

हिन्दी AI पुस्तकालय संस्करण

कृत्रिम बुद्धिमत्ता जानवरों की भाषा का अनुवाद करती है

व्हेल, डॉल्फ़िन, पक्षियों
और मधुमक्खियों की आवाज़
सुनने वाली एआई की कहानी

अधिवक्ता किम
क्यूंग-जिन

कृत्रुमि बुद्धमिर्ता जानवरों की भाषा का अनुवाद करती है

व्हेल, डॉल्फ़िन, पक्षियों और मधुमक्खियों की आवाज़ सुनने वाली एआई की कहानी

अधविक्ता कमि क्यूंग-जनि

यह कहानी बहामास की डॉल्फ़िन के साथ 40 साल बताने वाले शोधकर्ता, स्पर्म व्हेल की क्लिक आवाज़ों में स्वर जैसी बनावट खोजने वाले भाषाविद, और हम्पबैक व्हेल 'ट्वेन' (Twain) के साथ 20 मिनट तक आवाज़ों का आदान-प्रदान करने वाले वैज्ञानिकों से शुरू होती है। AI जानवरों की आवाज़ों को नक्शे की तरह बनाकर नयिम कैसे ढूँढता है, पक्षियों, मधुमक्खियों और बंदरों के संकेतों को समझने वाले नए उपकरण, जानवरों की संवेदी दुनिया का सम्मान करने का नज़रिया, नकली आवाज़ों से जानवरों को धोखा देने का खतरा और उसे रोकने के नयिम, तथा व्हेल को कानूनी व्यक्तित्व देने की चर्चा तक, इन सभी वषियों को बारह अध्यायों में समझाया गया है। बड़ी प्राथमिकी कक्षा के बच्चे भी इसे खुद पढ़ सकें, इसलिए इसे छोटे और आसान वाक्यों में लिखा गया है, हर उप-अध्याय में जांची गई संदर्भ सामग्री जोड़ी गई है, और यह दो AI समीक्षकों द्वारा सत्यापन के बाद सुधारा गया संस्करण है।

सार्वजनिक शोध सामग्री संग्रह

यह किताब AI द्वारा व्यवस्थित और तैयार की गई है। इसमें दुनिया भर की प्रमुख भाषाओं की सामग्री को इकट्ठा और संकलित किया गया है। NotebookLM सवाल पूछने की भाषा पर कोई रोक नहीं लगाता, चाहे सामग्री किसी भी भाषा में लिखी गई हो। पाठक इस किताब को तैयार करने में इस्तेमाल किए गए उसी शोध सामग्री संग्रह से अपनी भाषा में सवाल पूछ सकते हैं। भले ही कुछ सामग्री कोरियाई, अंग्रेज़ी या अन्य भाषाओं में हो, तब भी ऐसा किया जा सकता है। NotebookLM में सेटिंग्स खोलकर आउटपुट भाषा को अपनी मनपसंद भाषा में सेट करने पर, उत्तर और तैयार परिणाम उसी भाषा में मिलते हैं।

इस किताब को तैयार करने में इस्तेमाल किया गया सार्वजनिक NotebookLM संग्रह (AI पृथ्वी के जीवन का अनुवाद करता है):

<https://notebook.google.com/notebook/bb152971-e02c-4008-b56e-4076e1180e53>

इस किताब का मूल कोरियाई संस्करण kimkj.com की AI लाइब्रेरी में उपलब्ध है:

<https://kimkj.com/ai-library/?mod=document&uid=12655>

वर्षिय-सूची

प्रत्येक अध्याय के शीर्षक पर क्लिक करने से आप सीधे संबंधित पाठ पर पहुंच जाते हैं। शुरू से पढ़ने के लिए परिचय को चुनें।

प्रस्तावना: पृथ्वी समुदाय की आवाज़ों को खोलना • 4

अध्याय 1. Wild Dolphin Project और चित्तीदार डॉल्फ़िन की दुनिया • 10

अध्याय 2. स्पर्म व्हेल की ध्वन्यात्मक वर्णमाला और Project CETI • 16

अध्याय 3. हम्पबैक व्हेल ट्वेन से मुलाकात • 22

अध्याय 4. गहरे समुद्र का ऑर्केस्ट्रा • 28

अध्याय 5. धरती की भाषा • 33

अध्याय 6. आकाश की समिफनी और अदृश्य नृत्य • 39

अध्याय 7. जानवरों के दमिग से डीप लर्नगि कैनवास तक • 44

अध्याय 8. जानवरों का परविशी संसार और संकेत-वज्जिज्ञान की सीमाएं • 50

अध्याय 9. धोखे और संरक्षण का दोराहा • 56

अध्याय 10. गैर-मानव व्यक्तित्व और मशीन-सहायता प्राप्त अधिकार • 61

उपसंहार: इंटरस्पीशीज इंटरनेट (Interspecies Internet) और सह-अस्तित्व की पृथ्वी • 66

प्रस्तावना: पृथ्वी समुदाय की आवाजों को खोलना

0.1 हबल टेलीस्कोप द्वारा खोली गई एक नई आँख

दिसंबर 1995 में, वैज्ञानिकों ने एक अंतरिक्ष टेलीस्कोप को आकाश की ओर मोड़ा। उस टेलीस्कोप का नाम हबल स्पेस टेलीस्कोप (Hubble Space Telescope) है। टेलीस्कोप का रुख आकाश के सबसे अंधेरे और खाली दखिने वाले बटु की ओर था। कई वद्वानों को लगा कविहाँ कुछ भी नहीं है। टेलीस्कोप ने कई दनिों तक उस अंधेरे से प्रकाश इकट्ठा किया। तभी एक चौकाने वाला परणाम सामने आया। उस अंधेरे में हज़ारों आकाशगंगाएँ छपी हुई थीं। आकाशगंगा तारों का बना एक वशाल समूह होती है। वह जगह खाली नहीं, बल्कि पूरी तरह भरी हुई थी। इस घटना ने वज्जान को एक बड़ा सबक सखाया। जब नए उपकरण से देखा गया, तो खाली दखिने वाली जगह पर भी बहुत कुछ मौजूद था [1]।

आज AI कुछ ऐसा ही काम कर रही है। इस बार नशाना आकाश नहीं है। यह जीवों द्वारा आपस में साझा की जाने वाली आवाजों की दुनिया है। इस नए उपकरण को डजिटल जैव-ध्वनिकी (digital bioacoustics) कहा जाता है। यह जीवों की आवाजों को रकिॉर्ड करके कंप्यूटर से अध्थयन करने वाला वज्जान है। बहुत पहले पश्चिमी वज्जान जानवरों की आवाजों को कमतर आंकता था। उन्हें केवल स्वाभाविक प्रवृत्त से नकिला शोर समझा जाता था। लेकिन जब AI ने उन आवाजों की जांच की, तो सोच बदल गई। जंगल और समुद्र जानकारी से भरी बातचीत के वशाल मंच थे [2]।

इंसान की आँखें और कान दुनिया के बहुत छोटे से हसिसे को ही महसूस कर पाते हैं। हम आसानी से यह भूल कर बैठते हैं किजिसे हम महसूस नहीं करते, वह मौजूद ही नहीं है। डार्क-आइड जंको (dark-eyed junco) नाम की एक छोटी चड़िया होती है। इस पक्षी की आवाज़ इंसानी कानों को बलिकूल सपाट सुनाई देती है। लेकिन पक्षियों के कानों में समय की समझ इंसानों से तेज़ होती है। उस पक्षी के लिए वह आवाज़ दो परतों में बुनी एक बेहद बारीक धुन होती है। हाथी बहुत भारी और धीमी आवाज़ में दूर तक बातचीत करते हैं। यह आवाज़ 14 से 35 हर्टज़ (Hz) के बीच होती है। इनमें से नचिली श्रेणी की आवाज़ इंसान नहीं सुन सकते। यह आवाज़ कई कलिमीटर दूर तक फैल जाती है। यदपिरस्थितियाँ अनुकूल हों, तो यह और भी दूर तक जाती है। चमगादड़ इतनी ऊँची आवाज़ में आपस में झगड़ते हैं जिसे इंसान नहीं सुन सकते। यहाँ तक कि बिना कान वाले नन्हें मूंगे भी आवाज़ के सहारे जाने की जगह ढूँढते हैं [3]।

इंसान जो आवाज़ सुनता है, उसकी ऊपर और नीचे एक सीमा होती है। बहुत ऊँची आवाज़ जिसे हम नहीं सुन सकते, उसे पराश्रव्य (ultrasonic) कहते हैं। 'अल्ट्रा' का अर्थ पार करना होता है। यह इंसानी कान की ऊपरी सीमा को पार कर चुकी आवाज़ है। बहुत नीची आवाज़ जिसे हम नहीं सुन सकते, उसे अवश्रव्य (infrasonic) कहते हैं। यह इंसानी कान की नचिली सीमा से भी नीचे की आवाज़ है। चमगादड़ की आवाज़ अल्ट्रासोनिक श्रेणी में आती है। हाथी की आवाज़ का नचिला हसिसा इन्फ्रासोनिक में आता है। इंसान लंबे समय तक इन दोनों आवाज़ों से अनजान रहा। रकिॉर्डिंग उपकरण और वशिलेषण तकनीक के बेहतर होने से इन आवाज़ों का एक-एक करके अध्थयन करना संभव हो सका। केवल कान वाले जानवर ही आवाज़ सुनते हैं, ऐसा नहीं है। शोध बताते हैं कि बिना कान वाले पौधे भी आवाज़ पर प्रतिक्रिया देते हैं।

वैज्ञानिक ऐसी आवाज़ों को पानी के अंदर भी पकड़ लेते हैं। पानी में डाले जाने वाले माइक्रोफोन को हाइड्रोफोन (hydrophone) कहा जाता है। इस तरह जुटाई गई आवाज़ों का वशिलेषण डीप लर्नगि (deep learning) करता है। डीप लर्नगि बहुत सारे डेटा से खुद नयिम ढूँढने की एक तकनीक है। AI इस तरह इंसानी इंदरियों के सीमति दायरे को तोड़ रही है। पृथ्वी की प्राचीन बातचीत के भीतर जाने का दरवाज़ा खुल रहा है।

संबंधित संदर्भ सामग्री

- [1] Bakker, K. (2022). The Sounds of Life: How Digital Technology Is Bringing Us Closer to the Worlds of Animals and Plants. Princeton University Press.
- [2] Earth Species Project. (2024-11-11). Introducing NatureLM-audio: An audio-language foundation model for bioacoustics. Earth Species Project. <https://www.earthspecies.org/blog/introducing-naturelm-audio-an-audio-language-foundation-model-for-bioacoustics>
- [3] Poole, J. H., Payne, K., Langbauer, W. R., & Moss, C. J. (1988). The social contexts of some very low frequency calls of African elephants. Behavioral Ecology and Sociobiology, 22(6), 385-392. <https://doi.org/10.1007/BF00294975>

0.2 भाषा को नक्शे में बदलने वाली AI

AI इंसानी भाषा को कैसे संभालती है? AI शब्दों को एक प्रकार के नक्शे में बदल देती है। इस नक्शे को लेटेन्ट स्पेस (Latent Space) कहा जाता है। यह एक अदृश्य नक्शा है जो समान अर्थ वाले शब्दों को पास-पास रखता है। इस नक्शे को तारों से भरे रात के आकाश की तरह समझना आसान है। यहाँ हर एक शब्द एक तारा बन जाता है। मलिते-जुलते अर्थ वाले शब्द पास-पास आ जाते हैं। जनि शब्दों के अर्थ दूर होते हैं, वे एक-दूसरे से दूर रहते हैं [1]।

इस नक्शे की शक्ति केवल दूरी से ही नहीं आती। शब्दों के बीच का संबंध भी दर्शाता है। उदाहरण के लिए, 'राजा' और 'पुरुष' को जोड़ने वाला एक तीर होता है। 'रानी' और 'महिला' को जोड़ने वाला तीर भी बिल्कुल वैसा ही दिखता है। इसलिए यदि 'रानी' में से 'महिला' घटाकर 'पुरुष' जोड़ दिया जाए, तो वह 'राजा' के करीब पहुँच जाता है। कुछ शब्द-नक्शों में ऐसी गणनाएँ काफी सटीक बैठती हैं। 'कुत्ता' शब्द भी अकेला नहीं पड़ा रहता। वह 'इंसान', 'भौकना' और 'रोएँ' जैसे शब्दों से जुड़कर अपनी जगह बनाता है।

कंप्यूटर शब्दों के गहरे अर्थ को इंसानों की तरह महसूस नहीं कर सकता। वह उदासी या खुशी जैसी भावनाओं को भी नहीं समझता। कंप्यूटर का काम सिर्फ संबंधों की गणना करना है। वह देखता है कि शब्द एक-दूसरे से कैसे जुड़े हैं। यह काम खाली जगह भरने वाली एक बड़ी पहेली को हल करने जैसा है। जैसे खाली स्थान में आने वाले शब्दों को एक-एक करके सही बैठाया जाता है, वैसे ही AI शब्दों के स्थान को एक-एक करके जोड़ती जाती है। इस तरह भाषा का एक बारीक ढाँचा तैयार होता है [2]।

आवाज़ को भी इसी तरह नक्शे में बदला जा सकता है। AI आवाज़ को छोटे-छोटे टुकड़ों में काटकर संख्याओं में बदल देती है। वह उन संख्याओं को नक्शे पर तारों की तरह सजा देती है। मलिते-जुलती आवाज़ें पास-पास आ जाती हैं। अलग आवाज़ें एक-दूसरे से दूर हो जाती हैं। चाहे इंसानी बोली हो या जानवरों की आवाज़, उन्हें नक्शे पर रखने का तरीका लगभग एक जैसा ही है। बस, प्रजातियों की सीमाओं को पार करके क्या यह उनके अर्थ से भी जुड़ पाता है, इस पर अभी अध्ययन चल रहा है [3]।

शब्द दुनिया द्वारा बनाई गई परछाई की तरह होते हैं। AI उस परछाई के पीछे-पीछे उल्टी दर्शा में जाती है। इस तरह वह दुनिया की छपि हुई संरचना का एक नक्शा बना देती है। इस नक्शे के होने से ही आगे की बात संभव हो पाती है। क्योंकि यदि इंसानी बोली का नक्शा बनाया जा सकता है, तो जानवरों की आवाज़ों का भी नक्शा बनाया जा सकता है।

संबंधित संदर्भ सामग्री

[1] Mikolov, T., Yih, W., & Zweig, G. (2013). Linguistic regularities in continuous space word representations. Proceedings of NAACL-HLT 2013, 746-751.

[2] Bakker, K. (2022). The Sounds of Life: How Digital Technology Is Bringing Us Closer to the Worlds of Animals and Plants. Princeton University Press.

[3] Earth Species Project. (2026). What we do: Technology. Earth Species Project.
<https://www.earthspecies.org/what-we-do/technology>

0.3 बनी शब्दकोश के अनुवाद और Earth Species Project

2017 के आसपास, मशीन अनुवाद के क्षेत्र में एक बड़ी घटना घटी। तब दो भाषाओं को जोड़ने वाला कोई शब्दकोश मौजूद नहीं था। पुराने समय में रोसेटा स्टोन (Rosetta Stone) की आवश्यकता होती थी। रोसेटा स्टोन एक पुराना शिलालेख है जिस पर एक ही बात कई भाषाओं में उकेरी गई थी। इस शिलालेख की बदौलत प्राचीन मिस्र के अक्षरों को समझा जा सका था। लेकिन AI ने बनी किसी शब्दकोश के दोनो भाषाओं के शब्दों को आपस में जोड़ दिया। उसने यह कमाल अपने आप सीखने वाली मशीन लर्निंग के जरिये कर दिखाया [1]।

तरीका यह था कि दोनो भाषाओं के नक्शों को एक-दूसरे के ऊपर रखकर देखा जाए। AI ने अंग्रेज़ी का तारों वाला नक्शा अलग बनाया। उसने स्पैनिश का तारों वाला नक्शा भी अलग से बनाया। इसके बाद उसने दोनो नक्शों को बड़ी सावधानी से एक-दूसरे के ऊपर बठाया। कई वद्वानों को लगा था कि दोनो नक्शे आपस में मेल नहीं खाएंगे। ऐसा इसलिए क्योंकि हर भाषा का दुनिया को देखने का नज़रिया अलग होता है। लेकिन दोनो नक्शे बहुत सहजता से आपस में मेल गए। 'कुत्ते' का मतलब बताने वाला तारा लगभग एक ही जगह पर आकर बैठ गया। अलग आकार वाले दो नक्शों को खींचकर और सकोड़कर आपस में मलाने के इस तरीके को मैनफोल्ड अलाइनमेंट (manifold alignment) कहा जाता है।

यह बात केवल यूरोपीय भाषाओं में ही नहीं देखी गई। बिल्कुल अलग व्याकरण वाली भाषाओं में भी लगभग ऐसा ही रूप दिखाई दिया। इंसान द्वारा बनाई गई भाषा एस्पेरान्तो (Esperanto) में भी ऐसा ही हुआ। इंसानों की अलग-अलग भाषाएँ एक बड़े साझा ढाँचे से जुड़ी हुई लग रही थीं। ऐसा इसलिए माना जाता है क्योंकि इंसान नाम की एक ही प्रजाति एक जैसी पृथ्वी का अनुभव करती है [2]।

Earth Species Project के अज़ा रास्कनि (Aza Raskin) ने यहाँ एक सवाल उठाया। अगर इंसानी भाषाओं को इस तरह एक-दूसरे पर बठाया जा सकता है, तो क्या जानवरों की आवाज़ों के लिए भी यही तरीका इस्तेमाल नहीं किया जा सकता? यह संस्था 2017 में स्थापित एक गैर-लाभकारी शोध संस्थान है। यह सफेद व्हेल बेलुगा, कलिर व्हेल और ओरंगुटान की आवाज़ें जुटाती है। यह हर आवाज़ के लिए अलग से तारों वाला नक्शा बनाती है। फरि उस नक्शे को इंसानी भाषा के नक्शे पर सावधानी से बठाती है। इससे इंसान अपने आस-पास के साथी जीवों की आवाज़ को एक नई दृष्टि से समझ सकते हैं। उनका हालिया उपकरण NatureLM-audio है। यह जानवरों की आवाज़ सुनकर उनकी प्रजाति और उनकी पुकार का प्रकार बताने वाली AI है। 2025 में इसका एक डेमो सबके इस्तेमाल के लिए खोला गया। आवाज़ की फाइल अपलोड करके सवाल पूछने पर उत्तर मलि जाता है। 2026 में इसकी क्षमताएँ और बेहतर हो गईं। जुलाई 2026 में alp-data नाम का एक डेटा भंडार भी शुरू हुआ, जहाँ कई शोध संस्थान मलिकर डेटा साझा करते हैं। यह बदलाव इस वचिार से जुड़ता है कि अलग-अलग AI अंत में लगभग एक जैसे आकार के नक्शे तक पहुँचती है। इस सोच को प्लेटोनिक रपिरेजेंटेशन हाइपोथिसिस (Platonic Representation Hypothesis) कहा जाता है [3][4][5]।

संबंधित संदर्भ सामग्री

- [1] Lample, G., Conneau, A., Denoyer, L., & Ranzato, M. (2018). Unsupervised machine translation using monolingual corpora only. International Conference on Learning Representations (ICLR) 2018. arXiv:1711.00043.
- [2] Robinson, D., Miron, M., Hagiwara, M., et al. (2024/2025). NatureLM-audio: An audio-language foundation model for bioacoustics. International Conference on Learning Representations (ICLR) 2025. arXiv:2411.07186.
- [3] Earth Species Project. (2025-09). Hands-on with NatureLM-audio: A no-code demo. Earth Species Project. <https://earthspecies.org/blog/hands-on-with-naturelm-audio-a-no-code-demo>
- [4] Earth Species Project. (2026-07-22). Introducing alp-data: A shared data layer for animal language processing. Earth Species Project. <https://earthspecies.org/2026/07/22/introducing-alp-data-a-shared-data-layer-for-animal-language-processing/>
- [5] Huh, M., Cheung, B., Wang, T., & Isola, P. (2024). Position: The Platonic Representation Hypothesis. Proceedings of the 41st International Conference on Machine Learning (ICML). arXiv:2405.07987.

0.4 आवाज़ में छपि नियमों को खोजना

क्या जानवरों की आवाज़ें बनिा किसी मतलब का शोर हैं? या फरि इंसानी बोली की तरह उनमें भी नियम होते हैं? वैज्ञानिक आवाज़ों के आँकड़े मापकर इसका पता लगाते हैं। एक जाना-माना नियम यह है कि जसि शब्द का जतिना ज़्यादा इस्तेमाल होता है, वह उतना ही छोटा हो जाता है। इस नियम को जपि का संक्षेपण नियम (Zipf's law of abbreviation) कहते हैं। इसका नाम भाषाविद जॉर्ज किंग्सले जपि (George Kingsley Zipf) के नाम पर रखा गया है। इंसान अपनी ऊर्जा बचाने के लिए बार-बार बोले जाने वाले शब्दों को छोटा कर देते हैं [1]।

भाषाविद स्टुअर्ट सेम्पल (Stuart Semple) और रामोन फेरर-आई-कांचो (Ramon Ferrer-i-Cancho) के शोध दल ने कई जानवरों में इस नियम की जांच की। उन्होंने बंदरों सहित कई स्तनधारी जीवों की आवाज़ें इकट्ठी की और उनकी तुलना की। कंप्यूटर ने आवाज़ों को छोटे-छोटे हिस्सों में बाँटकर उनकी गनिती की। परिणाम एक जैसा निकला। जो आवाज़ जतिनी बार नकिली जाती थी, उसकी लंबाई उतनी ही छोटी थी। इंसानी बोली से मलिता-जुलता नियम जानवरों की आवाज़ों में भी मौजूद था। एक और नियम भी सामने आया कि आवाज़ जतिनी लंबी होती है, उसके भीतर के टुकड़े उतने ही छोटे होते हैं। इस नियम का नाम मेज़ेरथ-अल्टमैन नियम (Menzerath-Altmann law) है।

यह मापने का भी एक मान होता है कि अगली आवाज़ का अंदाज़ा लगाना कतिना कठिन है। यह इस बात को मापता है कि किसी एक आवाज़ के बाद कौन सी आवाज़ आएगी, इसका अनुमान लगाना कतिना मुश्किल है। इस मान को शैनन एन्ट्रॉपी (Shannon entropy) कहा जाता है। इसका नाम सूचना सिद्धांत की नींव रखने वाले क्लॉड शैनन (Claude Shannon) के नाम पर पड़ा है। यदि कोई आवाज़ बनिा किसी मतलब की चीख भर हो, तो यह मान एक तरफ झुक जाता है। लेकिन अगर उसमें नियमों वाली भाषा हो, तो उसका मान इनके बीच में रहता है [4]।

Whale-SETI की रसिर्च टीम ने इस मान की मदद से व्हेल की आवाज़ों की जांच की। इसमें ब्रेडा मैककोवन (Brenda McCowan), फ्रेड शार्प (Fred Sharpe) और लॉरेंस डॉयल (Laurance Doyle) साथ मलिकर काम कर रहे थे। उन्होंने हम्पबैक व्हेल (humpback whale) के गीतों का अध्ययन किया। व्हेल का गाना नियमों और आज़ादी के बीच बहुत खूबसूरती से संतुलित था। गीतों में यह बनावट भी देखी गई कि अक्सर आने वाली आवाज़ें छोटी थीं, जबकि कभी-कभार आने वाली आवाज़ें लंबी थीं। ऐसा करने से कम ऊर्जा खर्च करके वही संदेश भेजा जा सकता है। कंप्यूटर में भी साइज़ कम करने के लिए बार-बार इस्तेमाल होने वाली चीज़ों को छोटा और कभी-कभार आने वाली चीज़ों को लंबा लिखकर संपीड़ित (कंप्रेस) करने का एक तरीका होता है। इस तरीके का नाम हफ़मैन कोडिंग (Huffman coding) है [2]।

शोध दल ने एक जीवित व्हेल के साथ आवाज़ों का आदान-प्रदान करने का भी प्रयोग किया। वे अलास्का के समुद्र में ट्वेन (Twain) नाम की एक हम्पबैक व्हेल से मल्लि। शोध दल ने पहले से रिकॉर्ड की गई व्हेल के अभिवादन की आवाज़ पानी के नीचे लगे स्पीकर से बजाई। ऐसा नहीं

था कि AI ने तुरंत वास्तविक समय में बातचीत गढ़ी हो। इस अभिवादन की आवाज़ को whup call कहा जाता है। व्हेल ने इंसानों की तरह अपनी बारी का पालन करते हुए जवाब दिया। उसने आवाज़ खत्म होने का इंतज़ार किया और खाली जगह को भरा। यह बातचीत 20 मिनट तक चलती रही। यह बिल्कुल वैसा ही था जैसे लंच के समय दोस्त बारी-बारी से एक-दूसरे से बातें करते हैं [3]।

संबंधित संदर्भ सामग्री

- [1] Semple, S., Ferrer-i-Cancho, R., & Gustison, M. L. (2022). Linguistic laws in biology. *Trends in Ecology & Evolution*, 37(1), 53-66. <https://doi.org/10.1016/j.tree.2021.08.012>
- [2] McCowan, B., Hubbard, J., Walker, L., Sharpe, F., Frediani, J., & Doyle, L. (2023). Interactive bioacoustic playback as a tool for detecting and exploring nonhuman intelligence: "conversing" with an Alaskan humpback whale. *PeerJ*, 11, e16349. <https://doi.org/10.7717/peerj.16349>
- [3] Earth.com. (2024). AI helps humans have a 20-minute 'conversation' with a humpback whale named Twain. Earth.com. <https://www.earth.com/animals/whale-seti-ai-helps-humans-have-20-minute-conversation-with-humpback-whale-twain/>
- [4] Bakker, K. (2022). *The Sounds of Life: How Digital Technology Is Bringing Us Closer to the Worlds of Animals and Plants*. Princeton University Press.

0.5 स्पर्म् व्हेल के क्लक्स में मल्लि स्वर

यह तरीका जसि जगह सबसे ज़्यादा सफल साबित हुआ, वह गहरा समुद्र था। उसकी मुख्य पात्र स्पर्म् व्हेल है। स्पर्म् व्हेल के शरीर की लंबाई 20 मीटर तक होती है। उसके दमिग का वज़न लगभग 9 किलोग्राम होता है। यह इंसानी दमिग के वज़न से 6 गुना से भी ज़्यादा है। यह व्हेल बनी रोशनी वाले गहरे समुद्र में शिकार करती है। इसलिए वह आँखों के बजाय आवाज़ से दुनिया को देखती है। वह छोटी-छोटी क्लक्स जैसी आवाज़ें छोड़कर अपने आस-पास का पता लगाती है। क्लक्स की आवाज़ों के इस समूह को coda कहा जाता है [1]।

लंबे समय तक वदिवान coda को बनी किसी मतलब का संकेत मानते रहे। वे सोचते थे कि यह बस कुछ तय की गई चेतावनी भरी आवाज़ें हैं। 2020 में Project CETI ने इस सोच को पूरी तरह बदल दिया। यह कई विश्वविद्यालयों के AI वैज्ञानिकों और व्हेल विशेषज्ञों का एक समूह है। उन्होंने डोमनिका के समुद्र से रिकॉर्डिंग जुटाई। 10 से अधिक वर्षों में जमा की गई coda की संख्या 8,700 से भी ज़्यादा थी। ठीक-ठीक कहे तो वे 8,719 थीं। AI ने इन आवाज़ों का विश्लेषण किया। उसने पहले से ज्ञात 21 प्रकारों से कहीं ज़्यादा नए संयोजनों की खोज की। उसने यह भी पाया कि व्हेल आवाज़ की गति और उनके बीच के अंतराल को बदलती है। वे कभी-कभी आवाज़ के अंत में वाक्य के पूर्णवर्णम की तरह एक क्लक्स ध्वनि भी जोड़ देती थीं। वे coda की लंबाई में भी बहुत थोड़ा-थोड़ा बदलाव करती थीं। लय के इस बारीक बदलाव को रुबाटो (rubato) कहा जाता है। यह ठीक वैसा ही है जैसे कोई गायक गाते समय लय को थोड़ा धीमा या तेज़ कर देता है। यह आवाज़ के छोटे-छोटे टुकड़ों को परीकर कई अर्थ तैयार करने का तरीका है। इंसानी भाषा में देखने वाले इस तरीके को जानवरों में भी साफ़ तौर पर देखा गया। स्पर्म् व्हेल सुनने वाले की स्थिति के अनुसार coda का अलग तरह से इस्तेमाल भी करती थीं। सामने वाले के हिसाब से अपनी बात चुनने की इस क्षमता को ऑडियंस डिज़ाइन (audience design) कहा जाता है।

2025 में, एक नए अध्ययन ने एक और बड़ी खोज की जानकारी दी। यह अमेरिका के बर्कले विश्वविद्यालय और Project CETI का एक संयुक्त अध्ययन था। इसका नेतृत्व प्रोफेसर गास्पर बेगुश (Gašper Beguš) ने किया था। शोधकर्ताओं ने व्हेल की क्लक्स आवाज़ों के बीच के खाली समय को घटाया और उन्हें सामान्य से कई गुना तेज़ गति से सुना। ऐसा करने पर अलग-अलग सुनाई देने वाली आवाज़ें आपस में जुड़ गईं। वह आवाज़ इंसानों के स्वरों जैसी सुनाई दी। शोध दल ने 'आ' और 'ई' जैसी ध्वनियों में अंतर पहचान लिया। उन्होंने एक स्वर से दूसरे स्वर में बदलने वाले संयुक्त स्वरों (diphthongs) को भी खोज निकाला। शोधकर्ताओं ने AI के अंदर देखने वाले उपकरणों का भी उपयोग किया। उन्होंने देखा कि व्हेल की आवाज़ का विश्लेषण करते समय AI किस हिसाब से परतकिरिया देता है। इसके नतीजे से पता चला कि ध्वनि की गूँज के शिखर बहुत महत्वपूर्ण हैं। गूँज के इस शिखर को फॉर्मेट (formant) कहा जाता है। यह मुंह या सरि के गूँजने वाले हिस्से के आकार के अनुसार उभरने वाला ध्वनिकी शिखर है। इंसान भी स्वरों को पहचानते समय इसी मान का उपयोग करते हैं। 2026 में, इस पर आधारित एक अगला अध्ययन सामने आया। यह अध्ययन व्हेल की आवाज़ की संरचना की तुलना इंसानी ध्वनविज्ञान से करके और गहराई से जांच करता है [2][5][3]।

मार्च 2026 में, Project CETI ने एक और अध्ययन की घोषणा की। यह डोमनिका के समुद्र में स्पर्म् व्हेल (sperm whale) के बच्चे के जन्म को दर्ज करने वाला अध्ययन था। इसमें छह घंटे से अधिक की आवाज़ें और आसमान से ली गई वीडियो शामिल थे। समूह की दूसरी व्हेलों ने बच्चे के जन्म में मदद की। प्राइमेट्स (primates) के अलावा किसी अन्य जानवर में इस तरह के सहयोग को संख्याओं के ज़रिये साबित करने वाला यह पहला अध्ययन है। यह अध्ययन फरि से दिखाता है कि स्पर्म् व्हेल का समाज आवाज़ों के ज़रिये आपस में कितनी मज़बूती से जुड़ा हुआ है [4]।

स्पर्म व्हेल के पास इंसानों जैसा वोकल कॉर्ड (vocal cords) नहीं होता। इसके बजाय, वे अपने सरि के सामने के बड़े अंग और नाक के अंदर के ऊतकों का उपयोग करती हैं। वे इन्हें कंपति करके ध्वनिकी गूंज बनाती हैं। इंसानों से बिल्कुल अलग रास्ते पर वकिसति हुए एक जीव ने अपने आप इंसानों जैसे स्वरों का निर्माण किया है।

संबंधित संदर्भ सामग्री

- [1] Sharma, P., Gero, S., Rus, D., Goldwasser, S., & Andreas, J. (2024). Contextual and combinatorial structure in sperm whale vocalisations. *Nature Communications*, 15, 3617. <https://doi.org/10.1038/s41467-024-47221-8>
- [2] Beguš, G., Sprouse, R. L., Leban, A., Silva, M., & Gero, S. (2025). Vowel- and diphthong-like spectral patterns in sperm whale codas. *Open Mind*, 9. <https://doi.org/10.1162/OPMI.a.252>
- [3] Beguš, G., Sprouse, R. L., Leban, A., & Gero, S. (2026). The phonology of sperm whale coda vowels. *Proceedings of the Royal Society B*, 293. <https://doi.org/10.1098/rspb.2025.2994>
- [4] Aluma, A., Gero, S., et al. (2026). Description of a collaborative sperm whale birth and shifts in coda vocal styles during key events. *Scientific Reports*, 16, 9206. <https://doi.org/10.1038/s41598-025-27438-3>
- [5] University of California. (2025-11-13). UC Berkeley and Project CETI study shows sperm whales communicate in ways similar to humans. University of California. <https://www.universityofcalifornia.edu/news/uc-berkeley-and-project-ceti-study-shows-sperm-whales-communicate-ways-similar-humans>

0.6 नयित्रण नहीं, ध्यान से सुनना

AI अनुसंधान में एक पुराना सबक है। यह कंप्यूटर वैज्ञानिक रिचर्ड सटन (Richard Sutton) द्वारा बताया गया 'कड़वा सबक' (The Bitter Lesson) है। इसका मतलब है कि चितुर डज़ाइन के मुकाबले ज़्यादा डेटा और बड़ी गणना शक्त की जीत होती है। जानवरों की आवाज़ों को समझने वाले उपकरण भी इसी तरह वकिसति हुए हैं। उनमें से एक DolphinGemma है। इसे Google, जॉर्जिया टेक (Georgia Tech) और Wild Dolphin Project ने मलिकर बनाया है। Wild Dolphin Project 40 से अधिक वर्षों से चल रहा डॉल्फ़िन अध्ययन है। DolphinGemma आवाज़ों को समझने वाली लगभग 40 करोड़ आंतरिक संख्याओं वाला एक छोटा AI है। यह आकार में इतना छोटा है कि हाथ में रखे जाने वाले Pixel फ़ोन पर भी काम करता है। DolphinGemma ध्वनिकी साउंडस्ट्रीम (SoundStream) नामक तरीके से छोटे-छोटे टुकड़ों में बांटता है। वह उन टुकड़ों को देखकर अगली आवाज़ का पहले से अनुमान लगाता है। यह तरीका उसी AI जैसा है जो इंसानी भाषा में अगले शब्द का अनुमान लगाकर वाक्य पूरा करता है। इस उपकरण का अभी भी समुद्र में डॉल्फ़िन के पीछे जाकर परीक्षण किया जा रहा है। Google इस उपकरण को सभी के उपयोग के लिए खुला रखना चाहता है [3][1]।

लेकिन इस तेज़ी से बढ़ती तकनीक का एक डरावना पहलू भी है। एज़ा रास्कन (Aza Raskin) बार-बार इस बात पर ज़ोर देते हैं। AI अर्थ समझने से पहले ही आवाज़ की नकल कर सकता है। वह व्हेल की आवाज़ की बिल्कुल वैसी ही नकल बनाकर बातचीत को आगे बढ़ा सकता है। इसका मतलब है कि बिना अर्थ समझे समुद्र में नकली आवाज़ें उड़ेली जा रही हैं। स्पर्म व्हेल आवाज़ और व्यवहार के ज़रिये अपनी संस्कृति साझा करती हैं। अतीत में जब व्हेलों का शिकार करने वाले जहाज़ आए, तो स्पर्म व्हेलों ने खतरे से बचने के तरीके एक-दूसरे से सीखे। जहाज़ों के आने के सिर्फ़ ढाई साल के भीतर पकड़ी जाने वाली व्हेलों की संख्या में 58 प्रतिशत की कमी आई। इसे व्हेलों द्वारा एक-दूसरे को खतरे की सूचना देने का नतीजा माना जाता है। उनका संवाद हमारे सोचने से कहीं अधिक बड़े दायरे में फैला हुआ है। हंपबैक व्हेल के गीत भी समुद्र के पार फैलने वाली एक संस्कृति हैं। इंसानों द्वारा बनाई गई नकली आवाज़ें इस जुड़ाव को बग़ाड़ सकती हैं। अर्थ पूरी तरह समझने से पहले ही आवाज़ें बनाना खतरनाक है [2][4][5]।

इसलिए इस यात्रा की दशा बहुत महत्वपूर्ण है। यह तकनीक जानवरों पर अपनी मर्ज़ी चलाने का औज़ार नहीं बननी चाहिए। यह जानवरों की भीतरी भावनाओं को स्वीकार करने और उन्हें सुनने का प्रयास होना चाहिए। इसे अंतर-प्रजाति समझ (Interspecies Understanding) कहा जाता है। इंसान तब नहीं बदलता जब वह अपनी बातें थोपता है, बल्कि तब बदलता है जब वह अपनी आवाज़ धीमी करके दूसरों को स्वीकार करता है।

प्राचीन काल में लोग मानते थे कि पृथ्वी ही पूरे ब्रह्मांड का केंद्र है। दूरबीन ने साबित किया कि यह सोच गलत थी। अब आवाज़ों का नक्शा हमें एक बार फिर जागरूक कर रहा है। इंसान जीवन की दुनिया का केंद्र नहीं है। वह तो अपनी-अपनी आवाज़ रखने वाले जीवों के एक बड़े घेरे का सिर्फ़ एक हिस्सा है। यह अहसास ही पृथ्वी के समुदाय की आवाज़ को सुनने की दशा में पहला कदम है।

संबंधित संदर्भ सामग्री

- [1] Google. (2025). DolphinGemma: How Google AI is helping decode dolphin communication. Google Blog. <https://blog.google/innovation-and-ai/products/dolphingemma/>

- [2] Robinson, D., Miron, M., Hagiwara, M., et al. (2024/2025). NatureLM-audio: An audio-language foundation model for bioacoustics. International Conference on Learning Representations (ICLR) 2025. arXiv:2411.07186.
- [3] Sutton, R. S. (2019-03-13). The bitter lesson. incompleteideas.net.
<http://www.incompleteideas.net/IncIdeas/BitterLesson.html>
- [4] Whitehead, H., Smith, T. D., & Rendell, L. (2021). Adaptation of sperm whales to open-boat whalers: Rapid social learning on a large scale? Biology Letters, 17(3), 20210030. <https://doi.org/10.1098/rsbl.2021.0030>
- [5] Bakker, K. (2022). The Sounds of Life: How Digital Technology Is Bringing Us Closer to the Worlds of Animals and Plants. Princeton University Press.

अध्याय 1. Wild Dolphin Project और चित्तीदार डॉल्फ़िन की दुनिया

1.1 पन्ना जैसे हरे बहामास में लंबे समय का अवलोकन

समुद्र में ऐसी बुद्धिमत्ता रहती है जिसे इंसान अभी पूरी तरह नहीं समझ पाए हैं। डॉल्फ़िन उनमें से एक है। डॉल्फ़िन के मन को जानने के लिए क्या करना चाहिए? डॉल्फ़िन को पानी के टैंक में बंद करके देखने का एक तरीका है। लेकिन एक व्यक्ति ने दूसरा रास्ता चुना। वह है डॉ. डेनिस हर्ज़िंग (Denise Herzing)। डॉ. हर्ज़िंग सीधे उस समुद्र में गईं जहाँ डॉल्फ़िन रहती हैं [1]।

डॉ. हर्ज़िंग ने एक नयिम बनाया। वह नयिम है: "उनकी दुनिया में, उनके तरीके से।" इसका मतलब है डॉल्फ़िन को इंसानों की मर्ज़ी से काम पर न लगाना। इसका यह भी मतलब है कि डॉल्फ़िन को कोई नुकसान पहुँचाए बिना उनका अवलोकन करना। इस तरीके को गैर-आक्रामक अवलोकन (non-invasive observation) कहा जाता है। गैर-आक्रामक का मतलब है शरीर को सुई चुभाए या पकड़े बिना देखने का तरीका।

डॉल्फ़िन को पानी के टैंक में बंद करने पर उनका असली रूप देखना कठिन होता है। क्योंकि कैद में रहने वाली डॉल्फ़िन के जीने का तरीका बदल जाता है। इसलिए समुद्र में जाकर उनका अवलोकन करने का तरीका बहुत मूल्यवान है।

डॉ. हर्ज़िंग 'Wild Dolphin Project' का नेतृत्व करती हैं। यह शोध लगभग 40 वर्षों से चल रहा है। 40 साल एक प्राथमिक विद्यालय के बच्चे की उम्र से लगभग चार गुना लंबा समय है। शोध का स्थान बहामास का समुद्र है। यहाँ का पानी स्वमिगि पूल के पानी की तरह साफ़ और पारदर्शी है। इसलिए डॉल्फ़िन के हाव-भाव और आवाज़ों को एक साथ रिकॉर्ड करने के लिए यह बहुत अच्छा है [2][3]।

यह शोध 1985 में शुरू हुआ था। पानी के अंदर डॉल्फ़िन का पीछा करते हुए कथित जाने वाला यह दुनिया का सबसे पुराना शोध है। वे एक भी साल छोड़े बिना डॉल्फ़िन के उसी झुंड के पास गए। इसकी बदौलत वे एक डॉल्फ़िन के जन्म लेने, बड़े होने और बच्चे को जन्म देने के पूरे जीवन चक्र को देख सके। इस तरह जमा हुए रिकॉर्ड बाद में AI को सखाने की मजबूत नींव बनते हैं [4]।

इस समुद्र में अटलांटिक चित्तीदार डॉल्फ़िन रहती हैं। इनका वैज्ञानिक नाम स्टेनेला फ्रंटालिस (Stenella frontalis) है। वैज्ञानिक नाम वह आधिकारिक नाम होता है जिसे दुनिया भर के वैज्ञानिक मलिकर इस्तेमाल करते हैं। जहाँ यह शोध किया जाता है, वह उथले रेतीले किनारों वाला समुद्र है। पानी उथला होने के कारण सूरज की रोशनी तलहटी तक पहुँचती है। इसलिए ऊपर से भी डॉल्फ़िन की हरकतों को साफ़-साफ़ देखा जा सकता है।

यह डॉल्फ़िन बॉटलनोज़ डॉल्फ़िन से आकार में छोटी होती है। लेकिन यह बहुत बुद्धिमान जानवर है। शरीर के आकार की तुलना में दमिग कतिना बड़ा है, इसे नापने का एक मान होता है। इस मान को मस्तिष्क-करण गुणांक (Encephalization Quotient) कहा जाता है। यह शरीर की तुलना में मस्तिष्क के आकार की तुलना करने वाला मान है। डॉल्फ़िन प्रजातियों में यह मान काफी ऊँचा होता है। स्तनधारियों में यह इंसानों के बाद सबसे बड़े स्तर में आता है [5]।

शोधकर्ताओं ने डॉल्फ़िन को एक-एक करके पहचानना सीखा। क्योंकि उनके पंखों पर लगे घाव सबके अलग-अलग होते हैं। उनके शरीर पर बनने वाले धब्बों के पैटर्न भी एक-दूसरे से अलग होते हैं। चित्तीदार डॉल्फ़िन जैसे-जैसे बड़ी होती हैं, उनके शरीर पर धब्बे बढ़ते जाते हैं। बचपन में कोई धब्बे नहीं होते, लेकिन उम्र बढ़ने के साथ वे बढ़ते जाते हैं। इसलिए केवल धब्बों को देखकर भी उनकी उम्र का अंदाज़ा लगाया जा सकता है। जैसे हम अपनी कक्षा के दोस्तों के चेहरे एक-एक करके याद करते हैं, वैसे ही उन्होंने डॉल्फ़िन को याद कर लिया। इसकी बदौलत वे दादी, माँ और बच्चे तक जुड़ी तीन पीढ़ियों का अध्ययन कर सके।

लंबे समय तक अवलोकन करने का एक खास कारण है। डॉल्फ़िन की आदतें माँ से उसके बच्चे तक पहुँचती हैं। माँ शिकार कैसे खोजती है, यह बच्चा उसके पास रहकर सीखता है। सीखे गए व्यवहार का माँ से बच्चे तक पहुँचना 'सांस्कृतिक संचरण' कहलाता है। कम समय देखने से इस सलिसलै को नहीं समझा जा सकता। जब किसी एक झुंड को लंबे समय तक देखा जाए, तभी यह साफ़ दिखाई देता है।

डॉल्फ़िन का समाज इंसानी समाज की तरह ही जटिल होता है। नर डॉल्फ़िन जीवन भर के लिए दोस्ती करते हैं और साथ घूमते हैं। मादा डॉल्फ़िन लंबे समय तक अपने बच्चों की देखभाल करती हैं और उन्हें शिष्टाचार सिखाती हैं। इतना जटिल समाज बनाने के लिए आपस में संवाद करने का कोई तरीका होना चाहिए। डॉल्फ़िन की आवाज़ों को समझने की ज़रूरत यही से शुरू होती है।

संबंधित संदर्भ सामग्री

[1] Georgia Tech Research. (2025-04-14). DolphinGemma: How Google AI is helping decode dolphin communication. Georgia Tech Research News.

