter-click intervals are falling tones, and codas with "decreasing" inter-click intervals are rising tones. (Beguš et al. 2025, 2–3)

While the analogy between glottal pulses in human speech and click sequences in whale codas provides a useful acoustic and articulatory comparison, it should be approached with caution. Comparing coda types to lexical tones in tonal languages (e.g., Mandarin) illustrates potential syntactic patterning in animal vocalizations, but it does not by itself demonstrate semantic function.

From a zoosemiotic viewpoint, what matters is whether such structured variations are meaningfully interpreted by conspecifics in specific contexts. Syntax-like regularities, such as varying inter-click intervals, may reveal patterned signal production, but they do not prove symbolic reference or semantic encoding unless supported by evidence of contextual interpretation.

These acoustic structures may be useful for modeling, but cannot be taken as evidence of a full semiotic system in the Peircean sense. This distinction is essential for zoosemiotics, which seeks not only to analyze the formal features of animal signals, but to understand their role within communicative semiosis. Thus, while human tone languages offer a helpful heuristic, they should not be assumed to imply semantic or pragmatic equivalence with animal vocalizations.

The attempt to convert such an event into a typographic symbol constitutes a semiotic dispossession, a process in which relational meaning is sacrificed in favour of algorithmic representability.

> The significance of indexicality, within zoosemiotic study, is to be considered more central than in the case of anthroposemiotics, particularly its aspects related to human culture, whose logocentric nature predictably manifests a greater interest in symbols and icons. In a sense, one may go as far as to say that one of the main side-effects of the production of an artificial and/or cultural sign system is exactly its emancipation from indexicality, in favor of iconicity and (most of all) symbolicness. […] What instead is true is that indexicality, although requiring mental processes in most cases, does not necessarily requires complex and *multi-layered* mental processes, and that is primarily because an index tends to require temporal presence and spatial continuity, and its encoding is strongly bound to the *hic et nunc* of semiosis. (Martinelli 2010, 70)

As Sebeok (1990, 106) clearly argued, animal semiosis cannot be reduced either to linguistic models or to graphic formalization. This is not only because many animal signals are pre-linguistic or extra-linguistic, but above all because the sign cannot be separated from behaviour, intention, or interactive context. A vocalisation is not merely an acoustic output, but a complex communicative practice involving bodily positioning, emotional activation, interactive response, environmental conditions, and interspecific translations, it acquires meaning only within the species-specific Umwelt.

By proposing ISPA as an "interspecies phonetic alphabet", the authors thus perform an epistemologically powerful but also problematic operation: they transpose the human semiotic paradigm (based on textuality, phonemic interpretability, and discrete combinability) onto forms of life that do not share either the structure of linguistic signs or the communicative functions of human language. In this sense, ISPA aligns closely with the trajectory of Project CETI, contributing to the construction of a grammatical ideology of alterity, which seeks to understand animal communication only insofar as it can be transcribed, segmented, and symbolised.

A critical zoosemiotic reading, however, calls for a shift in perspective. Rather than transcribing animal sound into symbolic signs, it would be more appropriate to construct an interpretive model that recognises sound as an act, not a text[8], as a situated gesture, not a string to be encoded. What we are witnessing is a counter-revolution which, whereas in the 1960s had led researchers like Jane Goodall to develop field-based observation culminating in ecological ethology, is now shifting away from direct engagement toward impersonal data collection. This trend is, of course, partly due to human limitations: we cannot observe everything, nor listen to every whale song. However, the growing difficulty in doing so appears to stem less from an actual incapacity to manage data, and more from a broader performative pressure that prioritizes measurable output over situated knowledge. Transcription, in this regard, is always already an anthropocentric mediation, which risks masking semiotic otherness behind the familiarity of letters. The dream of an interspecies alphabet, for all its engineering elegance, thus reveals its conceptual limit: the attempt to universalise a form of the sign that is, in fact, the specific product of a phonocentric and scriptocentric linguistic culture. Animal communication, by contrast, continues to resist grammatization, asking not to be transcribed, but to be listened to.

2. SIMULATING THE SIGN: GENERATIVE AI AS AN EXPERIMENT IN EVOLUTIONARY HYPOTHESES

In an attempt to explore the evolution of animal signals through generative artificial intelligence, Héjja-Brichard et al. (2023; 2024) propose a computational approach that employs style transfer models to generate, manipulate, and visually test variations in animal body patterns. The experiment involves artificially modifying the appearance of organisms, such as fish, to simulate plausible chromatic and morphological changes, and observing whether such alterations are perceived as natural or ecologically coherent by conspecifics. The main objective is to use generative AI, specifically Neural Style Transfer (NST), to manipulate visual features of animal signals (e.g., male fish body patterns) in a biologically plausible manner, in order to test hypotheses about signal evolution and habitat congruence by assessing the visual preferences of conspecifics through behavioral experiments (Héjja-Brichard et al. 2023, 1). This is complemented by a broader comparative analysis of assortative mate preferences across darter species, highlighting interspecific variability beyond ecological or sex-based factors (Héjja-Brichard et al. 2024, 1).

From a semiotic perspective, this experimental design is conceptually stimulating, yet not without theoretical limitations. By relying on Neural Style Transfer to modify body patterns in relation to habitat textures, the

8 This stands in contrast to the attempt made in Marrone & Mangano (2018), which, despite its intentions, approaches the field of zoosemiotics in a reductive manner, ultimately misrepresenting or misunderstanding the core principles and epistemological stakes of the discipline. For a critique of such "zoosemiotics" approach see (Zengiaro 2024, 77).

model implicitly treats the animal signal as a separable visual artifact, an aesthetic feature that can be extracted, decomposed into "style" and "content" (Héjja-Brichard et al. 2023, 3), and recombined in novel forms. Such computational manipulation detaches the signal from its embodied origin, displacing the organism as a semiotic subject. What is evaluated is no longer the sign within a living Umwelt, but a visual surrogate, generated and assessed according to algorithmic plausibility and behavioral proxies. This shift exemplifies a semiotic disembodiment: communication is redefined not through lived interaction, but through a cascade of representations mediated by technological operations.

Zoosemiotics, however, show us that meaning-making in living beings cannot be reduced either to visual patterning or to isolated behavioural responses (Kull 2014; Lewis 2021; Zengiaro, Jaramillo 2025) (nor, as Pavlov would have it, to a one-to-one stimulus-response relation). It emerges from the concrete, situated interaction between body, environment, conspecifics, and ecological context. As Timo Maran (2020, 35) has emphasised, signs in animals are not merely "visual stimuli", but complex elements within a communicative dynamic that involves posture, rhythm, mutual positioning, emotional and affective states, ecological memory, and intercorporeality (Maran 2014, 85). Animal semiosis is not a form of visual design, it is an embodied and co-constructed process. The computational approach proposed by Héjja-Brichard et al. (2023; 2024) disables this pragmatic dimension by reducing the sign to a computable visual index, whose communicative value is inferred deductively by a system external to the animal's vital world.

The epistemological issue raised by this model is ultimately ontological: the signal is conceptualized not as a relational event between living organisms, but as a transferable and quantifiable property, amenable to manipulation and stylistic recombination. In Héjja-Brichard et al. (2023, 2), while the framework is situated within the sensory drive hypothesis, the notion of "signal" remains theoretically unexamined[9]. The authors treat it as a visual object to be processed, rather than as a communicative act embedded in ecological and social relations. This recalls the classic debate between Umberto Eco (1973) and Thomas A. Sebeok (1994) on the distinction between "signal" and "sign". While Eco emphasized the cultural coding and conventional nature of signs, reserving the term "sign" for semiotic systems characterized by syntactic and semantic articulation, Sebeok adopted a broader semiotic framework, arguing that animals, too, engage in sign processes, including icons and indexes, and that zoosemiotics must account for this variety. However, in the model analyzed here, the signal is reduced to a perceptual form whose validity is measured solely through behavioural responses or neural network recognition, without reference to an interpretant or to the semiotic agency of the organism. This results in a pseudo-pragmatics of effect, where perceptual response is equated with meaning, bypassing any genuine interpretive process. The living body disappears as a semiotic subject, replaced by computable traits optimized for algorithmic selection.

The logic of simulation underlying this approach generates a further epistemic risk: the construction of artificial semiotic worlds, in which manipulated images, though perceptually plausible, are disconnected from the ecological and biological dimension of animal signs. Maria Giulia Dondero (2025) analyses how, in the context of automated image generation (for example, with Midjourney), images are first transformed into numerical embeddings and then used as combinable elements within a latent space. This space, which Dondero (2025, 112) compares to a Saussurean *langue* as reinterpreted by Fontanille (2003), becomes a virtualised system of semiotic possibilities. The generation of new images occurs on this basis, through operations of stylistic and compositional manipulation which, as rightly noted, lose any biological, affective, or corporeal anchoring to the original situation of the sign.

When this model is transposed to animal communication, the epistemic risk becomes evident: the creation of "artificial semiotic worlds" in which animality is replaced by computational simulacra trained to satisfy human expectations. What emerges is an iconic and symbolic regime in which the animal sign becomes entirely detached from the body, from context, and from interspecific semiotic relations, replaced by a "plausible representation" that conforms to aesthetic recognisability, but not to ecological interpretation. Nevertheless, this very plausibility may give rise to ambiguous effects: artificial signals could be perceived by non-human animals not as meaningful signs, but as sufficiently patterned disturbances to trigger a communicative response, particularly of a *phatic*[10] nature. In such cases, the simulated signal functions as a semiotic approximation: not

9 From a semiotic perspective, the approach proposed by Héjja-Brichard et al. (2023) is of considerable interest in that it seeks to move beyond the traditional reduction of animal signals to isolated perceptual features, emphasizing the ecological congruence between visual patterns and habitat structures. However, it retains certain limitations. The signal is modelled as a stylisable and computable visual property, subject to algorithmic manipulation and measured via perceptual bias or behavioural response, yet it is not conceptualized as a communicative act embedded in a relational semiotic process. While the authors acknowledge the multimodal and multicomponent nature of animal signals, they do not explicitly reflect on the interpretive role of the receiver beyond preference or salience.

10 Roman Jakobson defines the phatic function (taken from Malinowski) as the use of language primarily aimed at establishing, maintaining, or discontinuing communication between interlocutors. It does not convey semantic content, but rather ensures that the channel of communication is open and operative. Typical examples include greetings, confirmations of attention, or ritualized exchanges such as "Hello?", "Are you listening?", or "Well!"; dialogues that serve only to maintain contact. Jakobson emphasizes that this is the first linguistic function acquired by infants, who engage in communication before being able to convey informative messages, and that it is also the only verbal function shared

"understood" in its content, but recognised as a possible attempt at interaction, akin to the perception of a defective or unfamiliar conspecific. This evokes what could be described as a kind of uncanny valley of semiosis, where the signal is neither fully interpretable nor completely ignored. Numerous examples of interspecific adaptation in the wild suggest that such responses may arise from minimal semiotic cues.

In this regard, it is useful to recall Roman Jakobson's (1960, 356–7) functional model of communication, which classifies six primary functions of semiosis: referential, expressive, conative, phatic, metalinguistic, and aesthetic. Although originally developed for linguistic contexts, this model has been successfully extended to animal communication (Cimatti 1998, 59–105), where it helps explain how even minimal or ambiguous signals can activate specific functions (particularly the phatic, which is oriented not toward content, but toward the maintenance or initiation of contact). Importantly, several of these functions, such as the referential, expressive, or aesthetic, do not presuppose a fully structured sender–receiver interaction, making them especially relevant for modelling interspecific encounters (Martinelli 2010, 77–81).

Nonetheless, this residual responsiveness does not validate the simulation as communication in the proper sense; rather, it underscores the precariousness and opacity of interspecies interpretation. In this light, the virtualisation of the sign does not merely obscure ecological meaning; it may also induce false positives in interspecies communication, provoking responses where none were intended, or concealing intentional signals under an aesthetic veil of recognisability aimed at the human observer.