<https://research.gatech.edu/dolphingemma-how-google-ai-helping-decode-dolphin-communication>

- [2] Herzing, D. L., Pack, A. A., Delfour, F., Starner, T., Mason, C., Gilliland, S., Ramsey, C., & Kohlsdorf, D. (2024). Imitation of computer-generated sounds by wild Atlantic spotted dolphins (*Stenella frontalis*). *Animal Behavior and Cognition*, 11(2), 136–166. <https://doi.org/10.26451/abc.11.02.02.2024>
- [3] Google. (2025-04-14). DolphinGemma: How Google AI is helping decode dolphin communication. The Keyword. <https://blog.google/innovation-and-ai/products/dolphingemma/>
- [4] Millman, R. (2025-04-15). Google launches AI that talks to dolphins. *Newsweek*. <https://www.newsweek.com/dolphingemma-google-ai-dolphins-2059955>
- [5] Marino, L. (1998). A comparison of encephalization between odontocete cetaceans and anthropoid primates. *Brain, Behavior and Evolution*, 51(4), 230–238. <https://doi.org/10.1159/00006540>

1.2 कठोर समुद्र और CHAT सस्टिम

डॉल्फ़िन की आवाज़ों को रिकॉर्ड करना आसान नहीं है। क्योंकि डॉल्फ़िन की आवाज़ों में बहुत सी ऐसी ऊँची ध्वनियाँ होती हैं जिनहे इंसान सुन नहीं सकते। इंसान के कान लगभग 20 किलोहर्ट्ज़ तक की आवाज़ सुन सकते हैं। किलोहर्ट्ज़ आवाज़ की ऊँचाई-नचाई मापने की इकाई है। डॉल्फ़िन 150 किलोहर्ट्ज़ के करीब तक की ऊँची आवाज़ें भी निकालती हैं। यह इंसानों द्वारा सुनी जा सकने वाली आवाज़ से कहीं अधिक ऊँची है [1]।

डॉल्फ़िन कई तरह की आवाज़ों का इस्तेमाल करती हैं। इनमें मुख्य रूप से तीन प्रकार सबसे प्रमुख हैं। पहला है ऊपर-नीचे होने वाली सीटी की आवाज़। दूसरा है खटखटाने वाली क्लिक की आवाज़। तीसरा है तेज़ी से फूटने वाली बर्स्ट पल्स आवाज़। बर्स्ट पल्स बहुत तेज़ी से लगातार निकलने वाली आवाज़ों की लड़ी होती है। डॉल्फ़िन जब आपस में झगड़ती हैं या पास रहकर बात करती हैं, तब इसे अक्सर इस्तेमाल करती हैं। क्लिक किसी वस्तु से टकराकर वापस लौटती है। डॉल्फ़िन उस गूँज से जान जाती हैं कि आगे क्या है। इसे इकोलोकेशन कहा जाता है। इकोलोकेशन आवाज़ की मदद से चीज़ों की स्थिति जानने की क्षमता है।

एक समस्या थी। पानी में सरिफ़ एक माइक्रोफ़ोन डालने से कोई फायदा नहीं था। क्योंकि कोई डॉल्फ़िन में से कसिने आवाज़ निकाली, यह पता लगाना मुमकिन नहीं था। इसलिए डॉ. हरज़गि एक इंजीनियर के पास गईं। वे जॉर्जिया टेक (Georgia Tech) के डॉ. थड स्टार्नर (Thad Starner) थे। डॉ. स्टार्नर शरीर पर पहने जाने वाले छोटे कंप्यूटर की तकनीक बनाने वाले विशेषज्ञ हैं।

डॉ. स्टार्नर को लगा था कि यह काम जल्दी पूरा हो जाएगा। लेकिन समुद्र एक कठिन चुनौती था। उपकरण को स्थिर रूप से चलाने में ही 4 साल लग गए। सबसे बड़ा दुश्मन खारा पानी था। खारा पानी दरारों से अंदर घुसकर मशीन को जला देता था [2]।

उपकरण को रखने वाला एल्युमिनियम का डबिबा भी परेशानी की वजह बना। वह तेज़ धूप को सोख लेता था जिससे अंदर का हसिसा बहुत गर्म हो जाता था। इससे कंप्यूटर काम करना बंद कर देता था। एल्युमिनियम वायरलेस सिग्नल को भी रोक देता था। शोध दल ने डबिबे के पछिले हसिसे को पारदर्शी प्लास्टिक से बदल दिया। तारों और सर्कट को एक विशेष रेज़िन से पूरी तरह सील कर दिया। तब जाकर यह उपकरण पानी के अंदर ठीक से काम करने लगा।

उपकरण को लंबे समय तक पानी में रखना भी आसान नहीं था। क्योंकि खारापन धातु को धीरे-धीरे खा जाता है। शोध दल हर बार इस्तेमाल के बाद उपकरण को धोता और उसकी जाँच करता था। एक छोटी सी दरार भी पूरी मशीन को बेकार कर सकती थी।

कई खतरे भी थे। गोताखोर द्वारा पहनी गई बड़ी बैटरी में पानी छू जाने पर वस्फोट हो सकता था। उपकरण से लीक होने वाला हल्का करंट शार्क को आकर्षित कर सकता था। क्योंकि शार्क में कमज़ोर बजिली के संकेतों को महसूस करने की इंद्री होती है। यहाँ तक कि पिरिक्षण के दौरान सचमुच एक शार्क सामने आ गई थी।

शार्क के पास कमज़ोर बजिली को महसूस करने वाला एक संवेदी अंग होता है। इस अंग को लोरेज़िनी का एंपुला (Ampullae of Lorenzini) कहा जाता है। यह शार्क की नाक के आसपास छोटे-छोटे छेदों के रूप में होता है। इसीलिए वे शिकार द्वारा पैदा की जाने वाली हल्की सी बजिली को भी महसूस कर लेती हैं। उन्हें डर था कि पानी के अंदर का उपकरण इस इंद्री को उत्तेजित न कर दे।

इन सारी मुश्किलों के बाद जो उपकरण बना, वह CHAT सस्टिम है। CHAT नाम का एक खास अर्थ है। यदि इसे सरल रूप में समझे तो यह 'सटिशियन हयिरिंग ऑगमेंटेशन टेलीमेट्री' (Cetacean Hearing Augmentation Telemetry) है। आसान शब्दों में कहे तो यह पानी के अंदर आवाज़ों का आदान-प्रदान करने और उन्हें पहचानने वाला, शरीर पर पहना जाने वाला कंप्यूटर है। सटिशियन डॉल्फ़िन और व्हेल को एक साथ मिलाकर कहा जाने वाला शब्द है [3]।

CHAT डॉल्फ़िन की आवाज़ों को सुनने और रिकॉर्ड करने वाला पानी के अंदर का कंप्यूटर है। यह मशीन आवाज़ों को बहुत बारीकी से रिकॉर्ड करती है। यह 1 सेकंड में 1,90,000 से अधिक बार ध्वनि को टुकड़ों में बाँटकर ग्रहण करती है। यह इंसानी कानों की तुलना में कहीं अधिक बारीकी से सुनती है। इतनी बारीकी से सुनने पर ही बहुत ऊँची आवाज़ें भी छूटने से बचती हैं। लेकिन इससे स्टोरेज स्पेस बहुत तेज़ी से भर जाता है। इसलिए इसके साथ एक बड़ा स्टोरेज ड्राइव इस्तेमाल किया जाता है।

गोताखोर इस उपकरण को अपनी पीठ पर पहनता है और कलाई पर एक छोटा कीबोर्ड बाँधता है। वह अस्थिचालन हेडफोन (bone conduction headphone) से आवाज़ सुनता है। अस्थिचालन कान के बजाय खोपड़ी की हड्डी को कंपति करके आवाज़ पहुँचाने का तरीका है। जब गोताखोर कीबोर्ड दबाता है, तो पानी में एक तय की गई आवाज़ निकलती है। यह वह आवाज़ नहीं है जो डॉल्फ़िनी आमतौर पर निकालती है। यह इंसानों द्वारा बनाई गई नई आवाज़ है। डॉल्फ़िनी की आवाज़ के साथ भ्रम न हो, इसलिए जानबूझकर एक नई आवाज़ चुनी गई। दूसरे गोताखोर का उपकरण इस आवाज़ को पहचान लेता है। यद्यपि पहले से तय आवाज़ से मेल खाती है, तो वह उस आवाज़ से जुड़ी वस्तु का नाम बता देता है।

आवाज़ का आदान-प्रदान करते समय यह बोलकर जानकारी देता है। जब कोई गोताखोर ध्वनि भेजता है, तो एक पुरुष की आवाज़ सुनाई देती है। उस आवाज़ को प्राप्त करने वाले की तरफ़ एक महिला की आवाज़ उसका अर्थ पढ़कर सुनाती है। दो अलग-अलग आवाज़ों का उपयोग इसलिए किया गया ताकि यह भ्रम न हो कि कौन बोल रहा है।

संबंधित संदर्भ सामग्री

- [1] Georgia Tech Research. (2025-04-14). DolphinGemma: How Google AI is helping decode dolphin communication. Georgia Tech Research News. <https://research.gatech.edu/dolphingemma-how-google-ai-helping-decode-dolphin-communication>
- [2] Herzing, D. L., Pack, A. A., Delfour, F., Starner, T., Mason, C., Gilliland, S., Ramsey, C., & Kohlsdorf, D. (2024). Imitation of computer-generated sounds by wild Atlantic spotted dolphins (*Stenella frontalis*). *Animal Behavior and Cognition*, 11(2), 136–166. <https://doi.org/10.26451/abc.11.02.02.2024>
- [3] Kohlsdorf, D., Gilliland, S., Presti, P., Starner, T., & Herzing, D. (2013). An underwater wearable computer for two way human-dolphin communication experimentation. *Proceedings of the 2013 International Symposium on Wearable Computers (ISWC)*, 147–148. <https://doi.org/10.1145/2493988.2494346>

1.3 साथ खेलकर सीखने वाला संवाद खेल

उपकरण बन जाने के बाद भी एक समस्या बची थी। वह समस्या थी कि डॉल्फ़िनी को आवाज़ें कैसे सिखाई जाएँ। उन्हें ज़बरदस्ती प्रशिक्षित करने का तरीका न अपनाने का फैसला किया गया। इसके बजाय, ऐसा तरीका चुना गया जिससे डॉल्फ़िनी खुद सीख सकें [1]।

डॉ. हर्ज़िंग ने जो तरीका चुना, उसका एक नाम है। यह तरीका है 'सामाजिक प्रतिद्वंद्विता मॉडलिंग' (social rival modeling)। यह दो इंसानों द्वारा आपस में सवाल-जवाब करने का प्रदर्शन दिखाने का तरीका है। इस तरीके का इस्तेमाल डॉ. आइरीन पेपरबर्ग (Irene Pepperberg) ने तोतों को सिखाते समय किया था। जब दो इंसान आपस में मुकाबला करते हुए दिखाते थे, तो जानवर पास से देखकर सीख जाता था। यह तरीका इस प्रकार है। दो इंसान डॉल्फ़िनी के सामने खेल का प्रदर्शन करते हैं। एक व्यक्ति वस्तु को रखता है, और दूसरा व्यक्ति उस वस्तु की आवाज़ निकालता है। आवाज़ निकालने वाला व्यक्ति वस्तु को लेता है और दोनों मलिकर खेलते हैं।

असल खेल को देखें तो यह कुछ ऐसा है। एक गोताखोर अपने हाथ में पीला स्कार्फ़ पकड़ता है। दूसरा गोताखोर कलाई के कीबोर्ड से स्कार्फ़ वाली आवाज़ निकालता है। तब स्कार्फ़ पकड़े हुए गोताखोर आवाज़ निकालने वाले गोताखोर को स्कार्फ़ सौंप देता है। फिर दोनों पानी के अंदर स्कार्फ़ को खींचते हुए मजे से खेलते हैं।

डॉल्फ़िनी के साथ खेलने के कई खिलौने होते हैं। वे केवल पीले स्कार्फ़ से ही नहीं, बल्कि सिमुद्र में तैरने वाले सरगासम (Sargassum) समुद्री शैवाल से भी खेलती हैं। सरगासम समुद्र की सतह पर तैरने वाला भूरा समुद्री शैवाल है। वे लंबी रस्सी को खींचकर भी खेलती हैं। वे नाव के आगे उठने वाली लहरों पर सवारी करने का खेल भी पसंद करती हैं। प्रत्येक खिलौने के लिए एक अलग आवाज़ तय की जाती है और उसका आदान-प्रदान किया जाता है।

डॉल्फ़िनी ने पास रहकर इस खेल को देखा। उन्होंने समझ लिया कि आवाज़ निकालने पर खिलौने का आदान-प्रदान होता है। डॉल्फ़िनी ने जल्द ही उस आवाज़ की नकल करना शुरू कर दिया। उन्होंने इंसानों द्वारा बनाई गई सीटी के उतार-चढ़ाव की हूबहू नकल की। इस प्रकार आवाज़ की नकल करने को ध्वनि अनुकरण कहते हैं। वैज्ञानिक इसे ध्वनिक अनुकरण कहते हैं। खेल का यह अवलोकन बाद में AI अनुसंधान का महत्वपूर्ण आधार बना [2][3]।

डॉल्फ़िनी ने कई बार खुद उस आवाज़ की नकल की। कभी-कभी उन्होंने एक बार सुनकर तुरंत नकल कर ली। तो कभी थोड़ी देर बाद नकल की। इससे पता चला कि वे नई आवाज़ों की लचीले ढंग से नकल कर सकती हैं। लेकिन एक बात ध्यान देने योग्य है। क्या वे यह जानकर उस आवाज़ का इस्तेमाल कर रही थीं कि वह किस वस्तु को दर्शाती है, यह अभी तक साबित नहीं हुआ है। आवाज़ की नकल करने में तो सफलता मिली। लेकिन वे आवाज़ को किसी वस्तु के नाम के रूप में इस्तेमाल कर रही थीं, इसका कोई पक्का सबूत नहीं मिल सका।

डॉल्फ़िन ने खुद आवाज़ क्यों नक़ाली? यह केवल भोजन पाने के लिए नहीं था। वे इंसानों के साथ खेलना चाहती थीं। जब खेल मज़ेदार होता था, तो डॉल्फ़िन खुद पहले पास आती थीं। यानी आनंद ही संवाद की वजह बना।

यह प्रयोग एक महत्वपूर्ण तथ्य दिखाता है। वह यह कि डॉल्फ़िन नई आवाज़ें सीख सकती हैं और उनकी नक़ल कर सकती हैं। साथ ही यह भी कि इंसानों की भाषा को ज़बरदस्ती सख़िने की ज़रूरत नहीं है। डॉल्फ़िन खुद आगे आई और उन्होंने इंसानों के साथ आवाज़ों का आदान-प्रदान किया।

संबंधित संदर्भ सामग्री

- [1] Georgia Tech Research. (2025-04-14). DolphinGemma: How Google AI is helping decode dolphin communication. Georgia Tech Research News. <https://research.gatech.edu/dolphingemma-how-google-ai-helping-decode-dolphin-communication>
- [2] Herzing, D. L., Pack, A. A., Delfour, F., Starner, T., Mason, C., Gilliland, S., Ramsey, C., & Kohlsdorf, D. (2024). Imitation of computer-generated sounds by wild Atlantic spotted dolphins (*Stenella frontalis*). *Animal Behavior and Cognition*, 11(2), 136–166. <https://doi.org/10.26451/abc.11.02.02.2024>
- [3] Google. (2025-04-14). DolphinGemma: How Google AI is helping decode dolphin communication. The Keyword. <https://blog.google/innovation-and-ai/products/dolphingemma/>

1.4 DolphinGemma द्वारा बनाया गया ध्वनिकी नक्शा

इंसानों द्वारा तय किए गए ध्वननियमों की एक सीमा थी। क्योंकि डॉल्फ़िन अपने झुंड के भीतर जो वास्तविक आवाज़ें नक़ालती हैं, वे कहीं अधिक जटिल थीं। इन जटिल आवाज़ों को समझने के लिए एक शक्तिशाली साधन की आवश्यकता थी। और वह साधन है AI [1]।

2025 में एक बड़ी ख़बर आई। Google DeepMind और Wild Dolphin Project ने मलिकर एक AI पेश किया। Google DeepMind, Google का AI अनुसंधान संगठन है। इस AI का नाम 'DolphinGemma' है। DolphinGemma, Google द्वारा बनाए गए Gemma नामक एक छोटे AI पर आधारित है। इसे 40 वर्षों के दौरान जुटाई गई डॉल्फ़िन की आवाज़ों और वीडियो को देकर सख़ाया गया [2]।

पुराने उपकरण पीठ पर लादे जाने वाले भारी बैकपैक कंप्यूटर थे। गोताख़ोर उस भारी मशीन को पीठ पर लादकर समुद्र में उतरते थे। अब हथेली के आकार का स्मार्टफ़ोन वही काम करता है। उपकरण बहुत छोटा हो गया है। DolphinGemma को छोटा बनाने का एक कारण है। इस AI के सीखने के लिए इस्तेमाल किए जाने वाले नयित्रण मान लगभग 40 करोड़ हैं। इन नयित्रण मानों को पैरामीटर कहा जाता है। बड़े AI की तुलना में यह काफी छोटा है। इसलिए यह हाथ में पकड़े जाने वाले स्मार्टफ़ोन पर भी चल जाता है। शोध दल ने इस AI को Google Pixel फ़ोन में इंस्टॉल किया। भारी कंप्यूटर को नाव पर रखे बिना ही समुद्र में सीधे इसका उपयोग करना संभव हो गया [4]।

DolphinGemma नाम का भी एक अर्थ है। यह डॉल्फ़िन शब्द के साथ Gemma को जोड़कर बना है। इसका मतलब है कि Gemma नामक AI के ऊपर डॉल्फ़िन की आवाज़ों को जोड़कर इसे तैयार किया गया है।

सबसे कठिन काम ध्वनिकी संख्याओं में बदलना था। पानी के अंदर की आवाज़ें बिना रुके लगातार बहती हैं। AI अलग-अलग टुकड़ों में बंटे डेटा को पसंद करता है। इसलिए शोधकर्ताओं ने ध्वनिकी छोटे-छोटे टुकड़ों में काटा। जैसे अक्षरों को व्यंजन और स्वर में बाँटा जाता है, वैसे ही आवाज़ को छोटे टुकड़ों में विभाजित किया गया। ध्वनिकी इन टुकड़ों को ध्वनिकी टोकन (acoustic tokens) कहा जाता है। इसका अर्थ है ध्वनिकी को बारीक हिस्सों में बाँटा गया टुकड़ा।

ध्वनिकी विभाजित करने के लिए इस्तेमाल किया गया एक साधन भी है। यह Google की SoundStream नामक ध्वनिकी-विभाजन तकनीक है। यह तकनीक पानी के अंदर की आवाज़ों को छोटे-छोटे टुकड़ों में बाँटती है। और उन टुकड़ों से DolphinGemma आगे सीखता है।

ध्वनिकी इन टुकड़ों की मदद से DolphinGemma ने अगली ध्वनिकी पहले से अनुमान लगाने का अभ्यास किया। यह पछिली आवाज़ को देखकर उसके बाद आने वाली आवाज़ का अंदाज़ा लगाने का अभ्यास था। नतीजतन, DolphinGemma अगली आने वाली आवाज़ का सही अनुमान लगाने में माहिर हो गया। यह ठीक वैसे ही है जैसे कोई इंसान भाषा के नियमों को जानकर अगले शब्द का अंदाज़ा लगा लेता है।

अगली आवाज़ का सही अनुमान लगाने का मतलब है कि वह उसके नियमों को समझता है। AI ने डॉल्फ़िन की आवाज़ों में ऐसे ही नियमों को खोज नक़ाला है।

DolphinGemma एक जैसी आवाज़ों को अपने आप इकट्ठा भी करता है। AI आवाज़ों को बड़ियों में बदलकर नक्शे की तरह सजाता है। एक जैसी आवाज़ें पास और अलग आवाज़ें दूर रखी जाती हैं। शोधकर्ताओं ने लंबे समय से अलग-अलग स्थितियों की आवाज़ों को इकट्ठा किया है। उदाहरण के लिए, एक सग्नेचर सीटी (signature whistle) होती है जिससे माँ और बच्चा एक-दूसरे को बुलाते हैं। सग्नेचर सीटी हर

डॉल्फिन के लिए अलग, नाम जैसी आवाज होती है। लड़ते समय नकिली जाने वाली बर्स्ट पल्स (burst pulse) भी होती है। साथी ढूँढते समय या शार्क का पीछा करते समय नकिलने वाली क्लिक (click) आवाज भी होती है। इन आवाजों को स्थितियों के साथ जोड़ने का एक फायदा होता है। यह AI को आवाजों के सुराग अधिक तेजी से खोजने में मदद करता है।

डॉ. हे जिंग (He Jing) ने इस काम को इस तरह समझाया। उन्होंने कहा कि अगर कोई इंसान हाथ से इस पूरे डेटा की जांच करे, तो 150 साल लग जाएंगे। उन्होंने आगे कहा कि AI यह काम बहुत तेजी से कर लेता है। 150 साल इंसान के दो पूरे जीवन जीने से भी लंबा समय है [3]।

संबंधित संदर्भ सामग्री

- [1] Google. (2025-04-14). DolphinGemma: How Google AI is helping decode dolphin communication. The Keyword. <https://blog.google/innovation-and-ai/products/dolphingemma/>
- [2] Google DeepMind. (2025). DolphinGemma. Google DeepMind. <https://deepmind.google/models/gemma/dolphingemma/>
- [3] Scientific American. (2025). DolphinGemma could enable AI communication with dolphins [Podcast]. Scientific American. <https://www.scientificamerican.com/podcast/episode/dolphingemma-could-enable-ai-communication-with-dolphins/>
- [4] Lanz, J. A. (2025-04-14). Google launches AI to decode dolphin talk, runs on Pixel phones. Decrypt. <https://decrypt.co/314766/google-launches-ai-decode-dolphin-talk>

1.5 समुद्र में शुरू हुई असली बातचीत

2025 की गर्मियों में एक नया अध्याय खुला। Wild Dolphin Project ने समुद्र में DolphinGemma का इस्तेमाल करना शुरू किया। उससे पहले तक, इसे केवल कंप्यूटर के अंदर ही सखाया जाता था। अब गोताखोर उस AI को पानी के अंदर ले जाते हैं। शोधकर्ताओं ने नया नकिला गूगल पिक्सल (Google Pixel) फोन इस्तेमाल किया [1][9]।

समुद्र में इस्तेमाल करने का तरीका ऐसा है। गोताखोर वाटरप्रूफ बॉक्स में रखे पिक्सल फोन को लेकर पानी में जाते हैं। AI पानी के अंदर तुरंत आवाजों की जांच करता है। और दोहराए जाने वाले आवाज के पैटर्न खोजकर शोधकर्ताओं को बताता है। नाव पर लौटकर डेटा देखे बिना भी, पानी के अंदर ही तुरंत इसकी पुष्टि हो जाती है।

इस समुद्री शोध में एक खास बात है। DolphinGemma ने खुद डॉल्फिन जैसी नई आवाजें बनाईं। यह ऐसी आवाज थी जैसी इंसानों द्वारा इस्तेमाल किए जाने वाले प्रोग्राम से बनाना मुश्किल था। शोधकर्ता भविष्य में डॉल्फिन के साथ आदान-प्रदान के शोध में इस आवाज का इस्तेमाल करने की योजना बना रहे हैं। वे नई आवाज से डॉल्फिन से बात करने की कोशिश करना चाहते हैं [2]।

इस तकनीक का इस्तेमाल बहुत सावधानी से किया जाता है। क्योंकि अगर इंसान बिना सोचे-समझे आवाज नकिले, तो डॉल्फिन का समाज अस्त-व्यस्त हो सकता है। डॉल्फिन के समूह लंबे समय से आवाजों के जरिए एक-दूसरे से जुड़े रहे हैं। इसलिए शोधकर्ता डॉल्फिन की मूल आवाजों में खलल नहीं डालना चाहते। डॉल्फिन को सखाई जाने वाली आवाज के रूप में भी उन्होंने एक नई सीटी चुनी, जो प्रकृति में नहीं होती। इसका मकसद पुरानी आवाजों के साथ कोई उलझन न पैदा करना है। शोधकर्ता डॉल्फिन के आवाज आदान-प्रदान के प्रवाह में भी बाधा न डालने की कोशिश करते हैं। वे उस शांत पल का ध्यान रखते हैं जब डॉल्फिन बोलना बंद कर देती है। यह भोजन की कतार में अपनी बारी का इंतजार करने जैसा शिष्टाचार है।

सिर्फ डॉल्फिन पर ही ऐसा शोध नहीं हो रहा है। Earth Species Project नाम का एक संगठन भी है। यह संगठन कई तरह के जानवरों की आवाजों का AI से अध्ययन करता है। पक्षी, वृहल और हाथी जैसे अलग-अलग जानवर इसके विषय हैं। ऐसा इसलिए है क्योंकि हर जानवर के आवाज नकिलने की वजह अलग होती है। कई जानवरों की तुलना करने पर आवाजों के नियम और साफ दिखाई देते हैं। 2026 में, जानवरों की आवाज के डेटा को एक साथ साझा करने वाला एक नया डेटा स्टोर भी खोला गया। इसका नाम alp-data है। यह नाम एनमिल लैंग्वेज प्रोसेसिंग (Animal Language Processing) के शुरुआती अक्षरों से लिया गया है। इसे इसलिए बनाया गया है ताकि कई शोधकर्ता डेटा बांटकर इस्तेमाल कर सकें [3][4]।

डेटा साझा करने का एक बड़ा फायदा है। कई देशों के शोधकर्ता जानवरों की आवाजों को एक जगह इकट्ठा करते हैं। जितना ज्यादा डेटा इकट्ठा होता है, AI उतना ही बेहतर सीखता है। यह संग्रह कई शोधकर्ताओं के लिए मिलकर काम करने का एक मजबूत आधार बनता है।

एक ही AI से कई जानवरों का अध्ययन भी किया जा सकता है। ऐसा इसलिए है क्योंकि आवाजों को समझने का तरीका हर जानवर में काफी मिला-जुला होता है। एक जानवर से सीखा गया तरीका दूसरे जानवर पर इस्तेमाल किया जाता है। इसलिए अलग-अलग शोध एक-दूसरे को आगे बढ़ाने में मदद करते हैं।

स्पर्म व्हेल (sperm whale) की भाषा को समझने का शोध भी आगे बढ़ रहा है। Project CETI नाम का एक संगठन यह काम करता है। स्पर्म व्हेल कई क्लिक (click) आवाजों को एक साथ मिलाकर निकालती है। इस समूह को coda कहा जाता है। शोधकर्ताओं ने coda के अंदर ध्वनिके नियम ढूँढ निकाले हैं। इन नियमों को स्पर्म व्हेल की ध्वनि वर्णमाला कहा जाता है। वर्णमाला शब्द एक तुलनात्मक अभिव्यक्ति है। यह इंसानों के अक्षर नहीं है, बल्कि क्लिक आवाजों के मलिन के नियमों की तुलना वर्णमाला से की गई है [5]।

2026 में एक नया तथ्य सामने आया। स्पर्म व्हेल के coda के अंदर इंसानी भाषा के स्वरों (vowels) जैसे पैटर्न मलिन। इंसानी भाषा में 'आ' और 'ई' जैसे स्वर होते हैं। स्पर्म व्हेल का coda भी ध्वनिके गुणों के आधार पर दो भागों में बंटा हुआ था। coda की लंबाई और उनके क्रम में भी नियम मौजूद थे। इसका मतलब है कि स्पर्म व्हेल का भी अपना एक ध्वनि ढाँचा होता है [6][7]।

यह शोध आगे भी बहुत सावधानी से जारी रहेगा। जानवरों की भाषा समझ लेने का मतलब यह नहीं है कि हम सब कुछ जान गए हैं। क्योंकि इंसान और जानवरों के जीने की दुनिया अलग है। फरि भी, ध्यान से सुनना बहुत सार्थक है। एक-दूसरे को जानने से साथ रहना आसान हो जाता है [8]।

यह सारा शोध एक ही दिशा की ओर इशारा करता है। यह जानवरों को जबर्दस्ती इंसानों की भाषा सिखाने का काम नहीं है। यह जानवरों की आवाजों को ठीक वैसे ही सुनना है जैसी वे हैं। कुछ लोग कहते हैं कि यह शोध जानवरों के अधिकारों में भी मददगार हो सकता है। क्योंकि अगर हम जान जाएँ कि जानवर क्या सोच रहे हैं, तो हम उनकी बेहतर रक्षा कर सकते हैं। समुद्र में शुरू हुई यह बातचीत इंसान और प्रकृति को जोड़ने वाला एक नया पुल बन रही है।

संबंधित संदर्भ सामग्री

- [1] Google. (2025-04-14). DolphinGemma: How Google AI is helping decode dolphin communication. The Keyword. <https://blog.google/innovation-and-ai/products/dolphingemma/>
- [2] Scientific American. (2025). DolphinGemma could enable AI communication with dolphins [Podcast]. Scientific American. <https://www.scientificamerican.com/podcast/episode/dolphingemma-could-enable-ai-communication-with-dolphins/>
- [3] Earth Species Project. (2026-04-02). Animal language processing: An AI convergence in animal communication. Earth Species Project. <https://earthspecies.org/2026/04/02/animal-language-processing-an-ai-convergence-in-animal-communication/>
- [4] Earth Species Project. (2026-07-22). Introducing alp-data: A shared data layer for animal language processing. Earth Species Project. <https://earthspecies.org/2026/07/22/introducing-alp-data-a-shared-data-layer-for-animal-language-processing/>
- [5] Project CETI. (2024). Sperm whale phonetic alphabet proposed for the first time. Project CETI Blog. <https://www.projectceti.org/blog-posts/sperm-whale-phonetic-alphabet-proposed-for-the-first-time>
- [6] Beguš, G., Dąbkowski, M., Sprouse, R. L., Gruber, D. F., & Gero, S. (2026). The phonology of sperm whale coda vowels. Proceedings of the Royal Society B, 293(2069), 20252994. <https://doi.org/10.1098/rspb.2025.2994>
- [7] National Geographic. (2026). We're one step closer to understanding the sperm whale 'alphabet'. National Geographic. <https://www.nationalgeographic.com/animals/article/sperm-whale-vowel-patterns-communication-discovery>
- [8] Boing Boing. (2026-06-11). Decoding whale language could unlock new legal rights for them. Boing Boing. <https://boingboing.net/2026/06/11/decoding-whale-language-could-unlock-new-legal-rights-for-them.html>
- [9] Schwartz, E. H. (2025-04-15). Google's new AI model could someday let you understand and talk to dolphins. TechRadar. <https://www.techradar.com/computing/artificial-intelligence/googles-new-ai-model-could-someday-let-you-understand-and-talk-to-dolphins>

अध्याय 2. स्पर्म व्हेल की ध्वन्यात्मक वर्णमाला और Project CETI

2.1 कोड सुलझाने वाले और गहरे समुद्र के विशाल जीव की मुलाकात

पृथ्वी पर सबसे बड़े दमिग वाला जानवर इंसान नहीं है। वह स्पर्म व्हेल (sperm whale) है। स्पर्म व्हेल का वैज्ञानिक नाम फसिटर मैक्रोसेफालस (Physeter macrocephalus) है। यह गहरे समुद्र में रहने वाली एक बहुत बड़ी व्हेल है। इसकी शरीर की लंबाई 20 मीटर तक हो सकती है। यह लंबाई दो बड़ी बसों को एक साथ जोड़ने जितनी है। इसका वजन 50 टन तक हो सकता है। यह वजन लगभग दस हाथियों के कुल वजन के बराबर है [1][2]।

इस व्हेल का दमिग लगभग 8 से 9 कलोग्राम का होता है। इंसान का दमिग लगभग 1.4 कलोग्राम का होता है। इसलिए स्पर्म व्हेल का दमिग इंसान के दमिग से लगभग 6 गुना बड़ा होता है। यह बड़ा दमिग समूह में मलि-जुलकर रहने और भावनाओं को संभालने के काम आता है। स्पर्म व्हेल अपने पूरे जीवन का 80 प्रतिशत हिस्सा गहरे और अंधेरे समुद्र में बिताती है। उस गहराई में रोशनी नहीं पहुंचती। इसलिए स्पर्म व्हेल आंखों के बजाय आवाज के जरिए दुनिया को देखती है।

स्पर्म व्हेल आवाज बाहर भेजती है और वापस लौटकर आने वाली गूंज को सुनती है। आवाज किसी वस्तु से टकराकर गूंज बनकर लौटती है। उस गूंज को सुनकर वह वस्तु की स्थिति और दूरी का पता लगाती है। इस तरीके को इकोलोकेशन (echolocation) कहा जाता है। इसका मतलब आवाज के जरिए आसपास का जायजा लेने का तरीका है। इस समय निकलने वाली क्लिक (click) आवाज रास्ते खोजने और शिकार का पता लगाने के काम आती है। लेकिन स्पर्म व्हेल बातचीत करते समय एक अलग आवाज का इस्तेमाल करती है। वह छोटी-छोटी क्लिक आवाजों को एक तय ताल में कई बार मिलाकर निकालती है। ध्वनियों के इस समूह को कोडा (coda) कहा जाता है। कोडा वह ध्वनिका टुकड़ा है जिसका उपयोग व्हेल आपस में बात करने के लिए करती है [3][5]।

लंबे समय तक वैज्ञानिक कोडा को कोई खास बात नहीं मानते थे। वे इसे केवल एक छोटा संकेत भर समझते थे। लेकिन 2017 में अमेरिका के हार्वर्ड विश्वविद्यालय के रैडक्लिफ़ संस्थान (Radcliffe Institute) में एक नया वचिर पैदा हुआ। समुद्री जीवों का अध्ययन करने वाले डॉ. डेविड ग्रुबर (David Gruber) ने व्हेल की आवाजों को रिकॉर्ड करके रखा था। एक व्यक्ति ने उस आवाज को इतनेफाक से सुना। वे अमेरिका के MIT की कोड शोधकर्ता प्रोफेसर शफी गोल्डवासर (Shafi Goldwasser) थीं। कोड गुप्त बातों को छपाने के लिए बनाए गए नियमों वाले संकेत होते हैं। प्रोफेसर गोल्डवासर को व्हेल की आवाज में एक नियम महसूस हुआ। वह इंसानों द्वारा बनाए गए गुप्त कोड जैसा देख रहा था। इसलिए उन्होंने कंप्यूटर के जरिए उस नियम को सुलझाने का सुझाव दिया [4]।

यह वचिर लंबे समय से व्हेल पर नजर रखने वाले एक वैज्ञानिक से जुड़ा। वे डॉ. शेन गेरो (Shane Gero) थे। डॉ. गेरो कैरेबियाई सागर में डोमनिका के तट पर 13 से अधिक वर्षों तक व्हेल परिवारों के पीछे-पीछे घूमे थे। और 2020 में, Project CETI (Cetacean Translation Initiative) की शुरुआत हुई। इसका अर्थ है 'व्हेल की भाषा का अनुवाद करने की परियोजना'। कई देशों के वैज्ञानिक एक साथ आए। उन्होंने डोमनिका के समुद्र में एक अनुसंधान केंद्र स्थापित किया। उन्होंने समुद्र में 1,000 मीटर की गहराई तक पानी के नीचे काम करने वाले माइक्रोफोन लगाए। व्हेल के शरीर पर आवाज और स्थिति को मापने वाले छोटे उपकरण लगाए गए। इस उपकरण को Project CETI के इंजीनियरों ने बनाया था [6][7][8]।

यह छोटा उपकरण बहुत खास है। इसके अंदर तीन माइक्रोफोन लगे हैं। कई माइक्रोफोन होने की वजह से यह शोर भरे समुद्र में भी आवाज को अलग-अलग पहचान लेता है। भले ही कई व्हेल एक साथ आवाज निकालें, यह सिर्फ एक व्हेल की आवाज को चुन सकता है। यह उपकरण केवल आवाज ही रिकॉर्ड नहीं करता। यह उस गहराई और हलचल को भी मापता है जहां व्हेल मौजूद होती है। यह पानी के तापमान और स्थान को भी दर्ज करता है। इसलिए यह सब पता चल जाता है कि कौन सी आवाज कब और कहां निकाली गई थी। शोध दल ने इस तरह इकट्ठे किए गए कोडा का कंप्यूटर से विश्लेषण किया। विश्लेषण किए गए कोडा की संख्या 8,719 तक पहुंच गई। ध्वनियों का यह समूह व्हेल की भाषा को समझने की कुंजी बन गया [9]।

संबंधित संदर्भ सामग्री

[1] Whitehead, H. (2003). Sperm Whales: Social Evolution in the Ocean. University of Chicago Press.

[2] Project CETI. (2024). About Sperm Whales. Project CETI. <https://www.projectceti.org>

[3] Madsen, P. T., Wahlberg, M., & Møhl, B. (2002). Male sperm whale acoustics in a high-latitude habitat: implications for echolocation and communication. Behavioral Ecology and Sociobiology, 53(1), 31-41. <https://doi.org/10.1007/s00265-002-0548-1>

[4] Radcliffe Institute. (2024). The CETI Moment. Harvard Radcliffe Institute. <https://www.radcliffe.harvard.edu/radcliffe-25/ceti-moment>

- [5] Andrews, R. (2024-05-07). Scientists Discover a 'Phonetic Alphabet' Used by Sperm Whales. Smithsonian Magazine. <https://www.smithsonianmag.com/smart-news/scientists-discover-a-phonetic-alphabet-used-by-sperm-whales-mov-ing-one-step-closer-to-decoding-their-chatter-180984326/>
- [6] Gero, S., Whitehead, H., & Rendell, L. (2016). Individual, unit and vocal clan level identity cues in sperm whale codas. Royal Society Open Science, 3(1), 150372. <https://doi.org/10.1098/rsos.150372>
- [7] Project CETI. (2020). Cetacean Translation Initiative. Project CETI. <https://www.projectceti.org>
- [8] National Geographic Society. (2024). Project CETI. National Geographic Society. <https://www.nationalgeographic.org/explore/programs/project-ceti>
- [9] Sharma, P., Gero, S., Payne, R., Gruber, D. F., Rus, D., Torralba, A., & Andreas, J. (2024). Contextual and combinatorial structure in sperm whale vocalisations. Nature Communications, 15, 3617. <https://doi.org/10.1038/s41467-024-47221-8>

2.2 ट्रैफिक लाइट से आगे की आवाज: कोडा के चार तत्व

पुरानी पाठ्यपुस्तकों में स्पर्म् व्हेल की आवाज को बहुत सीमिति नजरिए से देखा गया था। उनमें लिखा था कि कैरेबियाई सागर की व्हेल केवल 21 प्रकार के कोडा का उपयोग करती हैं। माना जाता था कि एक कोडा केवल एक ही अर्थ बताता है। इसे ट्रैफिक लाइट के समान माना गया था। जैसे लाल बत्ती का मतलब रुकना और हरी बत्ती का मतलब चलना होता है। लेकिन जब Project CETI के AI ने आवाजों की जांच की, तो यह सोच टूट गई [1]।

AI इंसानों द्वारा बनाया गया एक कंप्यूटर प्रोग्राम है। यह बहुत सारे डेटा के बीच छपि नियमों को खुद ढूँढ लेता है। इस AI ने व्हेल की बातचीत में 156 प्रकार के कोडा खोज निकाले। यह पहले ज्ञात 21 प्रकारों से 7 गुना से भी अधिक है। इसके अलावा, एक ही कोडा के अंदर भी कई छोटे टुकड़े छपि हुए थे। व्हेल चार तत्वों को बदल-बदल कर अलग-अलग आवाजें निकालती थीं।

पहला तत्व टेम्पो (tempo) है। इसका मतलब है कि कितनी आवाज कितनी तेजी से निकाली जाती है। इंसान भी जब जल्दी में होता है, तो उसकी बोली तेज हो जाती है। व्हेल भी परिस्थितियों के अनुसार आवाज की गति को बदलती है। दूसरा तत्व लय यानी रदम (rhythm) है। यह क्लिक आवाजों के बीच के अंतराल को बताता है। आवाजें पास-पास हैं या दूर-दूर, इसके आधार पर संकेत बदल जाते हैं। तीसरा तत्व अलंकरण यानी ऑर्नामेंटेशन (ornamentation) है। यह ध्वनि समूह के अंत में एक अतिरिक्त क्लिक आवाज जोड़ने जैसा है। यह वाक्य के अंत में लगने वाले प्रश्नचिह्न की तरह भाव को थोड़ा बदल देता है। चौथा तत्व रुबातो (rubato) है। संगीत की कक्षा में ताल को थोड़ा खींचने या दबाने की तरह, यह आवाज की लंबाई में सूक्ष्म बदलाव करना है। व्हेल बातचीत के प्रवाह के अनुसार इस लंबाई को नियंत्रित करती है [3][4]।

ये चारों मलिकर 156 प्रकार के कोडा बनाते हैं। वे कोडा समय के क्रम में जुड़कर और भी लंबे संकेत बनाते हैं। ध्वनिके एक छोटे टुकड़े का अपने आप में कोई अर्थ नहीं होता। लेकिन जब उन टुकड़ों को तरह-तरह से जोड़ा जाता है, तो अनगिनत संकेत बन जाते हैं। इंसानों की बोली भी इसी तरह बनती है। बिना अर्थ वाली ध्वनियाँ यानी व्यंजन और स्वर मलिकर शब्द और वाक्य बनाते हैं। शोधकर्ताओं ने कोडा बनाने वाले इन छोटे ध्वनिके टुकड़ों की तुलना इंसानों के व्यंजनों और स्वरों से की, और इसे ध्वन्यात्मक वर्णमाला (phonetic alphabet) कहा। भाषावर्ज्ञान में इस विशेषता को 'संरचना की द्वैतता (duality of patterning)' कहा जाता है। अब तक लोग मानते थे कि यह खूबी केवल इंसानी भाषा में ही होती है। लेकिन अब यह बात सामने आई है कि व्हेल की भाषा में भी ऐसी ही मलित्ती-जुलती संरचना हो सकती है [2]।

ऐसा लगता है कि व्हेल बात करने वाले साथी के अनुसार अपनी आवाज भी बदलती है। बच्चे के सामने निकाली जाने वाली आवाज और साथी व्हेल के सामने निकाली जाने वाली आवाज अलग होती है। शोधकर्ता अनुमान लगाते हैं कि व्हेल सुनने वाले साथी का ध्यान रखकर आवाज चुनती है। हालांकि यह अभी पूरी तरह स्पष्ट नहीं हुआ है। व्हेल कभी-कभी सामने वाले की आवाज खत्म होने तक इंतजार भी करती है। और जब बीच में खाली जगह मिलती है, तब अपनी आवाज जोड़ती है। यह वैसा ही है जैसे कक्षा में किसी दोस्त की बात खत्म होने का इंतजार करके बोलना। इस तरह अपनी बारी का पालन करते हुए आवाजों का आदान-प्रदान करने को टर्न-टेकिंग (turn-taking) कहा जाता है [6]।

व्हेल के हर समूह द्वारा इस्तेमाल किया जाने वाला कोडा थोड़ा अलग होता है। यह वैसा ही है जैसे अलग-अलग क्षेत्रों में इंसानों की बोली अलग होती है। वैज्ञानिक ऐसे समूहों को कबीला (clan) कहते हैं। कबीला (clan) एक बड़ा समूह है जो एक जैसे कोडा साझा करता है। एक ही कबीले की व्हेल आपस में मलित्ती-जुलती आवाजें निकालती हैं। बच्चा व्हेल बड़े होते समय अपने समूह के कोडा को सीखता है। यह वैसा ही है जैसे इंसान अपने माता-पिता की भाषा सीखता है। इसलिए व्हेल का कोडा जन्मजात नहीं होता, बल्कि सीखा जाता है [5]।

संबंधित संदर्भ सामग्री

- [1] Sharma, P., Gero, S., Payne, R., Gruber, D. F., Rus, D., Torralba, A., & Andreas, J. (2024). Contextual and combinatorial structure in sperm whale vocalisations. Nature Communications, 15, 3617. <https://doi.org/10.1038/s41467-024-47221-8>

- [2] Project CETI. (2024). Sperm Whale Phonetic Alphabet Proposed for the First Time. Project CETI Blog. <https://www.projectceti.org/blog-posts/sperm-whale-phonetic-alphabet-proposed-for-the-first-time>
- [3] Andrews, R. (2024-05-07). Scientists Discover a 'Phonetic Alphabet' Used by Sperm Whales. Smithsonian Magazine. <https://www.smithsonianmag.com/smart-news/scientists-discover-a-phonetic-alphabet-used-by-sperm-whales-movin-g-one-step-closer-to-decoding-their-chatter-180984326/>
- [4] Al Jazeera. (2024-05-08). Scientists discover sperm whale 'phonetic alphabet'. Al Jazeera. <https://www.aljazeera.com/news/2024/5/8/scientists-discover-sperm-whale-phonetic-alphabet>
- [5] Gero, S., Whitehead, H., & Rendell, L. (2016). Individual, unit and vocal clan level identity cues in sperm whale codas. Royal Society Open Science, 3(1), 150372. <https://doi.org/10.1098/rsos.150372>
- [6] Schulz, T. M., Whitehead, H., Gero, S., & Rendell, L. (2008). Overlapping and matching of codas in vocal interactions between sperm whales: insights into communication function. Animal Behaviour, 76(6), 1977-1988. <https://doi.org/10.1016/j.anbehav.2008.07.032>

2.3 गहरे समुद्र में मल्लि स्वर: प्रोफेसर बेगुश की खोज

लंबे समय तक व्हेल की आवाज के शोध में केवल एक ही चीज मापी गई थी। वह थी क्लिक आवाजों के बीच का समय अंतराल। आवाज के उतार-चढ़ाव पर ध्यान नहीं दिया जा सका था। इस बाधा को पार करने वाले एक व्यक्ति हैं। वे अमेरिका के UC बर्कले के प्रोफेसर गास्पेर बेगुश (Gašper Beguš) हैं। प्रोफेसर बेगुश ने आवाज के उतार-चढ़ाव पर ध्यान दिया। आवाज कई तरह के उतार-चढ़ावों के मिलने से बनती है। इस मशरूम को फ्रीक्वेंसी स्पेक्ट्रम (spectrum) कहा जाता है। जब सूरज की रोशनी कांच पर पड़ती है, तो वह इंद्रधनुष के रंगों में बखिर जाती है। आवाज को भी इसी तरह उसके उतार-चढ़ाव के हिसाब से फैलाया गया है [1][3]।