This dynamic, in which semiosis is absorbed into a latent space governed by stylistic plausibility rather than ecological embeddedness, finds a compelling counterpart in what Massimo Leone (2025, 26) describes as the emergence of digital agricultural environments, namely the "zoover" and "phytovers". These are immersive virtual spaces where animal and plant life are simulated through technologies such as virtual reality, AI, and digital twins. Leone's (2025, 27) analysis foregrounds how these systems are not merely technical supports, but semiotic artefacts that mediate our perception, understanding, and interaction with biological life. The zoover, in particular, refers to environments where animal behaviour is modelled, visualised, and manipulated for purposes such as veterinary training, animal welfare monitoring, or public education. While these applications offer valuable tools for knowledge and care, Leone shows that they also introduce a shift in the semiotic regime: animals are no longer perceived in direct ecological or affective relation, but are instead translated into programmable simulations operating within controlled symbolic frameworks.

In this sense, the zoover does not merely extend access to animal life, it reframes the way animality is signified, often favouring logics of visual legibility, control, and abstraction. The immersive qualities of these environments, along with their gamified and pedagogical interfaces, tend to reconfigure semiosis around user experience and cognitive simplification, potentially displacing the complex, embodied, and relational nature of animal communication. This transformation exemplifies a broader trend: the production of artificial semiotic ecologies, where signs are generated not through interspecies interaction, but through algorithms optimised to simulate plausibility and to satisfy human interpretability.

This distance between sign and life, image and body, data and mind, constitutes one of the deepest epistemological challenges posed by the generative paradigm when applied to animal communication. A zoosemiotic perspective, by contrast, insists on recognising the irreducibility of the sign to its external form, and the necessity of always situating it within the vital, affective, and sensory network that gives it meaning.

3. DECODING WITHOUT INTERPRETATION: THE ENGINEERING PROMISE OF MACHINE LEARNING IN ANIMAL COMMUNICATION

The third contribution examined, "Using Machine Learning to Decode Animal Communication" by Rutz et al. (2023), functions as a methodological manifesto for the application of machine learning (ML) in the study of communicative signalling among non-human species. The article outlines a systematic framework for applying machine learning techniques to key tasks in ethology and bioacoustics, including signal segmentation, classification, unsupervised repertoire discovery, cross-modal association, and functional decoding of communicative signals. The terminology is overtly informed by computer science and cognitive modeling, drawing on concepts such as "self-supervised learning", "foundation models", and "cross-modal associations" (Rutz et al. 2023, 153). In this framework, animal communication is reframed as a problem of computational modeling and pattern recognition, where meaning is inferred from correlations between signals and behavioral or environmental context. This approach tends to reduce animal semiosis to statistically tractable regularities, privileging prediction and classification over interpretation.

The epistemologically crucial move made by the article lies in the idea that communicative meaning can emerge from the automated identification of structural regularities in signals, an assumption that implies, albeit implicitly, a formalist conception of the sign, grounded in a correspondence between syntactic structure and biological function. This position, however, stands in stark contrast to the zoosemiotic tradition, which conceives animal communication as a relational, situated, and interpretive process. In the absence of a living, embodied

with talking birds, who mimic human speech to sustain interaction without necessarily grasping its meaning (Jakobson 1960, 356).

interpreter, engaged in the context of sign production and reception, semiosis cannot be regarded as "decipherable". The message of the animal is not a text to be translated[11] but a sign to be interpreted within its Umwelt; researchers of zoosemiotics are concerned with codes and messages much as linguists are concerned with competence, or language, and performance, or speech (Sebeok 1965); as Kull (2014, 47) stated: zoosemiotics is focused on the animal type of knowing. Rutz et al. (2023), by contrast, present animal signals as encrypted texts awaiting decomposition into their constituent elements through computational tools.

While acknowledging the difficulties associated with interspecies understanding, it is not necessary to adhere to a Wittgensteinian sceptical view, according to which "if a lion could speak, we could not understand him" (Wittgenstein 1953, 225), as this view may underestimate the possibilities for understanding between different species. From a biosemiotic perspective, communication between humans and animals is not based on translating animal signs as if they were signs of human verbal language, but on areas of overlap between different Umwelten. These shared areas (in terms of intensity of similarity, think of the relationship with pets, or between mammals) are shaped by common sensory modalities (such as sight, hearing and smell) and by bodily forms of interaction mediated by the environmental context (and, of course, from the four functional circles that animal species share in their relationship with other species, with themselves and with the environment [Zengiaro 2023]). Through these resonances, mutual attunement becomes possible: we do not completely "translate" animal communication, but rather respond to it within a shared semiotic landscape (or "medium" in von Uexküll [1929, 127] terminology, as a shared ground). In this way, interspecies understanding is not so much about decoding as it is about co-experiencing signs through embodied interaction.

Although algorithmically sophisticated, this approach systematically elides the pragmatic and agentive dimensions of animal communication. The concept of decoding, used repeatedly and without nuance, presupposes the existence of an "internal" meaning within the signal, one that is separable from the concrete communicative event and reconstructible through an external system of statistical inference. This idea is epistemologically weak: the sign, as such, never carries meaning autonomously, but constructs it through its relation with an interpreter, within a context that includes the body, the relational history (both ontogenetic and phylogenetic, in a regime of both endo- and exo-semiosis), objects and entities in the landscapes, and the ecological environment (one only needs to consider that the acidification of the oceans changes the communication channel of whales, and so the communication [Sehgal, Tumar, Schönwälder 2010, 3]). To speak of decoding implies a monadic conception of the sign, as something extractable and universalizable, whereas semiotics insists on the relational, processual, and situated nature of signification.

From this perspective, the notion of pattern discovery (the conceptual core of the paper) becomes ambiguous. If we are speaking of acoustic or visual patterns, the risk is to confuse the form of the signal with its function; if functional patterns are intended, then they cannot be identified merely through formal recurrences, but require an analysis of the pragmatic and relational conditions in which signals are produced. The semiotic approach to animal communication is not limited to a statistical mapping of behaviour (Scott-Phillips, Heintz 2023, 102), but entails the recognition of biological intentionality, interpretive selectivity, and the constitutive role of context (Farina, Villa 2023). Machine learning, on the other hand, constructs latent representations that are vectorial abstractions, disconnected from the body and from animal praxis. The "understanding" achieved by the system is a function of internal similarity within a representational space, not the result of an interspecific dialogic process. The algorithmic radicalisation of this epistemology entails a deeper theoretical problem: the expulsion of the living from the process of knowledge. The interpretive subject is no longer the animal, nor the ethologist, but the computational architecture itself. This gives rise to a semiotic paradox: animal communication is described as a natural process, but analysed as though it were a computational artefact. The result is a form of "dissemiotisation" of the living, in which communicative life is reduced to training data, and symbolic inference is replaced by numerical interpolation. The tension between these two regimes, that of life and that of the algorithm, is profound and far from resolved.

From a zoosemiotic standpoint, artificial intelligence should not be treated as a decoding device aimed at "translating" animal communication into human language. Instead, it ought to be understood as a mediating technology that enables a process of trans-specific interpretation. Such a process must acknowledge the structural opacity of animal signs and their resistance to reduction within human symbolic or linguistic frameworks. As Lotman (1990, 15) argues, genuine communication between different semiotic systems is often characterised not by perfect equivalence, but by a dynamic of partial translatability and irreducible untranslatability. The interplay between these two poles (what can and cannot be translated) defines the creative and interpretive potential of cross-system communication. In this light, AI-based models should not aim to render animal vocalizations into human-readable syntax or semantics, but rather support the emergence of relational understanding grounded in embodied interaction and contextual

11 From my biosemiotic perspective, textual semiotics has enormous limitations when it comes to analysing living beings, although this is debatable according to the author in question. Living beings are not pieces of stone engraved with a code to be deciphered (such as the Rosetta Stone) or a dead language that needs to be translated, but are living semiotic systems, constantly evolving and changing.

co-presence. A truly semiotic paradigm would shift focus from translation to interpretation, treating meaning as a dialogic, situated, and negotiated process, thus restoring to animal communication its ecological complexity and semiotic dignity.

The work of Rutz et al. (2023) represents one of the most articulate formulations of the computational dream of "translate" animal communication through automated signal analysis. Yet it is precisely this ambition, in its attempt to construct a formal universalism, that exposes the deeper limitations of an approach which conflates statistical evidence with semiotic understanding, and pattern extraction with the interpretation of signs. A critical zoosemiotics does not reject the contribution of AI, but rather deconstructs it, restoring centrality to the animal as a communicating subject and to the context as the vital ground of meaning.

3.1 NatureLM: The Ghost in the Animal

A paradigmatic example of this computational turn in the "translation" of animal communication is NatureLM-audio (Robinson et al. 2025), a project developed by the Earth Species Project[12] that introduces a foundational audio-language model (Large Audio-Language Model, LALM) designed to interpret, classify, and describe acoustic signals produced by animals of various species in natural language.

NatureLM-audio presents itself as the first audio-language model trained specifically on a multimodal corpus including animal vocalisations, human speech, music, and environmental sounds. Its stated goal is to create a generalisable and flexible system capable of species recognition, vocalisation classification, behavioural description, individual counting, and even life stage estimation, all from acoustic input alone, and in many cases through zero-shot inference, meaning the species or specific task has not been encountered during training. The audio input, such as a recording of animal vocalisations, is processed using BEATs, a pre-trained encoder capable of extracting structured representations of the audio signal. These are then aggregated by a Q-Former, an attention-based network that dynamically selects relevant features for language generation. Finally, a language model (LLaMA 3.1, fine-tuned via Low-Rank Adaptation) produces a natural language response to the prompt (e.g., "How many individuals are vocalising in this audio?", "Which species is it?"), called the "cocktail party problem".

The model was trained in two stages: an initial perception pretraining phase for vocal recognition across thousands of animal species, followed by a generalised fine-tuning phase across heterogeneous tasks (detection, captioning, life stage prediction, call-type classification, individual counting), which incorporated human-domain data (music, speech) to enable domain transfer. The methodological innovation is twofold: on one hand, the integration of unlabelled data via curriculum learning; on the other, the introduction of a new benchmark, BEANS-Zero, for evaluating zero-shot performance on unseen species and tasks.

NatureLM-audio is intended as a generic and scalable platform for analysing animal sounds, offering competitive performance compared to specialised models (e.g., BirdNET, Perch) but with greater taxonomic and functional flexibility. The model can classify unseen species (via zero-shot learning), using scientific or taxonomic names; recognise behavioural patterns (e.g., call-types) in heterogeneous recordings; and generate textual descriptions of bioacoustic events in a captioning style, a task thus far largely unexplored in zoological contexts. It achieves state-of-the-art results on multiple datasets (e.g., CBI, ENABirds, RFCx), with particularly strong performance in interspecies generalisation and narrative signal description. Unlike earlier models (BirdNET, Perch), which focused on supervised classification of bird vocalisations, NatureLM adopts a multitask and multimodal approach, capable of handling novel tasks not explicitly seen during training. This ability (known as zero-shot generalisation) derives from the logic of foundation models: systems trained on large semantic spaces to develop abstract capacities generalisable to new domains.

The stated potential of NatureLM-audio operates on multiple levels:

1. Biodiversity conservation and monitoring: automating the recognition of rare or endangered species from large-scale bioacoustic datasets collected via environmental sensors (passive acoustic monitoring).
2. Automated annotation of bioacoustic corpora: the model's ability to generate textual descriptions ("captioning") makes it possible to render vast sound archives accessible, indexable, and searchable, collections that would otherwise remain largely inaccessible to research.
3. Study of animal behaviour: new capabilities such as life stage estimation, call-type classification, and even the counting of vocalising individuals promises to transform acoustic ethology into a high-resolution computational science.

However, it is precisely in this translational and classificatory ambition that the most significant semiotic limitations emerge. This vision, although technically advanced, lacks two fundamental dimensions of animal semiosis:

I. The absence of an interpretive subject: NatureLM does not account for a subjective Umwelt, an embodied perspective, or a pragmatic situation. Vocalisation is treated as an object of classification, not as a relational expression.
II. The disembodiment of the sign: Animal semiosis is not merely acoustic information, but a situated performance embedded in ecological and social context. The caption

12 https://earthspecies.github.io/naturelm-audio-demo/ (visited 10/04/2025)

generated by the model is an anthropocentric projection, not an eco-relational interpretation of the sign (cf. Maran 2020).