प्रोफेसर बेगुश ने तब ध्यान से देखा जब AI व्हेल की आवाज की नकल करके उसे बना रहा था। AI न केवल समय अंतराल, बल्कि आवाज के उतार-चढ़ाव को भी बहुत महत्वपूर्ण मान रहा था। इसलिए प्रोफेसर ने व्हेल की क्लिक आवाज की गति को कई गुना धीमा कर दिया। उन्होंने इसे इंसानी कानों के अनुकूल गति में बदल दिया। तब क्लिक आवाज के अंदर कुछ नया दिखाई दिया। वह इंसानी स्वरों से मल्लि-जुलती आवाज थी [2][4]।

इंसान जब स्वर निकालता है, तो एक खास उतार-चढ़ाव जोर से गूंजता है। इस तरह जोर से गूंजने वाले उतार-चढ़ाव को फॉर्मेट (formant) कहा जाता है। यह खाली बोलत में फूंक मारने पर आने वाली गूंज जैसा होता है। प्रोफेसर बेगुश ने व्हेल की आवाज में से दो स्वरों को अलग-अलग पहचाना। एक है 'आ' कोडा (a-coda)। यह इंसानी स्वर 'आ' से मल्लि-जुलती उतार-चढ़ाव वाली संरचना दिखाता है। दूसरा है 'ई' कोडा (i-coda)। यह इंसानी स्वर 'ई' से मल्लि-जुलती उतार-चढ़ाव वाली संरचना दिखाता है।

व्हेल ने दोनों स्वरों को धुंधला मिलाकर नहीं निकाला। उसने 'आ' और 'ई' को साफ-साफ अलग करके निकाला। एक स्वर से दूसरे स्वर में फसिलती हुई जाने वाली आवाज भी थी। इंसानों के संदर्भ में कहे तो यह संयुक्त स्वर (diphthong) है। जैसे 'आई' शब्द में दो ध्वनियों आपस में जुड़ती हैं। इंसानी कानों को यह सरिफ लकड़ी पर दस्तक देने जैसी आवाज लगती थी। लेकिन जब गति को धीमा करके सुना गया, तो यह 'आ' और 'ई' का आदान-प्रदान करने वाली बातचीत में बदल गई।

इस खोज की बुनियाद में एक खास AI था। यह ऐसा AI है जिसमें इंसान यह देख सकता है कि उसने ऐसा जवाब क्यों दिया। यह तब की बात है जब यह AI व्हेल की आवाज के उतार-चढ़ाव का विश्लेषण कर रहा था। इसके अंदर का एक हिसा कसिी खास आवाज पर ही बहुत जोरदार प्रतिक्रिया दे रहा था। यह वैसा ही था जैसे कोई स्विच केवल कसिी खास आवाज पर ही चालू होता हो। एक ऐसी बात थी जिससे इंसान लंबे समय से नजरअंदाज करता आ रहा था। वह यह तथ्य था कि व्हेल की आवाज में स्वरों जैसे उतार-चढ़ाव छिपे हैं। AI ने इसे सबसे पहले पहचान लिया।

AI ने व्हेल की आवाज की खुद नकल करके बनाना सीखा। इंसान द्वारा सही जवाब बताए बिना ही उसने इसे खुद सीख लिया। इस तरह खुद सीखने के तरीके को अनसुपरवाइज्ड लर्निंग (unsupervised learning) कहा जाता है। AI ने आवाज बनाते समय खुद चुना कि कौन सा हिसा महत्वपूर्ण है। उसने जो चुना, वह आवाज का उतार-चढ़ाव ही था। अप्रैल 2026 में, यह शोध ब्रिटिश रॉयल सोसाइटी के जर्नल में प्रकाशित हुआ। इसमें बताया गया कि 'आ' कोडा और 'ई' कोडा इंसानी स्वरों की तरह नियमों का पालन करते हैं।

व्हेल के स्वरों में इंसानी भाषा से मल्लि-जुलती और भी समानताएं थीं। ध्वनिसमूह के अंत में जुड़ने वाली क्लिक आवाज को ही देखिए। यह आवाज अपने पास आने वाले स्वर जैसी होने लगी। इंसान भी जब बोलता है, तो आगे और पीछे की ध्वनियों एक-दूसरे पर असर डालती हैं। जैसे भाषा की कक्षा में सीखा जाता है कि 'गुकमुल' (gukmul) का उच्चारण 'गुंगमुल' जैसा सुनाई देता है। क्योंकि आगे और पीछे की ध्वनियों एक-दूसरे जैसी बनने लगती हैं। व्हेल की आवाज में भी ठीक यही बात हो रही थी।

संबंधित संदर्भ सामग्री

- [1] Beguš, G., Dąbkowski, M., Sprouse, R. L., Gruber, D. F., & Gero, S. (2026). The phonology of sperm whale coda vowels. *Proceedings of the Royal Society B*, 293(2069), 20252994. <https://doi.org/10.1098/rspb.2025.2994>
- [2] Beguš, G., Sprouse, R. L., Leban, A., Silva, M., & Gero, S. (2025). Vowel- and Diphthong-Like Spectral Patterns in Sperm Whale Codas. *Open Mind*, 9, 1849-1874. <https://doi.org/10.1162/OPMI.a.252>
- [3] University of California, Berkeley. (2024). Gašper Beguš, Speech and Computation Lab. UC Berkeley Department of Linguistics. <https://linguistics.berkeley.edu>
- [4] Zimmer, K. (2026-04-01). Sperm whale clicks follow similar rules to human speech. *Phys.org*. <https://phys.org/news/2026-04-sperm-whale-clicks-similar-human.html>

2.4 नाक के रास्ते से स्वर बनाने वाला शरीर

इंसान और स्पर्म् व्हेल अलग-अलग रास्तों पर चले हैं। वे लंबे समय तक अलग-अलग विकसित हुए हैं। फिर भी दोनों 'आ' और 'ई' जैसे स्वर निकालते हैं। इस प्रक्रिया को अभिसारी विकास (convergent evolution) कहा जाता है। यह वह घटना है जिसमें अलग-अलग जीव एक जैसी समस्या के सामने एक जैसा समाधान खोज लेते हैं। यह वैसा ही है जैसे पक्षी और चमगादड़ अलग-अलग विकसित हुए, फिर भी दोनों के पास पंख हैं। चूंकि रहने की जगह के भौतिक नियम एक जैसे हैं, इसलिए उनके समाधान भी एक जैसे हो जाते हैं [1][2]।

इंसान स्वर निकालने के लिए गले के वोकल कॉर्ड (vocal cords) का उपयोग करता है। वोकल कॉर्ड गले के अंदर की वे सलिवटें हैं जो सबसे पहले आवाज बनाती हैं। इंसान ने स्वर निकालने के लिए अपने स्वरयंत्र (larynx) को नीचे की ओर विकसित किया। स्वरयंत्र गले के बीच का वह गुंजने वाला बक्सा है जिसके अंदर वोकल कॉर्ड होते हैं। इंसान के होंठ और जीभ भी बहुत बारीक हलचल करने के लिए विकसित हुए हैं। लेकिन स्पर्म् व्हेल इंसान की तरह गले के वोकल कॉर्ड को कंपित करके आवाज नहीं निकालती। इसके बजाय व्हेल ने अपने सरि के अंदर नाक के रास्ते में अपना एक अनूठा ध्वनि अंग विकसित किया। क्योंकि उसे गहरे समुद्र के पानी के भारी दबाव को झेलते हुए आवाज निकालनी थी।

स्पर्म् व्हेल का ध्वनि अंग दो हस्सियों की जोड़ी से बना है। पहला हस्सिया फोनिक लिप्स (phonic lips) यानी ध्वनि होठ है। यह नाक के रास्ते के अंदर की एक झुर्रीदार संरचना है। यह इंसान के वोकल कॉर्ड जैसा ही काम करता है। जब हवा यहां से गुजरती है, तो यह कंपित होकर पहली आवाज बनाती है। यह वैसा ही है जैसे गुब्बारे के मुंह को हाथ से पकड़कर हवा निकालने पर सीटी जैसी आवाज आती है। दूसरा हस्सिया तेल की थैली और हवा की थैली है। यह व्हेल के बड़े माथे को भरने वाला अंग है। पहली आवाज इन थैलियों से गुजरते हुए कई बार टकराती है। इस दौरान कोई खास उतार-चढ़ाव जोर से गुंज उठता है। यह वैसा ही है जैसे इंसान का मुंह आवाज को तराशने का काम करता है [3]।

ये दोनों हस्सिसे मलिकर आवाज बनाते हैं। जहाँ से आवाज पहली बार निकलती है, उसे सोर्स (source) कहा जाता है। आवाज को तराशने वाले रास्ते को फिल्टर (filter) कहा जाता है। गटिर के तार को छेड़कर आवाज निकालने वाली जगह सोर्स है। उस आवाज को खूबसूरती से गुंजाने वाला बक्सा फिल्टर है। इन दोनों के मिलने से आवाज का रूप तय होता है। इस सदिधांत को सोर्स-फिल्टर सदिधांत (source-filter theory) कहा जाता है। स्पर्म् व्हेल ने वोकल कॉर्ड का उपयोग किए बिना ही इस सदिधांत को अपने आप हासिल कर लिया [4]।

इस ध्वनि अंग का एक छपा हुआ पहलू और भी है। यह वही अंग है जिससे व्हेल कोमल स्वर निकालती है। लेकिन यही अंग बहुत शक्तिशाली आवाज भी निकालता है। उसकी तीव्रता लगभग 230 डेसबिल मापी जाती है। यह पानी के भीतर मापी गई जानवरों की आवाजों में सबसे शक्तिशाली आवाजों में से एक है। हालांकि पानी के भीतर के डेसबिल और हवा के डेसबिल का पैमाना अलग होता है। इसलिए दोनों मानों की सीधे तौर पर तुलना नहीं की जानी चाहिए। इस आवाज से इंसान को नुकसान पहुंचने की बात अभी तक साबित नहीं हुई है। व्हेल इस बड़ी शक्ति का इस्तेमाल एक-दूसरे के खिलाफ नहीं करती। इसके बजाय वे कोमल आवाज से एक-दूसरे का साथ नभाती हैं। इस तरह उन्होंने लंबे समय से समुद्र में अपनी दोस्ती को बनाए रखा है [5]।

संबंधित संदर्भ सामग्री

- [1] Beguš, G., Dąbkowski, M., Sprouse, R. L., Gruber, D. F., & Gero, S. (2026). The phonology of sperm whale coda vowels. *Proceedings of the Royal Society B*, 293(2069), 20252994. <https://doi.org/10.1098/rspb.2025.2994>
- [2] Madsen, P. T., Wahlberg, M., & Møhl, B. (2002). Male sperm whale acoustics in a high-latitude habitat: implications for echolocation and communication. *Behavioral Ecology and Sociobiology*, 53(1), 31-41. <https://doi.org/10.1007/s00265-002-0548-1>
- [3] Cranford, T. W., Amundin, M., & Norris, K. S. (1996). Functional morphology and homology in the odontocete nasal complex. *Journal of Morphology*, 228(3), 223-285. [https://doi.org/10.1002/\(SICI\)1097-4687\(199606\)228:3<223::AID-JMOR1>3.0.CO;2-3](https://doi.org/10.1002/(SICI)1097-4687(199606)228:3<223::AID-JMOR1>3.0.CO;2-3)
- [4] Fant, G. (1960). *Acoustic Theory of Speech Production*. Mouton.

2.5 तीन करोड़ साल का समुद्री पुस्तकालय और व्हेल के अधिकार

Project CETI और प्रोफेसर बेगुश का शोध हमें सोचने के लिए एक बड़ा वषिय देता है। स्पर्म व्हेल केवल अपनी स्वाभाविक प्रवृत्तियों के अनुसार अंधाधुंध चलिलाने वाले जीव नहीं हैं। वे समुद्र में चारों ओर आवाज के माध्यम से जुड़े हुए समूह हैं। वे अपनी-अपनी संस्कृति और पारिवारिक बंधनों को साझा करने वाले प्राणी हैं [1]।

ऐसा जुड़ाव देखने वाला एक अध्ययन मार्च 2026 में सामने आया। इसमें Project CETI ने स्पर्म व्हेल (sperm whale) के बच्चे के जन्म को वसितार से दर्ज किया। बच्चे के जन्म के समय दूसरी मादा व्हेलें भी वहाँ मौजूद थीं। उन्होंने पास रहकर माँ व्हेल की मदद की। यहाँ तक कि जो व्हेलें आपस में रश्तिदार नहीं थीं, उन्होंने भी मदद में हाथ बँटाया। प्राइमेट (primate) के अलावा किसी जानवर द्वारा बच्चे के जन्म में मदद करने को आंकड़ों से साबित करने वाला यह एक दुर्लभ उदाहरण है। इस दौरान व्हेलों की आवाज़ भी बदल गई। हर महत्वपूर्ण मोड़ पर उनकी आवाज़ का अंदाज़ बदल जाता था। उनमें स्वरों (vowels) जैसी आवाज़ें भी शामिल थीं। यह अध्ययन फरि से दिखाता है कि व्हेलें एक-दूसरे की मदद करने वाले जीव हैं [2]।

कैरेबियन सागर (Caribbean Sea) की व्हेलें अक्सर एक खास आवाज़ निकालती हैं। वैज्ञानिक इसे 1 जमा 1 जमा 3 coda कहते हैं। यह दो क्लिक, एक छोटा ठहराव और फरि तीन तेज़ क्लिक से मिलकर बनती है। शोधकर्ता इस आवाज़ का अर्थ सावधानी से समझने की कोशिश कर रहे हैं। उनका मानना है कि इसका मतलब 'मैं यहाँ हूँ' हो सकता है। यह आवाज़ कैरेबियन समूह द्वारा आपस में बोली जाने वाली एक सांस्कृतिक बोली है। अलग-अलग इलाकों के समूहों की बोली अलग होती है। दूसरी ओर, स्पर्म व्हेल का वंश करोड़ों वर्षों से समुद्र में खुद को ढालते हुए जी रहा है। उन्होंने बहुत लंबे समय से अंधेरे समुद्र में आवाज़ के ज़रिए एक-दूसरे को पहचाना है [3]।

व्हेलों की आवाज़ खतरे से आगाह करने वाला एक जाल भी थी। 1800 के दशक में लोग जहाजों पर सवार होकर व्हेलों का शिकार करने जाते थे। उस समय बच निकली कुछ व्हेलों ने यह खबर फैलाई। उन्होंने आवाज़ के ज़रिए एक-दूसरे को सखाया कि जहाजों से कैसे बचा जाए। इसका नतीजा यह हुआ कि व्हेल शिकार की सफलता दर बहुत गिर गई। ढाई साल में यह लगभग 58 प्रतिशत तक कम हो गई। यह रिकॉर्ड दिखाता है कि व्हेलों की आवाज़ एक बड़ा सूचना नेटवर्क है [4]।

AI द्वारा समझाई गई व्हेलों की भाषा एक नया सवाल खड़ा करती है। वह यह कि इंसानों को व्हेलों के साथ कैसा व्यवहार करना चाहिए। इसलिए व्हेलों को कानूनी व्यक्तित्व (legal personhood) देने का आंदोलन बढ़ रहा है। कानूनी व्यक्तित्व का अर्थ है इंसानों की तरह कानून द्वारा सुरक्षा पाने का अधिकार। इसका मतलब है उन्हें किसी चीज़ की तरह नहीं, बल्कि अधिकारों के हकदार के रूप में देखना। मार्च 2024 में, न्यूजीलैंड और प्रशांत महासागर के द्वीपों के मूल निवासी नेताओं ने एक घोषणापत्र जारी किया। यह व्हेलों को कानूनी व्यक्तित्व के रूप में मान्यता देने वाली घोषणा है। इसका नाम 'हे व्हाकापुतांगा मोआना' (He Whakaputanga Moana) है। हालाँकि, यह घोषणा कोई कानूनी रूप से बाध्यकारी नियम नहीं है। यह आंदोलन Project CETI के साथ जुड़ गया। अमेरिका के न्यूयॉर्क यूनिवर्सिटी लॉ स्कूल का MOTH कार्यक्रम भी इसमें शामिल हुआ। MOTH इंसान से परे के जीवन को दर्शाने वाला नाम है [5][6]।

दूसरी तरफ, इंसानों द्वारा पैदा किया गया शोर व्हेलों की आवाज़ में रुकावट डालता है। यह जहाजों के प्रोपेलर और समुद्री खोजबीन का शोर है। यह शोर व्हेलों की आपसी बातचीत को दबा देता है। इससे छोटी व्हेलों के लिए अपने समूह की भाषा सीखना मुश्किल हो जाता है। उनकी संस्कृति अगली पीढ़ी तक नहीं पहुँच पाती है। Earth Species Project भी इस तरह के शोध में मदद कर रहा है। उन्होंने कई जानवरों की आवाज़ों को एक साथ समझने वाला AI बनाया है। इसका नाम NatureLM-audio है [7][8]।

हमें क्या करना है, यह बलिकुल स्पष्ट है। हमें व्हेलों की ओर लापरवाही से बनाई गई आवाज़ें भेजकर उनकी संस्कृति को नुकसान नहीं पहुँचाना चाहिए। इसके बजाय, हमें उन आवाज़ों को ध्यान से सुनना चाहिए जो उन्होंने इतने लंबे समय से समुद्र में सँजोई हैं। यह करोड़ों सालों से चली आ रही समुद्र की एक लाइब्रेरी है। सबसे पहले उसकी धुन को शांत से सुनना ज़रूरी है। यही इंसानों और व्हेलों के साथ मिलकर रहने का रास्ता है।

संबंधित संदर्भ सामग्री

[1] Whitehead, H. (2003). *Sperm Whales: Social Evolution in the Ocean*. University of Chicago Press.

[2] EurekAlert!. (2026-03). Studies documenting rare sperm whale birth and ancient cooperative care released. EurekAlert! / Project CETI. <https://www.eurekalert.org/news-releases/1120823>

[3] National Geographic. (2026). Whales could one day be heard in court and in their own words. National Geographic. <https://www.nationalgeographic.com/animals/article/whale-communication-legal-personhood>

- [4] Whitehead, H., Smith, T. D., & Rendell, L. (2021). Adaptation of sperm whales to open-boat whalers: rapid social learning on a large scale? *Biology Letters*, 17(3), 20210030. <https://doi.org/10.1098/rsbl.2021.0030>
- [5] Rodríguez-Garavito, C. (2026). Animal Communication, AI, and the Future of Nonhuman Rights. NYU Law MOTH Program. <https://www.law.nyu.edu/news/ideas/cesar-rodriguez-garavito-animal-communication-ai>
- [6] Te Ao Māori News. (2024-03-28). Māori king and other indigenous Pacific leaders sign up to granting whales legal personhood. Te Ao Māori News. <https://www.teaonews.co.nz/2024/03/28/maori-king-and-other-indigenous-pacific-leaders-sign-up-to-granting-whales-legal-personhood/>
- [7] Erbe, C., Marley, S. A., Schoeman, R. P., Smith, J. N., Trigg, L. E., & Embling, C. B. (2019). The Effects of Ship Noise on Marine Mammals—A Review. *Frontiers in Marine Science*, 6, 606. <https://doi.org/10.3389/fmars.2019.00606>
- [8] Earth Species Project. (2025). Using AI to Understand Animal Communication. Earth Species Project. <https://www.earthspecies.org/blog/using-ai-to-understand-animal-communication>

अध्याय 3. हम्पबैक व्हेल ट्वेन से मुलाकात

3.1 अलौकिक बुद्धिमत्ता की ओर पृथ्वी का पूर्वाभ्यास

सोचिए कि अंतरिक्ष में बहुत दूर से एक संकेत (signal) आ रहा है। क्या वह संकेत सिर्फ एक शोर है, या किसी का भेजा हुआ कोई संदेश? शोर बना किसी अर्थ वाली उलझी हुई आवाज़ होती है। हमारे पास उस भाषा को समझाने वाला कोई शब्दकोश नहीं है। यानी उसका अर्थ बताने वाली एक भी चाबी हमारे पास नहीं है।

इस सवाल का जवाब ढूंढने के लिए कुछ लोगों ने आसमान की जगह समुद्र की ओर रुख किया। एक ऐसा संस्थान है जो अंतरिक्ष से आने वाले बुद्धिमत्ता के संकेतों को खोजता है। इसका नाम SETI इंस्टीट्यूट (SETI Institute) है। इस संस्थान और अमेरिका के कैलिफ़ोर्निया विश्वविद्यालय, डेविस (UC Davis) के शोधकर्ताओं ने हाथ मिलाया। व्हेल की आवाज़ों का लंबे समय से अध्ययन करने वाले वैज्ञानिकों ने भी उनका साथ दिया। इन्होंने मलिकर 'Whale-SETI' नाम की एक रिसर्च टीम बनाई [1]।

इस शोध का नेतृत्व प्रोफ़ेसर ब्रेन्डा मैककोवन (Brenda McCowan) ने किया। SETI इंस्टीट्यूट के डॉ. लॉरेंस डॉयल (Laurance Doyle) भी उनके साथ थे। हम्पबैक व्हेल विशेषज्ञ डॉ. फ्रेड शार्प (Fred Sharpe) भी इस टीम में शामिल हुए। इन सभी की एक ही सोच थी। वह यह कि अंतरिक्ष के किसी अनजान जीव से मिलने से पहले, हमें पृथ्वी पर अपने पड़ोसियों को समझना चाहिए।

वह पड़ोसी बालीन व्हेल की एक प्रजाति, हम्पबैक व्हेल (Megaptera novaeangliae) है। हम्पबैक व्हेल लंबे गीत गाने वाला एक बड़ा समुद्री जीव है। ये गीत मुख्य रूप से नर व्हेल गाते हैं। एक ही झुंड के नर एक-दूसरे से गाना सीखते हैं। इस तरह यह गीत फैलता है और धीरे-धीरे बदलता जाता है। हम्पबैक व्हेल हर साल विशाल समुद्र को पार करते हुए आवाज़ों के ज़रिये आपस में संपर्क भी बनाए रखते हैं [2]।

शोधकर्ताओं की योजना इस प्रकार थी। अलास्का (Alaska) के समुद्र में एक छोटी नाव उतारी जाती है। नाव का इंजन पूरी तरह बंद कर दिया जाता है। पानी के अंदर माइक्रोफ़ोन और स्पीकर उतारे जाते हैं। फिर व्हेल की एक आवाज़ बजाई जाती है। पहले से रिकॉर्ड की गई आवाज़ को दोबारा सुनाकर यह देखने का एक तरीका है कि जानवर क्या करता है। इस तरीके को प्लेबैक प्रयोग (playback) कहा जाता है [3]।

सुनाई गई आवाज़ एक 'संपर्क कॉल' थी। यह वह आवाज़ है जो व्हेल अपने झुंड में अपनी जगह बताने के लिए निकालती है। इस आवाज़ को 'whup call' कहा जाता है। यह खतरे की नहीं, बल्कि एक शांत आवाज़ है। इंसानी भाषा में कहे तो इसे "मैं यहाँ हूँ" जैसी आवाज़ माना जाता है।

यह आवाज़ ठीक एक दिन पहले उसी समुद्र में रिकॉर्ड की गई एक असली व्हेल की आवाज़ थी। उस झुंड में 38 साल की एक मादा हम्पबैक व्हेल थी। उसका नाम 'ट्वेन (Twain)' था। ट्वेन वह व्हेल थी जिस पर वैज्ञानिकों का काफी समय से नज़र रख रहे थे।

शोधकर्ताओं ने व्हेल को ही क्यों चुना? व्हेल इंसानों की तरह लंबी उम्र जीती है। वे झुंड बनाकर एक साथ रहती हैं। वे आवाज़ों के ज़रिये एक-दूसरे से संपर्क में रहती हैं। वे सीखे हुए गीतों को दूसरी व्हेलों को भी सिखाती हैं। ये बातें इंसानों से बहुत मिलती-जुलती हैं। इसलिए शोधकर्ताओं ने माना कि किसी अनजान बुद्धिमत्ता को समझने के अभ्यास के लिए व्हेल सबसे सही साथी है।

अंतरिक्ष के संकेतों को समझने में एक बड़ी मुश्किल है। क्योंकि संकेत भेजने वाले और पाने वाले के बीच पहले से कोई नियम तय नहीं होता। अगर दो भाषाओं को आमने-सामने रखकर लिखी गई कोई किताब हो, तो अनुवाद करना आसान होता है। लेकिन अंतरिक्ष के संकेतों के लिए ऐसी कोई किताब नहीं है। इसलिए नियमों को बिल्कुल शुरुआत से ही खोजना पड़ता है।

प्रयोग से ठीक पहले नाव पर बिल्कुल सन्नाटा था। शोधकर्ताओं ने सांस रोककर पानी की सतह की ओर देखा। कोई नहीं जानता था कि ट्वेन कोई जवाब देगी या नहीं।

संबंधित संदर्भ सामग्री

[1] SETI Institute. (2023-12-12). Whale-SETI: Groundbreaking encounter with humpback whales reveals potential for non-human intelligence communication. SETI Institute. <https://www.seti.org/news/whale-seti-groundbreaking-encounter-with-humpback-whales-reveals-potential-for-non-human-intelligence-communication/>

[2] McCowan, B., Hubbard, J., Walker, L., Sharpe, F., Frediani, J., & Doyle, L. (2023). Interactive bioacoustic playback as a tool for detecting and exploring nonhuman intelligence: "Conversing" with an Alaskan humpback whale. PeerJ, 11, e16349. <https://doi.org/10.7717/peerj.16349>

[3] Ralls, E. (2026-04-29). AI helps humans have a 20-minute 'conversation' with a humpback whale named Twain. Earth.com. <https://www.earth.com/animals/whale-seti-ai-helps-humans-have-20-minute-conversation-with-humpback-whale-twain/>

3.2 ट्वेन के साथ 20 मिनट

स्पीकर से पहली आवाज़ गूँजी। कुछ ही देर बाद ट्वेन वहाँ आ पहुँची। ट्वेन चुपचाप नाव की ओर बढ़ी। उसने नाव के चारों ओर एक बड़ा चक्कर लगाया। और फिर स्पीकर की ओर अपनी आवाज़ नकाली।

ट्वेन एक बार आवाज़ नकालकर गायब नहीं हुई। जब इंसानों ने आवाज़ सुनाई, तो ट्वेन ने भी आवाज़ से जवाब दिया। जब इंसान रुकते, तो ट्वेन भी इंतज़ार करती। जब फिर से आवाज़ आती, तो वह फिर जवाब देती। यह बातचीत 20 मिनट तक चलती रही। 20 मिनट का समय एक छोटी कार्टून फ़िल्म जतिना होता है।

शोधकर्ताओं ने 20 मिनट में 38 बार whup call बजाई। ट्वेन ने whup call के ज़रिये 36 बार आवाज़ नकाली। उनमें से 33 बार उसने इंसानों द्वारा सुनाई गई आवाज़ के तुरंत बाद जवाब दिया था। बाकी 3 बार उसने तब आवाज़ नकाली जब इंसानों ने आवाज़ बंद कर दी थी। इंसान और कसिी जंगली व्हेल के बीच इतने लंबे समय तक आवाज़ों का आदान-प्रदान पहली बार हुआ था।

ट्वेन ने आवाज़ों के बीच के अंतराल का भी तालमेल बठाया। अगर इंसान आवाज़ों के बीच लंबा अंतराल रखते, तो ट्वेन भी लंबा अंतराल रखती। अगर अंतराल छोटा होता, तो ट्वेन भी जल्दी जवाब देती। उसने सामने वाले के तरीके का पूरी तरह पालन किया था।

शोध का नेतृत्व करने वाली डॉ. मैककोवन ने यह बात कही। उन्होंने कहा कि यह इंसान और व्हेल के बीच हुआ एक बेहद सार्थक आवाज़ों का आदान-प्रदान लगता है। हालाँकि, इसका मतलब यह नहीं है कि इंसानों ने व्हेल की भाषा का अर्थ समझ लिया। शोधकर्ताओं का मानना था कि यह दोनों द्वारा अपनी-अपनी बारी का ध्यान रखकर आवाज़ों को साझा करना था।

ट्वेन पूरे 20 मिनट सिर्फ़ खुशी-खुशी ही आवाज़ें नहीं नकाल रही थी। शुरुआत में उसने बातचीत में बहुत अच्छे से हसि़सा लिया। लेकिन जब शोधकर्ताओं ने बार-बार एक ही whup call दोहराई, तो उसकी प्रतिक्रिया बदल गई। ट्वेन धीरे-धीरे बेचैन दिखने लगी। वह तेज़ी से सांस लेने लगी और पानी की सतह पर थपकियाँ मारने लगी। और आखिरकार वह उस जगह से चली गई। उसने भागीदारी, बेचैनी और चले जाने के तीन चरण दिखाए।

इस मुलाकात की कुछ स्पष्ट सीमाएँ थीं। प्रयोग सिर्फ़ एक व्हेल, ट्वेन पर हुआ था। यह प्रयोग भी संयोग से मिला सिर्फ़ एक मौका था। इसलिए सिर्फ़ इस एक नतीजे से यह पक्के तौर पर नहीं कहा जा सकता कि व्हेल बातचीत कर रही थी। शोधकर्ताओं ने भी इस बात को बहुत सावधानी से स्पष्ट किया।

फिर भी इस मुलाकात ने प्लेबैक प्रयोग की शक्त को साबित कर दिया। यह रिकॉर्ड की गई आवाज़ सुनाकर प्रतिक्रिया देखने का एक तरीका है। इस तरीके से जंगली जानवरों की प्रतिक्रिया को करीब से देखा जा सकता है। शोधकर्ताओं ने इसे एक नया शोध उपकरण बताया।

संबंधित संदर्भ सामग्री

- [1] SETI Institute. (2023-12-12). Whale-SETI: Groundbreaking encounter with humpback whales reveals potential for non-human intelligence communication. SETI Institute. <https://www.seti.org/news/whale-seti-groundbreaking-encounter-with-humpback-whales-reveals-potential-for-non-human-intelligence-communication/>
- [2] McCowan, B., Hubbard, J., Walker, L., Sharpe, F., Frediani, J., & Doyle, L. (2023). Interactive bioacoustic playback as a tool for detecting and exploring nonhuman intelligence: "Conversing" with an Alaskan humpback whale. *PeerJ*, 11, e16349. <https://doi.org/10.7717/peerj.16349>
- [3] Frediani, J. (2024). "Conversing" with a humpback whale named Twain. The Safina Center. <https://www.safinacenter.org/blog/conversing-with-a-humpback-whale-named-twain>

3.3 बारी का पालन करने वाली बातचीत

इंसान जब आपस में बात करते हैं, तो नियमों का पालन करते हैं। वे सामने वाले की बात खत्म होने तक इंतज़ार करते हैं। जब सामने वाला रुकता है, तब वे अपनी बात शुरू करते हैं। इस तरह बारी-बारी से बोलने के नियम को टर्न-टेकिंग (turn-taking) कहते हैं। यह स्कूल में खाना लेने के लिए कतार में खड़े होकर अपनी बारी का इंतज़ार करने जैसा है।

जब सामने वाला बात खत्म करता है, तो एक पल के लिए सन्नाटा छा जाता है। उस छोटे से ठहराव को बातचीत का वरिम (conversational pause) कहते हैं। इस वरिम की वजह से ही दोनों की बातें एक-दूसरे के ऊपर नहीं टकराती हैं। हम्पबैक व्हेल भी खुले समुद्र में इस नियम का पालन करती हैं।

ट्वेन ने 20 मिनट की बातचीत में इस नियम को कभी नहीं तोड़ा। जब इंसानों की आवाज़ बज रही थी, तब वह बीच में नहीं बोली। उसने आवाज़ के पूरी तरह खत्म होने का इंतज़ार किया। उसने शांत होने वाले उसी पल को चुना। और उस खाली जगह में अपनी आवाज़ शामिल की।

शोधकर्ताओं ने ट्वेन द्वारा जवाब देने में लगे समय को मापा। उन्होंने देखा कि आवाज़ खत्म होने के कतिने सेकंड बाद जवाब आता है। इस काम के लिए AI का इस्तेमाल नहीं किया गया था। उन्होंने रिकॉर्ड की गई आवाज़ और सांख्यिकी (statistics) की मदद से एक-एक सेकंड की जाँच की। वह समय इंसानों की बातचीत के नियमों से बिल्कुल मेल खाता था।

यह नतीजा हमें एक महत्वपूर्ण बात बताता है। वह यह कि ट्वेन मशीन की आवाज़ से चढ़कर नहीं चली रही थी। ट्वेन ने सामने वाले की बारी और अपनी बारी का अंतर समझकर आवाज़ निकाली थी। शोधकर्ताओं ने माना कि ट्वेन सचमुच आवाज़ों के इस आदान-प्रदान में भाग ले रही थी।

इंसान इस नियम को बचपन से ही सीख लेते हैं। वे अपनी माँ से नज़रें मलाकर आवाज़ों का आदान-प्रदान करते हैं। बोलना सीखने से पहले ही वे इस क्रम को सीख लेते हैं। यह बातचीत की बुनियाद बनने वाला नियम है।

बातचीत का यह वरिष्ठ बहुत ही कम समय का होता है। कभी-कभी यह पलक झपकने से भी छोटा होता है। ट्वेन ने इस छोटे से अंतराल को कभी नहीं चूका। उसने बिल्कुल सही समय पर अपनी आवाज़ निकाली।

ऐसी सटीकता बर्ना अभ्यास के संभव नहीं है। ट्वेन खुले समुद्र में हमेशा से इसी तरह आवाज़ों को साझा करती आई थी। उसने इसी जाने-पहचाने तरीके को इंसानों की आवाज़ के साथ भी दोहराया। सामने कोई असली व्हेल न होने पर भी उसने नियमों का पालन किया था।

संबंधित संदर्भ सामग्री

- [1] McCowan, B., Hubbard, J., Walker, L., Sharpe, F., Frediani, J., & Doyle, L. (2023). Interactive bioacoustic playback as a tool for detecting and exploring nonhuman intelligence: "Conversing" with an Alaskan humpback whale. *PeerJ*, 11, e16349. <https://doi.org/10.7717/peerj.16349>
- [2] SETI Institute. (2023-12-12). Whale-SETI: Groundbreaking encounter with humpback whales reveals potential for non-human intelligence communication. SETI Institute. <https://www.seti.org/news/whale-seti-groundbreaking-encounter-with-humpback-whales-reveals-potential-for-non-human-intelligence-communication/>
- [3] Ralls, E. (2024-01-06). Scientists have a 20-minute 'conversation' with a humpback whale named Twain. *Earth.com*. <https://www.earth.com/news/scientists-have-20-minute-conversation-with-a-humpback-whale-named-twain/>

3.4 गणति द्वारा उजागर किया गया व्हेलों का व्याकरण

डॉ. डॉयल ने व्हेलों के गीतों को संख्याओं के ज़रिये समझने की कोशिश की। क्योंकि उनका मानना था कि व्हेलों के गीत सरिफ़ बेतरतीब आवाज़ें नहीं हैं। इसलिए उन्होंने सूचना सैद्धांत (information theory) नामक उपकरण का इस्तेमाल किया। सूचना सैद्धांत यह मापने का तरीका है कि किसी आवाज़ या अक्षर में कतिने नियम छिपे हैं।

शोधकर्ताओं ने जपिफ़ का नियम (Zipf's law) व्हेलों के गीतों पर लागू करके देखा। जपिफ़ का नियम भाषा में छिपे एक नियम को बताता है। इंसानी भाषा में अक्सर इस्तेमाल होने वाले शब्दों और कभी-कभार इस्तेमाल होने वाले शब्दों के बीच एक निश्चित अनुपात होता है। उदाहरण के लिए, हमारी भाषा में '은, 는, 이, 가' जैसे शब्द बहुत बार आते हैं। कठिनि शब्द कभी-कभार ही आते हैं। इस अधिकता और कमी के बीच हमेशा एक जैसा अनुपात बना रहता है।

2025 में एक ऐसा अध्ययन सामने आया जिसने व्हेलों के गीतों में यह नियम खोज निकाला। इनबाल अर्नोन (Inbal Arnon) की शोध टीम ने हम्पबैक व्हेल के गीतों का लंबे समय तक अध्ययन किया। ये गीत 8 साल तक दक्षिणी प्रशांत महासागर से इकट्ठे किए गए थे। वे गीत जपिफ़ के नियम के अनुसार बिल्कुल सही बैठते थे। इंसानी भाषा के अलावा किसी और जगह यह नियम इतनी स्पष्टता से पहली बार पाया गया था।

इस परिणाम का क्या मतलब है? यह इस बात का संकेत हो सकता है कि व्हेलों के गीतों में नियम मौजूद हैं। लेकिन सरिफ़ इससे यह साबित नहीं होता कि व्हेलों के गीतों का कोई अर्थ या व्याकरण है। शोधकर्ताओं ने भी यह नहीं कहा कि व्हेलों के गीत इंसानी भाषा जैसे हैं। बल्कि उन्होंने माना कि यह इंसानी संगीत के ज़्यादा करीब हो सकते हैं।

शोधकर्ताओं ने शैनन एन्ट्रॉपी (Shannon entropy) की भी जाँच की। शैनन एन्ट्रॉपी यह मापने का एक गणितीय तरीका है कि अगली आवाज़ का अंदाज़ा लगाना कतिना कठिन है। अगर आवाज़ पूरी तरह बेतरतीब हो, तो यह मान बहुत अधिक होता है। यदि केवल एक ही आवाज़ बार-बार दोहराई जाए, तो यह मान 0 के करीब होता है। व्हेलों के गीत इन दोनों छोरों के बिल्कुल बीच में थे। वे नियमों का पालन तो कर रहे थे, लेकिन हमेशा एक जैसे भी नहीं थे।

ऐसी संरचना दूसरी व्हेलों के अध्ययनों में भी दिखाई देती है। यहाँ थोड़ी देर के लिए दूसरी व्हेल का उदाहरण लेते हैं। Project CETI नाम की एक शोध टीम है। ये वैज्ञानिक स्पर्म् व्हेल का अध्ययन करते हैं। स्पर्म् व्हेल हम्पबैक व्हेल से अलग प्रजाति हैं। स्पर्म् व्हेल क्लिक जैसी आवाज़ों को छोटे समूहों में निकालती हैं। आवाज़ों के इस समूह को coda कहते हैं।

2024 में Project CETI के शोधकर्ताओं ने coda की संरचना का खुलासा किया। स्पर्म व्हेल लय, गति, रुबाटो और आभूषण स्वर जैसे चार तत्वों को मिलाकर आवाज़ें बनाती हैं। रुबाटो का मतलब ताल को थोड़ा लंबा या छोटा करना है। आभूषण स्वर का मतलब आवाज़ में एक और अतिरिक्त क्लिक जोड़ना है। वे इन चार तत्वों को अलग-अलग तरह से मिलाकर बहुत सारी आवाज़ें बनाती हैं। पहले माना जाता था कि केवल इंसान ही ऐसे संयोजन का इस्तेमाल करते हैं, लेकिन स्पर्म व्हेल भी ऐसा ही करती थी [4]।

संबंधित संदर्भ सामग्री

- [1] Doyle, L. R., & McCowan, B. (2011). Information theory, animal communication, and the search for extraterrestrial intelligence. *Acta Astronautica*, 68(3-4), 406-417.
<https://www.sciencedirect.com/science/article/abs/pii/S0094576509005682>
- [2] Arnon, I., et al. (2025). Whale song shows language-like statistical structure. *Science*, 387.
<https://doi.org/10.1126/science.adq7055>
- [3] Sharma, P., et al. (2024). Contextual and combinatorial structure in sperm whale vocalisations. *Nature Communications*, 15, 3617. <https://doi.org/10.1038/s41467-024-47221-8>
- [4] Project CETI. (2024-05-07). Sperm whale phonetic alphabet proposed for the first time. Project CETI.
<https://www.projectceti.org/blog-posts/sperm-whale-phonetic-alphabet-proposed-for-the-first-time>

3.5 अंतरिक्ष को भेजा जाने वाला पत्र

Whale-SETI शोध का एक अंतिम लक्ष्य है। यह सिर्फ व्हेलों को समझने तक ही सीमित नहीं है। इसका लक्ष्य एक ऐसा उपकरण बनाना है जो किसी दूरी अंतरिक्ष से आने वाले अनजान संकेतों की पहचान कर सके। क्योंकि शोर और अर्थपूर्ण संकेतों के बीच अंतर करने के लिए एक पैमाने की जरूरत होती है।

व्हेलों के गीतों को गणति से समझने के कई तरीके हैं। जैसे जपिफ का नियम या शैनन एन्ट्रॉपी जैसे तरीके। भविष्य में इन तरीकों का इस्तेमाल अंतरिक्ष के संकेतों पर किया जा सकता है। हालाँकि, अभी तक इनका सीधा इस्तेमाल अंतरिक्ष के संकेतों पर नहीं किया गया है। वैज्ञानिक इस बात पर शोध कर रहे हैं कि भविष्य में इनका इस्तेमाल कैसे किया जाए।

कई वैज्ञानिकों ने इंसानों के घमंड की ओर इशारा किया है। उनका कहना है कि हम उसी पृथ्वी पर रहने वाली व्हेल की आवाज़ तक नहीं समझ पाते। फिर हम अंतरिक्ष के उस पार से आने वाले संकेतों को कैसे समझ पाएंगे? यह सवाल इंसानों को विनम्र बनाता है।

प्राचीन समय में इंसान मानते थे कि पृथ्वी ब्रह्मांड का केंद्र है। लेकिन वह धारणा गलत साबित हुई। केवल इंसान ही बातचीत कर सकते हैं, यह विश्वास भी कुछ वैसा ही हो सकता है। समुद्र में भी आवाज़ों के ज़रिये संवाद करने वाले पड़ोसी मौजूद हो सकते हैं।

इंसानों ने लंबे समय तक व्हेलों के गीतों को सिर्फ शोर ही समझा। उन्होंने अनगिनत व्हेलों का शिकार भी किया। जहाजों और मशीनों से समुद्र को शोरगुल से भर दिया। फिर भी व्हेलों ने आपस में संवाद करना बंद नहीं किया। वे बहुत प्राचीन समय से ही एक-दूसरे को आवाज़ें पहुँचाती आ रही हैं।

इसलिए ट्वेन के साथ बतियाए गए वे 20 मिनट एक पूर्वाभ्यास थे। यह किसी अनजान बुद्धिमत्ता से मिलने पर इस्तेमाल किए जाने वाले तरीकों को पहले से सीखने का अवसर था। यह अंतरिक्ष में संदेश भेजने से पहले अपने पास की आवाज़ों को सुनने का प्रयास है। ध्यान से सुनने की शक्ति ही अनुवाद की शुरुआत है।

संबंधित संदर्भ सामग्री

- [1] SETI Institute. (2023-12-12). Whale-SETI: Groundbreaking encounter with humpback whales reveals potential for non-human intelligence communication. SETI Institute. <https://www.seti.org/news/whale-seti-groundbreaking-encounter-with-humpback-whales-reveals-potential-for-non-human-intelligence-communication/>
- [2] McCowan, B., Hubbard, J., Walker, L., Sharpe, F., Frediani, J., & Doyle, L. (2023). Interactive bioacoustic playback as a tool for detecting and exploring nonhuman intelligence: "Conversing" with an Alaskan humpback whale. *PeerJ*, 11, e16349. <https://doi.org/10.7717/peerj.16349>
- [3] Ralls, E. (2024-01-06). Scientists have a 20-minute 'conversation' with a humpback whale named Twain. *Earth.com*. <https://www.earth.com/news/scientists-have-20-minute-conversation-with-a-humpback-whale-named-twain/>

3.6 AI द्वारा खोला गया नया दरवाज़ा

ट्वेन से मुलाकात के बाद शोध तेज़ी से आगे बढ़ रहा है। क्योंकि AI व्हेलों की आवाज़ों का और गहराई से अध्ययन करने में मदद कर रहा है। जिस काम को करने में इंसानों को दशकों लग जाते, AI उसे बहुत तेज़ी से पूरा करने में मदद कर रहा है।

Earth Species Project नाम का एक संगठन है। वे जानवरों की आवाज़ों को समझने वाला AI बनाते हैं। उस AI का नाम NatureLM-audio है। यह उपकरण केवल आवाज़ सुनकर पहचान लेता है कि वह कौन सा जानवर है। यह आवाज़ों के प्रकारों को भी अलग-अलग करता है। यह काम यह पहले से ही बहुत अच्छे से कर रहा है।

इस AI में प्रायोगिक स्तर की क्षमताएँ भी हैं। जैसे कि रिकॉर्डिंग में कितने जानवर मौजूद हैं, यह गणिता। या किसी ऐसे जानवर की आवाज़ पहचानना जैसी इसने पहले कभी नहीं सीखा। ये काम अभी पूरी तरह सटीक नहीं हैं। यह सिर्फ़ अंदाज़ा लगाने से थोड़ा ही बेहतर अनुमान लगा पाता है।

AI में एक आधारभूत बड़ा AI होता है। ऐसे बड़े AI को फाउंडेशन मॉडल (foundation model) कहा जाता है। यह एक ऐसा AI है जिसका इस्तेमाल कई अलग-अलग कामों के लिए किया जा सकता है। Google द्वारा बनाया गया पर्च (Perch) इसका एक उदाहरण है।

AI ने व्हेल के शोध में कैसे मदद की, इस कहानी को समय के क्रम में बाँटकर देखा जा सकता है। सबसे पहले 2018 की बात है। Google और अमेरिकी राष्ट्रीय महासागरीय एवं वायुमंडलीय प्रशासन (NOAA) ने हाथ मिलाया। राष्ट्रीय महासागरीय एवं वायुमंडलीय प्रशासन समुद्र और मौसम का अध्ययन करने वाला एक अमेरिकी संस्थान है। दोनों ने 14 वर्षों के दौरान एकत्र की गई समुद्र के भीतर की आवाज़ों की मलिकर जाँच की। वह मात्रा 187,000 घंटों से भी अधिक थी। यह इतनी अधिक थी कि यदि कोई इंसान बना रुके लगातार सुने, तो भी लगभग 21 वर्ष लग जाते। AI ने इस विशाल डेटा में से हंपबैक व्हेल के गीतों को छोटकर निकाला।