In this light, NatureLM-audio appears as a system of representation rather than understanding. It does not "translate" animal communication, but replaces it with human linguistic equivalents. Rather than constructing an interspecific semiosis, it produces a simulation of sign behaviour. The ambition to build a "foundation model" for bioacoustics thus risks taking the form of a new kind of computational glossocentrism: an attempt to subsume the semiotic alterity of animals within the interpretive grids of human language and LLM architectures. A genuinely zoosemiotic approach should instead begin with the multiplicity of sign-forms and their irreducibility to a common code.

This critique, however, does not imply the wholesale rejection of technological mediation in the study of animal semiosis. On the contrary, as Pauline Delahaye (2021, 289) convincingly argues, computational tools can play a pivotal role in expanding human access to modalities of communication that would otherwise remain inaccessible. Machines, in her account, act not as interpreters but as epistemic prostheses, new eyes and ears that extend our perceptual field into domains such as ultrasonic vocalisations, infrasonic pulses, or rapid micro-movements, all of which escape human sensory resolution. In this sense, quantitative methods and machine learning architectures are not inherently reductive; they become problematic only when their outputs are mistaken for the semiotic phenomena themselves, rather than as transductions or translations mediated by a technical apparatus (Hernández-Fernández 2021, 378). Yet it is precisely at this junction that the divergence between instrumentalist and semiotic perspectives becomes clear. While Delahaye acknowledges the interpretive gap between signal and meaning, her approach tends to bracket the ontological shift that occurs when animal communication is parsed into computable features and classified according to human-designed taxonomies. A model like NatureLM-audio does not simply extend perception, it reformats semiosis through the constraints of foundation model architectures, which operate according to statistical regularities and generative probabilities rather than ecological embodiment or relational intention. The sign, in such systems, becomes a data point within an abstract latent space, stripped of the Umwelt that gives it sense within a living system.

4. RETHINKING THE "GENERATIVE" IN GENERATIVE ZOOSEMIOTICS

In contemporary artificial intelligence discourse, the term "generative" has acquired a specific and dominant meaning: the ability of a computational system to produce new content (texts, images, code, or sounds) based on probabilistic extrapolation from existing datasets. This form of generation is deeply rooted in a statistical logic of prediction, optimization, and reproduction. It presupposes that meaning can be reduced to the manipulation of formal features, abstracted from any bodily, ecological, or social context. In this view, the generative model is a machine of pattern recognition and recombination, operating within an infrastructure of quantifiable and computationally tractable signs.

This understanding, however, is at odds with the epistemological and ontological stakes of zoosemiotics. If, as Sebeok (2001, 14) and later Martinelli (2010, 137–38) argue, language must be seen not merely as a communicative tool but as a modeling system, one that enables the human species to construct and navigate possible worlds, then any attempt to describe nonhuman communication systems must begin with a critical reflection on what it means to "generate" meaning in the first place. The risk, otherwise, is to transpose a human-centric and computationally engineered model of semiosis onto forms of life that operate through radically different modalities of embodiment, perception, and relationality.

The challenge of thinking of a "Generative Zoosemiotics" (GZ) is thus to reclaim the term "generative" from its current techno-linguistic reduction, and to reposition it within a semiotic framework that accounts for life, embodiment, and interspecies relationality. In this reframed perspective, generativity refers not to the synthetic production of symbolic outputs by machines, but to the emergence of meaning through dynamic, context-dependent interactions among living beings, artifacts, and environments. This approach draws upon Sebeok's (& Danesi 2000) distinction between primary and secondary modeling systems. While human language, as a secondary modeling system, allows for the articulation of abstract and displaced meaning (including narrativity, metasemiosis, and distant space-time reference) zoosemiotics insists that animals, too, participate in modeling processes, albeit through different channels. These may include iconic and indexical relations, affective and behavioral cues, and Umwelt-specific semiotic loops that cannot be fully captured by linguistic analogy. And, at the same time, Generative Zoosemiotics borrows the term "generative" from Greimassian (1966) semiotics, where meaning is analyzed through a structured generative path. In this framework, analysis proceeds through distinct levels: a deep level of abstract semantic structures, a surface narrative level where oppositions are organized into actants and narrative programs, and finally the discursive level, where meaning is manifested in concrete texts through figurative and thematic strategies. By analogy, in Generative Zoosemiotics the study of animal communication adopts this layered approach: the zoosemiotic dimension is analyzed at the deep and surface levels, capturing fundamental oppositions (e.g. danger/safety, attraction/repulsion) and their organization into behavioral sequences. Meanwhile, the contribution of Artificial Intelligence corresponds to the discursive level, where these structures are modeled, simulated, and manifested through computational systems. In this sense, AI provides the "expression" plane that makes explicit, in discursive and technical form,

the semiotic structures identified at the biological and behavioral levels.

The notion of generation, then, must be reconceived not as artificial fabrication, but as "transductive emergence" or *semiotic transduction* (cf. Zengiaro 2025, 78–79). Drawing from Simondon's (2020) concept, semiotic transduction refers to the process by which meaning arises through the transformation and resonance of sign processes across heterogeneous domains, such as between organic and technological, or between human and nonhuman systems, without requiring their full commensurability. In this view, it is not the machine that generates meaning, but the relationship between human observer, technological interface, and animal sign that co-produces new interpretive possibilities. This process is inherently hybrid, situated, and partial. It resists the universalizing ambition of computational semiotics and foregrounds the structural opacity of animal signs, a point also emphasized by Lotman (1990), who argued that translation across codes (verbal to visual, human to nonhuman) is always bounded by zones of mutual untranslatability.