अगली बात 2025 की है। Google ने पर्च (Perch) को सुधारकर पर्च 2.0 (Perch 2.0) पेश किया। पर्च शुरू में पक्षियों की आवाज़ों से सीखने वाला AI है। लेकिन यह समुद्र के भीतर की आवाज़ों पर भी बिल्कुल सटीक बैठा। इसने व्हेल की आवाज़ों को पहचानने में बेहतरीन प्रणाम दिए। पक्षियों से सीखने वाले AI ने इस तरह व्हेल के शोध में भी मदद की।

AI की सबसे बड़ी ताकत उसकी गति में है। जिस आवाज़ को सुनने में इंसान को कई महीने लगते, AI उसे कुछ ही घंटों में जाँच लेता है। इससे शोध बहुत तेज़ हो गया। यह उन धीमी आवाज़ों को भी ढूँढ़ निकालता है जिनमें इंसान अक्सर छोड़ देते थे। इसी की बदौलत पुराने डेटा से भी नए तथ्य सामने आते हैं।

यहाँ एक बात पर ध्यान देना ज़रूरी है। AI व्हेल की आवाज़ों को पहचान लेता है, इसका यह मतलब बिल्कुल नहीं है कि उसने व्हेल की भाषा को पूरी तरह समझ लिया है। उसने आवाज़ के भीतर के नियम तो खोज लिए हैं, लेकिन उनके अर्थ को अभी नहीं समझा है। व्हेल की भाषा का अर्थ पूरी तरह सुलझ गया है, इसका कोई प्रमाण अभी तक नहीं मिला है। इंसान और व्हेल के बीच के शब्दकोश का अभी सिर्फ़ पहला पन्ना ही खुला है।

Earth Species Project सिर्फ़ व्हेलों का ही अध्ययन नहीं करता है। यह कौवों की आवाज़ों की भी जाँच करता है। यह हाथियों की आवाज़ों की भी जाँच करता है। यह बेलुगा व्हेल (beluga whale) की आवाज़ों की भी जाँच करता है। यह एक ही साधन से कई जानवरों की आवाज़ों का विश्लेषण करता है। भविष्य में, इंसानों द्वारा बनाई गई आवाज़ों को जानवरों को सुनाने की भी योजना है। इसका उद्देश्य यह देखना है कि जानवर उस आवाज़ पर कैसा जवाब देते हैं। ट्वेन (Twain) के साथ आवाज़ों का आदान-प्रदान इसका पहला कदम था।

हाल ही में व्हेल के गीतों की बनावट को उजागर करने वाला एक शोध भी सामने आया है। यह 2026 में अमेरिका के सरिक्यूज़ विश्वविद्यालय (Syracuse University) की जूलिया ज़े (Julia Zeh) की शोध टीम द्वारा किया गया। इन्होंने गाते हुए नर हंपबैक व्हेल के शरीर पर एक छोटा रिकॉर्डिंग उपकरण लगाया। यह संक्षेप की मदद से कुछ देर के लिए चपिकाया जाने वाला रिकॉर्डर और गति रिकॉर्डर था। यह शोध ट्वेन (Twain) के साथ आवाज़ों के आदान-प्रदान से अलग विषय पर है। इसमें बताया गया है कि व्हेल के पानी में ऊपर-नीचे तैरने की गति और उसके गीतों का आपस में संबंध है। गहराई में जाते समय और ऊपर आते समय गाए जाने वाले गीत अलग-अलग थे। इसका मतलब है कि गीत शरीर की गतिविधियों के अनुसार तैयार किए गए थे [5]।

इसलिए यह शोध आज भी जारी है। इंसान और व्हेल के बीच के शब्दकोश को एक-एक पन्ना करके भरा जा रहा है। ट्वेन (Twain) के साथ बीते 20 मिनटों ने उस शब्दकोश का पहला वाक्य शुरू किया था।

संबंधित संदर्भ सामग्री

[1] Earth Species Project. (2024-11-11). Introducing NatureLM-audio: An audio-language foundation model for bioacoustics. Earth Species Project. <https://www.earthspecies.org/blog/introducing-naturelm-audio-an-audio-language-foundation-model-for-bioacoustics>

- [2] Robinson, D., Miron, M., Hagiwara, M., Weck, B., Keen, S., Alizadeh, M., Narula, G., Geist, M., & Pietquin, O. (2025). NatureLM-audio: An audio-language foundation model for bioacoustics. International Conference on Learning Representations (ICLR) 2025. <https://arxiv.org/abs/2411.07186>
- [3] Google Research. (2018-10). Acoustic detection of humpback whales using a convolutional neural network. Google Research Blog. <https://research.google/blog/acoustic-detection-of-humpback-whales-using-a-convolutional-neural-network/>
- [4] Google Research. (2026-02-09). How AI trained on birds is surfacing underwater mysteries. Google Research Blog. <https://research.google/blog/how-ai-trained-on-birds-is-surfacing-underwater-mysteries/>
- [5] Syracuse University. (2026-08-31). How a whale's dive shapes its song. Syracuse University News. <https://news.syr.edu/2026/08/31/how-a-whales-dive-shapes-its-song/>

अध्याय 4. गहरे समुद्र का ऑर्केस्ट्रा

4.1 शोरगुल भरे समुद्र से आवाज़ें चुनना

समुद्र कोई शांत जगह नहीं है। पानी के अंदर की जगह आवाज़ों से भरा एक विशाल स्थान है। व्हेल के गीत, डॉल्फिन की सीटी और लहरो की आवाज़ें बना-बुके आपस में मलित रहती हैं। पानी के अंदर आवाज़ हवा की तुलना में बहुत तेज़ी से फैलती है। यह हवा से लगभग 4.5 गुना तेज़ होती है। इसलिए व्हेल अपनी आँखों से ज़्यादा कानों से दुनिया को समझती हैं। दूर समुद्र में एक-दूसरे को ढूँढ़ने के लिए वे आवाज़ का इस्तेमाल करती हैं। शिकार का पीछा करने और ख़तरे से बचने के लिए भी वे आवाज़ का ही इस्तेमाल करती हैं [1][2]।

पानी में आवाज़ 1 सेकंड में लगभग 1,500 मीटर की दूरी तय करती है। यह ज़मीन पर आवाज़ द्वारा तय होने वाली दूरी से कहीं ज़्यादा दूर है। इसलिए व्हेल बहुत दूर मौजूद अपने दोस्तों की आवाज़ भी सुन सकती हैं।

समस्या इंसानों द्वारा पैदा किया गया शोर है। बड़े जहाजों के प्रोपेलर की आवाज़ पानी को हलाकर रख देती है। समुद्र की गहराई की जाँच करने वाले सोनार (sonar) की आवाज़ भी बहुत तेज़ होती है। समुद्र तल के संसाधनों की खोज के लिए होने वाले वसिफोटों की आवाज़ें भी होती हैं। ऐसा शोर हजारों किलोमीटर दूर तक फैलता है। आज जहाजों की आवाज़ समुद्र में सबसे व्यापक रूप से फैला हुआ इंसानी शोर है। जब शोर बहुत तेज़ होता है, तो व्हेल की आवाज़ उसी में दब जाती है। वैज्ञानिक इसे आवाज़ का दब जाना कहते हैं। आवाज़ के सहारे जीने वाली व्हेल के लिए यह बहुत बड़ा ख़तरा है।

शोर व्हेल के शरीर को भी बीमार कर देता है। 2001 में अमेरिका में एक बड़ा हादसा होने के बाद, जहाजों की आवाज़ाही कुछ समय के लिए बहुत कम हो गई थी। उस समय समुद्र शांत हो गया था। रोज़ालिंड रोलैंड (Rosalind Rolland) की शोध टीम ने इस मौक़े का अध्ययन किया। उन्होंने कनाडा की फंडी की खाड़ी में रहने वाली राइट व्हेल के मल को इकट्ठा करके जाँच की। जैसे ही समुद्र शांत हुआ, व्हेल के शरीर में तनाव पैदा करने वाले तत्व कम हो गए। यह ठीक वैसा ही है जैसे किसी शोरगुल वाले निर्माण स्थल के पास रहने पर इंसान का दिल तेज़ी से धड़कने लगता है। यह इस बात का पहला सबूत था कि शोर व्हेल को थका देता है [3]।

कई आवाज़ों के घालमेल में से किसी एक व्हेल की आवाज़ को छानना बहुत मुश्किल होता है। वैज्ञानिक इस कठिनाई को 'कॉकटेल पार्टी समस्या' कहते हैं। यह लोगों से भरे कमरे में सिर्फ़ अपने एक दोस्त की बात को ध्यान से सुनने जैसा है। यह ठीक वैसा ही है जैसे स्कूल के कैफ़ेटरिया में शोरगुल होने पर भी पास बैठे दोस्त की बात सुनाई दे जाती है। पुरानी मशीनें यह काम ठीक से नहीं कर पाती थीं। जब जहाजों की धातु की खड़खड़ाहट व्हेल की आवाज़ से टकराती थी, तो वे अक्सर उलझन में पड़ जाती थीं।

AI ने इस समस्या को नए सरि से सुलझाया। यह आपस में मल्लि-जुली आवाज़ों में से मनचाही आवाज़ को अलग करने की तकनीक है। मशीन सबसे पहले आवाज़ को संख्याओं के एक नक्शे में बदलती है। इस नक्शे पर जहाजों की आवाज़ और व्हेल की आवाज़ अलग-अलग जगहों पर रखी जाती है। यह ठीक वैसा ही है जैसे रंग-बरिगे कंचे आपस में मलि जाने पर भी उन्हें रंग के हिसाब से अलग-अलग डबिबों में रखना। मशीन उन जगहों को बाँटकर केवल व्हेल की आवाज़ को फरि से इकट्ठा कर लेती है। Earth Species Project के शोधकर्ताओं ने साफ़ और सही आवाज़ का नमूना न होने पर भी शोर को हटाने का तरीका ढूँढ़ निकाला। इसकी बदौलत इंसान व्हेल की आवाज़ को और भी साफ़-साफ़ सुनने में सक्षम हो सके हैं [4]।

संबंधित संदर्भ सामग्री

- [1] Bakker, K. (2022). The sounds of life: How digital technology is bringing us closer to the worlds of animals and plants. Princeton University Press.
- [2] NOAA Fisheries. (2025). A whale's world of sound. National Oceanic and Atmospheric Administration. <https://www.fisheries.noaa.gov/feature-story/whales-world-sound>
- [3] Rolland, R. M., Parks, S. E., Hunt, K. E., et al. (2012). Evidence that ship noise increases stress in right whales. Proceedings of the Royal Society B, 279(1737), 2363-2368. <https://doi.org/10.1098/rspb.2011.2429>
- [4] Miron, M., Keen, S., Hoffman, B., Hagiwara, M., et al. (2024). Biodenoising: Animal vocalization denoising without access to clean data. Earth Species Project. arXiv:2410.03427. <https://arxiv.org/abs/2410.03427>

4.2 कलिर व्हेल परिवार की बोलियाँ

तेज़ शोर को हटाने के बाद, वैज्ञानिकों ने व्हेल की आवाज़ों में छपि रहस्यों को समझना शुरू किया। वह रहस्य है हर परिवार की अलग 'बोली'। कनाडा के पश्चिमी तट पर कलिर व्हेल का अध्ययन करने वाला एक शोध केंद्र है। इसका नाम OrcaLab है। यहाँ हेलेना साइमंड्स (Helena Symonds) और डॉ. पॉल स्पॉन्ग (Paul Spong) ने कई दशकों तक जंगली कलिर व्हेल की आवाज़ों को रिकॉर्ड किया है। व्हेल को पकड़े या परेशान किए बिना, उन्होंने पानी के नीचे के माइक्रोफ़ोन से चुपचाप केवल उनकी आवाज़ें इकट्ठा कीं [1]।

कलिर व्हेल अपनी माँ के इर्द-गिर्द रहने वाले जीव हैं। उनके बच्चे ज़िंदगी भर अपनी माँ के झुंड को छोड़कर कहीं नहीं जाते। इसलिए हर परिवार की अपनी एक अनोखी आवाज़ होती है, जो पीढ़ी-दर-पीढ़ी आगे बढ़ती है। वैज्ञानिक इसे माँ के परिवार द्वारा मलिकर बोली जाने वाली 'बोली' कहते हैं। यह ठीक वैसा ही है जैसे हमारी कक्षा और बगल की कक्षा के बच्चों के बोलने का अंदाज़ थोड़ा अलग होता है। इस बोली की मदद से कलिर व्हेल विशाल समुद्र में अपने परिवार को पहचान लेती हैं। यह नज़दीकी रश्तेदारों के बीच प्रजनन को रोकने का काम भी करती है।

OrcaLab ने बहुत लंबे समय तक आवाज़ों को इकट्ठा किया। यह रिकॉर्डिंग ही लगभग 20,000 घंटों की है। किसी इंसान के लिए यह सारी आवाज़ें सुन पाना बहुत मुश्किल है। इसलिए वैज्ञानिकों ने यह काम AI को सौंप दिया। जर्मनी और कनाडा के शोधकर्ताओं द्वारा बनाए गए प्रोग्राम का नाम ORCA-SPOT है। यह प्रोग्राम लंबी रिकॉर्डिंग में से कलिर व्हेल की आवाज़ को अपने आप ढूँढ़ निकालता है। इंसानों द्वारा तय किए गए सही उत्तरों से तुलना करने पर, इसने 100 में से 93 बार सही पहचान की। जिस काम में कई साल लग जाते, मशीन ने उसे केवल आठ दिनों में पूरा कर दिया [5]।

आवाज़ों को अलग-अलग करते समय AI महत्वपूर्ण हिससों पर खास ध्यान देता है। यह वैसा ही है जैसे कोई इंसान लंबे लेख में महत्वपूर्ण शब्दों के नीचे रेखा खींचता है। मशीन इस तरह बारीक लहजे के अंतर को भी पकड़ लेती है। कलिर व्हेल के बच्चे अपनी माँ और झुंड की आवाज़ें सुनकर अपनी बोली सीखते हैं। यह वैसा ही है जैसे इंसानी बच्चे अपने माता-पिता की बातों की नकल करके बोलना सीखते हैं। वैज्ञानिकों ने AI द्वारा बनाए गए आवाज़ के नक्शे को देखकर एक नया तथ्य भी खोजा। वह यह है कि एक ही परिवार की आवाज़ पीढ़ियों के साथ थोड़ी-थोड़ी बदलती रही है। यह वैसा ही है जैसे समय के साथ लंबे समय से बोली जाने वाली भाषा में थोड़ा बदलाव आ जाता है। इसका मतलब यह है कि व्हेल की आवाज़ सरिफ़ जन्मजात प्रवृत्ति नहीं है। यह इस बात का सबूत है कि यह इंसानी भाषा की तरह सीखी जाने वाली, आगे बढ़ाई जाने वाली और बदलने वाली एक संस्कृति है।

इन बोलियों को समझना व्हेल की रक्षा करने में भी काम आता है। क्योंकि आवाज़ से यह पता चल जाता है कि कौन सा परिवार कहाँ रहता है। जहाज़ों की लगातार आवाज़ाही वाले समुद्री इलाकों में कौन सा परिवार मौजूद है, इसकी नगिरानी की जा सकती है। आवाज़ के ज़रिए परिवारों की गिनती करने से यह जल्दी पता चल जाता है कि उनकी संख्या घट रही है या नहीं। यह शोध आज भी जारी है। 2026 में कलिर व्हेल की आवाज़ों को अपने आप खोजने और छांटने वाला एक नया प्रोग्राम orcaAI सामने आया। इस प्रोग्राम ने कनाडा की नहीं, बल्कि आइसलैंड के समुद्र की कलिर व्हेल की आवाज़ें सीखीं। वहाँ की कलिर व्हेल हेरिंग (herring) मछलियों का शिकार करके जीती हैं। इंसान को लंबे समय तक सुनने की ज़रूरत नहीं पड़ती, मशीन खुद ही लंबी रिकॉर्डिंग की जाँच कर लेती है [6]।

संबंधित संदर्भ सामग्री

- [1] Bakker, K. (2022). The sounds of life: How digital technology is bringing us closer to the worlds of animals and plants. Princeton University Press.
- [5] Bergler, C., Schröter, H., Cheng, R. X., et al. (2019). ORCA-SPOT: An automatic killer whale sound detection toolkit using deep learning. Scientific Reports, 9, 10997. <https://doi.org/10.1038/s41598-019-47335-w>
- [6] Bonhoeffer, S., Selbmann, A., et al. (2026). orcaAI: A machine learning tool to detect and classify acoustic signals of killer whales in audio recordings. Marine Mammal Science, 42(1), e70083. <https://doi.org/10.1111/mms.70083>

4.3 बर्फ़ के नीचे बेलुगा व्हेल की आवाज़ को वापस लाना

जहाँ कलिर व्हेल साफ़ तटीय पानी में अपनी बोलियों को सँवारती हैं, वहीं कुछ व्हेल ठंडे और मटमैले पानी में भी रहती हैं। वे आर्कटिक और कनाडा की सेंट लॉरेंस नदी (St. Lawrence River) में रहने वाली बेलुगा व्हेल (Beluga Whale) हैं। उनका शरीर सफ़ेद होता है, इसलिए उन्हें आर्कटिक की परी कहा जाता है। वे पक्षियों की तरह ऊँची और सुरीली आवाज़ें निकालती हैं, इसलिए उन्हें समुद्र की कैनरी का उपनाम भी दिया गया है। सेंट लॉरेंस नदी की बेलुगा व्हेल की संख्या आज बहुत तेज़ी से घट रही है। ऐसा कारख़ानों के शोर और बढ़ते जल प्रदूषण के कारण हुआ है। संरक्षित क्षेत्र घोषित करने के बाद भी उनकी संख्या ठीक से नहीं बढ़ रही है, जो एक बड़ी चिंता की बात है [1][7]।

कनाडा की वंडिसर यूनिवर्सिटी (University of Windsor) की शोधकर्ता जैकलिन ऑबनि (Jaclyn Aubin) भी इस काम में शामिल हुईं। उन्होंने बेलुगा व्हेल की आवाज़ों को समझने का फैसला किया। यह काम आसान नहीं था। बेलुगा व्हेल झुंड में रहने वाले सामाजिक जीव हैं। दर्जनों व्हेल एक साथ इकट्ठा होकर आवाज़ें निकालती हैं। वे एक साथ सीटियाँ और तेज़ क्लिक की आवाज़ें निकालती हैं। आवाज़ें इतनी आपस में उलझ जाती थीं कि किसी इंसान के लिए यह समझना मुश्किल था कि आवाज़ कहाँ से शुरू हुई और कहाँ ख़त्म हुई।

शोध दल ने इन आवाज़ों को AI को सौंप दिया। सबसे पहले आवाज़ ढूँढ़ने वाली मशीन ने लंबी रिकॉर्डिंग में से केवल बेलुगा व्हेल की आवाज़ों को छांटकर अलग किया। उसके बाद दूसरी मशीन ने उन आवाज़ों को प्रकार के अनुसार बाँटा। इस दौरान Earth Species Project द्वारा बनाए गए प्रोग्राम AVES का भी इस्तेमाल किया गया। AVES जानवरों की आवाज़ें सीखने वाला एक आधारभूत मॉडल (foundation model) है।

आधारभूत मॉडल एक ऐसा AI होता है जो पहले से ही ढेर सारे डेटा को सीख लेता है और फिर कई तरह के कामों में इसे माल किया जाता है। AVES इंसानों द्वारा सही उत्तर बताए बिना भी आवाजों की बनावट को अपने आप सीख लेता है [8]।

AI द्वारा छांटी गई आवाजों की तुलना इंसानों द्वारा की गई छंटाई से करने पर, दोनों परिणाम लगभग एक जैसे निकले। बिना इंसानी दखल के भी मशीन ने विशेषज्ञों की तरह ही बेलुगा व्हेल की आवाजों को अलग-अलग श्रेणियों में बाँट दिया। इसके कई बड़े फायदे हैं। इंसान को जिन लंबी रिकॉर्डिंग्स को सुनने में कई दिन लगते, मशीन उनकी तेज़ी से जाँच कर लेती है। नदी में रहने वाली बेलुगा व्हेल कब और कहाँ आवाज़ें निकालती है, इसका लगातार रिकॉर्ड रखा जा सकता है। यह रिकॉर्ड घटती जा रही बेलुगा व्हेल की रक्षा करने का एक अहम आधार बनता है।

बर्फ़ के टूटने और टकराने की आवाज़ एक बहुत बड़ी बाधा थी। शोधकर्ताओं ने इस शोर को मटाने के लिए शोर हटाने में माहिर एक विशेष AI का भी इस्तेमाल किया। यह मशीन तेज़ शोर के बीच से सफ़ि़ बेलुगा व्हेल की आवाज़ को वापस उभार लेती है। इसकी बदौलत हमेशा के लिए खो जाने वाली बेलुगा व्हेल की आवाज़ों को फिर से सुनना मुमकिन हो सका। इस तरह AI ने बर्फ़ के नीचे दबी आवाज़ों की रक्षा की।

संबंधित संदर्भ सामग्री

- [1] Bakker, K. (2022). The sounds of life: How digital technology is bringing us closer to the worlds of animals and plants. Princeton University Press.
- [7] Cotillard, T., Aubin, J., Vergara, V., et al. (2024). Automatic detection and classification of beluga whale calls in the St. Lawrence estuary. The Journal of the Acoustical Society of America, 156(6), 3723-3740.
<https://doi.org/10.1121/10.0030472>
- [8] Hagiwara, M. (2023). AVES: Animal vocalization encoder based on self-supervision. ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing.
<https://doi.org/10.1109/ICASSP49357.2023.10095642>

4.4 पक्षियों के गीतों से सीखकर समुद्र को सुनता AI

समुद्र की बात करते-करते अचानक पक्षी कहाँ से आ गए? इसकी एक खास वजह है। समुद्री जीवों की आवाज़ों को समझने में पक्षियों की आवाज़ से सीखे AI ने बहुत बड़ी मदद की। इसकी शुरुआत Google से हुई थी। यह कंप्यूटर वैज्ञानिक डॉ. टॉम डेंटन (Tom Denton) के नेतृत्व वाली टीम का 'पर्च (Perch)' प्रोजेक्ट था। इस टीम का सपना बहुत बड़ा है। वह है पृथ्वी से आने वाली प्रकृति की हर तरह की आवाज़ को एक साथ व्यवस्थित करना [10]।

Perch टीम ने पक्षियों की आवाज़ से AI को सिखाया। उन्होंने इसे दुनिया भर की 10,000 से ज़्यादा प्रजातियों के पक्षियों के गीतों को सिखाया। सीखने में बताया गया समय ही लाखों घंटों का था। इस तरह बना AI आवाज़ों को संख्यात्मक निर्देशांकों में बदलकर व्यवस्थित करता है। इसे आप हर आवाज़ के लिए जगह तय करने वाला एक बड़ा नक्शा बनाने जैसा समझ सकते हैं।

लेकिन तभी एक हैरान करने वाली बात हुई। पक्षियों की आवाज़ से सीखे AI ने पानी के भीतर के जानवरों की आवाज़ों को भी बहुत अच्छी तरह छांट दिया। इसका मतलब था कि एक जानवर से सीखे गए नियम दूसरे जानवरों की आवाज़ों पर भी काम कर गए। पक्षी हवा में चहकते हैं और व्हेल पानी में आवाज़ निकालती हैं। उनके रहने की जगह और आवाज़ दोनों अलग हैं। इसके बावजूद पक्षियों की आवाज़ से सीखा गया नक्शा व्हेल की आवाज़ों पर भी बिल्कुल सटीक बैठा।

शोधकर्ताओं ने इस क्षमता की पुष्टि मेढकों के ज़रिए की। इस दौरान इस्तेमाल किया गया डेटासेट AnuraSet था। यह ब्राज़ील के मेढकों की 42 प्रजातियों की आवाज़ों का संग्रह था। इसे अमेज़न में नहीं, बल्कि सैराडो और अटलांटिक वन कहे जाने वाले जंगलों में रिकॉर्ड किया गया था। यह डेटा Google ने नहीं, बल्कि कैनास (Cañas) की शोध टीम ने तैयार किया था। Google ने इस सार्वजनिक डेटा का उपयोग करके अपने AI का परीक्षण किया। मेढक की केवल कुछ आवाज़ें सुनाने पर भी AI ने 95% से अधिक सटीकता से आवाज़ों की पहचान कर ली। यह केवल कुछ उदाहरण देखकर नया काम सीख लेने के इस तरीके की वजह से संभव हुआ। यह क्षमता सफ़ि़ मेढकों तक ही सीमति नहीं रही। यह व्हेल की क्लिक आवाज़ों और डॉल्फ़िन की सीटियों पर भी पूरी तरह खरी उतरी [9]।

यह शोध लगातार आगे बढ़ता रहा। अगस्त 2025 में Google ने Perch का नया संस्करण (Perch 2.0) पेश किया। इस AI ने न केवल पक्षियों, बल्कि ज़मीन पर रहने वाले कई जानवरों की आवाज़ों से भी सीखा। इसने 14,000 से ज़्यादा प्रजातियों के जानवरों की आवाज़ें सीखीं। पक्षियों के साथ-साथ इसने स्तनधारियों, मेढकों और कीड़े-मकोड़ों की आवाज़ें भी सीखीं। हालाँकि, इसने समुद्र के भीतर की आवाज़ें लगभग नहीं सीखी थीं। फिर भी फ़रवरी 2026 में Google ने घोषणा की कि इस AI ने व्हेल की आवाज़ों के शोध में बेहतरीन प्रदर्शन किया है। व्हेल की आवाज़ के कई परीक्षणों में इस AI ने सबसे अच्छा या दूसरा सबसे अच्छा प्रदर्शन किया। Google ने डॉल्फ़िन के लिए विशेष AI मॉडल DolphinGemma भी बनाया। यह प्रोग्राम आकार में इतना छोटा है कि स्मार्टफ़ोन पर भी चल सकता है। इसने Wild Dolphin

Project द्वारा लंबे समय से इकट्ठा की गई आवाजों को सीखा। अभी तक इसने डॉल्फिन की भाषा का अर्थ स्पष्ट नहीं किया है। यह आवाजों की बनावट और उनके नयिमों को खोजने वाला एक शोध उपकरण है [11]।

DolphinGemma केवल आवाजों को छांटता ही नहीं है। यह डॉल्फिन की आवाजों की नकल करके अपने आप नई आवाजें भी बना लेता है। समुद्र में डॉल्फिन को यह आवाज सुनाने के प्रयोग भी किए जा रहे हैं। DolphinGemma ने जो आवाजें सीखी हैं, उन्हें बहुत लंबे समय में इकट्ठा किया गया था। Wild Dolphin Project 1985 से आवाजें जुटा रहा है। ये बहामास के समुद्र में रहने वाली अटलांटिक चित्तीदार डॉल्फिन की आवाजें हैं। उन्होंने संभोग या लड़ाई जैसे व्यवहारों के साथ-साथ इन आवाजों को रिकॉर्ड किया था। Google ने इस प्रोग्राम को हर किसी के इस्तेमाल के लिए खुला रखा है।

संबंधित संदर्भ सामग्री

[9] Cañas, J. S., Toledo, L. F., et al. (2023). AnuraSet: A dataset for benchmarking Neotropical anuran calls identification in passive acoustic monitoring. *PeerJ Computer Science*, 9, e1575. <https://doi.org/10.7717/peerj-cs.1575>

[10] Google Research. (2026, February 9). How AI trained on birds is surfacing underwater mysteries. Google Research Blog. <https://research.google/blog/how-ai-trained-on-birds-is-surfacing-underwater-mysteries/>

[11] Google. (2025, April 14). DolphinGemma: How Google AI is helping decode dolphin communication. Google Blog. <https://blog.google/innovation-and-ai/products/dolphingemma/>

4.5 मटमैले पानी में गुलाबी डॉल्फिन की रक्षा करता छोटा AI

पानी के भीतर का AI सरिफ वशाल समुद्र में ही नहीं, बल्कि नदियों में भी काम करता है। यह अमेज़न के उस मटमैले पानी में काम करता है जहाँ आँखों से कुछ भी दिखाई नहीं देता। यहाँ गुलाबी रंग के शरीर वाली नदी की डॉल्फिन बोटो (Boto) रहती है। ब्राजील का मामरिआ संस्थान (Mamirauá Institute) और राष्ट्रीय अमेज़न अनुसंधान संस्थान (INPA) बोटो की रक्षा के काम का नेतृत्व कर रहे हैं [12]।

बोटो की संख्या आज बहुत तेजी से घट रही है। वे इंसानों द्वारा नदी में फेंके गए जालों में फँस जाती हैं। सोना नकालने के दौरान बहाए गए पारे से वे प्रदूषित हो रही हैं। वे विकास कार्यों के शोर से भी परेशान हैं। अमेज़न नदी में मट्टि और पत्तियों की भरमार होने के कारण पानी हमेशा मटमैला रहता है। इसलिए कैमरों या ड्रोन की मदद से बोटो को ढूँढ़ पाना नामुमकिन है। केवल आवाज ही बोटो की रक्षा करने की एकमात्र चाबी है [13]।

शोधकर्ताओं ने नदी के पानी में आवाज सुनने वाले माइक्रोफोन लगा दिए। ये माइक्रोफोन बनी रुके लंबे समय तक आवाजों को रिकॉर्ड करते हैं। किसी इंसान के लिए यह इतनी लंबी रिकॉर्डिंग सुनना बहुत मुश्किल है। इसलिए AI उनकी जगह इनकी जाँच करता है। यहाँ इस्तेमाल किया गया AI आवाज के चित्रों को देखकर नरिणय लेने वाली मशीन है। यह जंगल की बारिश, तूफान और टहनियों की आवाजों को छानकर अलग कर देता है। और बोटो द्वारा नकाली जाने वाली क्लिक की आवाजों को ढूँढ़ निकालता है।

इस AI ने 95% सटीकता के साथ बोटो की आवाजों को ढूँढ़ निकाला। इसने 92% सटीकता के साथ नावों के गुज़रने की आवाज को भी अलग पहचान लिया। उस मटमैले पानी के नीचे जहाँ आँखें बिल्कुल काम नहीं करती, यह मशीन सरिफ आवाज के दम पर बोटो के रहने की जगह की रक्षा करती है। इसने नदी में रहने वाली एक अन्य डॉल्फिन तुकुक्सी की आवाज को भी अलग से पहचान लिया। इस तरह AI न देखने वाली डॉल्फिन को आवाज के ज़रिए ढूँढ़ निकालता है।

यह रिकॉर्ड बोटो (boto) की रक्षा करने में सीधे काम आता है। शोध दल ने आवाज की मदद से उन रास्तों का नक्शा बनाया जहाँ से बोटो अक्सर गुज़रती है। उन्होंने यह भी पता लगाया कि किस मौसम में वे कहाँ इकट्ठा होती हैं। जब जलस्तर बढ़ने का मौसम आता है, तो बोटो बड़ी नदी से निकलकर पानी में डूबे जंगलों में चली जाती है। आवाजों की जाँच करने पर उनकी यह हलचल साफ़-साफ़ समझ में आ गई। यह नक्शा नावों के चलने के रास्तों और जाल डालने की जगहों को बदलने में मदद करता है। इस तरह आवाज बोटो के जीवन की रक्षा करने वाला एक दशिया-सूचक बोर्ड बन गई। बड़े कंप्यूटरों के बनिा भी जंगल के दूर-दराज इलाकों में लंबे समय तक आवाजें सुनने और उनकी रक्षा करने के तरीके ढूँढ़ने का शोध लगातार जारी है [2]।

संबंधित संदर्भ सामग्री

[2] NOAA Fisheries. (2025). A whale's world of sound. National Oceanic and Atmospheric Administration. <https://www.fisheries.noaa.gov/feature-story/whales-world-sound>

[12] Erbs, F., Gaona, M., van der Schaar, M., Zaugg, S., & André, M. (2023). Towards automated long-term acoustic monitoring of endangered river dolphins: A case study in the Brazilian Amazon floodplains. *Scientific Reports*, 13, 10801. <https://doi.org/10.1038/s41598-023-36518-1>

[13] Kane, S. (2023, August 1). AI unlocks secrets of Amazon river dolphins' behavior, no tagging required. Mongabay. <https://news.mongabay.com/2023/08/ai-unlocks-secrets-of-amazon-river-dolphins-behavior-no-tagging-required/>

4.6 समुद्र को ध्यान से सुनना

अध्याय 4 में देखा गया पानी के भीतर AI का शोध एक पुरानी सोच को बदल देता है। इंसान लंबे समय से दूसरे जानवरों को सिर्फ अपनी मूल प्रवृत्तियों के अनुसार जीने वाला जीव मानता आया है [1]। लेकिन हाल के शोध एक अलग ही तस्वीर दिखाते हैं [1]। कलर व्हेल की बोलियों पीढ़ियों से चली आ रही संस्कृतियों की दर्शाती हैं [1]। नदी में सफेद व्हेल (बेलुगा व्हेल) का एक-दूसरे से आवाजों का आदान-प्रदान करना इस बात का सबूत है कि वे आपस में जुड़े रश्तों का एक समाज बनाती हैं [1]।

व्हेल की भाषा के बारे में नए तथ्य भी लगातार सामने आए हैं। यह Project CETI और कई विश्वविद्यालयों के शोधकर्ताओं का अध्ययन है। 2024 में उन्होंने स्पर्म् व्हेल की आवाज़ की बनावट को स्पष्ट किया। स्पर्म् व्हेल कई क्लिक्स को एक साथ जोड़कर संकेत बनाती हैं। आवाजों के इस समूह को कोडा (coda) कहा जाता है। शोधकर्ताओं ने पाया कि कोडा में गति और लय जैसे नियम होते हैं। इन नियमों को मिलाकर अनगिनत अलग-अलग आवाज़ें बनती हैं। इसका मतलब है कि वे संकेत उम्मीद से कहीं अधिक सुव्यवस्थित हैं। शोध दल ने ध्वनियों के इन टुकड़ों को व्यवस्थित करके एक 'ध्वनि वर्णमाला' भी पेश की। 2026 में वे एक कदम और आगे बढ़ेंगे। उन्होंने स्पर्म् व्हेल की आवाज़ में इंसानी भाषा के स्वरों से मिलते-जुलते दो प्रकार के ध्वनिरूप खोज निकाले। यह ठीक वैसा ही रूप है जैसे इंसानों की 'आ' की आवाज़ और 'ई' की आवाज़ अलग-अलग होती है। हालाँकि, इसका यह मतलब बिल्कुल नहीं है कि उन्होंने स्पर्म् व्हेल की भाषा का अर्थ समझ लिया है [14][15]।

ऐसा AI भी विकसित हो रहा है जो जानवरों की आवाजों के बारे में इंसानी भाषा में सवाल पूछने और जवाब पाने में मदद करता है। यह Earth Species Project द्वारा बनाया गया NatureLM-audio है। यह जानवरों की आवाजों के लिए पहला ध्वनि-भाषा आधारभूत मॉडल है। जब कोई इंसान सामान्य भाषा में सवाल पूछता है, तो AI आवाज़ की जाँच करके जवाब देता है। यह प्रजातियों को अलग करता है, आवाजों को ढूँढता है और यहाँ तक कि कितने जीव आवाज़ निकाल रहे हैं, उनकी गिनती भी करता है [16]।

व्हेल की भाषा को समझने की तकनीक कानून के सामने भी नए सवाल खड़े करती है। अक्टूबर 2025 में, कानून के वद्वानों और Project CETI के शोधकर्ताओं ने मिलकर एक शोध पत्र प्रकाशित किया। उन्होंने व्हेल को संस्कृत साझा करने वाला और आपस में संवाद करने वाला जीव माना। इसलिए उन्होंने व्हेल को भी कठोर व्यवहार से मुक्त रहने का अधिकार देने का प्रस्ताव रखा। न्यूयॉर्क यूनिवर्सिटी लॉ स्कूल का 'मानव से परे जीवन (MOTH)' कार्यक्रम इस काम का नेतृत्व कर रहा है। AI द्वारा खोले गए नए कानून और नैतिकता के सवालों से जुड़ गए हैं [17]।

हमने अब तक समुद्र में भारी शोर उड़ेल दिया है। बड़े जहाजों की आवाजों और खोज संबंधी वसिफोटों से हमने व्हेलों के संवाद के रास्ते को तोड़ा है। AI अब समुद्र की आवाज़ को सावधानी से सुनने वाला एक सटेथोस्कोप बन गया है। यह उपकरण इंसानों को एक बात सिखाता है। वह यह कि हम अपनी आवाज़ को धीमा करें और प्रकृति की आवाज़ को वैसे ही सुनें जैसी वह है। जब हम समुद्र पर राज करने की इच्छा को छोड़ देंगे, तभी हम व्हेलों के पुराने गीतों को ठीक से सुन सकेंगे। गहरे समुद्र के ऑर्केस्ट्रा ने अभी तक बजना बंद नहीं किया है। हम जितना ध्यान से सुनेंगे, वह आवाज़ उतनी ही हमारे पास वापस लौटकर आएगी [2][1]।

संबंधित संदर्भ सामग्री

- [1] Bakker, K. (2022). The sounds of life: How digital technology is bringing us closer to the worlds of animals and plants. Princeton University Press.
- [2] NOAA Fisheries. (2025). A whale's world of sound. National Oceanic and Atmospheric Administration. <https://www.fisheries.noaa.gov/feature-story/whales-world-sound>
- [14] Sharma, P., Gero, S., Payne, R., et al. (2024). Contextual and combinatorial structure in sperm whale vocalisations. Nature Communications, 15, 3617. <https://doi.org/10.1038/s41467-024-47221-8>
- [15] Beguš, G., Dąbkowski, M., Sprouse, R. L., Gruber, D. F., & Gero, S. (2026). The phonology of sperm whale coda vowels. Proceedings of the Royal Society B, 293(2069). <https://doi.org/10.1098/rspb.2025.2994>
- [16] Robinson, D., Miron, M., Hagiwara, M., Pietquin, O., et al. (2024). NatureLM-audio: An audio-language foundation model for bioacoustics. arXiv:2411.07186. <https://arxiv.org/abs/2411.07186>
- [17] Rodríguez-Garavito, C., Gruber, D. F., Nemeth, A. O., & Beguš, G. (2025, October). What if we understood what animals are saying? The legal impact of AI-assisted studies of animal communication. Ecology Law Quarterly, 52(1). <https://www.ecologylawquarterly.org/print/what-if-we-understood-what-animals-are-saying-the-legal-impact-of-ai-assisted-studies-of-animal-communication/>

अध्याय 5. धरती की भाषा

5.1 चपिंजी की आवाज़ों में नयिम होते हैं

इंसानों की भाषा कुछ आवाज़ों को आपस में जोड़कर बनती है। बहुत कम आवाज़ों से अनगणित अर्थ बन जाते हैं। लंबे समय तक लोग मानते थे कि यह हुनर सिर्फ इंसानों के पास है। लेकिन जानवरों की आवाज़ों का लंबे समय तक अध्ययन करने वाले वैज्ञानिकों ने इस सोच पर फरि से विचार किया।

इसकी शुरुआत पश्चिमी अफ्रीका के आइवरी कोस्ट (Côte d'Ivoire) के ताई नेशनल पार्क (Taï National Park) से हुई। यहाँ जंगली चपिंजियों का एक झुंड रहता है। जर्मनी के मैक्स प्लैंक इंस्टीट्यूट फॉर इवोल्यूशनरी एंथ्रोपोलॉजी (Max Planck Institute for Evolutionary Anthropology) और फ्रांस के राष्ट्रीय वैज्ञानिक अनुसंधान केंद्र (CNRS) के वैज्ञानिकों ने लंबे समय तक इन पर नज़र रखी। डॉ. सेड्रिक गिरार्ड-बुटोज़ (Cédric Girard-Buttoz) और डॉ. रोमन वटिग (Roman Wittig) ने इस शोध का नेतृत्व किया [1]।

यह अध्ययन एक या दो साल में खत्म नहीं हुआ। यह कई वर्षों तक एक ही झुंड को देखने का परणाम है। वैज्ञानिकों ने 46 वयस्क चपिंजियों की आवाज़ें इकट्ठी कीं। रिकॉर्ड की गई आवाज़ों की संख्या 4,826 तक पहुँच गई। पहले लोग इन आवाज़ों को सिर्फ भावनाएं जताने वाली चीख-पुकार मानते थे। लेकिन जब एक-एक करके उनकी गणिती की गई, तो उनके अंदर एक व्यवस्था दिखाई दी।

वैज्ञानिकों ने इन सभी आवाज़ों को लंबे समय तक सुना और उन्हें अलग-अलग वर्गों में बाँटा। उन्होंने आवाज़ के उतार-चढ़ाव और लंबाई को बारीकी से नापा। इसके परणामस्वरूप चपिंजियों द्वारा इस्तेमाल की जाने वाली बुनियादी आवाज़ें 12 प्रकारों में बाँट गईं। इनमें कराहना, भौकना और चीखना जैसी आवाज़ें शामिल हैं। यह इंसानी भाषा के स्वर और व्यंजन की तरह ध्वनिकों के सबसे छोटे टुकड़े हैं।

चपिंजी इन 12 आवाज़ों को यूँ ही बिना सोचे-समझे नहीं निकालते थे। वे इन्हें एक तय क्रम के अनुसार आपस में जोड़ते थे। इस तरह बनाए गए ध्वनिसमूह 390 से भी अधिक थे। कुछ क्रम बार-बार सुनाई देते थे, जबकि कुछ क्रम शायद ही कभी आते थे। इसका मतलब है कि आवाज़ों को जोड़ने का एक नश्चिती ढाँचा होता है।

आजकल इतनी सारी आवाज़ों का अध्ययन करने में मशीन लर्निंग (machine learning) नामक AI वैज्ञानिकों की मदद करती है। मशीन लर्निंग अनगणित आवाज़ों में बार-बार दोहराए जाने वाले पैटर्न को तेज़ी से ढूँढ़ निकालती है। जसि काम में एक अकेले इंसान को कई साल लग जाते, उसे यह कुछ ही दिनों में पूरा कर देती है। नयिमों की खोज वैज्ञानिक करते हैं, और AI उस काम में सहायता करने वाला एक औज़ार है।

ध्वनिकों के ये समूह परतों में बने हुए थे। दो आवाज़ों को जोड़कर बना समूह तीन आवाज़ों वाले समूह के अंदर भी आ जाता था। इंसानों की भाषा भी इसी तरह परतों में बनती है। यह बिल्कुल वैसा ही है जैसे शब्द मलिकर वाक्यांश बनते हैं, और वाक्यांश मलिकर वाक्य बनते हैं।

वैज्ञानिक इंसानी भाषा की एक बड़ी विशेषता बताते हैं। वह है बिना अर्थ वाले छोटे ध्वनिकों को जोड़कर नया अर्थ बनाने की शक्ति। एक अकेले ध्वनिकों का कोई अर्थ नहीं होता। उन्हें आपस में जोड़ने का क्रम ही अंततः अर्थ का निर्माण करता है। इस शक्ति को ध्वनिकी दोहरी संरचना (duality of patterning) कहा जाता है। हालाँकि, चपिंजी की 12 बुनियादी आवाज़ें इंसानों के स्वर और व्यंजन के बिल्कुल समान नहीं हैं। वैज्ञानिक अभी भी इसका बहुत सावधानी से अध्ययन कर रहे हैं [2]।

समय के साथ नए अध्ययन सामने आए। 2025 में वैज्ञानिकों ने पता लगाया कि चपिंजी दो आवाज़ों को जोड़कर उनका अर्थ बदल देते हैं। दो आवाज़ों के ऐसे संयोजनों की जांच करने पर 16 प्रकार मिले। कुछ संयोजनों में दोनों आवाज़ों के अर्थ सीधे जुड़ गए थे। जबकि कुछ संयोजन पूरी तरह से नए अर्थ में बदल गए थे।

उदाहरण के लिए, किसी एक आवाज़ के पीछे दूसरी आवाज़ जुड़ने से उसका अर्थ बदल जाता था। यह पेड़ पर सोने की जगह बनाने का संकेत भी बन जाता था। खतरे की चेतावनी देने वाली आवाज़ और मदद बुलाने वाली आवाज़ को जोड़कर वे साँप से सावधान रहने का संदेश भी देते थे। एक आवाज़ होने और दो आवाज़ें मिलने पर उनके अर्थ अलग-अलग होते थे [3]।

यह खोज हमें बताती है कि इंसानी भाषा की शुरुआत कैसे हुई होगी। चपिंजी इन ध्वनिसंयोजनों से खतरे की सूचना देते हैं और एक-दूसरे के साथ मलिकर काम करते हैं। इसका मतलब है कि आवाज़ों को आपस में परीने का हुनर सिर्फ इंसानों तक सीमति नहीं है। चपिंजी की आवाज़ों के नयिम सुदूर अतीत से चले आ रहे संवाद के नशान दिखाते हैं [4]।

संबंधित संदर्भ सामग्री

[1] Girard-Buttoz, C., Wittig, R. M., et al. (2022). Chimpanzees produce diverse vocal sequences with ordered and recombinatorial properties. *Communications Biology*, 5, 410. <https://doi.org/10.1038/s42003-022-03350-8>

- [2] Girard-Buttoz, C., et al. (2025). Versatile use of chimpanzee call combinations promotes meaning expansion. *Science Advances*, 11(19), eadq2879. <https://doi.org/10.1126/sciadv.adq2879>
- [3] Leroux, M., et al. (2023). Call combinations and compositional processing in wild chimpanzees. *Nature Communications*, 14, 2225. <https://doi.org/10.1038/s41467-023-37816-y>
- [4] Sun, W. (2025, May 9). Origins of language: Wild chimps mirror linguistic structures in human language. *Phys.org*. <https://phys.org/news/2025-05-language-wild-chimps-mirror-linguistic.html>