Within this paradigm, artificial intelligence is not conceived as a transparent window into animal communication, but as a mediating apparatus, one that actively shapes the conditions under which certain signals are rendered legible while others are silenced. Models that claim to "translate" whale codas or primate gestures into human language risk reifying a form of epistemic extractivism, wherein animal expressivity is reformatted according to human cognitive and syntactic norms. Instead, a GZ calls for a shift from translation to relation, from decoding to dialogue. The generative act lies not in producing symbolic replicas of animal signs, but in co-constructing new imaginaries, models, and narratives that enrich our understanding of interspecies semiosis. This includes reimagining whales not as linguistic subjects in a human sense, but as semiotic agents embedded in acoustic ecologies; it includes acknowledging that what is perceived as "noise" in a dataset may in fact be the trace of a living body's resonance. This mode of generation is not neutral: it is an ethical and epistemic stance. It recognizes the limitations of our modeling systems while refusing to retreat into agnosticism. It affirms the possibility of interspecies understanding, not through the projection of human categories onto animal signs, but through the cultivation of attentiveness to difference, similarity, and resonance across Umwelten.

GZ is not about replacing human language with AI-driven reconstructions of animal communication. Rather, it is about generating new ways of thinking about communication itself, ways that are open to alterity, sensitive to material and ecological embeddedness, and reflexive about the mediating infrastructures that shape knowledge production. By integrating insights from Sebeok, Lotman, Martinelli, Maran and contemporary AI critique, this approach offers a path forward: one in which the generative is not artificially imposed but ecologically emergent, not extractive but participatory, not predictive but somehow poetic and generative. It is in this spirit that GZ affirms the generativity of interpretation itself, not as a final decoding of the animal other, but as an ongoing semiotic encounter with life's expressive multiplicity.

CONCLUSION

The analysis of the three case studies – Project CETI, ISPA, and machine learning frameworks for decoding animal communication – has revealed how the application of artificial intelligence to non-human semiosis entails a radical epistemological shift: the transition from the sign as an embodied relation to a computational model of symbolic representation. In all three projects, one observes a tendency to transpose onto animal communication categories drawn from human language (syntax, phonetics, structure, translation), and to reduce signification to statistically extracted patterns from acoustic data. Within this perspective, the sign is treated as a decodable, isolable, and segmentable object, while the animal is reduced to a sound source or an emission agent devoid of interpretive subjectivity.

This algorithmic drift in zoosemiotics exposes a profound theoretical and semiotic limitation. As highlighted in the preceding chapters, it replaces the attentive listening to animal alterity with a regime of symbolic simulation, in which the living being is modelled according to the expectations and cognitive infrastructures of the human and its technical apparatus. In this context, artificial intelligence is not a neutral tool for interpretation, but rather a device of forced translation, operating under a grammatical, linguistified, and phonocentric regime.

However, it is precisely within this tension between symbolic representation and semiotic relation that a new space opens for contemporary zoosemiotics. Following the classification proposed by Dario Martinelli (2007; 2010), the discipline can be divided into two major branches: on the one hand, Ethological Zoosemiotics (EZ) (Martinelli 2010, 65–111), which studies animal semiosis through the models of the natural sciences, from ethology to biosemiotics, and which is itself divided into an early phase (eEZ), more descriptive and behaviourist in orientation, and a modern phase (mEZ), which treats semiosis as its own central object; on the other hand, Anthropological Zoosemiotics (AZ) (Ivi., 121–163), which focuses on the forms of representation and interaction between humans and animals, and which in turn comprises a communicative variant (cAZ), centred on interspecific interaction, and a representational one (srAZ), focused on the cultural, narrative, and symbolic ways in which animals are signified and depicted.

On the basis of this distinction, we propose here the introduction of a third (a "zoosemiotics 3.0"?), emerging field within the discipline: Generative Zoosemiotics (GZ). The use of the term "generative" here does not imply that AI models can generate meaning in a semiotic sense. Rather, it refers to a specific class of machine learning architectures, known as generative AI, which are trained on massive datasets and capable of producing statistically plausible outputs such as text, audio, or symbolic

representations. This usage of generative aligns with its technical meaning in AI literature, but is critically reframed within this paper to signal a disjunction: while these models generate structured outputs, they do so without interpretive agency or Umwelt-situated semiosis. Generative Zoosemiotics thus aims to expose the epistemological tension between pattern recognition and embodied meaning-making. This area arises from the need to critically examine the relations that emerge between animal signs, computational technologies, and forms of algorithmic mediation. GZ defines a transdisciplinary space that brings together (bio)semiotics, bioacoustics, artificial intelligence, philosophy of technology, and generative art, with the aim of exploring how animal semiosis is represented, simulated, transformed, and at times distorted by contemporary algorithmic models. It is also necessary to distinguish between natural language processing (NLP) techniques (often used for classification, clustering, and semantic parsing) and truly generative models such as GPT, T5, or diffusion-based audio generators. While some systems, like ISPA, apply NLP tools analytically to animal vocalizations, others envision generative use: e.g., synthesizing new vocal sequences or attempting predictive outputs. GZ addresses this ambiguity by asking not only what AI does, but how its operations reconfigure the sign systems it engages with.

Unlike classical zoosemiotics (EZ and AZ), GZ questions the forms of technical mediation that arise when the animal meets the machine (and vice versa, that is, how the machine is thought about animal communication). It does not merely critique the application of AI to animal communication, but analyzes the hybrid semiotic configurations that emerge between living beings and automated symbolic systems. GZ can be a critical tool that analyzes how a sperm whale is represented by a sequence of ISPA tokens (an animal-machine-human meta-analysis), how a bird is segmented by an audio transformer (and why precisely according to such dynamics), or how a vocalization is reinterpreted by a large linguistic model as a "message" and how to reintegrate it into an ecology of meaning (intervening in the analysis of the animal-body-vocalization-ecosystem chain).

The idea of giving shape to the GZ approach is not merely a theoretical space of analysis, but a political and epistemic proposal for zoosemiotics in relation to AI. It invites us to construct interpretive practices that acknowledge the distance between the vital worlds of different species and that resist the temptation to subsume animal alterity within models of intelligibility that erase relation. Such interpretive practices must also interrogate the theoretical slippage by which algorithmic outputs, often the result of statistical regularity and token manipulation, are treated as if they possessed semiotic intentionality. What GZ exposes is the risk of a technosemiotic reduction: a model of communication that conflates representation with interpretation, and that replaces dialogical engagement with machinic reproduction. To avoid this, GZ insists on the irreducibility of meaning to computability, and on the situated, relational, and embodied nature of semiosis. Where AI seeks to "translate", Generative Zoosemiotics proposes to interpret the chain of meaning; where machine learning recognises patterns, GZ seeks forms and relations, content and expression; where the algorithm classifies, GZ listens.