5.2 गेलाडा बबून की इंसानों जैसी बोलने की लय

अगर चंपैंजियों ने जंगल में आवाज़ के नयिम दिखाए, तो एक दूसरा बंदर इंसानों से और भी मलिती-जुलती आवाज़ निकालता है। यह इथियोपिया के ऊँचे पहाड़ों में रहने वाला गेलाडा बबून (*Theropithecus gelada*) है। ये खड़ी चट्टानों वाले इलाके में बड़े-बड़े झुंड बनाकर रहते हैं। आपस में की जाने वाली इनकी बातचीत ऐसी सुनाई देती है जैसे लोग बैठकर आराम से गपशप कर रहे हों [1][2]।

गेलाडा की आवाज़ इंसानों के कानों को बहुत प्रिय और जानी-पहचानी लगती है। ऐसा लगता है मानो कई लोग धीमी आवाज़ में फुसफुसा रहे हों। वह आवाज़ दनि भर चट्टानों के बीच गूँजती रहती है। यही कारण है कि वैज्ञानिक लंबे समय से इस आवाज़ की ओर आकर्षित रहे हैं।

अमेरिका की मशिगिन यूनिवर्सिटी (University of Michigan) के प्रोफेसर थोर बर्गमैन (Thore Bergman) ने इस आवाज़ का लंबे समय तक अध्ययन किया। गेलाडा जब आपस में दोस्ताना व्यवहार करते हैं, तो अपने होंठों को तेज़ी से फड़फड़ाते हैं। इस आवाज़ को होंठ फड़फड़ाना (lip-smacking) कहते हैं। इसके साथ वे गले को कंपकंपाकर एक भारी आवाज़ भी जोड़ते हैं। इस आवाज़ को वॉबल ध्वनि (wobble sound) कहा जाता है।

प्रोफेसर बर्गमैन ने इस आवाज़ की लय को नापा। इंसान बोलते समय एक सेकंड में अपने मुँह को लगभग तीन से आठ बार खोलते और बंद करते हैं। गेलाडा की वॉबल ध्वनि एक सेकंड में छह से नौ बार कंपकंपी पैदा करती है। ये दोनों लय एक-दूसरे से काफी मेल खाती हैं। एक सेकंड में कोई क्रिया कतिनी बार दोहराई जाती है, इसे नापने की इकाई को हर्ट्ज़ (Hz) कहते हैं।

यह खोज 2013 में दुनिया के सामने आई। इंसान और गेलाडा एक-दूसरे से बहुत दूर रहते हैं। फिर भी उनके आवाज़ निकालने की लय बहुत मलिती-जुलती थी। मानो उन्होंने आपस में पहले से कोई वादा कर रखा हो।

गेलाडा सरिफ़ होंठ फड़फड़ाने तक ही सीमिति नहीं रहते। वे दो आवाज़ों को मिलाकर उसमें और भी अधिक अर्थ जोड़ते हैं। वैज्ञानिकों ने गेलाडा की आवाज़ों की तुलना चकमा बबून (*Chacma baboon*) की आवाज़ों से की। साथ रहने वाला झुंड जतिना बड़ा था, उनकी आवाज़ों के प्रकार भी उतने ही अधिक थे। यानी बात करने के लिए जतिने अधिक साथी होंगे, बातें भी उतनी ही अधिक होंगी।

आजकल ऐसी आवाज़ों का बहुत सारा डेटा इकट्ठा हो चुका है। इसलिए वैज्ञानिकों ने AI को एक औज़ार के रूप में इस्तेमाल करना शुरू किया। AI अनगिनत आवाज़ों को बहुत तेज़ी से वर्गीकृत कर देता है। सरिफ़ कानों से अंदाज़ा लगाई जाने वाली लय को यह संख्याओं में साफ़-साफ़ दिखा देता है। सच्चाई की खोज वैज्ञानिक करते हैं, और AI उस काम में मदद करने वाला हाथ है [3]।

यहाँ एक सवाल उठता है। इंसान और गेलाडा के पूर्वज अलग-अलग हैं, फिर भी उनकी लय एक जैसी क्यों है? अलग-अलग बनावट वाले जानवरों के एक ही दिशा में विकसित होने को अभिसारी विकास (convergent evolution) कहते हैं। यह वैसा ही है जैसे दो अलग-अलग कक्षाओं द्वारा अलग से बनाए गए जयकारे की ताल एक जैसी हो जाए। समान जीवन जीते हुए जीव समान हुनर तक पहुँच जाते हैं।

संबंधित संदर्भ सामग्री

- [1] Bergman, T. J. (2013). Speech-like vocalized lip-smacking in geladas. *Current Biology*, 23(7), R268-R269. <https://doi.org/10.1016/j.cub.2013.02.038>
- [2] Gustison, M. L., le Roux, A., & Bergman, T. J. (2012). Derived vocalizations of geladas (*Theropithecus gelada*) and the evolution of vocal complexity in primates. *Philosophical Transactions of the Royal Society B*, 367(1597), 1847-1859. <https://doi.org/10.1098/rstb.2011.0218>
- [3] Nature. (2025, September 17). AI is helping to decode animals' speech. *Nature News*. <https://doi.org/10.1038/d41586-025-02917-9>

5.3 ज़मीन के नीचे नेकेड मोल-रेट की बोली

अगर सूरज की रोशनी वाली ज़मीन पर बंदर आवाज़ों का आदान-प्रदान करते हैं, तो अंधेरी ज़मीन के नीचे भी आवाज़ों का एक समाज मौजूद है। यह बर्ना रोशनी वाली मटिटी की सुरंगों की दुनिया है। यहाँ नेकेड मोल-रेट (*Heterocephalus glaber*) रहते हैं। ये जीव चींटियों या मधुमक्खियों की तरह झुंड बनाकर रहते हैं। एक रानी बच्चे पैदा करती है, और बाकी सदस्य आपस में काम बाँट लेते हैं। रानी और कामगारों की तरह भूमिकाएँ बाँटने वाले ऐसे समाज को पूर्ण-सामाजिकता (eusociality) कहा जाता है [1]।

जर्मनी के बर्लिन में मैक्स डेलब्रुक सेंटर (Max Delbrück Center) के शोधकर्ताओं ने इनकी आवाज़ें इकट्ठी कीं। डॉ. एलिसन बार्कर (Alison Barker) ने इसका नेतृत्व किया। नेकेड मोल-रेट जब एक-दूसरे से मिलते हैं, तो धीमी आवाज़ में चहचहाते हैं। शोध दल ने ऐसी 30,000 से अधिक आवाज़ें इकट्ठी कीं और एक-एक करके उनकी जांच की। उन्होंने आवाज़ के उतार-चढ़ाव और लंबाई जैसी कई विशेषताओं को बारीकी से नापा [2]।

उन्होंने जो सच्चाई पता लगाई, वह इस प्रकार है। हर झुंड की आवाज़ नकालने की अपनी अलग आदत थी। इंसानों के हिसाब से देखें तो यह अलग-अलग क्षेत्रों की बोलियों की तरह है। इसे भूमिगत झुंड की बोली कहा जाता है। नेकेड मोल-रेट अंधेरी सुरंगों में आँखों से देख नहीं सकते। वे केवल आवाज़ से ही अपने और पराए की पहचान करते हैं।

सुरंगों के अंदर हमेशा अंधेरा रहता है। रोशनी न होने के कारण चेहरा पहचानने का कोई उपाय नहीं होता। इसलिए आवाज़ ही नाम की तख्ती का काम करती है। अगर वे एक ही बोली बोलते हैं, तो वे एक ही झुंड के हैं। यदि कोई अपरिचित बोली सुनाई देती है, तो उसे दूसरे झुंड का मान लिया जाता है। इस तरह एक अकेली आवाज़ ही झुंड की रक्षा करने वाले दरबान का काम करती है।

यह बोली जन्मजात नहीं होती थी। नया जन्मा बच्चा बड़ों की आवाज़ की लगातार नकल करता है। इस तरह वह अपने झुंड की आवाज़ की आदतें सीखता है। दूसरे झुंड से नए आए चूहे के साथ भी ऐसा ही था। उसने नए झुंड की आवाज़ की नकल करके उनकी बोली सीख ली। यानी बोली खून के रश्ति से तय नहीं होती, बल्कि सीखने से तय होती है।

सीखने का यह तरीका इंसानी बच्चों के बोलना सीखने के तरीके से मिला-जुलता है। जसि तरह बच्चा बड़ों की बातें सुनकर धीरे-धीरे उनकी नकल करता है, उसी तरह छोटे चूहे भी आवाज़ें सीखते चले गए [3]।

सबसे ध्यान देने योग्य बात रानी की शक्ति है। जब तक रानी जीवित रहती है, झुंड की बोली एक जैसी बनी रहती है। लेकिन जब रानी की अचानक मृत्यु हो जाती है, तो बोली बखिर जाती है। आवाज़ के नियम एक साथ डगमगा जाते थे। इसका मतलब है कि रानी ही झुंड की आवाज़ को एक सूत्र में बांधने वाला मुख्य स्तंभ थी।

इतनी सारी आवाज़ों को वर्गीकृत करने में AI बहुत मददगार साबित हुआ। AI ने हर झुंड की अलग-अलग आवाज़ की आदतों को तेज़ी से पहचान लिया। लेकिन बोलियाँ सीखने से आती हैं और रानी उन्हें बनाए रखती है, यह बात लंबे समय के अवलोकन से ही सामने आई। AI तो केवल उस अवलोकन में सहायता करने वाला एक औज़ार था।

रानी के न रहने पर आवाज़ का भी बखिर जाना एक गहरा अर्थ रखता है। आवाज़ों को एक साथ जोड़े रखने के लिए एक केंद्र का होना ज़रूरी था। यह हमारी कक्षा द्वारा मलिकर बनाए गए नियमों जैसा ही है। अगर नियमों का पालन करवाने वाला केंद्र डगमगा जाता है, तो नियम भी डगमगा जाते हैं।

संबंधित संदर्भ सामग्री

[1] Barker, A. J., et al. (2021). Cultural transmission of vocal dialect in the naked mole-rat. *Science*, 371(6528), 503-507. <https://doi.org/10.1126/science.abc6588>

[2] Max Delbrück Center for Molecular Medicine. (2021, January 28). Naked mole-rats 'chirp' in colony-specific, culturally transmitted dialects. *EurekAlert!* <https://www.eurekalert.org/news-releases/510214>

[3] Vaidyanathan, G. (2021, February). The queen sets the tone: Deciphering the dialects of naked mole-rats. *Mongabay News*. <https://news.mongabay.com/2021/02/the-queen-sets-the-tone-deciphering-the-dialects-of-naked-mole-rats/>

5.4 मीरकैट की चेतावनी भरी आवाज़ और लकड़बग्घे की हँसी

ज़मीन पर रहने वाले कई अन्य जानवर भी अपनी आवाज़ों में बहुत से अर्थ छुपाकर रखते हैं। मीरकैट (*Suricata suricatta*) इसका एक अच्छा उदाहरण है। मीरकैट आवाज़ से बताते हैं कि दुश्मन आसमान से आ रहा है या ज़मीन से। चील आ रही है या साँप, इसके लिए वे अलग आवाज़ नकालते हैं। खतरा कतिना नज़दीक और गंभीर है, यह भी वे अपनी आवाज़ में बता देते हैं [1]।

यह बात स्वट्रिज़रलैंड की प्रोफेसर मार्था मानसर (Marta Manser) ने काफी समय पहले खोजी थी। प्रोफेसर मानसर ने दक्षिणी अफ्रीका के कालाहारी रेगसिस्तान में मीरकैट्स का लंबे समय तक अध्ययन किया। उन्होंने चेतावनी की आवाज़ों को एक-एक करके सुना और उन्हें छाँटा। खतरा जतिना नज़दीक होता, आवाज़ उतनी ही कर्कश और तेज़ हो जाती थी। खतरे का प्रकार बताने वाली आवाज़ और उसकी गंभीरता बताने वाली आवाज़ अलग-अलग थी। मीरकैट एक अकेली आवाज़ से पूरे झुंड को एक साथ हरकत में ले आते थे।

आजकल ऐसी चेतावनी भरी आवाज़ों को AI भी छाँटता है। AI ध्वनिकी उन बारीक परतों को तेज़ी से पहचान लेता है जिन्हें इंसान आसानी से नहीं सुन पाते। लेकिन आवाज़ में खतरे का प्रकार और उसकी गंभीरता शामिल होती है, यह सच्चाई सबसे पहले लंबे मैदानी अवलोकन से ही सामने आई थी।

अफ्रीकी घास के मैदानों की रातों पर राज करने वाला धब्बेदार लकड़बग्घा (*Crocuta crocuta*) भी ऐसा ही है। लकड़बग्घा हँसने जैसी आवाज़ निकालता है। इस खलिखलिहट भरी आवाज़ में कई जानकारीयें छिपी होती हैं। फ्रांस के निकोलस माथेवोन (Nicolas Mathevon) के शोध दल ने इस आवाज़ का अध्ययन किया। यह सुनने में बेतरतीब शोर जैसा लगता है, लेकिन ऐसा नहीं था। उसमें उम्र और झुंड के अंदर का स्थान शामिल था। उसमें प्रत्येक सदस्य की अलग पहचान बताने वाला नशान भी मौजूद था। यह इंसानों के आवाज़ के फगिरप्रटि जैसा ही है [2]।

माथेवोन के शोध दल ने अमेरिका के बर्कले में मौजूद 17 लकड़बग्घों की आवाज़ें रिकॉर्ड कीं। और उन्होंने हँसी की आवाज़ के दो मुख्य पहलुओं को अलग निकाला। पहला मुंह और गले के आकार से बनने वाला ध्वनिकी पैटर्न है। इसे फॉर्मैट (formant) कहा जाता है। दूसरा आवाज़ का उतार-चढ़ाव है। इसे पिच (pitch) कहा जाता है। ऊँचे दर्जे वाले लकड़बग्घे और नचिले दर्जे के लकड़बग्घे में ये दोनों पहलू अलग-अलग थे। यानी सिर्फ आवाज़ सुनकर ही पता चल जाता था कि कौन ज़्यादा ताकतवर है।

इंसानों के करीबी कुत्ते (*Canis lupus familiaris*) के मामले में भी ऐसा ही है। अमेरिका की मशिगिन यूनिवर्सिटी के शोधकर्ताओं ने कुत्तों की आवाज़ों का अध्ययन किया। उन्होंने भौंकने, गुर्राने और कराहने जैसी आवाज़ों को इकट्ठा किया। और इंसानी भाषा का विश्लेषण करने वाले AI को कुत्तों की आवाज़ों पर आजमाकर देखा। इस AI का नाम वेव2वेक2 (Wav2Vec2) है। परिणामस्वरूप इसने पहचान लिया कि आवाज़ खेल-कूद भरी है या आक्रामक। इसने कुत्ते की नस्ल, नर या मादा होने और उसकी उम्र का भी काफी हद तक सही अंदाज़ा लगाया। सही अंदाज़ा लगाने की दर 70% तक पहुँच गई। इंसानी भाषा का विश्लेषण करने वाला औज़ार कुत्तों की आवाज़ पर भी काम कर गया [3]।

यहाँ एक सवाल उठता है। पहले हमने देखा कि गेलाडा और इंसान की लय का मेलना अभिसारी विकास (convergent evolution) कहलाता है। उनके पूर्वज अलग थे, लेकिन एक जैसा जीवन जीने के कारण वे एक जैसे हो गए। लेकिन इंसान और कुत्ते द्वारा आवाज़ निकालने का आधार एक अलग कारण से मेलित-जुलता है। इंसान और कुत्ते दोनों ही स्तनधारी जीव (mammals) हैं। सुदूर अतीत में वे एक ही पूर्वज से अलग हुए थे। इसलिए गले और स्वरयंत्र से आवाज़ निकालने वाली शरीर की बनावट शुरू से ही मिलती-जुलती थी। पूर्वज एक होने के कारण जब बुनियादी संरचना मिलती है, तो इसे समजातता (homology) कहते हैं। यह वैसा ही है जैसे हमारे हाथ और कुत्ते के आगे के पैर एक ही हड्डी से बने हैं। समानता के इस तरह दो रास्ते होते हैं। पहला अलग रहकर एक जैसा बन जाने का रास्ता, और दूसरा एक ही जड़ से अलग होने का रास्ता।

संबंधित संदर्भ सामग्री

- [1] Manser, M. B. (2001). The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. *Proceedings of the Royal Society B*, 268(1483), 2315-2324. <https://doi.org/10.1098/rspb.2001.1773>
- [2] Mathevon, N., et al. (2010). What the hyena's laugh tells: Sex, age, dominance and individual signature in the giggling call of *Crocuta crocuta*. *BMC Ecology*, 10, 9. <https://doi.org/10.1186/1472-6785-10-9>
- [3] Abzaliev, A., Pérez Espinosa, H., & Mihalcea, R. (2024). Towards dog bark decoding: Leveraging human speech processing for automated bark classification. *Proceedings of LREC-COLING 2024*. <https://aclanthology.org/2024.lrec-main.1432/>

5.5 एक ही AI कई जानवरों की आवाज़ें सीखता है

अब तक वैज्ञानिक हर जानवर की आवाज़ का अलग-अलग अध्ययन करते थे। चिंपी का अलग और गेलाडा का अलग अध्ययन किया जाता था। लेकिन आजकल एक ही AI कई जानवरों की आवाज़ें एक साथ सीख रहा है [1]।

अर्थ स्पीशीज प्रोजेक्ट (Earth Species Project) इसमें सबसे आगे है। यह संस्था AI से जानवरों की आवाज़ों को समझने की कोशिश कर रही है। इन्होंने NatureLM-audio नाम का एक मॉडल बनाया है। यह जानवरों की आवाज़ सुनकर इंसानों के सवालों के जवाब देने वाला AI है। इसे नवंबर 2024 में पहली बार पेश किया गया, और 2025 में सभी के उपयोग के लिए खोल दिया गया।

इस मॉडल ने हज़ारों प्रजातियों की आवाज़ें सीखी हैं। यह सरिफ़ आवाज़ सुनकर पहचान लेता है कि वह कौन सा जानवर है। यह नर है या मादा, वयस्क है या बच्चा, इसका भी फ़र्क़ कर लेता है। यह रिकॉर्डिंग में कितने जानवर मौजूद हैं, उनकी गिनती भी कर लेता है। यहाँ तक कि इसने वे क्षमताएँ भी अपने आप दिखाईं जो इसे सिखाई नहीं गई थीं।

इस संस्था को अक्टूबर 2024 में बहुत बड़ी सहायता मिली। इसे कई दानदाताओं से 1.7 करोड़ डॉलर का शोध अनुदान मिला। हमारे पैसों में यह 23 अरब डॉलर से भी अधिक की बड़ी रकम है। इस पैसे से शोध दल का काफी विस्तार किया गया। संस्था का मानना है कि वे 2030 तक जानवरों की आवाज़ों के बड़े रहस्यों को उजागर कर सकते हैं [2]।

2026 की गर्मियों में, कई शोध दलों के एक साथ उपयोग के लिए एक डेटा भंडार भी खोला गया। इस भंडार का नाम पशु भाषा प्रसंस्करण डेटा (alp-data) है। इसमें यहाँ-वहाँ बखिरे जानवरों की आवाज़ों के डेटा को एक जगह इकट्ठा किया गया है। अभी इसमें कौवे, बेलुगा व्हेल और हाथी की आवाज़ों का एक साथ अध्ययन किया जा रहा है। जितना अधिक डेटा इकट्ठा होता है, AI उतनी ही तेज़ी से सीखता है [5]।

इस बदलाव को बढ़ावा देने के लिए एक पुरस्कार भी शुरू हुआ है। यह कॉलर-डोलिटल पुरस्कार (Coller-Dolittle Prize) है। यह जानवरों की आवाज़ों को समझने वाले शोध को बढ़ावा देता है। 2026 की विजेता अमेरिका के बर्कले की डॉ. जूली एली हैं। डॉ. एली ने ज़ेब्रा फ़िच (zebra finch) नाम की एक छोटी चड़िया की आवाज़ के रहस्य को सुलझाया। उन्होंने इस चड़िया द्वारा इस्तेमाल की जाने वाली 11 बुनियादी आवाज़ें खोजीं। इनमें भूख, डर और अभिवादन जैसे अर्थ शामिल थे। डॉ. एली ने भी मशीन लर्निंग नामक AI की मदद ली थी [3]।

AI ने जानवरों की आवाज़ों में इंसानी भाषा जैसा 'बचत का नियम' भी ढूँढ़ निकाला। इंसान अक्सर इस्तेमाल होने वाले शब्दों को छोटा करके बोलते हैं। यानी जो आवाज़ जितनी बार इस्तेमाल होती है, वह उतनी ही छोटी होती जाती है। 2025 में एक वैज्ञानिक ने व्हेल की 16 प्रजातियों की आवाज़ों का अध्ययन किया। और इन आवाज़ों की तुलना इंसानों की 51 अलग-अलग भाषाओं से की। हंपबैक व्हेल और नीली व्हेल के गीतों में भी आवाज़ की यह बचत दिखाई दी। जो आवाज़ जितनी बार गाई जाती थी, वह उतनी ही छोटी थी। आवाज़ को बचाकर इस्तेमाल करने की समझदारी सरिफ़ इंसानों तक सीमिति नहीं थी। हालाँकि, सभी व्हेल ऐसा नहीं करती, इसलिए वैज्ञानिक इसका और अधिक अध्ययन कर रहे हैं [4]।

चमिपांजी की आवाज़ के नियमों, गेलाडा (gelada) की बोली की लय और नग्न मोल-चूहे (naked mole-rat) की बोलियों को याद कीजिए। हमारे इन पड़ोसी जीवों की आवाज़ें हमें सोचने का मौका देती हैं। इंसानी भाषा आसमान से अचानक टपका कोई चमत्कार नहीं है। यह कई जीवों के एक साथ खड़े होने की विशाल नींव पर खलिकर सामने आई है।

AI आवाज़ों के भीतर छपे नियमों को ढूँढ़ निकालता है। जिस काम में इंसान को अकेले कई साल लग जाते, यह उसे कुछ ही दिनों में पूरा कर देता है। इसकी बदौलत वैज्ञानिक कहीं अधिक आवाज़ों का अध्ययन करने में सक्षम हो गए हैं। लेकिन AI जानवरों की बातों का हूबहू अनुवाद नहीं कर देता। यह तो केवल आवाज़ों को अलग-अलग वर्गों में बाँटने और उनके नियम ढूँढ़ने में मदद करता है।

जानवरों की आवाज़ों को समझना उन्हें सुरक्षित रखने में भी मददगार होता है। अगर हमें पता हो कि कौन-सी आवाज़ डर की ओर इशारा करती है, तो हम जान सकते हैं कि जानवर कब परेशानी में है। आवाज़ सुनकर यह अंदाज़ा भी लगाया जा सकता है कि कौन-सी चीज़ उन्हें परेशान कर रही है। आवाज़ को समझने की यह शक्ति ही हमारे इन पड़ोसी जीवों की देखभाल करने की शक्ति बन जाती है।

हम अब तक इन आवाज़ों को जंगल का सरिफ़ शोर ही समझते रहे। हमने उनके रहने की ज़मीन को भी बेपरवाही से उजाड़ दिया। लेकिन ये जीव बहुत लंबे समय से आवाज़ों के ज़रिए अपने झुंड के इतिहास को ज़िदा रखे हुए हैं। अब हमें जो करना है, वह यह नहीं कि हम पहले उन्हें अपनी बनावटी आवाज़ें सुनाएँ। हमें तो वनिमृता से उनकी आवाज़ों को सुनने से शुरुआत करनी होगी।

AI इस सुनने में हमारी मदद करने वाले एक अच्छे कान की तरह काम करता है। यह उन छोटी-छोटी बारीकियों को भी सुन लेता है जो इंसानों से छूट जाती थी। लेकिन क्या सुनना है और क्यों सुनना है, यह इंसान ही तय करता है। एक अच्छे औज़ार के साथ-साथ वनिमृता से सुनने का भाव भी साथ चलना चाहिए। तभी हम इस धरती के अपने पड़ोसी जीवों के साथ लंबे समय तक साथ-साथ रह सकते हैं।

संबंधित संदर्भ सामग्री

- [1] Robinson, D., Miron, M., Hagiwara, M., & Pietquin, O. (Earth Species Project). (2024). NatureLM-audio: An audio-language foundation model for bioacoustics. arXiv:2411.07186. <https://arxiv.org/abs/2411.07186>
- [2] Earth Species Project. (2024, October 17). Earth Species Project secures \$17M in grants to apply frontier AI to decoding animal communication. <https://earthspecies.org/2024/10/17/earth-species-project-secures-17m-in-grants-to-apply-frontier-ai-to-decoding-animal-communication/>
- [3] Jeremy Coller Foundation. (2026, July 3). Scientists win US\$100,000 in 2026 Coller Dolittle Prize. <https://www.jeremycollerfoundation.org/news-and-insights/press-releases/scientists-win-us100-000-in-2026-coller-dolittle-prize>
- [4] Youngblood, M. (2025). Language-like efficiency in whale communication. Science Advances, 11(6), eads6014. <https://doi.org/10.1126/sciadv.ads6014>

[5] Earth Species Project. (2026, July 22). Introducing alp-data: A shared data layer for animal language processing.
<https://earthspecies.org/2026/07/22/introducing-alp-data-a-shared-data-layer-for-animal-language-processing/>

अध्याय 6. आकाश की समिफनी और अदृश्य नृत्य

6.1 पक्षियों और AI के बीच हुई बातचीत

इंसानों की बातचीत पहले से लखी हुई कोई पटकथा नहीं होती। हम सामने वाले की बात और उसके बोलने के अंदाज़ को सुनते हैं। हम इस बात पर भी ध्यान देते हैं कि वह कब थोड़ा रुकता है। फिर उस बहाव के अनुसार अपना जवाब जोड़ते हैं। पक्षियों के गीतों में भी ऐसा ही लेन-देन होता है।

अब तक वैज्ञानिक पक्षियों की आवाज़ों का अध्ययन करते समय उन्हें रिकॉर्डिंग सुनाते थे। इस तरीके को दोबारा आवाज़ सुनाना, यानी प्लेबैक (playback) कहा जाता है। पुराना प्लेबैक पहले से तय छोटी आवाज़ों या तय रिकॉर्डिंग को बजाने का तरीका था। पक्षियों के बीच लंबी बातचीत के बहाव को इसमें शामिल करना मुश्किल था। इसलिए लोगों ने लंबे समय तक पक्षियों के गीत को जन्म से तय आवाज़ ही माना [1]।

इस बाधा को पार करने के लिए दो शोध टीमों ने हाथ मिलाया। ये थीं जानवरों के संवाद को AI से समझने वाला अर्थ स्पीशीज प्रोजेक्ट (Earth Species Project) और कनाडा का मैकगिल विश्वविद्यालय (McGill University)। इसका नेतृत्व डॉ. लोगन जेम्स (Logan James) और प्रोफेसर सारा वूली (Sarah Woolley) ने किया। उन्होंने गौरैया जतिने छोटे पक्षी, मादा ज़ेब्रा फचि (zebra finch) की आवाज़ें इकट्ठी कीं। उनके आपस में बातचीत की यह आवाज़ 1,000 घंटे से भी अधिक की थी [2]।

शोध दल ने इन आवाज़ों से एक AI मॉडल ZF-AIM बनाया। जब पक्षी एक धुन गाता है, तो यह मॉडल तुरंत अगली धुन बनाकर जवाब देता है। यह वैसा ही है जैसे स्मार्टफोन में टाइप करते समय अगला शब्द अपने आप दिखने लगता है। सामने वाले की आवाज़ सुनकर तुरंत जवाब देने के इस तरीके को बातचीत का लेन-देन कहते हैं।

शोध दल ने प्रयोग को बहुत सावधानी से तैयार किया। उन्होंने मादा ज़ेब्रा फचि को एक ध्वनि-रोधी बक्से में अलग रखा। बक्से के अंदर मौजूद पक्षी को माइक्रोफोन और स्पीकर के ज़रिए आवाज़ों का आदान-प्रदान करने दिया गया। जब पक्षी कोई आवाज़ निकालता, तो ZF-AIM उस आवाज़ के अनुसार अगली आवाज़ बनाकर सुनाता था। इस तरह पक्षी और AI ने वास्तविक समय में आवाज़ों का आदान-प्रदान किया।

शोध दल ने तीन स्थितियों की तुलना की। जब पक्षी किसी जीवित पक्षी से बात करता है, जब सिर्फ रिकॉर्डिंग बजाई जाती है, और जब वह वास्तविक समय के AI के साथ बातचीत करता है। स्वाभाविक विशेषताएं केवल तभी सामने आईं जब वह जीवित पक्षी या AI के साथ आवाज़ों का लेन-देन कर रहा था। सिर्फ रिकॉर्डिंग बजाने पर ऐसी विशेषताएं नहीं दिखाई दीं।

नतीजों ने पुरानी सोच को पूरी तरह बदल दिया। पाठ्यपुस्तकों में लिखा था कि मादा ज़ेब्रा फचि की पुकार आनुवंशिकी रूप से तय होती है और कभी नहीं बदलती। लेकिन AI के साथ आवाज़ों का लेन-देन करने वाले पक्षी अलग थे। उन्होंने सामने वाले की आवाज़ के अनुसार अपनी आवाज़ को बदल लिया। इससे साबित हुआ कि पक्षियों की बातचीत सिर्फ तय आवाज़ों का दोहराव नहीं है।

शोध दल ने ज़ेब्रा फचि को ही क्यों चुना? ज़ेब्रा फचि अकेले नहीं रहती हैं। वे कई पक्षियों के झुंड में रहकर बिना रुके आवाज़ों का लेन-देन करती हैं। वे एक-दूसरे को पुकारती हैं, जवाब देती हैं और जगह बदलती हैं। यह सक्रिय बातचीत इस शोध के लिए बिल्कुल सही थी। AI ने भी उनके इस तेज़ लेन-देन की पूरी तरह नकल की।

शोध दल को पक्षियों की बातचीत में तीन विशेषताएं मिलीं। पहली, दोनों पक्षियों की आवाज़ें एक जैसी होने लगीं। दूसरी, उन्होंने बहुत तेज़ी से अपनी आवाज़ बदली। तीसरी, उन्होंने हर किसी आवाज़ पर जवाब नहीं दिया। उन्होंने चुनकर केवल उन्हीं आवाज़ों का जवाब दिया जिनका वे देना चाहते थे। यह वैसा ही है जैसे कई बच्चों के शोरगुल वाली कक्षा में आप केवल अपना नाम पुकारे जाने पर ही मुड़कर देखते हैं।

शोध दल ने मॉडल के एक-एक हिस्से को रोककर देखा। जब उन्होंने लय मिलाने वाले हिस्से को रोका, तो बातचीत बगड़ गई। जब उन्होंने आवाज़ का आकार बनाने वाले हिस्से को रोका, तो कोई दूसरा हिस्सा गड़बड़ा गया। दोनों हिस्से अलग-अलग भूमिका निभा रहे थे।

शोध दल का मानना था कि यह तरीका इंसानों की बातचीत को समझने में भी मदद करता है। इंसान भी बातचीत करते समय अगली बात का अंदाज़ा लगाते हैं। पक्षियों और AI के आदान-प्रदान में भी ऐसा ही अंदाज़ा दिखाई दिया। आवाज़ों के लेन-देन का यह नियम कई जानवरों पर भी लागू हो सकता है।

शोध दल ने दो ZF-AIM मॉडलों को आमने-सामने रखकर बातचीत कराई। इंसानों के देखल के बिना, AI मॉडलों ने आपस में आवाज़ों का आदान-प्रदान किया। उस बातचीत में असली पक्षियों के बीच होने वाली बातचीत की विशेषताएं फिर से उभर आईं। इसका मतलब था कि मॉडल ने नयिमों को अच्छी तरह सीख लिया था।

यह शोध एक और बड़े लक्ष्य की ओर ले जाता है। वह है AI द्वारा बनाई गई आवाज़ों से जानवरों से बात करना। अभी हम उस चरण तक पहुँच चुके हैं जहाँ पक्षी AI की आवाज़ का जवाब देते हैं। अगला चरण इंसानों द्वारा अर्थपूर्ण आवाज़ें भेजने का है। जब वह दनि आएगा, तब शायद इंसान और पक्षी सचमुच आपस में बातचीत कर सकेंगे [3]।

संबंधित संदर्भ सामग्री

- [1] Earth Species Project. (2023, December 3). Using AI to understand animal communication. Earth Species Project. <https://www.earthspecies.org/blog/using-ai-to-understand-animal-communication>
- [2] James, L. S., Hoffman, B., Woolley, S. C., et al. (2026). Leveraging AI-powered interactive playbacks to decipher rules of communication in zebra finches. bioRxiv. <https://doi.org/10.64898/2026.02.12.705387>
- [3] Earth Species Project. (2026, April 2). Animal Language Processing: An AI convergence in animal communication. Earth Species Project. <https://earthspecies.org/2026/04/02/animal-language-processing-an-ai-convergence-in-animal-communication/>

6.2 गीतों में छपि नियमों की खोज

जब जेब्रा फचि AI के साथ आवाज़ों का लेन-देन करती है, तो पक्षी के मस्तिष्क में भी सीखने की प्रक्रिया चलती है। इंसान की तरह पक्षी भी बचपन में बड़ों की आवाज़ सुनकर गाने की नकल करते हैं। इस समय को संवेदनशील अवधि (sensitive period), यानी आवाज़ सीखने का सबसे अच्छा समय कहा जाता है। यह सीखना केवल आनुवंशिकी से तय नहीं होता। साथ रहने वाले झुंड का माहौल भी इस पर गहरा असर डालता है [1]।

छोटी जेब्रा फचि बड़े पक्षियों के गीत सुनकर बड़ी होती है। शुरुआत में वह अनाड़ी तरीके से गाने की नकल करती है। धीरे-धीरे उसका गाना बड़े पक्षियों जैसा होने लगता है। यह इंसानी बच्चे के बोलना सीखने जैसा ही है। झुंड में जितनी बार वह सुनती है, उतनी ही अच्छी तरह सीखती है।

पक्षी के दमिग में एक ऐसी जगह होती है जो गीत के क्रम को नियंत्रित करती है। इसे HVC कहा जाता है, जिससे हम गीत का कमांड सेंटर कह सकते हैं। यह जगह तय करती है कि कौन सी आवाज़ कब और किस क्रम में निकालनी है। यह किसी संगीत निर्देशक द्वारा ताल निर्देशित करने जैसा है।

अमेरिका के यूसी सैन डिएगो (UC San Diego) के प्रोफेसर टिमोथी जेंटनर (Timothy Gentner) की टीम ने यूरोपीय मैना (European starling) का अध्ययन किया। यह पक्षी सैकड़ों छोटी-छोटी आवाज़ों को जोड़कर लंबा गीत गाता है। सरिफ कानों से सुनकर इंसानों के लिये इसके नियमों को समझना बहुत कठिन है। इसलिए शोध दल ने बना लेबल वाले गीतों को AI में डाला [2]।

AI ने मलित-जुलती आवाज़ों के टुकड़ों को एक जगह इकट्ठा किया। उसने गीतों को एक नक्शे की तरह फैला दिया। उस नक्शे पर अलग-अलग आवाज़ों के समूह बन गए। इससे वह बनावट साफ़ दिखाई देने लगी जसि इंसान नहीं देख पाते थे। इस तरीके से कई जानवरों की आवाज़ों की एक साथ तुलना की जा सकी।

शोध दल ने इस तरीके का इस्तेमाल कई प्रजातियों के जानवरों की आवाज़ों पर किया। उन्होंने पक्षियों, चूहों, बंदरों, इंसानों और वृहेल की आवाज़ों को एक ही नक्शे पर रखा। प्रजातियाँ अलग होने के बावजूद उनकी तुलना एक ही तरीके से की जा सकती थी। इससे यह भी पता चला कि आवाज़ की किस विशेषता से यह पहचान होती है कि वह किसने निकाली है।

इंसान लंबे समय तक पक्षियों के गीत सुनकर भी सारे नियम नहीं ढूँढ पाते। ऐसा इसलिए है क्योंकि आवाज़ें बहुत तेज़ होती हैं और उनके बहुत सारे टुकड़े होते हैं। लेकिन AI बना थके हज़ारों घंटों के गीतों की जाँच कर सकता है। वह इंसानों से छूटे हुए क्रम को एक-एक करके पकड़ लेता है।

कैनरी (canary) पक्षी के गीत में भी छपि हुए नियम मिले। अमेरिका के बोस्टन विश्वविद्यालय (Boston University) के शोध दल ने कैनरी के लंबे गीत का विश्लेषण किया। पहले वैज्ञानिक सरिफ पछिली आवाज़ को देखकर अगली आवाज़ का अंदाज़ा लगाने का तरीका अपनाते थे। इस सीधे-सादे तरीके को मार्कोव मॉडल कहा जाता है। लेकिन कैनरी का गीत इतना सरल नहीं था [3]।

शोध दल ने पाया कि गीत का शुरुआती हिस्सा काफी बाद में आने वाली आवाज़ को तय करता है। शुरुआत में गाई गई आवाज़ ने बहुत बाद की आवाज़ का मार्गदर्शन किया। इस गुण को दीर्घकालिक निर्भरता (लॉन्ग-टर्म डिपेंडेंसी) कहा जाता है। यह किसी लंबे वाक्य के पहले और आखिरी शब्द के आपस में जुड़े होने जैसा है। देखने में भले ही लगे कि वे अपनी मर्ज़ी से गा रहे हैं, पर उसमें एक ढाँचा छपि हुआ था।

ऐसे शोध बढ़ते-बढ़ते एक बड़ी तस्वीर सामने आई। अर्थ स्पीशीज प्रोजेक्ट (Earth Species Project) ने कई जानवरों की आवाज़ें एक साथ सीखने वाला AI बनाया। इस मॉडल का नाम NatureLM-audio है। यह एक ऐसा मॉडल है जिसने इंसानी भाषा, संगीत और प्रकृति की आवाज़ों को एक साथ सीखा है। यह मॉडल सरिफ आवाज़ सुनकर पहचान लेता है कि वह किस जानवर की है। यह उन जानवरों की आवाज़ों का

भी अंदाज़ा लगा लेता है जनिहें इसने कभी नहीं सीखा। जब कोई इंसान लिखकर सवाल पूछता है, तो यह आवाज़ सुनकर जवाब देता है। अगर आप पूछें कि यह आवाज़ किस जानवर की है, तो यह उसकी प्रजाति बता देता है। यह यह भी गनि लेता है कि रिकॉर्डिंग में कतिने जीव बोल रहे हैं [4]।

2026 में, अर्थ स्पीशीज प्रोजेक्ट (Earth Species Project) ने इस नई दशा को एक नाम दिया। जानवरों की आवाज़ों को डेटा के रूप में अध्ययन करने वाला यह क्षेत्र 'पशु भाषा प्रसंस्करण' (Animal Language Processing) कहलाया। यह नाम पहली बार 2024 में इस्तेमाल किया गया था। यह इंसानी भाषा पर काम करने वाले प्राकृतिक भाषा प्रसंस्करण (Natural Language Processing) से लिया गया नाम है। पक्षियों के गीतों के नयिमों को समझने का काम भी इसी के तहत आगे बढ़ रहा है [5]।

संबंधित संदर्भ सामग्री

- [1] Earth Species Project. (2023, December 3). Using AI to understand animal communication. Earth Species Project. <https://www.earthspecies.org/blog/using-ai-to-understand-animal-communication>
- [2] Sainburg, T., Thielk, M., & Gentner, T. Q. (2020). Finding, visualizing, and quantifying latent structure across diverse animal vocal repertoires. PLOS Computational Biology, 16(10), e1008228. <https://doi.org/10.1371/journal.pcbi.1008228>
- [3] Markowitz, J. E., Ivie, E., Kligler, L., & Gardner, T. J. (2013). Long-range order in canary song. PLOS Computational Biology, 9(5), e1003052. <https://doi.org/10.1371/journal.pcbi.1003052>
- [4] Earth Species Project. (2024, November 11). Introducing NatureLM-audio: An audio-language foundation model for bioacoustics. Earth Species Project. <https://www.earthspecies.org/blog/introducing-naturelm-audio-an-audio-language-foundation-model-for-bioacoustics>
- [5] Earth Species Project. (2026, April 2). Animal Language Processing: An AI convergence in animal communication. Earth Species Project. <https://earthspecies.org/2026/04/02/animal-language-processing-an-ai-convergence-in-animal-communication/>

6.3 अंधेरे छत्ते में 8 के आकार का नृत्य

आकाश की व्यवस्था सरिफ़ बड़े पक्षियों में ही नहीं होती। यह छोटी मधुमक्खी के पंखों की फड़फड़ाहट में भी मौजूद है। इंसानों द्वारा खोजी गई जानवरों की पहली जटिल बातचीत मधुमक्खियों का नृत्य ही है। ऑस्ट्रियाई विद्वान कार्ल वॉन फ्रिश (Karl von Frisch) ने इसकी खोज की थी और इसके लिए उन्हें नोबेल पुरस्कार मिला था [3]।

जब मधुमक्खी को रस वाली जगह मलि जाती है, तो वह छत्ते में लौट आती है। फिर वह अपने शरीर को हलिते हुए 8 के आकार में घूमती है। इस नृत्य को 8 के आकार का नृत्य या वैगल डांस (waggle dance) कहा जाता है। छत्ते के अंदर बनिा किसी रोशनी के बलिकुल अंधेरा होता है। मधुमक्खी छत्ते की दीवार पर ऊपर की दशा को अपना आधार बनाती है। ऊपर की दशा गुरुत्वाकर्षण के वपिरीत होती है।

मधुमक्खी 8 के आकार के नृत्य के बीच से सीधी रेखा में गुजरती है। इस दौरान वह जसि कोण पर सीधे आगे बढ़ती है, उससे दशा का पता चलता है। ऊपर की दशा के मुकाबले उसका झुकाव कतिना है, वही रस की दशा होती है। शरीर को हलिते के समय से दूरी का पता चलता है। वह जतिनी देर तक हलिती है, जगह उतनी ही दूर होती है। कई किलोमीटर दूर की जानकारी भी सरिफ़ इस एक नृत्य से दी जाती है।

अगर रस बहुत पास हो, तो वह 8 के आकार के बजाय गोल नृत्य करती है। उस समय वह दशा की पूरी जानकारी नहीं देती। क्योंकि वह जगह पास होती है, इसलए ज़ूयादा जानकारी देने की ज़रूरत नहीं होती। जैसे-जैसे दूरी बढ़ती है, वह 8 के आकार के नृत्य में बदल जाती है।

मधुमक्खियों का संकेत सरिफ़ शारीरिक हरकतों तक सीमति नहीं है। उनके सीने का हल्का कंपन और पंखों की आवाज़ भी साथ मलिकर काम करते हैं। छत्ते की दीवार से होकर गुजरने वाला कंपन भी होता है। कई संकेतों को एक साथ इस्तेमाल करने के इस तरीके को मल्टीमॉडल (multimodal) कहा जाता है। यह हाव-भाव, आवाज़ और कंपन को एक साथ इस्तेमाल करने का तरीका है। छत्ते में अंधेरा होने पर भी साथी मक्खियाँ अपने शरीर से इन संकेतों को महसूस कर लेती हैं।

ये संकेत इंसानों की कई इंद्रियों से मलिते-जुलते हैं। हम जब किसी दोस्त से बात करते हैं, तो सरिफ़ आवाज़ का इस्तेमाल नहीं करते। हम आँखों के इशारे और हाथों के हाव-भाव का भी इस्तेमाल करते हैं। मधुमक्खियाँ भी शारीरिक हरकतों, कंपन और आवाज़ का एक साथ उपयोग करती हैं।

वॉन फ्रिश ने इस नृत्य को लंबे समय तक देखकर इसके नयिमों को समझा। उन्होंने एक-एक मधुमक्खी पर रंग से नशान लगाया। उन्होंने यह हसिब लगाया कि कौन सी मधुमक्खी कहाँ जाकर आई थी। यह ऐसा काम था जसि इंसानों को हाथ से करने में बहुत समय लगता था। AI ने वही काम बहुत तेज़ी से कर दिखाया [1]।

जर्मनी के बर्लिन मुक्त विश्वविद्यालय (Free University of Berlin) के प्रोफेसर टिम लैंडग्राफ (Tim Landgraf) की टीम ने इस नृत्य को अपने आप पढ़ने का एक तरीका विकसित किया। शोध दल ने छत्ते में कैमरे लगाए। यह ऐसा तरीका था जिससे मधुमक्खियों को कोई नुकसान नहीं पहुँचता था। AI ने वीडियो में मधुमक्खियों के नृत्य को खुद ही पहचान लिया।

इस AI ने उन हरकतों को भी तेज़ी से पढ़ लिया जो इंसानी आँखों से छूट जाती थीं। उसने मधुमक्खी के सीधे आगे बढ़ने की दिशा को पढ़ा। उसने शरीर हलाने के समय को भी मापा। इन दोनों को मिलाकर उसने रस वाली जगह का अंदाज़ा लगाया। जो काम इंसान अपने हाथों से करते थे, उसे मशीन ने संभाल लिया।

शोध दल यही नहीं रुका। उन्होंने समझी गई नृत्य की जानकारी को एक रोबोट मधुमक्खी में डाला। यह रोबोट छत्ते के अंदर 8 के आकार के नृत्य की नकल करने वाली एक मशीन है। यह हवा में उड़ने वाली हार्वर्ड विश्वविद्यालय की छोटी रोबोट मधुमक्खी से अलग है। जब रोबोट ने छत्ते के बीच में नृत्य किया, तो असली कामकाजी मधुमक्खियों ने उस नृत्य का अनुसरण किया [2]।

कामकाजी मधुमक्खियों ने रोबोट द्वारा बताई गई दिशा में अपनी उड़ान का रुख बदल दिया। शोध दल ने कई वर्षों तक रोबोट की हरकतों और कंपन को और बेहतर बनाया। उन्होंने उसमें फूलों की खुशबू भी जोड़ी। इस तरह उन्होंने ऐसा प्रबंध किया जिससे मधुमक्खियाँ रोबोट के नृत्य का और बेहतर ढंग से पालन कर सकें।

यह प्रयोग दो बातें साबित करता है। पहली बात यह कि इंसानों ने मधुमक्खियों के नृत्य को बिल्कुल सही समझा। अगर रोबोट ने गलत संकेत भेजे होते, तो मधुमक्खियाँ उसका अनुसरण नहीं करती। दूसरी बात यह कि मधुमक्खियाँ एक-दूसरे के नृत्य पर भरोसा करके चलती हैं। एक छोटे से कीट की हरकत में दिशा और दूरी दोनों की जानकारी छिपी थी।

संबंधित संदर्भ सामग्री

[1] Wario, F., Wild, B., Rojas, R., & Landgraf, T. (2017). Automatic detection and decoding of honey bee waggle dances. PLOS ONE, 12(12), e0188626. <https://doi.org/10.1371/journal.pone.0188626>

[2] Landgraf, T., et al. (2018). Dancing honeybee robot elicits dance-following and recruits foragers. arXiv:1803.07126. <https://arxiv.org/abs/1803.07126>

[3] Frisch, K. von. (1967). The dance language and orientation of bees. Harvard University Press.