It thus presents itself as a new paradigm, attuned to contemporary technologies, yet capable of engaging with them critically and creatively. At a time when life is increasingly rewritten in code, GZ reminds us that not everything that communicates can be codified, and that perhaps, as Sebeok suggested, true understanding begins with the recognition that every living sign always exceeds its representation, in a chain of interpretants that bind to one another in an infinite semiosis of life.

The proposal for a GZ manifesto, then, could be: *Generative Zoosemiotics was not created to explain animals to machines, but to restore the animal's right to be other, in a shared world of signs between nonhuman bodies immersed in a space full of meaning.*

Generative Zoosemiotics does not celebrate the creative capacities of machines, but invites a rethinking of how AI participates in shaping the ways we perceive, simulate, and interpret animal meaning. Rather than reducing communication to computable structure, it foregrounds the relational and embodied dimensions of semiosis, without which no true interpretation is possible. At the same time, it opens a space for new forms of inquiry: where AI becomes not just a tool, but a partner in exploring the boundaries of sense-making across species. In this unfolding dialogue between biological life and artificial mediation, the challenge is not to translate the animal, but to remain responsive to the signs through which life invents its own meanings.

Funding: This work is part of the project "San Francesco e l'IA", PI: Prof. Massimo Leone, which received funding from the CRT Foundation (LEOM_CRT_25_01-110992/2024.1420).

There are no conflicts of interest

REFERENCES

Andreas, J. et al., 2022. Toward understanding the communication in sperm whales. *iScience*, 25(6), 104393, available at: https://doi.org/10.1016/j.isci.2022.104393.

Beguš, G., Dabkowski, M., Sprouse, R. L. et al., 2025. The phonology of sperm whale codas vowels. *bioRxiv*, 1–13, available at: https://doi.org/10.1101/2025.06.09.658556.

Bernardi della Rosa, S., Zengiaro, N., 2025. The Biosemiotic Glossary Project: Habit. *Biosemiotics*, 18, 5–30, available at: https://doi.org/10.1007/s12304-025-09594-1.

Cerrone, M., Mäekivi, N., 2021. A zoosemiotic approach to the transactional model of communication. *Semiotica*, 242, 39–62,

available at: https://doi.org/10.1515/sem-2020-0052.
Cimatti, F., 1998. *Mente e linguaggio negli animali. Introduzione alla zoosemiotica cognitiva*. Roma: Carocci.
Delahaye, P., 2021. Use of Quantitative Measures in Zoosemiotics: How Machines Are Becoming a New Pair of Ears and Eyes for Researchers. *Biosemiotics*, 14, 287–294, available at: https://doi.org/10.1007/s12304-021-09440-0.
Dondero, M. G., 2025. Semiotics of artificial intelligence: enunciative praxis in image analysis and generation. *Semiotica*, 262, 111–146, available at: https://doi.org/10.1515/sem-2024-0195.
Eco, U., 1973. Segno. Milano: ISEDI.
Farina, A., 2014. *Soundscape Ecology. Principles, Patterns, Methods and Applications*. Dordrecht: Springer.
Farina, A., Villa, A., 2023. On the semantics of ecoacoustic codes. *BioSystems*, 232, 105002, available at: https://doi.org/10.1016/j.biosystems.2023.105002.
Fontanille, J., 2003. *Semiotics of discourse*. Berlin: Peter Lang.
Genette, G., 1982. *Palimpseste. La littérature au second degré*. Paris: Seuil.
Gero, S., Whitehead, H., Rendell, L., 2016. Individual, unit, and vocal clan level identity cues in sperm whale codas. *Royal Society Open Science*, 3(11), 150372, available at: https://doi.org/10.1098/rsos.150372.
Goldwasser, S., Gruber, D. F., Tauman Kalai, A., Paradise, O., 2023. A Theory of Unsupervised Translation Motivated by Understanding Animal Communication. *arXiv*, 37th Conference on Neuronal Processing Systems (NeurIPS), 1–35, available at: https://doi.org/10.48550/arXiv.2211.11081.
Greimas, A. J., 1966. *Sémantique structurale: Recherche de méthode*. Paris: Larousse.
Greimas, A. J., Courtés, J., 1979. *Sémiotique. Dictionnaire raisonné de la théorie du langage*. Paris: Hachette.
Hagiwara, M., Miron, M., Liu, Y., 2024. ISPA: Inter-Species Phonetic Alphabet for Transcribing Animal Sounds. In: *IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW)*. Seoul, Korea, pp. 828–832, available at: https://doi.org/10.48550/arXiv.2402.03269.
Héjja-Brichard, Y., Million, K., Renoult, J., Mendelson, T. C., 2023. Using generative artificial intelligence to test hypotheses about animal signal evolution: A case study in an ornamented fish. *HAL Id: hal-04276927*, available at: https://doi.org/10.1101/2023.03.13.532060.
Héjja-Brichard, Y., Million, K., Renoult, J., Mendelson, T. C., 2024. Using neural style transfer to study the evolution of animal signal design: A case study in an ornamented fish. *Ecological Informatics*, 84, 102881, 1–11, available at: https://doi.org/10.1016/j.ecoinf.2024.102881.
Helbo, A., 1997. *L'adaptation. Du théâtre au cinéma*. Paris: Colin.
Hernández-Fernández, A., 2021. Qualitative and Quantitative Examples of Natural and Artificial Phenomena. *Biosemiotics*, 14, 377–390, available at: https://doi.org/10.1007/s12304-021-09423-1.
Hjelmslev, L., 1961. *Prolegomena to a theory of language*. Madison: University of Wisconsin Press.
Jakobson, R., 1959. On linguistic aspects of translation. In Brower, R. (Ed.), *On translation*. Cambridge, MA: Harvard University Press, pp. 232–239.
Jakobson, R., 1960. Closing Statement: Linguistics and Poetics. In Sebeok, T. A. (Ed.), *Style in Language*. Massachusetts: MIT Press, pp. 350–77.
Kull, K., 2010. Umwelt and modeling. In Cobley, P. (Ed.), *The Routledge companion to semiotics*. London, NY: Routledge, pp. 43–56.
Kull, K., 2014. Zoosemiotics is the study of animal forms of knowing. *Semiotica*, 198, 47–60, available at: https://doi.org/10.1515/sem-2013-0101.
Kull, K., Torop, P., 2003. Biotranslation: Translation between umwelten. In Petrilli, S. (Ed.), *Translation Translation*. Amsterdam: Rodopi, pp. 313–328.
Leone, M., 2025. Sémiotique de l'agriculture numérique. In Chatenet, L., Thierry Giuliana, G. (Eds.), *Semiovers. Pour une sémiotique des mondes virtuels et numériques*. Roma: Aracne, pp. 25–46.
Lewis, A., 2021. A Semiotic Modern Synthesis: Conducting Quantitative Studies in Zoosemiotics and Interpreting Existing Ethological Studies through a Semiotic Framework. *Biosemiotics*, 14, 295–327, available at: https://doi.org/10.1007/s12304-021-09435-x.
Lotman, Y., 1990. *Universe of the Mind: A Semiotic Theory of Culture*. Bloomington: Indiana University Press.
Maran, T., 2014. Place and sign. Locality as a foundational concept for ecosemiotics. In Siewer, K. A. (Ed.), *Re-Imagining Nature: Environmental Humanities and Ecosemiotics*. Lewinsburg: Bucknell University Press, pp. 79–89.
Maran, T., 2020. *Ecosemiotics: The Study of Signs in Changing Ecologies*. Cambridge: Cambridge University Press.
Maran, T., Martinelli, D., Turovski, A., 2011. Introduction. In Id. (Eds.), *Readings in Zoosemiotics*, Berlin: Mouton De Gruyter, pp. 51–61.
Marrone, G., Mangano, D. (Eds.), 2018. *Semiotics of Animals in Culture. Zoosemiotics 2.0*. Cham: Springer.
Martinelli, D., 2007. *Zoosemiotics: Proposals for a Handbook*. Helsinki: International Semiotics Institute.
Martinelli, D., 2010. *A Critical Companion to Zoosemiotics: People, Paths, Ideas*. Dordrecht: Springer.
Martinelli, D., 2011. Let's Get Physical! – On the Zoosemiotics of Corporeality. *Biosemiotics*, 4, 259–279, available at: https://doi.org/10.1007/s12304-010-9098-5.
Metz, C., 1971. *Language et cinema*. Paris: Larousse.
Robinson, J., Miron, M., Hagiwara, M., Pietquin, O., 2025. NatureLM-audio: An Audio-Language Foundation Model for Bioacustic. *arXiv*, Published as a conference paper at ICLR 2025, 1–15, available at: https://doi.org/10.48550/arXiv.2411.07186.