6.4 हवा में कौवों की तुरि-आयामी बातचीत

उड़ने वाले बुद्धिमत्ता जीवों में कौवे सबसे अनोखे हैं। कौवे अपनी समझदारी के लिए जाने जाते हैं। पक्षियों का दमिग इंसानों के दमिग से छोटा होता है। लेकिन इसमें तंत्रिका कोशिकाएं बहुत घनी रूप से जुड़ी होती हैं। इसीलिए लोग कौवे को 'पंखों वाला वानर' कहते हैं [1]।

स्पेन के लियोन विश्वविद्यालय (University of León) और अर्थ स्पीशीज प्रोजेक्ट (Earth Species Project) के शोध दल ने कौवों का अध्ययन किया। उनका अध्ययन उत्तरी स्पेन में रहने वाले कैरियन कौवे (carrion crow) पर था। शोध दल ने पक्षी के पूँछ के पंखों पर एक बहुत छोटा सेंसर लगाया। इसका वजन 12.5 ग्राम था, यानी बहुत हल्का। इसे ऐसा बनाया गया था कि कुछ समय बाद यह अपने आप गिर जाए। यह ऐसा उपकरण था जिससे पक्षी को कोई चोट नहीं पहुँचती थी।

इस सेंसर में एक छोटा माइक्रोफोन और एक मोशन सेंसर लगा था। माइक्रोफोन ने पक्षी की आवाज़ को रिकॉर्ड किया। मोशन सेंसर ने पंखों और शरीर की हलचल को दर्ज किया। पक्षी ने कब, कौन सी आवाज़ निकाली और वह कैसे हिला, यह सारा डेटा एक साथ इकट्ठा हुआ। यह इंसानों द्वारा कदमों की गिनती करने वाली स्मार्टवॉच जैसा ही उपकरण था।

शोध दल ने तीन वर्षों तक इस सेंसर का इस्तेमाल किया। उन्होंने तब आवाज़ें इकट्ठी की जब बच्चे लगभग दस से पंद्रह दिनों के हो चुके थे। AI ने इकट्ठी की गई आवाज़ों को तीन श्रेणियों में बांटा। सेंसर पहने वयस्क पक्षी की आवाज़, पास मौजूद दूसरे वयस्क पक्षी की आवाज़, और बच्चों की आवाज़। इस तरह यह पता लगाया गया कि कौन कब बोल रहा था।

इकट्ठी की गई आवाज़ों को AI ने अलग-अलग छांट। अर्थ स्पीशीज प्रोजेक्ट (Earth Species Project) ने वोक्सबॉक्स (Voxaboxen) नाम का एक उपकरण बनाया। यह लंबी रिकॉर्डिंग में से एक-एक पुकार को ढूँढ़ निकालने वाला उपकरण है। इस उपकरण ने कौवों की 1,20,000 से अधिक आवाज़ें ढूँढ़ निकालीं। इनमें कुछ ऐसी आवाज़ें भी थीं जो 25 से अधिक वर्षों से कौवों का अध्ययन करने वाले विद्वानों ने भी पहली बार सुनी थीं। ये ऐसी आवाज़ें थीं जो इंसानी कानों के लिए बहुत धीमी थीं और छूट जाती थीं। जब आवाज़ें एक-दूसरे पर चढ़ी हुई थीं, तब भी इस टूल ने उन्हें अलग-अलग कर दिया। कई पक्षियों के एक साथ बोलने वाली रिकॉर्डिंग में भी इसने अपना काम बखूबी निभाया [4]।

शोध टीम ने आवाज़ और गतिविधियों के डेटा की अलग-अलग जांच की। दोनों डेटा को एक साथ रखकर देखना भविष्य का काम था। अभी वे ऐसा डेटा तैयार करने के चरण में हैं जैसे एक साथ देखा जा सके। इस डेटा से कौवों की बातचीत को और गहराई से समझने का रास्ता खुला है।

AI ने कौवों की बात का पूरा मतलब नहीं निकाला है। AI ने सरिफ़ छोटी आवाज़ों को अलग-अलग प्रकारों में बांटने में मदद की। कौवों के मन को समझने का काम बहुत पहले जीववर्ज्ञानियों ने किया था [2]।

अमेरिका के वाशिंगटन विश्वविद्यालय (University of Washington) के प्रोफ़ेसर जॉन मार्ज़लफ़ (John Marzluff) की टीम ने दिखाया कि कौवे चेहरे याद रखते हैं। शोध टीम ने रबर का मुखौटा पहनकर कौवों को पकड़ा। इसके बाद, उसी मुखौटे को पहनने वाले इंसान पर कौवे चलि़लाते हुए झपटने लगे। उन्होंने खतरनाक इंसान का चेहरा याद रखा था। उन्होंने यह बात दूसरे कौवों को भी बताई।

उसी विश्वविद्यालय की केली स्विफ़्ट (Kaeli Swift) और प्रोफ़ेसर मार्ज़लफ़ ने एक दूसरे व्यवहार का अध्ययन किया। कौवे अपने मरे हुए साथी के पास इकट्ठा होकर जोर-जोर से चलि़लाते हैं। लोगों ने इसे उनका अंतिम संस्कार कहा। लेकिन यह सरिफ़ दुख मनाने के लिए नहीं था। यह सीखने का मौका था कि साथी किस वजह से मरा। उन्होंने उस खतरे के बारे में अपने झुंड को खबर दी और इसे लंबे समय तक याद रखा [3]।

पहले लोग कौवे की आवाज़ को सरिफ़ एक शोर मानते थे। AI की मदद से जब इन आवाज़ों को अलग-अलग श्रेणियों में बांटा गया, तो बात कुछ और ही निकली। हर आवाज़ का एक अलग मतलब लग रहा था। ऐसा लगा कि उसमें किसी से सावधान रहने और कहीं जाने जैसी बातें मलि़ी हुई थीं। पक्षियों की बोली इंसानों की भाषा के काफी करीब लग रही थी।

दूसरे पक्षियों पर हुए शोध में वैज्ञानिक एक कदम और आगे बढ़े। 2026 में Earth Species Project ने एक ऐसा AI पेश किया, जो केवल आवाज़ सुनकर किसी खास जानवर को पहचान लेता है। इस शोध में कौवे नहीं, बल्कि ज़ेब्रा फ़िच (zebra finch) शामिल थे। झुंड में हर एक पक्षी की आवाज़ को अलग से पहचानने का रास्ता खुल गया। अब इंसान यह पता लगा सकते हैं कि किस पक्षी ने कब क्या कहा। वे देख सकते हैं कि कौन ज़्यादा बोलता है और कौन शांत रहता है। इस तरह के उपकरण कौवों के अध्ययन में भी इस्तेमाल किए जा सकते हैं [5]।

डेटा जतिना ज़्यादा साझा किया जाता है, शोध उतनी ही तेज़ी से आगे बढ़ता है। जुलाई 2026 में Earth Species Project ने एक नया टूल पेश किया। इसका नाम अल्प-डेटा (alp-data) है। यह इधर-उधर बखिरे जानवरों की आवाज़ के डेटा को एक साथ जोड़ने वाला टूल है। यह 35 से अधिक ध्वनि संग्रहों को एक ही तरीके से लोड करता है। पहले हर शोध टीम डेटा को अलग-अलग तैयार करती थी। अब वे एक ही टूल से मलि़कर इसका इस्तेमाल करते हैं। इस साझे डेटा की मदद से कौवों की आवाज़ का शोध भी तेज़ हो रहा है। AI द्वारा खोले गए आसमान के पीछे, पक्षी हर दिन ऐसी बातचीत कर रहे थे जसिके बारे में इंसान कुछ नहीं जानते थे [6]।

संबंधित संदर्भ सामग्री

- [1] Baglione, V., Canestrari, D., Cusimano, M., Hoffman, B., Moreno, V., & Trapote, E. (2025). Capturing vocal communication in a free-living corvid: high-resolution data from low-impact miniaturized tags. *Animal Cognition*, 28, 85. <https://doi.org/10.1007/s10071-025-02018-0>
- [2] Marzluff, J. M., Walls, J., Cornell, H. N., Withey, J. C., & Craig, D. P. (2010). Lasting recognition of threatening people by wild American crows. *Animal Behaviour*, 79(3), 699-707. <https://www.sciencedirect.com/science/article/abs/pii/S0003347209005806>
- [3] Swift, K. N., & Marzluff, J. M. (2015). Wild American crows gather around their dead to learn about danger. *Animal Behaviour*, 109, 187-197. <https://www.sciencedirect.com/science/article/abs/pii/S0003347215003188>
- [4] Mahon, L., et al. (2025). Robust detection of overlapping bioacoustic sound events. *arXiv:2503.02389*. <https://arxiv.org/abs/2503.02389>
- [5] Earth Species Project. (2026, May 5). From what to who? AI learns to recognize individual animals. Earth Species Project. <https://earthspecies.org/2026/05/05/from-what-to-who-ai-learns-to-recognize-individual-animals/>
- [6] Earth Species Project. (2026, July 22). Introducing alp-data: A shared data layer for Animal Language Processing. Earth Species Project. <https://earthspecies.org/2026/07/22/introducing-alp-data-a-shared-data-layer-for-animal-language-processing/>

अध्याय 7. जानवरों के दमिग से डीप लर्नगि कैनवास तक

7.1 आवाजें खुद सीखने वाला AI

पुराने समय में वैज्ञानिक जानवरों की आवाजों को कानों से सुनते थे और हाथों से नापते थे। वे आवाज के उतार-चढ़ाव और उसकी लंबाई को एक-एक करके लिखते थे। यह तरीका किसी शांत कमरे में तो उपयोगी था। लेकिन समुद्र तट या जंगल में यह ठीक से काम नहीं करता था। ऐसा इसलिए था क्योंकि वहां हवा और लहरों की आवाजें घुल-मिल जाती थीं। जानवरों की ऐसी भी बहुत सी आवाजें थीं जिनका मतलब अभी तक कोई नहीं जानता था। सरिफ़ इंसानी हाथों से करने के लिए यह काम बहुत बड़ा था [1]।

आजकल 'डीप लर्नगि' (deep learning) नाम का AI यह काम संभालता है। डीप लर्नगि के कई तरीके होते हैं। एक तरीका ऐसा है जिसमें इंसान सही उत्तर बताता है। वहीं एक तरीका ऐसा भी है जिसमें सही उत्तर के बनिा ही अपने आप नयिम खोजे जाते हैं। जानवरों की आवाजों पर काम करने वाला AI बनिा सही उत्तर के सीखने वाले तरीके का खूब इस्तेमाल करता है। यह जानवरों की अनगनित आवाजें सुनता है और उनके भीतर छपि नयिमों को खुद खोज निकालता है।

इस तरह व्यापक रूप से सीखने के बाद कई अलग-अलग कामों में इस्तेमाल होने वाला AI भी है। ऐसे AI को जानवरों की आवाज का बुनयिादी मॉडल कहा जाता है। अंग्रेजी में इसे बायोअकाउस्टिक फाउंडेशन मॉडल (Bioacoustic Foundation Model) कहते हैं। फाउंडेशन का मतलब नींव या बुनयिादी आधार होता है। यदानीव मजबूत बना ली जाए, तो किसी नई समस्या के सामने आने पर भी इसे तुरंत इस्तेमाल किया जा सकता है [2, 4]।

बुनयिादी मॉडल एक बार सीखी गई नींव को कई कामों में बांटकर इस्तेमाल करता है। यह आवाजों को अलग-अलग छांटने के काम भी आता है। यह गैर-जरूरी शोर को हटाने के काम भी आता है। कोई नई समस्या आने पर इसे शुरू से दोबारा नहीं सीखना पड़ता। पहले से सीखी गई नींव पर बस थोड़ी सी नई जानकारी जोड़नी होती है। इस तरह नींव के ऊपर किए जाने वाले इन छोटे कामों को व्यावहारिक कार्य कहा जाता है। अंग्रेजी में इन्हें 'डाउनस्ट्रीम टास्क' (downstream task) कहते हैं [2, 4]।

इस क्षेत्र की शुरुआत करने वाले शुरुआती AI में से एक एव्स (AVES) है। AVES ने इंसानों द्वारा दी गई सही उत्तरों की सूची के बनिा सीखा। इसने वभिन्नि जगहों से इकट्ठा की गई लगभग 5,000 घंटों की आवाजें सुनीं। यह इतनी बड़ी मात्रा है कि अगर कोई दनि के पूरे 24 घंटे लगातार सुने, तब भी 200 से ज्यादा दनि लग जाएंगे।

AVES के सीखने का तरीका खाली जगह भरने के खेल जैसा है। AVES आवाज की मूल तरंगों को सीधे स्वीकार करता है। वह इन तरंगों को छोटे-छोटे टुकड़ों में काटता है। फिर कुछ जगहों को जानबूझकर मटिा देता है। इसके बाद वह आगे-पीछे की आवाजों को देखकर मटिाई गई जगहों का सही अंदाजा लगाता है। वह इस खेल को अनगनित बार दोहराता है। वह आवाज को किसी चित्र में बदले बनिा, खुद आवाज से ही सीखता है।

AVES जो सीखता है, वह आवाज की बनावट (पैटर्न) है। वह देखता है कि आवाज कतिनी नयिमतिता से दोहराई जा रही है। वह आवाज के ऊपर चढ़ने और नीचे उतरने के बहाव को देखता है। वह आवाज द्वारा बनने वाले आकार को देखता है। यदिस बनावट को समझ लिया जाए, तो किसी भी जानवर की आवाज को समझा जा सकता है। ऐसा इसलिए है क्योंकि प्रजातियां अलग होने पर भी आवाजों की बुनयिादी बनावट आपस में जुड़ी होती है।

AVES के आने से पहले, सही उत्तरों के लेबल ही सबसे बड़ी रुकावट थे। हर आवाज पर जब इंसान नाम लिखता था, तभी AI सीख पाता था। बनिा लेबल वाली आवाजों से वह सीख नहीं पाता था। AVES ने बनिा लेबल के सीखने का रास्ता खोला। इसी वजह से वह दुनिया भर की तमाा आवाजों को सीखने की सामग्री बना सका।

इस प्रक्रिया से गुजरने के बाद AI आवाज की खूबियों को संख्याओं में बदल देता है। संख्याओं के इस समूह को एम्बेडगि (embedding) कहते हैं। यह आवाज को ऐसे संख्यात्मक नरिदेशांकों में बदलना है जिन्हें कंप्यूटर समझ सके। एम्बेडगि की मदद से किसी भी जानवर की आवाज को एक ही तरीके से समझा जा सकता है।

ये नरिदेशांक कक्षा के सीटगि चार्ट जैसे होते हैं। सीटगि चार्ट देखकर पता चलता है कि कौन कहां बैठा है। एम्बेडगि को देखकर यह जाना जा सकता है कि कौन सी आवाज कहां स्थिति है। मलिती-जुलती आवाजें पास-पास की जगह पर जमा होती हैं। जो आवाजें दूर होती हैं, वे एक-दूसरे से अलग होती हैं।

इस तरीके का एक बहुत बड़ा फायदा है। इंसानों को एक-एक करके सही उत्तर के लेबल नहीं लगाने पड़ते। सही उत्तर लगाना बहुत ज्यादा समय लेने वाला काम है। किसी एक आवाज पर नाम लगाने के लिए भी किसी विशेषज्ञ को उसे खुद सुनना पड़ता है। AI बनिा सही उत्तर के भी आवाज के नयिमों को खोज लेता है। इसलिए वह ऐसी आवाजों को भी समझ सकता है जिनका अर्थ अभी तक किसी को नहीं मालूम। यही बात पुराने तरीकों से बहुत अलग है।

संबंधित संदर्भ सामग्री

- [1] Hagiwara, M. (2023). AVES: Animal vocalization encoder based on self-supervision. ICASSP 2023 - IEEE International Conference on Acoustics, Speech and Signal Processing. arXiv:2210.14493. <https://arxiv.org/abs/2210.14493>
- [2] Robinson, D., Robinson, A., & Akrapongpisak, L. (2023). Transferable models for bioacoustics with human language supervision (BioLingual). arXiv:2308.04978. <https://arxiv.org/abs/2308.04978>
- [4] Robinson, D., Miron, M., Hagiwara, M., & Pietquin, O. (2025). NatureLM-audio: An audio-language foundation model for bioacoustics. ICLR 2025. arXiv:2411.07186. <https://arxiv.org/abs/2411.07186>

7.2 कई जानवरों की आवाजें एक साथ

बुनियादी मॉडल तैयार करने के लिए बहुत सारे ध्वनि डेटा की आवश्यकता होती है। लेकिन अलग-अलग डेटा का प्रारूप (फॉर्मेट) अलग होने के कारण इसमें बहुत मेहनत लगती थी। जुलाई 2026 में, Earth Species Project ने एक मददगार साधन पेश किया। इसका नाम alp-data है। यह 35 से भी ज्यादा जगहों पर बखिरे जानवरों की आवाजों के डेटा को एक साथ जोड़ता है। यह आवाज की गति और लेबलों को एक समान प्रारूप में ढाल देता है। इससे शोधकर्ताओं की डेटा सुधारने की मेहनत बचती है और वे केवल AI को सखिने पर ध्यान लगा पाते हैं [10]।

Earth Species Project ने डेटा को देखने वाला एक वजुअल टूल भी जारी किया। यह कई प्रजातियों के ध्वनि डेटा को एक ही नज़र में देखने वाला साधन है। इस प्रकार डेटा को एक जगह जुटाना और चुनना ही AI को वकिसति करने का पहला कदम है। अच्छा डेटा जितना अधिक होगा, AI उतना ही समझदार बनेगा।

डेटा तैयार होने पर एक और भी बड़ा AI सामने आया। यह डेविड रॉबिन्सन की शोध टीम द्वारा बनाया गया BioLingual है। इस AI ने ज़मीनी और समुद्री जानवरों की सैकड़ों प्रजातियों की आवाजें सीखीं। इसने जसि सामग्री से सीखा, वह आवाजों और उनके वविरण वाले लेखों के जोड़ों का संग्रह था। ऐसे जोड़ों की संख्या 10 लाख से भी अधिक है [2]।

BioLingual आवाज और शब्दों को आपस में जोड़ता है। यह ज्ञानकोश में लिखे जानवरों के वविरण और असल में रिकॉर्ड की गई आवाजों की जोड़ी बनाता है। इस प्रक्रिया में इस्तेमाल होने वाले तरीके को कंट्रास्टिवि लर्निंग (contrastive learning) कहते हैं। यह समान अर्थ वाली आवाज और लिखे गए शब्दों को कंप्यूटर के नक्शे पर पास-पास लाता है। जनिका अर्थ अलग होता है, उन आवाजों और शब्दों को यह एक-दूसरे से दूर कर देता है।

आवाज और शब्दों को जोड़ने का यह वचार CLAP नाम के AI से आया था। इसे कार्नेगी मेलन यूनिवर्सिटी और माइक्रोसॉफ्ट के शोधकर्ताओं ने मलिकर बनाया था। CLAP ने इंसानों की तमाम आवाजों और उन आवाजों को बयां करने वाले शब्दों के जोड़े बनाए थे। BioLingual ने इसी सोच को जानवरों की आवाजों पर लागू किया [3]।

इसकी बदौलत इंसान शब्दों के जरिए आवाजें खोज सकता है। उदाहरण के लिए, यदि कोई लिखे, "बसंत में अपने साथी को पुकारती रॉबिनी पक्षी की ऊंची चहचहाहट"। तो AI अनगणित रिकॉर्डिंग्स में से उस आवाज को ढूँढ निकालता है। इसके लिए उसे पहले से अलग से सखिने की भी जरूरत नहीं होती। इस तरह बना सीखे किसी नए काम को तुरंत कर लेने को जीरो-शॉट (zero-shot) कहते हैं। यह बना अलग प्रशिक्षण के भी एक ही बार में सटीक उत्तर देने की कला है।

BioLingual अकेले ही कई काम कर लेता है। वह पहचानता है कि आवाज निकालने वाला जानवर कसि प्रजाति का है। वह लंबी रिकॉर्डिंग में से उन जगहों को खोज निकालता है जहां जानवरों की आवाजें हैं। यदि लिखकर पूछा जाए, तो वह उसके अनुकूल आवाज ढूँढ देता है। नया डेटा थोड़ा और सीखने पर इसकी क्षमता और बढ़ जाती है। इंसानों में जो काम कई लोगों को बांटकर करना पड़ता, उसे यह अकेला AI संभाल लेता है।

BioLingual ने जो आवाजें सीखीं, वे कई डेटा भंडारों से आई थीं। इनमें पक्षियों की आवाजें सहेजने वाला Xeno-canto शामिल है। साथ ही iNaturalist भी है, जहां लोग प्रकृति के दृश्य और आवाजें साझा करते हैं। यह शोधकर्ताओं द्वारा कई सालों तक एक-एक करके रिकॉर्ड करके जमा किया गया डेटा है। AI ने इन भंडारों को अपनी नींव बनाकर सीखा।

ऐसा AI शोधकर्ताओं के लिए बहुत मददगार होता है। पहले किसी एक आवाज को खोजने के लिए कई-कई दिनों तक रिकॉर्डिंग्स सुननी पड़ती थी। अब केवल एक छोटी सी पंक्ति लिखकर उसे खोजा जा सकता है। इससे ध्वनि के अध्ययन की गति बहुत तेज हो गई है [2, 3]।

संबंधित संदर्भ सामग्री

- [2] Robinson, D., Robinson, A., & Akrapongpisak, L. (2023). Transferable models for bioacoustics with human language supervision (BioLingual). arXiv:2308.04978. <https://arxiv.org/abs/2308.04978>

- [3] Elizalde, B., Deshmukh, S., Al Ismail, M., & Wang, H. (2023). CLAP: Learning audio concepts from natural language supervision. ICASSP 2023. arXiv:2206.04769. <https://arxiv.org/abs/2206.04769>
- [10] Earth Species Project. (2026-07-22). Introducing alp-data: A shared data layer for animal language processing. earthspecies.org. <https://earthspeices.org/2026/07/22/introducing-alp-data-a-shared-data-layer-for-animal-language-processing/>

7.3 आवाज सुनकर शब्दों में समझाने वाला AI

Earth Species Project जानवरों की भाषा को AI से समझने की कोशिश करने वाला एक शोध समूह है। इस समूह द्वारा विकसित किया गया AI है NatureLM-audio। यह AI पक्षियों, ज़मीनी जानवरों और समुद्री जीवों की आवाजों को व्यापक रूप से समझता है [4]।

NatureLM-audio आवाज सुनकर शब्दों में जवाब देता है। जब कोई इंसान इसे जानवर की आवाज सुनाता है, तो यह लिखकर बताता है कि वह जानवर क्या कर रहा है। यह पहचान लेता है कि वह किस प्रजाति का जानवर है। यह रिकॉर्डिंग में कितने जानवर हैं, उनकी गिनती भी कर लेता है। यह यह भी बता देता है कि आवाज निकालने वाला जानवर छोटा बच्चा है या पूरा बड़ा हो चुका है। ये सारे काम इसने बिना अलग से सिखाए खुद ही कर दिखाए।

यह AI आवाज सुनकर लिखने वाला AI है। यह नई आवाजें पैदा करने वाला AI नहीं है। अब तक इंसानी बात समझकर शब्दों में जवाब देने की क्षमता मौजूद थी। अब वही शक्ति जानवरों की आवाजों तक आ पहुंची है। यह जानवरों के लिए बनाया गया पहला बड़ा ऑडियो भाषा मॉडल है।

इस AI की काबिलियत को एक परीक्षा के जरिए परखा जाता है। इस परीक्षा का नाम BEANS-Zero है। यह सिर्फ ऐसे सवालों की परीक्षा है जो इसे पहले कभी नहीं सिखाए गए थे। इसमें इस AI ने प्रजातियों को सही-सही पहचाना। इसने तरह-तरह की आवाजों को श्रेणियों में बांटने का काम भी किया। इसने आवाज सुनकर क्या घट रहा है, यह शब्दों में समझाने का काम भी बखूबी नभाया।

जिस AI ने जितनी व्यापकता से सीखा, वह अनजान कामों में उतना ही माहिर निकला। सिर्फ जानवरों की आवाजें सीखने वाले AI की तुलना में इसका परिणाम बेहतर रहा। ऐसा इसलिए हुआ क्योंकि इसने इंसानों की बोली और पर्यावरण की तमाम आवाजें भी सुनी थीं। सीमिति दायरे में सीखने वाला छोटा AI नए काम के सामने उलझ गया। जब आधार बड़ा हो, तभी नई समस्याओं को भी हल किया जा सकता है।

Earth Species Project इस AI का इस्तेमाल कई जानवरों के शोध में करता है। वे इसका इस्तेमाल कौवों की आवाज का अध्ययन करने में करते हैं। वे बेलुगा व्हेल की आवाज की भी जांच करते हैं। वे हाथियों की आवाज पर भी काम करते हैं। एक ही AI से वे कई जानवरों का एक साथ अध्ययन करते हैं। कई प्रजातियों की आवाजों को एक ही तरीके से संभालने के इस विचार को एनमिल लैंग्वेज प्रोसेसिंग (animal language processing) कहते हैं [11]।

इस AI का इस्तेमाल करना मुश्किल नहीं है। इसे रिकॉर्ड की गई आवाज सुनाई जाती है और सवाल पूछा जाता है। पूछा जाता है कि यह आवाज निकालने वाला जानवर कौन सा है। तब शब्दों में उत्तर मिलता है। यह आवाज को समझने वाले उपकरण से शब्दों के जरिए बात पूछने जैसा है [4, 12]।

जब AI आवाज सुनकर शब्दों में जवाब देता है, तो एक सावधानी बरतनी होती है। AI कभी-कभी ऐसी बातें गढ़ लेता है जो सच नहीं होती, फिर भी सच जैसी लगती हैं। वह किसी गलत प्रजाति का नाम भी ले लेता है। वह ऐसी हरकत भी लिख देता है जो जानवर ने की ही नहीं थी। इस तरह झूठे या मनगढ़ंत जवाब बनाने को हैलुसिनेशन (hallucination) कहते हैं।

ऐसी गलतियों को कम करने का एक तरीका है। वह यह है कि AI जवाब देने से पहले असली रिकॉर्ड में खोज करे। केवल वही बातें लिखने दी जाएं जिनकी पुष्टि जानवरों के वास्तविक रिकॉर्ड से हुई हो। इस तरीके को रिट्रिवल-ऑगमेंटेड (retrieval-augmented) तरीका कहते हैं। यह बिल्कुल वैसा ही है जैसे लाइब्रेरी में किताब ढूँढ़कर पुष्टि करने के बाद कोई बात कही जाए।

Earth Species Project जानवरों की भाषा समझकर प्रकृति की रक्षा करना चाहता है। यदि हम आवाज समझ लें, तो जानवरों की परेशानियों को पहले ही जान सकते हैं। उनके झुंड या संख्या घटने के संकेत आवाजों से पढ़े जा सकते हैं। AI उन संकेतों को इंसानों से कहीं अधिक तेजी से पकड़ लेता है [4, 11]।

संबंधित संदर्भ सामग्री

- [4] Robinson, D., Miron, M., Hagiwara, M., & Pietquin, O. (2025). NatureLM-audio: An audio-language foundation model for bioacoustics. ICLR 2025. arXiv:2411.07186. <https://arxiv.org/abs/2411.07186>
- [11] Earth Species Project. (2026-04-02). Animal language processing: An AI convergence in animal communication. earthspecies.org. <https://earthspeices.org/2026/04/02/animal-language-processing-an-ai-convergence-in-animal-communication/>

7.4 बनिा शब्दकोश के अनुवाद करने का तरीका

दो अलग-अलग भाषाओं का अनुवाद करने के लिए आमतौर पर शब्दकोश (डिक्शनरी) की जरूरत होती है। किसी को यह बताना पड़ता है कि अंग्रेजी का 'dog' हमारी भाषा में 'कुत्ता' है। लेकिन जानवरों की भाषा का कोई शब्दकोश नहीं होता। क्योंकि किसी ने यह लिखकर नहीं रखा कि व्हेल किस आवाज से क्या कहती है। बनिा शब्दकोश के अर्थ का अनुवाद कैसे किया जा सकता है [5]?

2017 में कंप्यूटर वैज्ञानिकों ने एक तरीका ढूंढा। यह दो भाषाओं के बीच एक भी अनुवादित उदाहरण वाक्य न होने पर भी अनुवाद करने का तरीका है। इसे अनसुपरवाइज्ड मशीन ट्रांसलेशन (unsupervised machine translation) कहते हैं। इसका मतलब है कि बिना सही उत्तर सखाए भी AI अपने आप अनुवाद कर लेता है।

व्हेल के लिए अनुवादित वाक्यों की जोड़ियाँ बनाना लगभग असंभव है। ऐसा इसलिए है क्योंकि कोई नहीं जानता कि व्हेल किस आवाज से क्या कह रही है। जोड़ी बनाने के लिए पहले अर्थ पता होना चाहिए। लेकिन अर्थ जानने के लिए पहले से अनुवाद होना जरूरी है। इस बंद रास्ते से नकिलने का सुराग अनसुपरवाइज्ड अनुवाद देता है।

इस तरीके की कुंजी शब्दों द्वारा बनाए गए आकार में छिपी है। जब AI में किसी एक भाषा के बहुत सारे शब्द डाले जाते हैं, तो वह एक नक्शा बनाता है। मलिते-जुलते अर्थ वाले शब्द पास-पास जमा होते हैं। शब्दों के बीच की दूरी और दिशा में एक नियम बन जाता है। नक्शे में पुरुष और महिला के बीच जतिनी दूरी होती है, राजा और रानी के बीच भी ठीक उतनी ही दूरी होती है। शब्दों द्वारा बनाए जाने वाले इस नक्शे को लेटेन्ट स्पेस (latent space) कहते हैं। यह शब्दों के अर्थ को स्थान के रूप में देखने वाला नक्शा है।

यहाँ एक कमाल की बात होती है। जब अलग-अलग भाषाओं के नक्शे बनाए जाते हैं, तो उनकी बनावट लगभग एक जैसी होती है। अंग्रेजी का नक्शा और फनिशि भाषा का नक्शा आपस में मलिते-जुलते हैं। ऐसा इसलिए है क्योंकि इंसानी दमिग दुनिया को लगभग एक ही तरह से बांटता है। इसलिए यदि एक नक्शे को घुमाकर दूसरे नक्शे पर रख दिया जाए, तो शब्द एक-दूसरे के ऊपर सटीक बैठ जाते हैं। अंग्रेजी के 'dog' की जगह पर फनिशि का 'koira' आ बैठता है। इस तरह बनिा शब्दकोश के ही अनुवाद की डिक्शनरी बन जाती है।

शोधकर्ता इस तरीके को व्हेल पर आजमा रहे हैं। Project CETI यही काम कर रहा है। वे स्पर्म् व्हेल और बॉटलनोज डॉल्फिन की आवाज के नक्शे अलग-अलग बनाते हैं। फिर वे दोनों नक्शों को घुमाकर एक-दूसरे के ऊपर बैठाने के तरीके खोजते हैं। दोनों नक्शों को आसानी से एक-दूसरे से मलिते के लिए वे गणति का इस्तेमाल करते हैं। अभी तक सारा अर्थ पूरी तरह सामने नहीं आया है। फलिहाल यह आवाज की बनावट को ढूँढकर नक्शों को एक-दूसरे पर बैठाने की जगह तलाशने का दौर है।

Project CETI ने स्पर्म् व्हेल की आवाजों में इंसानी स्वरों से मलिते-जुलते पैटर्न खोज निकाले। इस पैटर्न का पता लगाने वाले भाषावदि गास्पर बेगुश (Gašper Beguš) हैं। उन्होंने यूनिवर्सिटी ऑफ कैलिफोर्निया, बर्कले में शोध किया। स्पर्म् व्हेल की आवाज के टुकड़ों यानी coda में दो रूप पाए गए। coda उन छोटी-छोटी चटकने जैसी आवाजों का समूह है जो व्हेल बातचीत के समय निकालती हैं। इनमें से एक 'आ' coda है और दूसरा 'ई' coda है। जसि तरह इंसान अपने मुँह के आकार से 'आ' और 'ई' में भेद करते हैं, उसी तरह व्हेल आवाज की गूँज से इनमें अंतर करती हैं। यह शोध 2026 में रॉयल सोसाइटी के जर्नल में प्रकाशित हुआ [6]।

Project CETI ने पहले व्हेल की आवाजों की बनावट का भी पता लगाया था। उन्होंने AI की मदद से स्पर्म् व्हेल के 8,700 coda की जाँच की। इसके नतीजे से पता चला कि coda लेगो (Lego) ब्लॉक की तरह आपस में जुड़ते हैं। गति और लय जैसी कुछ विशेषताएँ कई तरीकों से आपस में मलिते हैं। इस बनावट की तुलना इंसानी भाषा के अक्षरों से करते हुए इसे ध्वनि वर्णमाला (sound alphabet) कहा गया। हालाँकि, प्रत्येक आवाज का क्या अर्थ है, यह अभी तक पता नहीं चला है। यह अभी बनावट के नियमों को देखने वाला चरण है [7]।

उसी साल मार्च में, Project CETI ने स्पर्म् व्हेल के बारे में एक और खोज की जानकारी दी। उन्होंने बच्चे के जन्म के पल को वसितार से दर्ज किया। उस समय उन्होंने कई व्हेलों को मलिकर बच्चे की देखभाल करते देखा। यह अध्ययन व्हेलों की मलिकर देखभाल करने की आदत को दिखाता है। यह आवाजों के स्वर पैटर्न को बताने वाले अध्ययन से अलग एक दूसरा शोध है [13]।

आवाजों की बनावट को ढूँढना बनिा शब्दकोश वाले अनुवाद में मदद करता है। अगर व्हेल की आवाज में इंसानी भाषा जैसी बनावट हो, तो दोनों नक्शों को मलिते की जगह ढूँढना आसान हो जाता है। आवाज को जतिना बारीक बांटा जाता है, नक्शा उतना ही वसितृत बनता है। वसितृत नक्शे आपस में एक-दूसरे के ऊपर बेहतर तरीके से बैठते हैं [5, 6, 7]।

- [5] Goldwasser, S., Gruber, D. F., Kalai, A. T., & Paradise, O. (2023). A theory of unsupervised translation motivated by understanding animal communication. NeurIPS 2023. arXiv:2211.11081. <https://arxiv.org/abs/2211.11081>
- [6] Beguš, G., Dąbkowski, M., Sprouse, R. L., Gruber, D. F., & Gero, S. (2026). The phonology of sperm whale coda vowels. Proceedings of the Royal Society B. <https://doi.org/10.1098/rspb.2025.2994>
- [7] Sharma, P., Gero, S., et al. (2024). Contextual and combinatorial structure in sperm whale vocalisations. Nature Communications, 15, 3617. <https://doi.org/10.1038/s41467-024-47221-8>
- [13] Project CETI / Harvard SEAS. (2026-03). Studies documenting rare sperm whale birth and ancient cooperative care released. EurekAlert! <https://www.eurekalert.org/news-releases/1120823>

7.5 प्रजातियों की सीमा पार करती AI और पार न होने वाली दीवार

कभी-कभी केवल एक जानवर की आवाज सीखने वाली AI दूसरे जानवरों की आवाजें भी पहचान लेती है। इसे प्रजातियों की सीमा पार करने वाली अनुकूलन क्षमता कहते हैं। अंग्रेजी में इसे क्रॉस-स्पीशीज जनरलाइजेशन (cross-species generalization) कहा जाता है। जैसे पक्षियों की आवाज सीखने वाली AI मेंढक की आवाजों को भी छंट लेती है [8, 9]।

इसका सबसे स्पष्ट उदाहरण Google का Perch है। साल 2025 में आया Perch 2.0 मुख्य रूप से पक्षियों की आवाजों से सीखा था। इसने पक्षियों सहित जमीन पर रहने वाले जानवरों की बहुत सारी आवाजें सुनी थीं। इसमें 14,500 से अधिक प्रजातियों के जानवर शामिल हैं। इस AI ने समुद्र के भीतर की आवाजें लगभग नहीं सीखी थीं। फिर भी इसने व्हेल की आवाजों को अच्छी तरह पहचान लिया। थोड़े से डेटा से बने क्लासिफायर में इसने बहुत अच्छा प्रदर्शन किया। पक्षियों की आवाज सीखने वाली AI ने व्हेल को पहचान लिया [8, 9] [8]।

Perch ने सरिफ व्हेल की ही दीवार पार नहीं की। पक्षियों की आवाज से सीखने वाले Perch ने मेंढकों की आवाजों को भी सुलझाया। ब्राजील के अमेज़न के मेंढकों की आवाज के डेटासेट को AnuraSet कहा जाता है। Perch ने इन मेंढकों की आवाजों को अलग-अलग प्रकारों में बांट दिया। मेंढक की आवाज कभी न सीखने के बावजूद उसने यह कर दिखाया [8, 9]।

यह इसका उज्ज्वल पहलू है। लेकिन ऐसी बड़ी दीवारें भी हैं जिनमें AI पार नहीं कर सकती।

पहला, डेटा बहुत कम है। इंसानी भाषा सीखने वाली AI लाखों घंटों से भी अधिक की रिकॉर्डिंग सुनकर बड़ी हुई है। जंगली जानवरों की सारी आवाजें जमा कर लें, तो भी उससे बहुत कम है। इसके अलावा, पुराने रिकॉर्डिंग उपकरण सरिफ तेज आवाजें ही पकड़ पाते थे। व्हेलों द्वारा पास में की जाने वाली धीमी फुसफुसाहट छूट जाती थी। इस धीमी आवाज को सॉफ्ट क्लिक (soft click) कहते हैं। इस तरह महत्वपूर्ण आवाजें डेटा से गायब रह गई [6]।

सॉफ्ट क्लिक वह धीमी आवाज है जो व्हेल एक-दूसरे के करीब होने पर निकालती हैं। इंसानों के हिसाब से यह कानाफूसी जैसा है। इस धीमी आवाज में दिल की बात बांटने का भाव छिपा हो सकता है। पुराने रिकॉर्डर इस आवाज को नहीं पकड़ पाए, इसलिए जरूरी बातचीत छूट गई।

दूसरा, उनके रहने की दुनिया बहुत अलग है। स्पर्म व्हेल बर्ना रोशनी वाले गहरे समुद्र में रहती हैं। वे अपने माथे की तेल की थैली से अल्ट्रासाउंड भेजकर अपने आसपास की दुनिया को आवाज से देखती हैं। इंसान आंखों से दुनिया को देखता है। आवाज से दुनिया देखने वाली व्हेल और आंखों से देखने वाले इंसान के महसूस करने का तरीका अलग है। इसलिए दोनों नक्शे पूरी तरह एक-दूसरे के ऊपर नहीं बैठ सकते।

हर प्राणी द्वारा अलग तरह से महसूस की जाने वाली दुनिया को Umwelt कहते हैं। इसका अर्थ 'आसपास का संसार' है। यह शब्द जीवविज्ञानी याकोब फॉन उएक्सकुएल (Jakob von Uexküll) ने बनाया था। इसका मतलब है कि हर जानवर अपनी इंद्रियों से जसि दुनिया को महसूस करता है, वह अलग होती है। दार्शनिक लुडविग वट्टिगेन्स्टाइन (Ludwig Wittgenstein) ने भी ऐसी ही बात कही थी। उन्होंने कहा था कि अगर शेर बोल भी सकता, तब भी इंसान उसे समझ नहीं पाता। ऐसा इसलिए है क्योंकि शेर और इंसान के रहने की दुनिया अलग है।

तीसरा, लापरवाही से इस्तेमाल करने पर यह खतरनाक हो सकता है। AI मादा पक्षी की आवाज की नकल करके नर पक्षी को बुला सकती है। यह साथी ढूँढने और बच्चों को पालने वाले पक्षियों की व्यवस्था को हिला देने वाली बात है। नकली आवाजों से जंगली जानवरों को धोखा देने की इस हरकत को पारस्थितिक धोखा (ecological deception) कहा जाता है [14]।

खतरा सरिफ नकल करने से भी आगे जाता है। AI जानवरों की प्रतिक्रिया देखकर आवाज को बदल सकती है। जब जानवर जवाब देता है, तो वह उस जवाब के अनुसार आवाज को सुधारती है। इस तरह जानवरों की प्रतिक्रिया को तुरंत देखते हुए आवाज बदलने के प्रयोग को प्रतिक्रियाशील ध्वनि प्रयोग कहा जाता है। सही तरीके से इस्तेमाल करने पर यह बातचीत की चाबी बन सकता है। लेकिन गलत इस्तेमाल होने पर यह जानवरों को धोखा देने का जरिया बन जाता है।

व्हेल करोड़ों सालों से अपनी-अपनी ध्वनिसंस्कृतिको नभाती आ रही हैं। नकली आवाज इस पुरानी संस्कृतिको बगिड़ सकती है। आवाजों की व्यवस्था एक बार बगिड़ जाए, तो उसे वापस लाना मुश्किल होता है। इसलिए उनकी आवाज की नकल करने से पहले हमें बहुत सावधान रहना होगा।

बड़ी AI को जंगल और समुद्र में ले जाने में भी परेशानी होती है। बड़ी AI भारी होती है और बहुत अधिक ऊर्जा खर्च करती है। जंगल के रिकॉर्डर में इतनी बड़ी चीज को रखना मुश्किल होता है। इसलिए शोधकर्ता बड़ी AI के जरूरी हिस्से को ही छोटी AI में डालते हैं। यह एक बड़े बर्तन की मुख्य सामग्री को छोटे बर्तन में डालने जैसा है।

इसलिए शोधकर्ता कहते हैं कि हमें कुछ नियम बनाने चाहिए जिनका पालन किया जाए। वह नियम जानवरों की आवाज में बना सोचे-समझे दखल न देने का है। इसे सबसे पहले माना जाने वाला सदिधांत कहा जाता है। हमें जानवरों की आवाज की नकल करने से पहले चुपचाप सुनना सीखना चाहिए।

व्हेल की भाषा समझने से कानून भी बदल सकता है। साल 2026 में व्हेलों को कानूनी अधिकार देने की चर्चा हुई। जब हमें आवाजों से यह पता चलेगा कि जानवर क्या महसूस करते हैं और किस दौर से गुजरते हैं, तब इंसान उनके साथ मनमाना व्यवहार नहीं कर पाएगा। आवाजों को समझना जानवरों की रक्षा करने से जुड़ा हुआ है [15]।

AI ने जानवरों की भाषा की ओर जाने वाला नया रास्ता खोला है। जिस काम में इंसानों को सैकड़ों साल लगते, उसे इसने कुछ ही वर्षों में समेट दिया। लेकिन रास्ता खुल जाने का यह मतलब नहीं कि हम मंजिल पर पहुंच गए हैं। डेटा अब भी बहुत कम है, और जानवरों की दुनिया इंसानों की दुनिया से अलग है। AI सिर्फ एक उपकरण है। उस उपकरण का इस्तेमाल कैसे करना है, यह इंसान के हाथ में है। वनिमृता से सुनने का भाव तकनीक से पहले आना चाहिए।

संबंधित संदर्भ सामग्री

- [6] Beguš, G., Dąbkowski, M., Sprouse, R. L., Gruber, D. F., & Gero, S. (2026). The phonology of sperm whale coda vowels. *Proceedings of the Royal Society B*. <https://doi.org/10.1098/rspb.2025.2994>
- [8] Burns, A., et al. (2025). Perch 2.0: The bittern lesson for bioacoustics. arXiv:2512.03219. <https://arxiv.org/abs/2512.03219>
- [9] Choi, C. Q. (2025). AI trained on birdsong can recognize whale calls. *IEEE Spectrum*. <https://spectrum.ieee.org/foundation-models-google-birds-whales>
- [14] Beguš, G. (2024). New patterns in animal communication with AI: Ethical and legal implications. *Interspecies Internet Lectures*.
- [15] Bolling, R. (2026-06-11). Decoding whale language could unlock new legal rights for them. *Boing Boing*. <https://boingboing.net/2026/06/11/decoding-whale-language-could-unlock-new-legal-rights-for-them.html>

अध्याय 8. जानवरों का परविशी संसार और संकेत-वज्जिज्ञान की सीमाएं

8.1 सबकी अपनी अलग दुनिया, परविशी संसार

दुनिया केवल एक नहीं है। हर जीव अपनी एक अलग दुनिया में जीता है। इस विचार को सबसे पहले स्पष्ट रूप से बताने वाले एक व्यक्ति थे। वे एस्टोनिया में जन्मे जीववज्जिज्ञानी याकॉब फॉन उएक्सकुल (Jakob von Uexküll) थे। वे ऐसे विद्वान थे जिनका मानना था कि हर जानवर दुनिया को अलग तरह से महसूस करता है [1]।

उएक्सकुल ने इस दुनिया को परविशी संसार कहा। जर्मन भाषा में इसे उमवेल्ट (Umwelt) कहा जाता है। परविशी संसार (Umwelt) वह अपनी अलग दुनिया है जसि कोई जीव अपनी इंद्रियों से महसूस करता है और बनाता है। इंसान आँखों से रंग देखते हैं और कानों से आवाज़ सुनते हैं। लेकिन सभी जानवर इंसानों की तरह महसूस नहीं करते।

परविशी संसार शब्द समझना कठिन नहीं है। एक ही खेल के मैदान में होने पर भी कुत्ता और इंसान अलग-अलग चीजें महसूस करते हैं। कुत्ता सूंघकर दुनिया को समझता है और इंसान आँखों से देखकर समझता है। हर जीव अपनी इंद्रियों के अनुसार अपनी दुनिया बनाता है।

स्पर्म व्हेल (Physeter macrocephalus) को ही लीजिए। स्पर्म व्हेल सैकड़ों मीटर गहरे समुद्र में उतरती है। वहाँ रोशनी की एक करिण भी नहीं होती, बस घना अंधेरा होता है। उस अंधेरे में रंगों का कोई मतलब नहीं होता। इसके बजाय स्पर्म व्हेल अपने माथे से तेज़ आवाज़ छोड़ती है। वह उस आवाज़ के टकराकर वापस आने को सुनकर दुनिया को समझती है। स्पर्म व्हेल की दुनिया रंगों से नहीं, बल्कि आवाज़ों से बनी दुनिया है [2]।

इंसानों ने लंबे समय तक एक भ्रम पाले रखा। इंसान सोचते थे कि जो वे महसूस नहीं कर सकते, उसका कोई अस्तित्व ही नहीं है। इंसान अपनी ही इंद्रियों को एकमात्र पैमाना मानते थे। इस नज़रिए को मानव-केंद्रितता (anthropocentrism) कहा जाता है। मानव-केंद्रितता का अर्थ है केवल इंसानी इंद्रियों और विचारों को ही दुनिया का मानक मानना। इसी सोच के कारण इंसान जानवरों की आवाज़ों को केवल रोना या साधारण प्रतिक्रिया भर मानते रहे।

जानवरों की आवाज़ों और उनके हाव-भाव का अध्ययन करने वाला एक विज्ञान है। इसे पशु संकेत-विज्ञान (Zoosemiotics) कहते हैं। पशु संकेत-विज्ञान वह अध्ययन है जो जानवरों द्वारा आपस में आदान-प्रदान किए जाने वाले संकेतों के अर्थ की खोज करता है। संकेत का मतलब है अर्थ से भरी आवाज़, गंध या शारीरिक हलचल। अमेरिकी विद्वान थॉमस सेबेओक (Thomas Sebeok) ने इस अध्ययन का नाम रखा था। यह विज्ञान उएक्सकुल के परविशी संसार (Umwelt) के विचार पर आधारित है। अर्थपूर्ण संकेत केवल इंसानी भाषा में ही नहीं होते। सभी जीव अपनी-अपनी दुनिया में संकेतों का आदान-प्रदान करते हैं [3]।

उएक्सकुल का विचार नए सवाल की ओर ले जाता है। इतालवी विद्वान निकोला जेंगियारो (Nicola Zengiaro) ने 'उत्पादक पशु संकेत-विज्ञान' (generative zoosemiotics) का विचार प्रस्तुत किया। उनका कहना है कि AI द्वारा जानवरों के संकेतों की मशीनी नकल करने भर से उनका अर्थ नहीं समझा जा सकता। यानी जानवर जसि पर्यावरण में रहते हैं और जो हाव-भाव दिखाते हैं, उन्हें साथ देखकर ही सही अर्थ समझा जा सकता है [4]।

अब AI इस विषय को नए तरीके से खोल रहा है। AI गणना के ज़रिये इंसानों से कहीं अधिक विस्तृत ध्वनिकी क्षेत्रों को पढ़ता है। इंसान जसि खाली सन्नाटा समझते थे, वह जगह आवाज़ों से भरी हुई थी। समुद्र, जंगल और ज़मीन के नीचे लगातार संकेत आ-जा रहे थे। AI उन छपे हुए नयियों को खोजने में मदद करता है।

यह मुद्दा इंसानी समाज से भी जुड़ा हुआ है। Earth Species Project की स्थापना करने वाले अज़ा रस्कनि (Aza Raskin) कहते हैं कि तकनीक इंसानों को एक-दूसरे से अलग कर रही है। इंसान प्रकृति से भी दूर हो गए हैं। AI की मदद से जानवरों की आवाज़ें सुनना इस टूटे हुए रिश्ते को फेर से जोड़ने जैसा है [5]।

संबंधित संदर्भ सामग्री

- [1] Bakker, K. (2022). The sounds of life: How digital technology is bringing us closer to the worlds of animals and plants. Princeton University Press. <https://doi.org/10.2307/j.ctv2hnkcc3>
- [2] Bakker, K. (2022). The sounds of life: How digital technology is bringing us closer to the worlds of animals and plants. Princeton University Press. <https://doi.org/10.2307/j.ctv2hnkcc3>
- [3] Sebeok, T. A. (1972). Perspectives in zoosemiotics. Mouton.
- [4] Zengiaro, N. (2025). AI and animal communication: A generative zoosemiotics perspective. Linguistic Frontiers, 8(2). <https://doi.org/10.2478/lf-2025-0021>

8.2 इंसानी इंद्रियों से परे की आवाज़ें और कंपन

इंसानी कानों की एक सीमा होती है। इंसान केवल 20 हर्ट्ज़ (Hz) से 20,000 हर्ट्ज़ के बीच की आवाज़ें ही सुन सकते हैं। हर्ट्ज़ वह इकाई है जिससे यह नापा जाता है कि आवाज़ 1 सेकंड में कतिनी बार कांपती है। इससे अधिक या कम की आवाज़ें इंसानों को सुनाई नहीं देती हैं। लेकिन वे आवाज़ें गायब नहीं होती हैं। AI इन आवाज़ों को पकड़कर इंसानों के सामने प्रस्तुत करता है [1]।

पहले पौधों की आवाज़ को देखते हैं। लंबे समय तक माना जाता था कि पौधों का आवाज़ से कोई लेना-देना नहीं है। लेकिन इज़राइल के तेल अवीव विश्वविद्यालय के प्रोफेसर योसी योवेल (Yossi Yovel) और प्रोफेसर लीलाख हदानी (Lilach Hadany) के शोध दल ने इस सोच को बदल दिया। पानी की कमी होने पर या तना कटने पर टमाटर और तंबाकू के पौधों ने पराश्रव्य तरंगें (अल्ट्रासाउंड) निकालीं। अल्ट्रासाउंड ऐसी बहुत ऊँची आवाज़ होती है जिसे इंसान नहीं सुन सकते। वह आवाज़ इंसानों की सामान्य बातचीत जतिनी तीव्रता से हवा में फैली।

पौधे आवाज़ क्यों निकालते हैं? जब पानी सूखता है, तो तने के भीतर की पतली जल नलिकाओं में हवा के छोटे बुलबुले बन जाते हैं। जब ये बुलबुले फूटते हैं, तो 'क्लिक' जैसी आवाज़ आती है। यह इंसानी कानों को सुनाई नहीं देती, लेकिन मशीनें इसे पकड़ लेती हैं।

इस आवाज़ को डीप लर्निंग AI में डाला गया। डीप लर्निंग ऐसा AI है जो आवाज़ की विशेषताओं को अपने आप पहचान लेता है। तब पानी की कमी वाली आवाज़ और चोट लगने वाली आवाज़ अलग-अलग समूहों में बंट गईं। दोनों आवाज़ों की तरंगों का आकार (वेवफॉर्म) बिल्कुल अलग था। वेवफॉर्म पानी की लहरों की तरह हलने वाली ध्वनि का रूप होता है। कीड़ों जैसे जानवर जो अल्ट्रासाउंड सुन सकते हैं, वे इस आवाज़ को सुन सकते हैं।

पौधे आवाज़ सुनते भी हैं। प्रोफेसर योवेल और प्रोफेसर हदानी के शोध दल ने शाम की प्रमिरोज़ (evening primrose) के फूल का अध्ययन किया। उन्होंने फूल के पास मधुमक्खी के उड़ने की आवाज़ बजाई। तब फूल ने 3 मिनट के भीतर अपने रस (अमृत) को और अधिक मीठा बना दिया। फूल ने अन्य अवांछित आवाज़ों पर कोई प्रतिक्रिया नहीं दी। यानी फूल की पंखुड़ियों ने आवाज़ इकट्ठा करने वाले कान की तरह काम किया [2]।

अमेरिका के मसिसौरी विश्वविद्यालय की प्रोफेसर हेइडी एप्पल (Heidi Appel) और प्रोफेसर रेक्स कॉक्रॉफ्ट (Rex Cocroft) के शोध दल ने भी ऐसा ही एक प्रयोग किया। उन्होंने अरेबिडोप्सिस (Arabidopsis) नाम के एक छोटे पौधे को कई तरह की आवाज़ें सुनाईं। हवा और बारिश की आवाज़ पर वह पौधा शांत रहा। लेकिन जब उन्होंने इल्ली (कैटरपिलर) के पतले चबाने का कंपन सुनाया, तो परणाम अलग था। पौधे ने तुरंत कीड़ों को भगाने वाला पदार्थ बनाना शुरू कर दिया। पौधे ने पतले चबाने के कंपन को अन्य कंपनों से अलग पहचान लिया था [3]।

2026 में सामने आए एक अध्ययन ने इस तथ्य को एक कदम और आगे बढ़ाया। प्रोफेसर योवेल और प्रोफेसर हदानी के शोध दल ने पतंगों (moths) का अवलोकन किया। मादा पतंगों ने अंडे देने की जगह चुनते समय पौधों की आवाज़ सुनी। पतंगों ने सूखे पौधों की आवाज़ से बचते हुए स्वस्थ पौधों को चुना। यह पहला सबूत था कि जानवर पौधों की आवाज़ पर प्रतिक्रिया करते हैं [4]।

जमीन के नीचे और छोटे जीवों की दुनिया में भी ऐसा ही होता है। मधुमक्खियों अंधेरे छतों के भीतर '8' के आकार का नृत्य करती हैं। यह नृत्य फूलों की दिशा और दूरी बताने वाला संकेत होता है। लेकिन यह संकेत सरिफ़ आँखों से देखने वाला नृत्य ही नहीं था। इसमें पंखों का फड़कना और छतों में फैलने वाला कंपन भी शामिल था। प्रोफेसर टिम लैंडग्राफ (Tim Landgraf) के शोध दल ने AI की मदद से इस नृत्य को समझा। रोबोट और कैमरों ने एक महीने से अधिक समय तक छतों पर नज़र रखी। AI ने वीडियो में देखने वाले हर एक नृत्य का अपने आप विश्लेषण कर दिया [5]।

चूहे भी ऐसी आवाज़ों में बात करते हैं जिन्हें इंसान नहीं सुन सकते। चूहे घबराहट या खुशी के समय अल्ट्रासाउंड निकालते हैं। वाशिंगटन विश्वविद्यालय के केविन कॉफी (Kevin Coffey) के शोध दल ने डीपस्क्वीक (DeepSqueak) नाम का AI बनाया। इस AI ने चूहों की आवाज़ों को छोटे-छोटे टुकड़ों में बांटकर समझा। नतीजतन, आवाज़ के ज़रिये चूहों की मानसिक स्थितियों को समझा जा सका [6]।

चूहों की आवाज़ इंसानी कानों से नहीं पकड़ी जा सकती। डरने पर चूहे लगभग 22 किलोहर्ट्ज़ (kHz) की आवाज़ निकालते हैं। खुश होने पर वे 50 किलोहर्ट्ज़ जैसी ऊँची आवाज़ निकालते हैं। किलोहर्ट्ज़ का मतलब 1,000 हर्ट्ज़ होता है। DeepSqueak ने इस आवाज़ को चित्र की तरह बदलकर उनकी भावनाओं का एक नक्शा तैयार किया।

ये सभी खोजें मलिकर एक ही बात कहती हैं। इंसान नहीं सुन सकते इसका मतलब यह नहीं है कि आवाज़ मौजूद नहीं है। धरती, घास और छोटे जीव हमेशा आपस में संकेतों का आदान-प्रदान कर रहे थे। AI इन छिपी हुई आवाज़ों को इंसानों के सामने खोलकर रख देता है।

संबंधित संदर्भ सामग्री

- [1] Khait, I., Lewin-Epstein, O., Yovel, Y., & Hadany, L. (2023). Sounds emitted by plants under stress are airborne and informative. *Cell*, 186(7), 1328-1336. <https://doi.org/10.1016/j.cell.2023.03.009>
- [2] Veits, M., et al. (2019). Flowers respond to pollinator sound within minutes by increasing nectar sugar concentration. *Ecology Letters*, 22(9), 1483-1492. <https://doi.org/10.1111/ele.13331>
- [3] Appel, H. M., & Coccoft, R. B. (2014). Plants respond to leaf vibrations caused by insect herbivore chewing. *Oecologia*, 175(4), 1257-1266. <https://doi.org/10.1007/s00442-014-2995-6>
- [4] Seltzer, R., Zer Eshel, G., Hadany, L., & Yovel, Y. (2026). Female moths incorporate plant acoustic emissions into their oviposition decision-making process. *eLife*, 14, e104700. <https://doi.org/10.7554/eLife.104700>
- [5] Landgraf, T., et al. (2024). Autonomous tracking of honey bee behaviors over long-term periods with cooperating robots. *Science Robotics*, 9(95). <https://doi.org/10.1126/scirobotics.adn6848>
- [6] Coffey, K. R., Marx, R. G., & Neumaier, J. F. (2019). DeepSqueak: A deep learning-based system for detection and analysis of ultrasonic vocalizations. *Neuropsychopharmacology*, 44(5), 859-868. <https://doi.org/10.1038/s41386-018-0303-6>

8.3 कई इंद्रियों को एक साथ समझने वाला AI

जानवर सिर्फ आवाज़ से ही बात नहीं करते। आवाज़, शारीरिक हाव-भाव और गंध मलिकर अर्थ बनाते हैं। इस तरह कई माध्यमों का एक साथ इस्तेमाल करने वाले संचार को मल्टीमॉडल (multimodal) कहा जाता है। यह सिर्फ आवाज़ ही नहीं, बल्कि निज़रों, हाव-भाव और गंध को भी मिलाकर किया जाने वाला संचार है [1]।

सिर्फ एक आवाज़ को देखने पर उसका अर्थ छूट जाना आसान होता है। एक ही आवाज़ का अर्थ भी हाव-भाव के अनुसार बदल जाता है। इंसान भी जब मुस्कराते हुए वही बात कहते हैं और गुस्से में कहते हैं, तो उसका मतलब अलग होता है। इसलिए कई संकेतों को एक साथ देखने पर ही बात समझ आती है।

बहामास के समुद्र में चित्तीदार डॉल्फ़िन पर किया गया अध्ययन इसका एक अच्छा उदाहरण है। डॉ. डेनिस हरज़िंग (Denise Herzing) और प्रोफेसर थैड स्टार्नर (Thad Starner) ने CHAT नाम की एक मशीन बनाई। यह मशीन डॉल्फ़िन और इंसानों के बीच आवाज़ों के आदान-प्रदान में मदद करती है। लेकिन सिर्फ आवाज़ से बातचीत को पूरी तरह समझना संभव नहीं था। डॉल्फ़िन द्वारा इंसान से आँख मलाने का कृष्ण और उसके शरीर का कोण भी साथ में दर्ज करना पड़ा। कई संकेतों को एक साथ देखने पर ही अर्थ समझ में आया।

कौओं पर किया गया अध्ययन भी दिखाता है कि कई संकेत कितने महत्वपूर्ण हैं। अमेरिका के वाशिंगटन विश्वविद्यालय के प्रोफेसर जॉन मारज़लुफ़ (John Marzluff) के शोध दल ने कौओं का अध्ययन किया। शोधकर्ताओं ने एक विशेष मुखौटा पहनकर कौओं को पकड़ा और फरि छोड़ दिया। उसके बाद जब उसी मुखौटे को पहने हुए कोई व्यक्ति आया, तो कौवे ज़ोर-ज़ोर से चिल्लाते हुए इकट्ठा हो गए। कौवों ने खुद को परेशान करने वाले व्यक्ति का चेहरा याद रखा था। वह व्यक्ति जब भी आता, वे चेतावनी की आवाज़ निकालकर अपने साथियों को खतरे से आगाह करते थे [2]।

यह याद लंबे समय तक बनी रही। कौवे 2 साल से अधिक समय तक उस चेहरे को नहीं भूले। सिर्फ आवाज़ सुनकर इस सच्चाई का पता लगाना मुश्किल था। कौवों की आवाज़ और उनके व्यवहार को एक साथ देखने पर ही बात समझ में आई।

इस तरह के शोध को आगे बढ़ाने के लिए Earth Species Project आगे आया। यह संस्था AI की मदद से जानवरों की भाषा को समझने वाला एक शोध समूह है। इनका कहना है कि हमें सिर्फ आवाज़ पर ध्यान देने के तरीके से आगे बढ़ना चाहिए। क्योंकि जानवरों के शरीर पर लगाए गए सेंसर द्वारा जुटाया गया डेटा मौजूद है। सेंसर आवाज़, हरकत और पानी की गहराई जैसी चीज़ों को मापते हैं [3][4]।

जानवरों के शरीर पर लगे सेंसर कई चीज़ों को मापते हैं। वे तैरने की दूरी, पानी की गहराई और दिल की धड़कन की लय को इकट्ठा करते हैं। आवाज़ के साथ इस डेटा को जोड़ने पर अर्थ और अधिक स्पष्ट हो जाता है। यह भूख की आवाज़ है या डर की आवाज़, इसे पहचाना जा सकता है। AI इस तरह के कई डेटा को समय के क्रम के अनुसार एक साथ समझता है।

Earth Species Project ने 2026 के एक लेख में एक नया नाम पेश किया। वह नाम है 'पशु भाषा प्रसंस्करण' (Animal Language Processing)। इस शब्द का प्रयोग पहली बार 2024 की एक प्रस्तुति में किया गया था। यह तरीका इंसानी भाषा के व्याकरण को जानवरों पर जबरन नहीं थोपता। चाहे आवाज़ हो या हाव-भाव, यह वभिन्न संकेतों को नियमों वाले डेटा के रूप में देखकर उनका विश्लेषण करता है। यह इंसानी भाषा को पैमाना न बनाकर जानवरों की दुनिया को जैसी है वैसी ही समझने का नज़रिया है।

इस शोध में मदद करने वाला एक डेटा संग्रह भी सामने आया। Earth Species Project ने 2026 में alp-data नाम का एक उपकरण जारी किया। इसमें पक्षियों, वृहलों और कीड़ों की आवाज़ों से जुड़ा 35 से अधिक प्रकार का डेटा एक जगह इकट्ठा किया गया। फलिहाल इसमें

ध्वनि डेटा मुख्य है, और हरकत तथा वीडियो डेटा आगे चलकर जोड़ा जाएगा। इसे इस तरह खुला रखा गया है ताकि कोई भी उसी डेटा से शोध कर सके। तभी शोध तेजी से आगे बढ़ सकता है।

संबंधित संदर्भ सामग्री

- [1] Herzing, D. L., Starner, T., et al. (2025). CHAT Jr: A wearable bioacoustics system for two-way communication research between humans and cetaceans. NeurIPS 2025. <https://neurips.cc/virtual/2025/131501>
- [2] Marzluff, J. M., Walls, J., Cornell, H. N., Withey, J. C., & Craig, D. P. (2010). Lasting recognition of threatening people by wild American crows. Animal Behaviour, 79(3), 699-707. <https://doi.org/10.1016/j.anbehav.2009.12.022>
- [3] Earth Species Project. (2026, April 2). Animal language processing: An AI convergence in animal communication. Earth Species Project. <https://earthspecies.org/2026/04/02/animal-language-processing-an-ai-convergence-in-animal-communication/>
- [4] Earth Species Project. (2026, July 22). Introducing alp-data: A shared data layer for animal language processing. Earth Species Project. <https://earthspecies.org/2026/07/22/introducing-alp-data-a-shared-data-layer-for-animal-language-processing/>

8.4 स्पर्म् व्हेल की ध्वनि विवर्णमाला और AI

स्पर्म् व्हेल 'क्लिक' जैसी आवाजों में बात करती है। कई क्लिक ध्वनियों के छोटे समूह को कोडा (coda) कहा जाता है। कोडा स्पर्म् व्हेलों द्वारा आपस में साझा किए जाने वाले छोटे क्लिक ध्वनियों का एक गुच्छा होता है। Project CETI इस कोडा का अध्ययन करता है। CETI, AI की मदद से व्हेल की भाषा को समझने की एक शोध योजना है [1]।

2024 में CETI और अमेरिका के मैसाचुसेट्स इंस्टीट्यूट ऑफ टेक्नोलॉजी (MIT) के शोधकर्ताओं ने एक बड़ी खोज की। उन्होंने AI की मदद से व्हेल की आवाज के 8,700 से अधिक टुकड़ों के डेटा का विश्लेषण किया। तब यह सामने आया कि कोडा नयिमों से युक्त एक ध्वनि प्रणाली है। शोधकर्ताओं ने इन नयिमों को 'स्पर्म् व्हेल ध्वनि विवर्णमाला' कहा [2]।

इस वर्णमाला में चार मुख्य तत्व थे। पहला, क्लिक ध्वनियों के बीच की नियमिति ताल यानी रदिम (rhythm) है। दूसरा, समग्र गति यानी टेम्पो (tempo) है। तीसरा, बातचीत के प्रवाह के अनुसार गति को बढ़ाने या घटाने वाला रुबाटो (rubato) है। चौथा, अंत में एक अतिरिक्त क्लिक जोड़ने वाला अलंकरण है। इन तत्वों को मिलाकर कई प्रकार के कोडा बनते हैं। शोधकर्ताओं ने अक्सर सामने आने वाले 143 से अधिक संयोजनों की पहचान की। यह इंसानी भाषा में व्यंजन और स्वर होने जैसी ही एक संरचना है।

2025 के एक शोध में यह बात एक कदम और आगे बढ़ी। अमेरिका के बर्कले विश्वविद्यालय और CETI के गाशपर बेगुश (Gašper Beguš) के शोध दल ने कोडा का बारीकी से अध्ययन किया। क्लिक ध्वनियों के भीतर इंसानी स्वरों जैसी आवाजें मौजूद थीं। शोधकर्ताओं ने व्हेलों के दो अलग-अलग स्वर खोज निकाले। वे इंसानों की 'आ' और 'ई' की ध्वनि की तरह एक-दूसरे से भिन्न आवाजें थीं। इसका मतलब है कि स्पर्म् व्हेल इंसानों की तरह आवाजों को बहुत बारीकी से अलग करती है [3]।

बातचीत की स्थिति के अनुसार कोडा का रूप बदल जाता था। आगे वाली व्हेल की आवाज के अनुसार पीछे वाली व्हेल ने अपनी आवाज बदल ली। यह वैसा ही था जैसे इंसान एक-दूसरे की बात का जवाब देते हुए बातचीत करते हैं। यह सलिसला विशाल डेटा और कम्प्यूटेशनल विश्लेषण के कारण और भी स्पष्ट रूप से सामने आया।

Project CETI ने 2026 में एक नया समाचार भी साझा किया। शोधकर्ताओं ने स्पर्म् व्हेल के बच्चे के जन्म के दृश्य को वस्तुतः दर्ज किया। उन्होंने पानी के भीतर की 6 घंटे से अधिक की आवाजें रिकॉर्ड कीं और आकाश से लिए गए वीडियो भी इकट्ठा किए। उस डेटा से पता चला कि पूरा समूह मलिकर बच्चे के जन्म में सहायता कर रहा था। इंसानों और बंदरों जैसे प्राइमेट्स से अलग जानवरों में इस प्रकार की सामूहिक देखभाल को आंकड़ों के साथ सिद्ध करने वाला यह पहला मामला था [4]।

व्हेल की भाषा को इंसानी भाषा में बदलना आसान नहीं है। क्योंकि दोनों भाषाओं को एक साथ दर्ज करने वाला कोई शब्दकोश नहीं है। बेगुश का शोध दल एक नया तरीका खोज रहा है। यह बिना किसी उत्तर-पुस्तिका के, अपने आप आवाजों के नयिम खोजकर अर्थ का अनुमान लगाने का तरीका है [5]।

AI के बिना ऐसा शोध करना कठिन था। इंसान के कानों से उन ढेर सारी क्लिक ध्वनियों के नयिम खोजना मुश्किल है। AI अनगिनत ध्वनियों की तेजी से तुलना करके छपि हुए नयिमों को ढूंढ निकालता है।

एक बात हमें नहीं भूलनी चाहिए। coda के नयिमों को खोज लेने का मतलब यह नहीं है कि हम सारा अर्थ समझ गए हैं। ध्वनि की बनावट को जानना और उसके अर्थ को समझना अलग बातें हैं। स्पर्म् व्हेल (sperm whale) क्या कह रही है, यह अभी तक पुष्ट नहीं हुआ है। शोध का दरवाजा अभी-अभी खुला है।

- [1] Project CETI. (2024, May 7). Sperm whale phonetic alphabet proposed for the first time. Project CETI. <https://www.projectceti.org/blog-posts/sperm-whale-phonetic-alphabet-proposed-for-the-first-time>
- [2] Sharma, P., Gero, S., Payne, R., Gruber, D. F., Rus, D., Torralba, A., & Andreas, J. (2024). Contextual and combinatorial structure in sperm whale vocalisations. *Nature Communications*, 15, 3617. <https://doi.org/10.1038/s41467-024-47221-8>
- [3] Beguš, G., Sprouse, R. L., Leban, A., Silva, M., & Gero, S. (2025). Vowel- and diphthong-like spectral patterns in sperm whale codas. *Open Mind*, 9. <https://doi.org/10.1162/OPMI.a.252>
- [4] Project CETI. (2026, March). Studies documenting rare sperm whale birth and ancient cooperative care released. *EurekAlert!* <https://www.eurekalert.org/news-releases/1120823>
- [5] Beguš, G., Leban, A., & Gero, S. (2026). Approaching an unknown communication system by latent space exploration and causal inference. *Royal Society Open Science*, 13(8), 250829. <https://doi.org/10.1098/rsos.250829>

8.5 सेमियोटिक सीमा और अंतर-प्रजात कूटनीति

AI इंसानों के लिए व्यापक संवेदनाओं के नए द्वार खोल देता है। तो फिर समुद्र के भीतर जीवों की आवाज़ें किस चीज़ से इकट्ठा की जाती हैं? इंसान हाइड्रोफोन (hydrophone) का इस्तेमाल करते हैं। हाइड्रोफोन पानी के भीतर की आवाज़ों को पकड़ने वाला माइक होता है। व्हेल या डॉल्फिन के शरीर पर छोटे रिकॉर्डिंग उपकरण लगाए जाते हैं। समुद्र में जगह-जगह आवाज़ पकड़ने वाले बॉय (buoy) सेंसर नेटवर्क तैराए जाते हैं। इस तरह इकट्ठा की गई आवाज़ों को AI पढ़ता है [1]।

भविष्य में आने वाली तकनीक को लेकर भी कुछ विचार सामने आए हैं। यह जल्द आने वाले 6th जनरेशन नेटवर्क (6G) को सेंसर के रूप में इस्तेमाल करने का प्रस्ताव है। 6G आज के नेटवर्क से कहीं अधिक तेज़ वायरलेस संचार नेटवर्क है। विचार यह है कि इस नेटवर्क की रेडियो तरंगें जब किसी वस्तु से टकराती हैं, तो उस हलचल को पढ़कर गति का पता लगाया जा सकता है। हालाँकि, यह अभी केवल प्रस्ताव के स्तर का विचार है। इसके अलावा, रेडियो तरंगें समुद्र के पानी से होकर लगभग नहीं गुज़र पाती हैं। इसलिए इस तरीके का इस्तेमाल ज़मीन के ऊपर और हवा में निगरानी करने के लिए किया जाना है। समुद्र के भीतर की आवाज़ें इकट्ठा करने के लिए यह उपयुक्त नहीं है [2]।

यहाँ एक महत्वपूर्ण सीमा रेखा सामने आती है। सेमियोटिक सीमा (semiotic boundary) वह रेखा है जो यह तय करती है कि संकेतों और उनके अर्थ को किस हद तक समझा जाना चाहिए। मान लीजिए कि AI ने जानवरों के सभी संकेत एकत्र कर लिए हैं। लेकिन क्या इसका मतलब यह है कि उन संकेतों को ज़बरदस्ती इंसानी भाषा के ढाँचे में बठिया जाना चाहिए? उदाहरण के लिए, क्या व्हेल की आवाज़ों को एक-एक करके इंसानी संगीत से जोड़ना ठीक होगा? ऐसा मेल बठाना इंसानी सोच को जानवरों पर थोपने जैसा है [3]।

आइए गलत मेल बठाने का एक उदाहरण देखें। कोई यह कह सकता है कि 10 में से 3 डॉल्फिन संज्ञा (nouns) का इस्तेमाल करती हैं। या कोई व्हेल की क्लिक ध्वनियों को इंसानी संगीत के साथ एक-एक करके मिला सकता है। लेकिन यह इंसानी ढाँचे को ज़बरदस्ती थोपना है। जानवरों की आवाज़ों को जानवरों की दुनिया के भीतर ही समझा जाना चाहिए।

ध्वनियों का अर्थ समझने के काम में नैतिक सवाल भी जुड़े होते हैं। वैज्ञानिकों का कहना है कि इंसानों को बना सोचे-समझे जानवरों की बोली की नकल नहीं करनी चाहिए। अगर इसका गलत इस्तेमाल हुआ, तो इसका उपयोग जानवरों को धोखा देने या उन्हें परेशान करने के लिए किया जा सकता है। इसलिए वे कहते हैं कि हमें पहले नियम बनाने चाहिए [4]।

उक्सकुल (Uexküll) का उमवेल्ट (Umwelt) का विचार आज भी हमारा मार्गदर्शन करता है। जानवर और पौधे इंसानी शब्दकोश के अनुवादित शब्द बनने के लिए नहीं जीते हैं। उनके संकेत सदियों से चली आ रही उनकी अपनी संस्कृति हैं। आधुनिक शोधकर्ता उक्सकुल की सोच को अपनाते हुए कहते हैं कि AI से जानवरों की बना सोचे-समझे नकल नहीं की जानी चाहिए। AI का लक्ष्य बोलने वाले जानवरों के खिलौने बनाना नहीं है।

जानवरों की आवाज़ को रिकॉर्ड करके दोबारा सुनाने का एक प्रयोग होता है। इसे प्लेबैक (playback) कहते हैं। यह रिकॉर्ड की गई आवाज़ को स्पीकर के ज़रिए सुनाकर यह देखने का प्रयोग है कि जानवर कैसी प्रतिक्रिया देते हैं। अगर इस उपकरण का बना सोचे-समझे इस्तेमाल किया जाए, तो यह जानवरों की दुनिया को अस्त-व्यस्त कर सकता है। प्लेबैक की आवाज़ से हैरान जानवर असली और नकली संकेतों में उलझ सकते हैं। इसलिए शोधकर्ता कहते हैं कि हमें सावधान रहना चाहिए।

हमारा रास्ता जीत हासिल करने का नहीं, बल्कि ध्यान से सुनने का है। इसका अर्थ है इंसान का पृथ्वी के एकमात्र व्याख्याकार होने के पद से नीचे उतरना। यह इस बात की रक्षा करना है कि जानवरों की आवाज़ें उनकी ज़मीन और उनके समुद्र में हमेशा गूँजती रहें। शोधकर्ता इसे अंतर-प्रजात कूटनीति (interspecies diplomacy) कहते हैं। अंतर-प्रजात कूटनीति इंसान और अन्य जीवों के एक-दूसरे को नुकसान पहुँचाए बिना साथ रहने का एक वादा है [5]।

ध्यान से सुनना केवल चुपचाप सुनने तक ही सीमिति नहीं है। यह जानवरों के रहने वाले समुद्र और जंगलों की रक्षा करने तक जाता है। आवाज़ों को बचाने के लिए हमें उस जगह की भी रक्षा करनी होगी जहाँ से वे आवाज़ें आती हैं। यही अंतर-प्रजात कूटनीतिका सच्चा अर्थ है।

संबंधित संदर्भ सामग्री

- [1] Bakker, K. (2022). The sounds of life: How digital technology is bringing us closer to the worlds of animals and plants. Princeton University Press. <https://doi.org/10.2307/j.ctv2hnkcc3>
- [2] Ericsson. (2025). Integrated sensing and communication (ISAC) explained. Ericsson. <https://www.ericsson.com/en/6g/isac>
- [3] Zengiaro, N. (2025). AI and animal communication: A generative zoosemiotics perspective. Linguistic Frontiers, 8(2). <https://doi.org/10.2478/lf-2025-0021>
- [4] Alcantara, L., & Andrews, K. (2026). Can we talk to the animals? The ethics of using machine learning to decode animal communication. Topoi. <https://doi.org/10.1007/s11245-026-10409-2>
- [5] Earth Species Project. (2026, April 2). Animal language processing: An AI convergence in animal communication. Earth Species Project. <https://earthspecies.org/2026/04/02/animal-language-processing-an-ai-convergence-in-animal-communication/>

अध्याय 9. धोखे और संरक्षण का दौरा

9.1 नकली आवाज़ों से होने वाला भ्रम

AI अब जानवरों की आवाज़ों की बहुत मलिटि-जुलती नकल कर सकता है। यह पक्षियों की चहचहाहट, व्हेल के गीत और हाथियों की चघिड़ भी बना लेता है। जैसे इंसान बोलना सीखते हैं, वैसे ही AI भी आवाज़ें सुनकर उनके नयिम सीखता है। इस तरह सीखने के बाद, वह अपने आप नई आवाज़ें बना लेता है। इस तकनीक का उपयोग जानवरों की मदद करने के लिए किया जा सकता है। लेकिन अगर इसका गलत इस्तेमाल किया जाए, तो यह प्रकृति में भारी गड़बड़ी पैदा कर सकता है [1]।

जानवरों के अध्ययन में 'प्लेबैक (playback)' नाम का एक तरीका होता है। यह रिकॉर्ड की गई आवाज़ को दोबारा सुनाकर जानवरों की प्रतिक्रिया देखने का प्रयोग है। पहले के समय में, केवल इंसानों द्वारा रिकॉर्ड की गई असली आवाज़ों का ही इस्तेमाल किया जाता था। पुरानी रिकॉर्डिंग में शोर मला होता था, इसलिए कभी-कभी जानवर इसे समझ जाते थे। अब AI बिल्कुल साफ और स्वाभाविक नकली आवाज़ें बनाता है। इसलिए जानवरों का धोखा खाना और भी आसान हो गया है। यह नकली आवाज़ जानवरों के जोड़े बनाने और बच्चों की देखभाल करने में भ्रम पैदा करती है [7]।

कई देशों के वैज्ञानिकों ने इस खतरे के बारे में चेतावनी दी है। प्रोफेसर क्रिश्चियन रुटज़ (Christian Rutz) और उनके साथी इसके प्रमुख उदाहरण हैं। उनका मानना था कि नकली आवाज़ें जंगली जानवरों के व्यवहार को बदल देती हैं। आवाज़ से जानवरों को इस तरह धोखा देने को 'पारस्थितिक छल' कहा जाता है। इसका मतलब है प्रकृति में जानवरों को धोखा देकर उनके व्यवहार को बदलना। साथी की नकली पुकार सुनने वाला पक्षी बेकार ही साथी की तलाश में भटकता रहता है। ऐसा करते हुए वह अपनी ताकत और समय गंवा देता है और बच्चों की देखभाल करने का अवसर भी खो बैठता है।

प्रोफेसर करेन बेकर (Karen Bakker) ने भी लंबे समय तक इस खतरे के बारे में आगाह किया। प्रोफेसर बेकर डिजिटल तकनीक से जानवरों की आवाज़ें सुनने के शोध का नेतृत्व करने वाली विद्वान थीं। उन्होंने 2022 में आई अपनी किताब 'द साउंड्स ऑफ लाइफ (The Sounds of Life)' में इस विचार को दर्ज किया था। प्रोफेसर बेकर का 2023 में निधन हो गया। उनके द्वारा दी गई चेतावनी आज भी कायम है [2]।

इंसान अक्सर जानवरों की आवाज़ों पर इंसानी भावनाएं थोप देते हैं। व्हेल के गीतों को प्रेम गीत मान लेना बहुत आसान होता है। लेकिन उस आवाज़ का असली मतलब अभी तक कोई नहीं जानता। बर्ना सोचे-समझे अंदाज़ा लगाने से गलतियाँ हो जाती हैं। वैज्ञानिक इस आदत से भी सावधान रहने को कहते हैं।

हाथियों पर लंबे समय तक शोध करने वाली डॉ. जॉयस पूल (Joyce Poole) से जुड़ा एक कसिसा है। यह घटना केन्या के अंबोसेली (Amboseli) में हुई थी। शोध दल ने झुंड को एक ऐसी माँ हथिनी की पुकार सुनाई जो पहले ही मर चुकी थी। हाथियों ने उस याद आने वाली आवाज़ को सुना और चारों तरफ दौड़कर उसे ढूँढने लगे। उनमें से उस मृत माँ की बेटी ने बार-बार माँ को पुकारा और जवाब का इंतज़ार किया। लेकिन उस आवाज़ की मालिकनी कहीं नहीं थी। हाथी अपने मृत परजिनों की आवाज़ तक याद रखते हैं। हाथी ऐसी धीमी और गंभीर आवाज़ में बहुत दूर तक बातचीत करते हैं, जो इंसानी कानों को ठीक से सुनाई नहीं देती। इसी वजह से वे दर्जनों साथियों की आवाज़ों को अलग-अलग याद रख पाते हैं। एक ही आवाज़ ने पूरे झुंड के मन को गहराई से झकझोर दिया [3][4]।

व्हेल के मामले में भी सावधानी बरतनी चाहिए। हंपबैक व्हेल (Humpback whale) गीत गाने वाली व्हेल होती है। जब किसी एक समुद्र में कोई नया गीत बनता है, तो वह किसी लोकप्रिय गाने की तरह फैल जाता है। ऑस्ट्रेलिया के तट का गीत दक्षिण प्रशांत महासागर के पार तक भी फैल जाता है। यह गीत कुछ महीनों से लेकर कई सालों में धीरे-धीरे फैलता है। यह धीरे-धीरे लगभग 6,000 किलोमीटर की दूरी पार करता है। यह सयोल से किसी बहुत दूर देश तक जाने जितनी लंबी दूरी है। व्हेल के बच्चे भी बड़ी व्हेल के गीत सुनकर सीखते हैं। यह स्मार्टफोन के ज़रिए लोकप्रिय गानों के फैलने जैसा ही है। इंसानों के अलावा इस तरह एक-दूसरे से गीत सीखना बहुत दुर्लभ है। अगर AI द्वारा बनाए गए नकली व्हेल गीत समुद्र में बजाए जाएं, तो बड़ी समस्या खड़ी हो सकती है। लंबे समय से चली आ रही उनकी गीतों की परंपरा बखिर सकती है। यह जानवरों की भाषा को समझने से पहले ही उनकी संस्कृतिको नष्ट करने जैसा होगा [5]।

इसलिए शोधकर्ता भी यह प्रयोग बहुत सावधानी से करते हैं। वे इसे केवल बेहद ज़रूरी होने पर और थोड़े समय के लिए ही करते हैं। प्रयोग खत्म होने के बाद वे देखते हैं कि जानवर ठीक है या नहीं। एक ही तकनीक इस बात पर निर्भर करती है कि उसका इस्तेमाल कैसे किया जाए—वह मददगार भी हो सकती है और नुकसानदेह भी।

संबंधित संदर्भ सामग्री

[1] Rutz, C., Bronstein, M., Raskin, A., Vernes, S. C., Zacarian, K., & Blasi, D. E. (2023). Using machine learning to decode animal communication. *Science*, 381(6654), 152-155. <https://doi.org/10.1126/science.adg7314>

- [2] Bakker, K. (2022). The sounds of life: How digital technology is bringing us closer to the worlds of animals and plants. Princeton University Press.
- [3] Poole, J. (1996). Coming of age with elephants: A memoir. Hyperion.
- [4] McComb, K., Moss, C., Sayialel, S., & Baker, L. (2000). Unusually extensive networks of vocal recognition in African elephants. *Animal Behaviour*, 59(6), 1103-1109. <https://doi.org/10.1006/anbe.2000.1406>
- [5] Garland, E. C., Goldizen, A. W., Rekdahl, M. L., et al. (2011). Dynamic horizontal cultural transmission of humpback whale song at the ocean basin scale. *Current Biology*, 21(8), 687-691. <https://doi.org/10.1016/j.cub.2011.03.019>
- [7] Küçükuncular, A. (2025). Ethical implications of AI-mediated interspecies communication. *AI and Ethics*, 5, 6379-6391. <https://doi.org/10.1007/s43681-025-00828-z>

9.2 हर जगह मौजूद नगिरानी तंत्र और जानवरों की नजिता

पुराने समय में घने जंगल और दूर-दराज के समुद्र जानवरों के छपने की सुरक्षा जगह थे। इंसान की आँखें और कान वहाँ तक नहीं पहुँच पाते थे। अब ऐसा नहीं है। जंगलों और समुद्रों में जगह-जगह छोटे और स्मार्ट ध्वनि सेंसर लगाए गए हैं। ये उपकरण 24 घंटे आवाज़ें सुनते रहते हैं। AI उन आवाज़ों का तुरंत विश्लेषण करता है। जंगलों और समुद्रों में लगाए जाने वाले ऐसे स्वचालित रिकॉर्डिंग उपकरणों को 'नभिक्रिय ध्वनिकि नगिरानी नेटवर्क' कहा जाता है [6]।

अब जंगली जानवर हर जगह पहरा देने वाली आँखों के सामने आ गए हैं। यह ऐसा है मानो कोई अदृश्य नगिरानी मीनार हर समय चालू हो। Earth Species Project के सह-संस्थापक अज़ा रास्कनि (Aza Raskin) ने इस बदलाव की ओर ध्यान दिलाया। उनका कहना है कि AI की मदद से अब जंगलों और समुद्रों पर बना रुके लगातार नज़र रखी जा सकती है। इसके साथ ही वे चेतावनी देते हैं कि यह शक्ति जानवरों को जकड़ने वाली नगिरानी न बन जाए, इसके लिए नियम बनाए जाने चाहिए [8]।

इन उपकरणों का इस्तेमाल अच्छे कामों के लिए भी किया जाता है। कई बार व्हेल बड़े जहाजों से टकराकर अपनी जान गंवा बैठती हैं। ध्वनि सेंसरों की मदद से व्हेल की मौजूदगी का पता लगाकर जहाजों का रास्ता बदल दिया जाता है। इससे बड़े जहाज व्हेल से टकराने से बच जाते हैं। ये सेंसर जंगलों में चोरी-छिपे पेड़ काटने वाली आरी की आवाज़ भी पकड़ लेते हैं। इस तरह यह लुप्तप्राय जानवरों की रक्षा करता है और अवैध शिकार को रोकता है।

लेकिन बुरे लोगों के हाथों में पड़ने पर यह एक हथियार बन सकता है। यह ठीक वैसा ही है जैसे चाकू से खाना भी बनाया जा सकता है और चोट भी पहुँचाई जा सकती है। साथी की पुकार या भोजन की आवाज़ की नकली नकल करके जानवरों को बहला-फुसलाकर बुलाया जा सकता है। हालाँकि यह बात अभी बहुत बड़े पैमाने पर नहीं फैली है, फरि भी वैज्ञानिक इस खतरे को लेकर चिंतित हैं। जानवरों की मदद के लिए बनाई गई तकनीक का इस्तेमाल जानवरों को पकड़ने के लिए किया जा सकता है [7]।

जानवरों की नजिता (privacy) भी एक अहम मुद्दा है। नजिता वह अपना दायरा है जिसमें कोई दूसरों को देखाना नहीं चाहता। प्रोफेसर करेन बेकर ने कहा था कि व्हेल और डॉल्फिन को अधिकारों वाले प्राणियों के रूप में देखा जाना चाहिए। माँ और बच्चे के बीच होने वाली धीमी और शांत बातचीत को भी इंसान कभी-कभी चुपके से सुन लेते हैं। वे उन आवाज़ों को मनमाने ढंग से इकट्ठा करके किसी के भी इस्तेमाल के लिए सार्वजनिक कर देते हैं। वैज्ञानिकों का मानना है कि इसे नजिता का उल्लंघन माना जाना चाहिए। यह वैसा ही है जैसे कोई आपकी डायरी चुपके से पढ़कर इंटरनेट पर डाल दे [2]।

जंगलों और समुद्रों की लंबे समय से रक्षा करने वाले मूल नविसियों के अधिकार भी हैं। कभी-कभी बड़ी कंपनियाँ उनकी ज़मीन से रिकॉर्ड की गई आवाज़ों को बना बताए ले जाती हैं। वे उन आवाज़ों का उपयोग AI को सिखाने के लिए करती हैं। ध्वनि डेटा भी उस ज़मीन की अपनी संपत्ति है। मालिक की अनुमति के बिना इसे नहीं लिया जाना चाहिए। उस क्षेत्र से मलिन डेटा के बारे में स्थानीय लोगों द्वारा खुद फैसला करने के अधिकार को 'मूल नविासी डेटा संप्रभुता' कहा जाता है।

इंटरनेट पर ऐसी कहानियाँ भी चलती हैं कि जानवर नगिरानी का वरिध कर रहे हैं। ऐसी बात कही जाती है कि स्प्रूम व्हेल के एक पूरे झुंड ने एक साथ आवाज़ निकालना बंद कर दिया। ऐसी भी कहानी है कि कौवों ने नई आवाज़ें निकालकर AI को भ्रमति कर दिया। लेकिन ऐसी बातें अभी तक वैज्ञानिकों द्वारा प्रमाणित नहीं हुई हैं। ये केवल इंटरनेट वीडियो की कहानियाँ हैं, इसलिए इन्हें सच नहीं मान लेना चाहिए।

अच्छे और बुरे दोनों तरह के उपयोग एक ही तकनीक से निकलते हैं। इसलिए यह बात सबसे महत्वपूर्ण है कि कौन किस भावना से इसका इस्तेमाल करता है। जानवरों की भी अपनी एक नज़ी जगह होती है, जिसकी रक्षा की जानी चाहिए।

संबंधित संदर्भ सामग्री

- [2] Bakker, K. (2022). The sounds of life: How digital technology is bringing us closer to the worlds of animals and plants. Princeton University Press.