Rutz, C., Bronstein, M., Raskin, A. et al., 2023. Using machine learning to decode animal communication. *Science*, 381(6654), 152–155, available at: https://doi.org/10.1126/science.adg7314.

Scott-Phillips, T., Heintz, C., 2023. Animal Communication in Linguistic and Cognitive Perspective. *Annual Review of Linguistics*, 9, 93–111, available at: https://doi.org/10.1146/annurev-linguistics-030421-061233.

Sebeok, T. A., 1965. Animal Communication. *Science*, 147 (3661), 1006–14, available at: https://doi.org/10.1126/science.147.3661.1006.

Sebeok, T. A., 1972. *Perspective in Zoosemiotics*. Berlin: Mouton de Gruyter.

Sebeok, T. A., 1977. *How Animals Communicate*. Bloomington: Indiana University Press.

Sebeok, T. A. (Ed.), 1990. *Essays in Zoosemiotics*. Toronto: Toronto Semiotic Circle.

Sebeok, T. A., 1994. *Signs: An Introduction to Semiotics*. Toronto: University of Toronto Press.

Sebeok, T. A., 2001. *The Swiss pioneer in nonverbal communication Studies: Heini Hediger* (1908–1992). New York: Legas.

Sebeok, T. A., Danesi, M., 2000. *The Forms of Meaning: Modeling Systems Theory and Semiotic Analysis*. Berlin: De Gruyter Mouton.

Sehgal, A., Tumar, I., Schönwälder, J., 2010. Effects of climate change and anthropogenic ocean acidification on underwater acoustic communications. In: *OCEANS'10 IEEE SYDNEY*. Sydney, NSW, Australia, pp. 1–6.

Simondon, G., 2020. *Individuation in light of notions of form and information*. Minneapolis: University of Minnesota Press.

Stjernfelt, F., 2013. The generality of signs: The actual relevance of anti-psychologism. *Semiotica*, 194, 77–109.

Uexküll von, J., 1929. *Theoretical Biology*. New York: K. Paul, Trench, Trubner & Co.

Wittgenstein, L., 1953. *Philosophical Investigations*. Oxford: Blackwell.

Yovel, Y., Rechavi, O., 2023. AI and the Doctor Dolittle challenge. *Current Biology*, 33(15), 783–787, available at: https://doi.org/10.1016/j.cub.2023.06.063.

Zengiaro, N., 2023. The circulation of meaning: a biosemiotic perspective on the functional circle. *Punctum. International Journal of Semiotics*, 9(2), 209–225.

Zengiaro, N., 2024. En búsqueda de los límites de la semiótica: desde el umbral inferior de Eco hasta la ecosemiótica materialista. *Metatheoria – Revista De Filosofía E Historia De La Ciencia* 14(2), 75–92, available at: https://doi.org/10.48160/18532330me14.365.

Zengiaro, N., 2025. Towards an Ecosystemic Biosemiotics: the semiotic threshold between life and non-life. In Dentico, M. (Ed.), *Semiotics, Life, Ecology and Sustainability. Biosemiotic Critics*. Bogotà: E. Editorial, pp. 67–96.

Zengiaro, N., Jaramillo, T. J., 2025. Mythology and Zoosemiotics: Exploring Snake Narratives in Greek, Aztec, and Amazonian Cultures. *Biosemiotics*, 18, 1–15, available at: https://doi.org/10.1007/s12304-025-09598-x.