- [6] Earth Species Project. (2024). Introducing NatureLM-audio: An audio-language foundation model for bioacoustics. <https://www.earthspecies.org/blog/introducing-naturelm-audio-an-audio-language-foundation-model-for-bioacoustics>
- [7] Küçükuncular, A. (2025). Ethical implications of AI-mediated interspecies communication. AI and Ethics, 5, 6379-6391. <https://doi.org/10.1007/s43681-025-00828-z>
- [8] Faunalytics (Hanes, A.). (2025, October 30). The ethics of using AI to talk to animals. Faunalytics. <https://faunalytics.org/the-ethics-of-using-ai-to-talk-to-animals/>

9.3 ध्वनि में छपाई इंसानी नशान: ऑडियो वॉटरमार्किंग

नकली आवाज़ों की पहचान करने का एक तरीका होना जरूरी है। वैज्ञानिकों और इंजीनियरों ने इसका एक उपाय निकाला है। वह उपाय है AI द्वारा बनाई गई आवाज़ में एक 'नशान' छपाकर रखना। इस तरीके को 'ऑडियो वॉटरमार्किंग (audio watermarking)' कहा जाता है [9]।

वॉटरमार्किंग ऑडियो फ़ाइल के अंदर एक अदृश्य नशान दर्ज करने की तकनीक है। यह नशान ध्वनि फ़ाइल में छपा हुआ एक कंप्यूटर सग्नल होता है। इंसान इस सग्नल को आसानी से नहीं पकड़ पाते। लेकिन जाँच करने वाला प्रोग्राम उस सग्नल को तुरंत पढ़ लेता है। Google द्वारा बनाया गया SynthID ऐसे ही एक नशान का उदाहरण है। इस नशान को आवाज़ के कई हिस्सों में फैलाकर डाला जाता है। इसलिए इसे काटना या मटिना बहुत कठिन होता है। आवाज़ को धीमा करने या उसकी गति बदलने पर भी यह नशान बना रहता है।

यह नशान आवाज़ के स्रोत का पता देता है। मशीन यह जाँच सकती है कि कौन सी आवाज़ इंसान द्वारा बनाई गई नकली आवाज़ है। इससे असली जानवरों की आवाज़ों और नकली आवाज़ों को अलग-अलग करके संभाला जा सकता है। असली और नकली में अंतर करना प्रकृति की रक्षा की दृष्टि में पहला कदम है। इस नशान के बिना यह जानना बहुत मुश्किल है कि क्या सच है और क्या नकली [1]।

इस नशान का उपयोग शोध डेटा की सुरक्षा के लिए भी किया जा सकता है। जंगलों और समुद्रों में आवाज़ें इकट्ठा करने वाले हज़ारों उपकरण लगे हैं। यदि उनमें नकली आवाज़ें मलि जाएँ, तो शोध के परिणाम पूरी तरह गड़बड़ा जाएँगे। असली पक्षियों की आवाज़ और मशीनों की बनाई आवाज़ को एक साथ गनने पर उत्तर गलत हो जाएगा। नशान होने पर नकली आवाज़ों को छंटकर अलग किया जा सकता है। तब शोध डेटा पर पूरा भरोसा करके उसका उपयोग किया जा सकता है [7]।

इसका उपयोग अवैध शिकार रोकने के लिए करने का विचार भी सामने आया है। यह शोध में इस्तेमाल होने वाली आवाज़ों में नशान डालकर यह जाँचने का तरीका है कि आवाज़ कहाँ से आई है। इससे प्रमाणित शोध आवाज़ों और अज्ञात स्रोतों वाली नकली आवाज़ों के बीच अंतर किया जा सकता है। लेकिन इस तरीके की अपनी सीमाएँ हैं। यदि कोई बुरा व्यक्ति बिना नशान वाली आवाज़ बना ले, तो उसे छंटा नहीं जा सकता। अगर कोई नशान मटिाकर आवाज़ इस्तेमाल करे, तो उसे ढूँढना कठिन होता है। इसलिए सिर्फ एक तकनीक से शिकार को पूरी तरह नहीं रोका जा सकता [8]।

यह तरीका अभी केवल सुझाव के स्तर पर है। यह एक अच्छा विचार है, लेकिन अपने आप पूरा नहीं हो सकता। जब सभी देश और कंपनियाँ नशान लगाने पर सहमत होगी, तभी यह कारगर होगा। यदि किसी एक जगह से भी बिना नशान वाली नकली आवाज़ आ जाती है, तो व्यवस्था में कमी रह जाएगी। यह जाल में एक भी छेद होने पर मछलियों के नकिल भागने जैसा है। इसलिए सिर्फ तकनीक ही काफी नहीं है। मलिकर पालन किए जाने वाले नयिमों की भी आवश्यकता है।

इस तकनीक का उपयोग इंसानों द्वारा बनाई गई अन्य आवाज़ों के लिए भी किया जा सकता है। यानी AI द्वारा बनाई गई आवाज़ या गानों में नशान डालना। इससे नकली आवाज़ों के धोखे में आने से बचा जा सकता है। जानवरों की रक्षा के लिए बनाया गया तरीका इंसानों की भी रक्षा करता है। आजकल इंटरनेट पर AI द्वारा बनाए गए गीतों और आवाज़ों की भरमार है। इसलिए ऐसे नशान आगे चलकर और भी उपयोगी साबित होंगे।

संबंधित संदर्भ सामग्री

- [1] Rutz, C., Bronstein, M., Raskin, A., Vernes, S. C., Zaccarian, K., & Blasi, D. E. (2023). Using machine learning to decode animal communication. Science, 381(6654), 152-155. <https://doi.org/10.1126/science.adg7314>
- [7] Küçükuncular, A. (2025). Ethical implications of AI-mediated interspecies communication. AI and Ethics, 5, 6379-6391. <https://doi.org/10.1007/s43681-025-00828-z>
- [8] Faunalytics (Hanes, A.). (2025, October 30). The ethics of using AI to talk to animals. Faunalytics. <https://faunalytics.org/the-ethics-of-using-ai-to-talk-to-animals/>
- [9] Google DeepMind. (2024). SynthID: Identifying AI-generated content. <https://deepmind.google/models/synthid/>

9.4 इंसानों और जानवरों के बीच नयिम बनाना

AI की मदद से जानवरों की आवाजों पर काम करने के लिए नयिमों की ज़रूरत है। क्योंकि ताकतवर लोगों को अपनी मर्ज़ी से प्रकृति को असंत-व्यस्त करने की छूट नहीं दी जा सकती। आवाज़ सीमाओं के पार दूर तक फैलती है। इसलिए केवल एक देश का नयिमों पर चलना काफी नहीं है। वैज्ञानिकों का कहना है कि कई देशों को मलिकर ऐसे नयिम बनाने चाहिए जिनका सभी पालन करें। इसका मतलब है इंसानों और जानवरों के बीच नभिया जाने वाला एक वादा [10]।

अमेरिका के न्यूयॉर्क विश्वविद्यालय (New York University) लॉ स्कूल में 'इंसान से परे जीवन (More Than Human Life)' नाम का एक शोध समूह है। इसे संक्षेप में MOTH कहा जाता है। यह समूह 2022 में शुरू हुआ था। इसमें कानून के वदिवान और वैज्ञानिक एक साथ जुड़े। इस समूह का कहना है कि जानवरों को केवल डेटा नहीं, बल्कि अपनी-अपनी ज़िंदगी जीने वाले प्राणियों के रूप में देखा जाना चाहिए। उन्होंने चार वादों वाले नयिम पेश किए हैं। वे वादे हैं: पहले से तैयारी करना, मलिकर सलाह-मशवरा करना, नुकसान को रोकना और सुरक्षा करना।

पहला, पहले से जाँच-परख करने का नयिम। कुछ ऐसे प्रयोग होते हैं जिनमें AI और जंगली जानवर आपस में आवाजों का आदान-प्रदान करते हैं। ऐसे प्रयोग करने से पहले जोखिमों की पूरी जाँच की जाती है। यह देखा जाता है कि कहीं इससे उनके जोड़े बनाने और बच्चों की देखभाल करने की संस्कृति को कोई नुकसान तो नहीं पहुँच रहा। उसके बाद ही प्रयोग की अनुमति दी जाती है। यह किसी निर्माण कार्य को शुरू करने से पहले सुरक्षा जाँच कराने जैसा है।

दूसरा, अंधाधुंध इस्तेमाल पर रोक लगाने का नयिम। जानवरों की आवाजों की नकल करने वाली तकनीक को कोई भी यूँ ही इस्तेमाल न कर सके। शिकार में इसके इस्तेमाल पर तो पूरी तरह प्रतिबंध लगाने की बात कही गई है। केवल कड़े परीक्षणों को पास करने वाले संस्थानों को ही इसके सावधानीपूर्वक इस्तेमाल की अनुमति दी जानी चाहिए। यह खतरनाक औज़ारों को यूँ ही कहीं भी न छोड़ने जैसा है। वे इस मुख्य तकनीक का इस्तेमाल केवल पैसे कमाने के लिए करने पर भी रोक लगाने की वकालत करते हैं।

तीसरा, बढ़ा-चढ़ाकर की जाने वाली खबरों को रोकने का नयिम। "AI ने जानवरों से खुलकर बातचीत करने वाला अनुवादक बना लिया है," जैसी अतिशयोक्ति भरी खबरें देखने को मिलती हैं। वदिवानों का मानना है कि ऐसी खबरें नुकसानदेह होती हैं। क्योंकि इससे लोग जानवरों को किसी खलौने जैसा समझने लगते हैं। इसलिए वे कहते हैं कि तकनीक अभी क्या नहीं कर सकती, उसे पूरी ईमानदारी से बताया जाना चाहिए [11]।

नयिमों का मकसद शोध को रोकना नहीं है। ये तो शोध को सही तरीके से आगे बढ़ाने में मदद करने वाला सुरक्षा घेरा है। यह वैसा ही है जैसे खेल के मैदान में लकीरें खींची होने पर बच्चे बेफकिर होकर खेल-कूद सकते हैं। नयिम बनाने के लिए कई लोगों के वचिारों की ज़रूरत होती है। केवल वैज्ञानिक ही नहीं, बल्कि कानून के जानकार, जंगलों और समुद्रों में रहने वाले मूल निवासी और जानवरों से प्यार करने वाले लोग मलिकर इन्हें तय करते हैं। कई नज़रों से देखने पर कमियाँ छूटने की गुंजाइश कम हो जाती है। इन नयिमों का लक्ष्य केवल एक ही है: जानवरों पर अपना हुक्म चलाना नहीं, बल्कि उनका सम्मान करना [7][8]।

संबंधित संदर्भ सामग्री

[7] Küçükuncular, A. (2025). Ethical implications of AI-mediated interspecies communication. *AI and Ethics*, 5, 6379-6391. <https://doi.org/10.1007/s43681-025-00828-z>

[8] Faunalytics (Hanes, A.). (2025, October 30). The ethics of using AI to talk to animals. Faunalytics. <https://faunalytics.org/the-ethics-of-using-ai-to-talk-to-animals/>

[10] Rodríguez-Garavito, C. (2025). Listening to the more-than-human world: Legal and ethical principles for nonhuman animal communication technologies. More-Than-Human Life (MOTH) Program, NYU School of Law. <https://www.law.nyu.edu/news/ideas/cesar-rodriguez-garavito-animal-communication-ai>

[11] Kluskens, C. (2025). The responsible media representation of artificial intelligence for non-human animal communication. NeurIPS 2025 Workshop on AI for Non-Human Animal Communication.

9.5 ताज़ा खबरें और हमारे सामने खड़ा दोराहा

वर्ष 2026 में भी जानवरों की भाषा पर शोध बहुत तेज़ी से आगे बढ़ा। Earth Species Project ने इस अध्ययन को 'पशु भाषा प्रसंस्करण (Animal Language Processing)' के रूप में परिभाषित किया। जैसे कंप्यूटर इंसानों की बातें समझता है, वैसे ही यह जानवरों की बहुत सारी आवाज़ें इकट्ठा करके उनका विश्लेषण करने का तरीका है। यह एक-एक प्रजाति को अलग से देखने के पुराने तरीके से आगे बढ़ता है। इसमें कई जानवरों की आवाजों को एक साथ देखा जाता है। उन्होंने यह नाम पहली बार नहीं बनाया था, बल्कि इधर-उधर बखिरे शोधों को एक साथ जोड़ा था। यह घोषणा अप्रैल 2026 में सामने आई [12]।

उसी साल जुलाई में एक नया डेटा संग्रह केंद्र भी खोला गया। इसका नाम 'alp-data' है। यह एक ऐसी जगह है जहाँ कई शोधकर्ताओं के एक साथ उपयोग के लिए जानवरों की आवाज़ों का डेटा जमा किया गया है। इसमें 35 से अधिक डेटासेट शामिल हैं। डेटा साझा करने से शोध और तेज़ हो जाता है। यह वैसा ही है जैसे कक्षा के सहपाठी आपस में सामान बाँटकर काम करें, तो काम जल्दी पूरा हो जाता है [13]।

इस शोध के आधार में 'NatureLM-audio' है। यह जानवरों की आवाज़ों के लिए बनाया गया एक बड़ा AI है। यह AI केवल आवाज़ सुनकर पहचान लेता है कि वह कौन-सा जानवर है। यह आवाज़ में मौजूद जानवरों की संख्या भी गनि लेता है। यह यह भी पता लगा लेता है कि कोई पक्षी बच्चा है या पूरी तरह बड़ा हो चुका है। इसने उन जानवरों की आवाज़ों को भी पहचानने की क्षमता दिखाई है जिन्हें इसने पहले कभी नहीं सीखा था। ऐसा AI उस काम को कुछ ही मिनटों में कर देता है जिसमें इंसानों को कई दिन लग जाते हैं। इसलिए शोधकर्ता आवाज़ों का अध्ययन बहुत तेज़ी से कर सकते हैं [6]।

व्हेल पर शोध भी आगे बढ़ा। Project CETI ने स्पर्म् व्हेल की आवाज़ के टुकड़ों, यानी coda को इकट्ठा किया। coda "टकि, टकि, टकि" जैसी लगातार आने वाली क्लिक आवाज़ों का एक समूह है। शोध दल ने 8,700 से अधिक coda का विश्लेषण किया। इसके परिणामस्वरूप उन्होंने आवाज़ बनाने वाले बुनियादी टुकड़े खोज निकाले। शोधकर्ताओं ने इन बुनियादी टुकड़ों को चार प्रकारों में बाँटा। शोधकर्ताओं ने इसे 'स्पर्म् व्हेल के ध्वनिअक्षर' कहा। यह इंसानी भाषा में व्यंजनों और स्वरों को अलग करने जैसा है। व्हेल आवाज़ की गति और ताल को बदलकर अर्थ अलग करती है। वे अतिरिक्त क्लिक जोड़कर भी अर्थ बदलती हैं। वदिवानों ने इंसानी स्वरों की तरह आवाज़ों को जोड़ने के नियम भी खोजे। इसका यह अर्थ नहीं है कि व्हेल इंसानों की तरह 'आ' या 'ई' जैसी आवाज़ें निकालती है। जैसे इंसान व्यंजनों और स्वरों को जोड़कर कई शब्द बनाते हैं, वैसे ही व्हेल भी आवाज़ों को नियमों के अनुसार जोड़ती है, यह केवल इसकी एक उपमा है [14][15][16]।

AI का उपयोग वलुप्त हो रहे जानवरों की रक्षा के लिए भी किया जाता है। हवाईयन कौआ (Hawaiian crow) इसका एक उदाहरण है। यह पक्षी लंबे समय तक कैद में रहने के बाद जंगल में वापस लौटा है। शोधकर्ता यह जांच रहे हैं कि क्या इस बीच वह अपने झुंड की कुछ आवाज़ें भूल तो नहीं गया। वे AI से पुरानी आवाज़ों और आज की आवाज़ों की तुलना करते हैं। इसका उद्देश्य भूली हुई आवाज़ों को वापस लौटाना है। Earth Species Project कौआ, बेलुगा व्हेलों और हाथियों की आवाज़ों का भी एक साथ अध्ययन करता है। कई जानवरों का एक साथ अध्ययन करने पर आवाज़ों के सामान्य नियम दिखाई देते हैं।

हम अभी एक दोराहे पर खड़े हैं। एक रास्ता धोखे का है। यह नकली आवाज़ों से जानवरों को धोखा देने और उन्हें परेशान करने का रास्ता है। दूसरा रास्ता संरक्षण का है। यह चुपचाप ध्यान से सुनकर जानवरों की रक्षा करने का रास्ता है। ये दोनों रास्ते एक ही तकनीक पर बने हैं। कसि रास्ते पर चलना है, यह इंसान के चुनाव पर निर्भर करता है। AI केवल एक उपकरण है। आवाज़ सुनने की ताकत जितनी बढ़ती है, उस ताकत की रक्षा करने वाले नियम भी उतने ही बढ़े होने चाहिए। ताकत जितनी बढ़ती है, उसे संभालने का तरीका भी साथ-साथ सीखना चाहिए। जब हम अपना घमंड छोड़कर कान खोलेंगे, तब इंसान और जानवर साथ-साथ मलिकर रह सकते हैं [10][11]।

संबंधित संदर्भ सामग्री

- [1] Rutz, C., Bronstein, M., Raskin, A., Vernes, S. C., Zacarian, K., & Blasi, D. E. (2023). Using machine learning to decode animal communication. *Science*, 381(6654), 152-155. <https://doi.org/10.1126/science.adg7314>
- [6] Earth Species Project. (2024). Introducing NatureLM-audio: An audio-language foundation model for bioacoustics. <https://www.earthspecies.org/blog/introducing-naturelm-audio-an-audio-language-foundation-model-for-bioacoustics>
- [10] Rodríguez-Garavito, C. (2025). Listening to the more-than-human world: Legal and ethical principles for nonhuman animal communication technologies. *More-Than-Human Life (MOTH) Program*, NYU School of Law. <https://www.law.nyu.edu/news/ideas/cesar-rodriquez-garavito-animal-communication-ai>
- [12] Earth Species Project. (2026, April 2). Animal language processing: An AI convergence in animal communication. <https://earthspecies.org/2026/04/02/animal-language-processing-an-ai-convergence-in-animal-communication/>
- [13] Earth Species Project. (2026, July 22). Introducing alp-data: A shared data layer for animal language processing. <https://earthspecies.org/2026/07/22/introducing-alp-data-a-shared-data-layer-for-animal-language-processing/>
- [14] Project CETI. (2024). Sperm whale phonetic alphabet proposed for the first time. <https://www.projectceti.org/blog-posts/sperm-whale-phonetic-alphabet-proposed-for-the-first-time>
- [15] Sharma, P., Gero, S., Payne, R., et al. (2024). Contextual and combinatorial structure in sperm whale vocalisations. *Nature Communications*, 15, 3617. <https://doi.org/10.1038/s41467-024-47221-8>
- [16] National Geographic. (2024). We're one step closer to understanding the sperm whale 'alphabet'. <https://www.nationalgeographic.com/animals/article/sperm-whale-vowel-patterns-communication-discovery>

अध्याय 10. गैर-मानव व्यक्तित्व और मशीन-सहायता प्राप्त अधिकार

10.1 मौन से गवाही तक

इंसानों के कानून ने लंबे समय तक केवल इंसानों को ही अधिकार दिए। बाद में, कंपनी की तरह लोगो द्वारा मलिकर बनाए गए संगठनों को भी अधिकार दिए गए। ऐसे संगठनों को कानूनी व्यक्ति (legal person) कहा जाता है। कानूनी व्यक्ति कानून द्वारा बनाया गया एक काल्पनिक इंसान होता है। कंपनियों या स्कूल इसमें शामिल हैं। कानूनी व्यक्ति असली इंसान नहीं होते, लेकिन कानून के सामने उनके साथ इंसान जैसा व्यवहार किया जाता है [3]।

पुराने कानून का एक ही मानना था। वह यह विश्वास करता था कि केवल वही पक्ष अधिकार रख सकता है जो अपने कर्तव्यों को जानता है और वादे या समझौते कर सकता है। जानवर बोलकर अपना बचाव नहीं कर सकते। इसलिए कानून ने जानवरों को अधिकारों का हकदार नहीं माना। उन्हें इंसानों द्वारा खरीदी और बेची जा सकने वाली वस्तु और संपत्ति समझा गया।

यह पुरानी दीवार अब हल रही है। ऐसा इसलिए है क्योंकि AI और जैव-ध्वनिकी (bioacoustics) ने हाथ मिला लिया है। जैव-ध्वनिकी वह विज्ञान है जो जानवरों द्वारा निकाली जाने वाली आवाजों का अध्ययन करता है। इन दोनों तकनीकों ने मलिकर जानवरों के मन की बात को संख्याओं के जरिए दिखाना शुरू कर दिया है [2]।

पुराने समय में इंसान केवल अपने कानों से जानवरों की आवाजें सुनते थे। इसलिए बहुत-सी आवाजें छूट जाती थीं। अब समुद्र के भीतर बड़े पैमाने पर साउंड रिकॉर्डर लगाए जाते हैं। Project CETI (Cetacean Translation Initiative) यही काम करता है। इसका मतलब व्हेल भाषा अनुवाद पहल है। इस परियोजना ने लंबे समय से स्पर्म व्हेल (sperm whale) की आवाजों को इकट्ठा किया है। इस तरह जमा की गई आवाजों की AI एक-एक करके जांच करता है। इंसान जिन नियमों को नहीं पकड़ पाए, उन्हें भी कंप्यूटर ढूंढ निकालता है [4]।

स्पर्म व्हेल ध्वनिकी के टुकड़ों, यानी coda का आदान-प्रदान करती है। Coda छोटी क्लिक (click) आवाजों को कई बार जोड़कर बनाई गई एक ध्वनि है। उसमें अपने समूह की पहचान बताने वाला पहचान coda (identity coda) भी होता है। यह एक ही परिवार के सदस्यों द्वारा साझा किए जाने वाले नशान जैसी आवाज है। स्पर्म व्हेल इस आवाज से अपने समूह को पहचानती है। इसे इंसानों के व्यक्तिगत नाम के बिल्कुल समान मानना मुश्किल है। यह पूरी कक्षा द्वारा मलिकर बोले जाने वाले नारे के अधिक निकट है।

स्पर्म व्हेल की आवाज में नियम छपे होते हैं। वे कुछ ध्वनिकाइयों को आपस में जोड़कर कई रूप बनाती हैं। इंसानी भाषा भी इसी तरह काम करती है। हम कुछ व्यंजनों और स्वरों को मिलाकर अनगिनत शब्द बनाते हैं। हालांकि, प्रत्येक coda का सटीक अर्थ क्या है, इस पर अभी शोध चल रहा है। ध्वनिकी संरचना तो सामने आ चुकी है, लेकिन उसका अर्थ अभी पूरी तरह नहीं सुलझा है।

इंसानी दमाग का अध्ययन करने वाले वैज्ञानिक भाषा की नए सरि से व्याख्या करते हैं। यह मैसाचुसेट्स इंस्टीट्यूट ऑफ टेक्नोलॉजी (MIT) की प्रोफेसर एवेलिना फेडोरेनको (Evelina Fedorenko) की शोध टीम है। इन्होंने 2024 में वैज्ञानिक पत्रिका नेचर (Nature) में अपना शोध प्रकाशित किया। वे मानते हैं कि भाषा गहरे विचारों के निर्माण के लिए कोई अनिवार्य शर्त नहीं है। इसका मतलब है कि भाषा मन की बात दूसरों तक पहुंचाने का केवल एक साधन है। यानी सोचना और बोलना दो अलग-अलग काम हैं। बोल न पाने का अर्थ यह नहीं है कि कोई विचार ही नहीं है [1]।

यह विचार कानून की पुरानी जड़ को हिला देता है। जानवर इंसानों जैसी भाषा नहीं बोल सकते, केवल इसलिए उनकी बुद्धिमत्ता को अनदेखा नहीं किया जा सकता। जानवर भी अपने तरीके से परिष्कृत आवाजों का आदान-प्रदान कर रहे थे। AI ध्वनिकी उन नियमों को खोजने में मदद करता है जिनसे इंसान नहीं देख पाए थे।

यहाँ हमें एक बात को अलग करके देखना होगा। आवाजों में नियम होने की बात जानना और कानून बदलना दो अलग बातें हैं। भले ही विज्ञान नए तथ्य उजागर कर दे, कानून तुरंत नहीं बदलता। समाज की सहमति, लंबी चर्चा और नए न्यायिक फैसलों के बाद ही कानून आगे बढ़ता है। फिर भी, अगर यह साबित हो जाए कि जानवरों की आवाजें अर्थपूर्ण संचार हैं, तो कानून भी बदलेगा। तब कानून के लिए जानवरों को केवल वस्तु मानकर रखना मुश्किल हो जाएगा।

संबंधित संदर्भ सामग्री

[1] Fedorenko, E., Piantadosi, S. T., & Gibson, E. A. F. (2024). Language is primarily a tool for communication rather than thought. *Nature*, 630, 575-586. <https://doi.org/10.1038/s41586-024-07522-w>

[2] Sharma, P., Gero, S., Payne, R., Gruber, D. F., Rus, D., Torralba, A., & Andreas, J. (2024). Contextual and combinatorial structure in sperm whale vocalisations. *Nature Communications*, 15, 3617. <https://doi.org/10.1038/s41467-024-47221-8>

- [3] Rodríguez-Garavito, C., Gruber, D. F., Nemeth, A. O., & Beguš, G. (2025). What if we understood what animals are saying? The legal impact of AI-assisted studies of animal communication. *Ecology Law Quarterly*, 52(1). <https://doi.org/10.15779/Z383X83N5Q>
- [4] Project CETI. (2024, May 7). Sperm whale phonetic alphabet proposed for the first time. Project CETI. <https://www.projectceti.org/blog-posts/sperm-whale-phonetic-alphabet-proposed-for-the-first-time>

10.2 NYU MOTH कार्यक्रम और मशीन-सहायता प्राप्त अधिकार

इस बदलाव को कानून में लागू करने का एक मुख्य केंद्र है। यह न्यूयॉर्क यूनिवर्सिटी (NYU) स्कूल ऑफ लॉ का 'मानव से परे जीवन कार्यक्रम' है। अंग्रेजी में इसे MOTH (More Than Human Life) कहा जाता है। इस कार्यक्रम का नेतृत्व पर्यावरण कानूनविद् प्रोफेसर सेसर रोड्रिगुज़-गराविति (César Rodríguez-Garavito) कर रहे हैं [1]।

प्रोफेसर गराविति कई विद्वानों के साथ मलिकर शोध करते हैं। कैलिफोर्निया विश्वविद्यालय, बर्कले (UC Berkeley) के डॉ. गास्पेर बेगुश (Gašper Beguš) भी उनके साथ हैं। उन्होंने एक नया विचार पेश किया है। वह यह है कि AI द्वारा सुलझाई गई जानवरों की आवाज़ को अदालत में सबूत के तौर पर इस्तेमाल किया जाए। इसे मशीन-सहायता प्राप्त अधिकार (Machine-Assisted Rights) कहा जाता है। यह अभी व्यापक रूप से इस्तेमाल होने वाला शब्द नहीं है, बल्कि उनका एक नया प्रस्ताव है।

मशीन-सहायता प्राप्त अधिकारों को आसान शब्दों में समझाएं तो यह इस प्रकार है। AI जानवरों के संचार और उनके पर्यावरण को समझता है। इंसानों की वजह से जानवरों को होने वाले नुकसान को संख्याओं में मापता है। फिर उन आंकड़ों को अदालत में सबूत के तौर पर पेश किया जाता है। एक तरह से यह बोल न सकने वाले जानवरों का अनुवादक और प्रतिनिधि बनने जैसा है। यह वैसा ही है जैसे जब बच्चे खुद दस्तावेज़ जमा नहीं कर पाते, तो बड़े उनकी जगह मदद करते हैं।

यह तरीका कानून की एक पुरानी शर्त से आगे निकलने का प्रयास करता है। पुराने कानून ने मुकदमा दायर करने का अधिकार केवल इंसानों को दिया था। मुकदमा दायर करने की इस योग्यता को कानूनी पात्रता (वादी का अधिकार) कहा जाता है। जानवरों को यह पात्रता नहीं मिली थी। इसलिए विद्वान AI द्वारा मापे गए नुकसान को सबूत बनाना चाहते हैं। इस तरह वे जानवरों के लिए अदालत का दरवाज़ा खटखटाना चाहते हैं।

प्रोफेसर बेगुश की शोध टीम ने एक बड़ी खोज की। उन्होंने स्पर्म् व्हेल की आवाज़ में फॉर्मैट (formant) संरचना पाई। फॉर्मैट आवाज़ का वह हिस्सा है जहाँ ऊर्जा केंद्रित होकर ध्वनिविशेष रूप से स्पष्ट गूंजती है। यह वैसा ही है जैसे इंसान मुंह का आकार बदलकर स्वर निकालते हैं। शोध टीम ने 2025 की एक घोषणा में बताया कि स्पर्म् व्हेल के coda में 'a' coda और 'i' coda मौजूद होते हैं। इंसानी स्वरों की तरह ये एक-दूसरे से अलग आवाज़ें हैं [2]।

लेकिन AI द्वारा बनाए गए ध्वनि भिन्नता ने एक दुखद सच्चाई भी उजागर की। बड़े जहाजों के प्रोपेलर की आवाज़ें और संसाधन खोजने वाले धमाके समुद्र को भर रहे हैं। ऐसी आवाज़ों को मानव-निरमिति समुद्री शोर कहा जाता है। 2021 में वैज्ञानिक पत्रिका साइंस (Science) में प्रकाशित एक शोध में इस पर बात की गई थी। इसके अनुसार, उद्योग के विकास के बाद से समुद्र में बहुत शोर बढ़ गया है। ऐसे अध्ययन भी हैं जो बताते हैं कि जहाजों का शोर स्पर्म् व्हेल की coda संरचना को प्रभावित कर सकता है [3]।

MOTH कार्यक्रम का दावा है कि इस नुकसान को अदालत में सबूत के रूप में इस्तेमाल किया जा सकता है। उनका मानना है कि आवाज़ के अर्थ का पूरी तरह अनुवाद करना ज़रूरी नहीं है। बस आंकड़ों में यह दिखाना काफी है कि स्पर्म् व्हेल के समूह का संचार कितना नष्ट हुआ है। वक़्त आवाज़ का ग्राफ ही सबूत बन सकता है। इससे अदालत जहाजों का मार्ग बदलने का आदेश दे सकती है या सैन्य सोनार उपकरणों को रोक सकती है। यह दावा स्पर्म् व्हेल के दो अधिकारों की बात करता है: पीड़ा से मुक्ति का अधिकार और अपनी संस्कृति की रक्षा का अधिकार [4]।

यह विचार पहले से मौजूद कानूनों को भी मज़बूत बनाता है। अमेरिका में लुप्तप्राय जानवरों की रक्षा के लिए कानून है। समुद्री स्तनधारियों की सुरक्षा के लिए भी एक अलग कानून मौजूद है। AI द्वारा मापे गए नुकसान के आंकड़े इन कानूनों को वास्तव में लागू करने की ताकत देते हैं। इसका मतलब है कि नए अधिकार स्थापित करने से पहले भी मौजूदा कानूनों का अधिक प्रभावी ढंग से उपयोग किया जा सकता है।

यहाँ भी हमें बातों को अलग करके देखना होगा। आवाज़ के खराब होने को जानना और कानून बदलना दो अलग बातें हैं। क्या अदालत में ऐसे सबूत पेश करके कोई मुकदमा लड़ा गया है, इसकी पुष्टि अभी तक नहीं हुई है। वैज्ञानिक और विद्वान अभी केवल यह सुझाव दे रहे हैं कि इसे इस तरह इस्तेमाल किया जा सकता है।

संबंधित संदर्भ सामग्री

- [1] Rodríguez-Garavito, C., Gruber, D. F., Nemeth, A. O., & Beguš, G. (2025). What if we understood what animals are saying? The legal impact of AI-assisted studies of animal communication. *Ecology Law Quarterly*, 52(1). <https://doi.org/10.15779/Z383X83N5Q>

- [2] Beguš, G., Sprouss, C., Lebanidze, T., Gero, S., & Andreas, J. (2025). Vowel- and diphthong-like spectral patterns in sperm whale codas. *Open Mind*, 9, 1849-1874. <https://doi.org/10.1162/OPMI.a.252>
- [3] Duarte, C. M., Chapuis, L., Collin, S. P., et al. (2021). The soundscape of the Anthropocene ocean. *Science*, 371(6529), eaba4658. <https://doi.org/10.1126/science.aba4658>
- [4] Surma, K. (2025, October 29). AI is decoding whales' communications. Could that be a turning point in the push for their rights? *Inside Climate News*. <https://insideclimatenews.org/news/29102025/ai-sperm-whale-communications-legal-rights/>

10.3 कैरेबियाई सागर की गवाही और मूल नविसियों का ज्ञान

गैर-मानव व्यक्तित्व पर चर्चा केवल पश्चिमी वदिवानों के कंप्यूटरो तक ही सीमति नहीं है। गैर-मानव व्यक्तित्व का अर्थ उन प्राणियों से है जो इंसान नहीं हैं, फरि भी कानून उनकी रक्षा कर सकता है। यह चर्चा लंबे समय से समुद्र की रक्षा करने वाले मूल नविसियों के ज्ञान से मलिकर और समृद्ध हुई है [2]।

28 मार्च 2024 को प्रशांत महासागर के कुक द्वीपसमूह (Cook Islands) के रारोटोंगा में एक ऐतिहासिक घटना घटी। माओरी राजा और कुक द्वीपसमूह के नेता एक साथ इकट्ठा हुए। ताहिती, टोंगा और हवाई के नेता भी इसमें शामिल हुए। उन्होंने व्हेल को अधिकारों से संपन्न एक स्वतंत्र इकाई के रूप में मान्यता दी। उन्हें ऐसा प्राणी घोषित किया गया जो इंसान नहीं है, लेकिन कानून द्वारा संरक्षित है। इस संयुक्त घोषणा का नाम हे व्हाकापुतंगा मोआना (He Whakaputanga Moana), यानी महासागर घोषणा है। यह घोषणा फरवरी 2026 में न्यूजीलैंड के वैतांगी में फरि से दुनिया के सामने आई [1][4]।

यह घोषणा व्हेल को संवेदनशील प्राणी के रूप में स्वीकार करने वाली, मूल नविसियों के नेतृत्व वाली पहली घोषणा है। इसमें व्हेल के आज़ादी से तैरने के अधिकार और स्वस्थ समुद्र में जीने के अधिकार की बात कही गई है।

यह घोषणा पारंपरिक पारिस्थितिक ज्ञान पर आधारित है। पारंपरिक पारिस्थितिक ज्ञान का अर्थ वह प्राचीन ज्ञान है जो प्रकृति को अपने परिवार जैसा मानता है। मूल नविसियों ने इस ज्ञान को सदियों से पीढ़ी-दर-पीढ़ी आगे बढ़ाया है। वे जानते थे कि व्हेल का गीत कोई संयोगवश होने वाला शोर नहीं है। उनका मानना था कि वह गीत समुद्र की यादों और कुल के इतिहास को संजोकर रखता है।

आधुनिक AI भी इस पुरानी समझ की दशा में ही संकेत करता है। Project CETI ने कैरेबियाई सागर में डोमिनिका के तट पर रहने वाली व्हेल का लंबे समय तक अध्ययन किया। शोध टीम ने AI की मदद से इन व्हेल के coda की जांच की। जब कुछ विशेषताओं को आपस में जोड़ा गया, तो पहचानी जा सकने वाली coda की संख्या बहुत बढ़ गई। पहले केवल 21 प्रकार के coda माने जाते थे। लेकिन जब संयोजन का तरीका सामने आया, तो अलग पहचाने जा सकने वाले रूप 100 से भी ज्यादा हो गए। यह वैसा ही है जैसे इंसान थोड़े से अक्षरों से अनगिनत शब्द बना लेते हैं। यह मूल नविसियों के उस विश्वास से भी मेल खाता है कि गीतों से अर्थ व्यक्त होते हैं [3]।

कुक द्वीपसमूह के नेता बहुत पहले से व्हेल को अपना रश्तेदार मानते आए हैं। वे व्हेल को समुद्र के बड़े बुजुर्ग की तरह सम्मान देते थे। यह सोच आधुनिक विज्ञान के आने से बहुत पहले से मौजूद थी। अब AI उस भावना को संख्याओं और आंकड़ों का प्रमाण दे रहा है।

इसलिए मूल नविसियों और वैज्ञानिकों ने मलिकर एक नया तरीका विकसित किया। यह तरीका AI की गणनाओं और मूल नविसियों के समुद्री ज्ञान को एक साथ रखता है। यहाँ मूल नविसी केवल सहायक नहीं हैं। वे AI को तैयार करने की शुरुआत से ही नरिणय लेने वाले मुख्य भागीदार हैं। वैज्ञानिक और मूल नविसी बराबरी से हाथ मलिकर समुद्र की रक्षा कर रहे हैं।

यह सहमत केवल एक देश तक सीमति नहीं रहती। ऐसा इसलिए है क्योंकि महासागर कई देशों के समुद्रों में फैले हुए हैं। व्हेल सीमाओं को नहीं जानती और विशाल समुद्र में तैरती रहती हैं। इसलिए सिर्फ एक देश के कानून से व्हेल की रक्षा नहीं हो सकती। महासागर घोषणा कई देशों और जनजातियों के एकजुट होने का पहला कदम है।

संबंधित संदर्भ सामग्री

- [1] He Whakaputanga Moana. (2024, March 28). Whale protection and legal personhood declaration. Rarotonga, Cook Islands. <https://www.whaledclaration.com/>
- [2] Rodríguez-Garavito, C., Gruber, D. F., Nemeth, A. O., & Beguš, G. (2025). What if we understood what animals are saying? The legal impact of AI-assisted studies of animal communication. *Ecology Law Quarterly*, 52(1). <https://doi.org/10.15779/Z383X83N5Q>
- [3] Sharma, P., Gero, S., Payne, R., Gruber, D. F., Rus, D., Torralba, A., & Andreas, J. (2024). Contextual and combinatorial structure in sperm whale vocalisations. *Nature Communications*, 15, 3617. <https://doi.org/10.1038/s41467-024-47221-8>

10.4 प्रजात के वकिसति होने का अधिकार

AI द्वारा खोले गए इस रास्ते के अंत में एक और बड़ा वचिार है। यह जहाजों का रास्ता बदलने के आदेश से भी आगे की बात है। यह गैर-मानव बुद्धिमत्ता का खुद अपने भवष्य की ओर बढ़ने का अधिकार है। कुछ विद्वान इसे 'प्रजात के वकिसति होने का अधिकार' कहते हैं। यह अभी कानून में दर्ज अधिकार नहीं है, बल्कि विद्वानों का एक प्रस्ताव है [1]।

इंसानी भाषा लंबे समय में वकिसति हुई है। यह कुछ ध्वनियों से शुरू होकर जटिल व्याकरण में बदली। जानवरों की ध्वनिसंस्कृत भी कोई ठहरी हुई बेजान चीज़ नहीं है। यह पीढ़ियों के दौरान सीखी जाने वाली, दोहराई जाने वाली, बदलने वाली और फलने-फूलने वाली एक जीवति धरोहर है। यही कारण है कि अलग-अलग क्षेत्रों में व्हेल के गीत थोड़े अलग होते हैं। यह बलिकुल वैसा ही है जैसे इंसानी भाषा क्षेत्रों के हिसाब से बोलियों में बदल जाती है [3]।

स्पर्म व्हेल मां को देखकर इसे समझा जा सकता है। मां अपने नवजात बच्चे को समूह की आवाज़ें सखाती है। बच्चा लंबे समय तक तुतलाने जैसी कच्ची आवाज़ों में उसकी नकल करता है। यह वैसा ही दखिता है जैसे मानव शिशु बोलना सीखते समय अधूरी नकल करता है। इसका अर्थ है कि स्पर्म व्हेल की आवाज़ जन्म से मली केवल एक सहज वृत्तनिही है। यह एक पीढ़ी से दूसरी पीढ़ी तक पहुंचने वाली सांस्कृतिक वरिसत है। जैसे इंसान कतिबे और गीत सौपते हैं, वैसे ही व्हेल अपनी आवाज़ें सौपती है।

लेकिन इंसानों द्वारा समुद्र में डाला गया शोर इस सीख को रोकता है। औद्योगिक शोर और दनि-रात होने वाले सैन्य वसिफोट समुद्र को भर देते हैं। नन्ही व्हेल अपने माता-पति की आवाज़ ठीक से नहीं सुन पाती हैं। ऐसा इसलए होता है क्योकि आवाज़ सीखने के लए जरूरी पचि और फ्रीक्वेंसी आपस में उलझ जाती है। यह वैसा ही है जैसे शोरगुल वाली कक्षा में शकिषक की आवाज़ सुनाई न दे। यह केवल एक व्हेल को नुकसान पहुंचाने की बात नहीं है। यह पूरे समूह की संस्कृति और यादों के सलिसलि को तोड़ना है [2]।

इसलए कुछ पारसिथतिकि दारशनकि और कानूनविद एक बात का तर्क देते हैं। वह यह है कि जानवरों को सुरक्षति रूप से आवाज़ सीखने और उसे आगे बढ़ाने का अधिकार मलिना चाहए। यहाँ उस दुनयि की बात होती है जसि हर जानवर अपनी इंद्रियों से महसूस करता है। चमगादड़ जसि दुनयि को महसूस करता है और व्हेल जसि महसूस करती है, वे दोनों अलग हैं। यह प्रस्ताव इंसानों के शोर या नकली आवाज़ों के बनिा खुद वकिसति होने का अधिकार देने की बात करता है। यह अभी कानून नहीं बना है, बल्कि विद्वानों के बीच वचिार-वमिर्श का चरण है।

AI द्वारा बनाई गई नकली व्हेल आवाज़ों से भी सावधान रहना चाहए। इंसानों द्वारा बनिा सोचे-समझे बजाई गई नकली आवाज़ें नन्हे बच्चों के सीखने में भ्रम पैदा करती हैं। इसलए इस प्रस्ताव में ऐसी नकली आवाज़ों को रोकने का नयिम भी शामिल है। गलत इस्तेमाल होने पर अच्छी तकनीक भी एक नया शोर बन सकती है। अनुवादक यंत्र को व्हेल की दुनयि को ध्यान से सुनने का साधन होना चाहए। उसे व्हेल की दुनयि में बाधा डालने वाला औज़ार नहीं बनना चाहए।

समुद्र के शोर को कम कयिा जाए, तो वह बहुत जल्दी घट जाता है। 2021 के साइंस (Science) शोध ने इस बात को स्पष्ट कयिा था। शोर सीधे स्रोत से नकिलता है, इसलए उसके कारण को कम करते ही समुद्र तुरंत शांत हो जाता है। अगर हम व्यापारिक शोर को कम कर दें, तो व्हेल एक बार फरि एक-दूसरे की आवाज़ को साफ़-साफ़ सुन सकेंगी।

संबंधति संदर्भ सामग्री

- [1] Rodríguez-Garavito, C., Gruber, D. F., Nemeth, A. O., & Beguš, G. (2025). What if we understood what animals are saying? The legal impact of AI-assisted studies of animal communication. *Ecology Law Quarterly*, 52(1).
<https://doi.org/10.15779/Z383X83N5Q>
- [2] Duarte, C. M., Chapuis, L., Collin, S. P., et al. (2021). The soundscape of the Anthropocene ocean. *Science*, 371(6529), eaba4658. <https://doi.org/10.1126/science.aba4658>
- [3] Sharma, P., Gero, S., Payne, R., Gruber, D. F., Rus, D., Torralba, A., & Andreas, J. (2024). Contextual and combinatorial structure in sperm whale vocalisations. *Nature Communications*, 15, 3617.
<https://doi.org/10.1038/s41467-024-47221-8>

10.5 नए सबूत और आगे की कहानी

यह क्षेत्र आज भी बहुत तेज़ी से आगे बढ़ रहा है। हर साल नए शोध सामने आते हैं और नए उपकरण वकिसति होते हैं। ये ऐसी चीज़ें हैं जिनकी कुछ साल पहले कल्पना करना भी मुश्किल था [1]।

बर्कले (Berkeley) और Project CETI का शोध इसका एक अच्छा उदाहरण है। जैसा कहिमने पहले देखा, शोध टीम ने स्पर्म व्हेल के coda में स्वरों (vowels) जैसी आवाज़ें खोजी थीं। 2026 में, प्रोसीडिंग्स ऑफ द रॉयल सोसाइटी बी (Proceedings of the Royal Society B) में एक अगला शोध प्रकाशित हुआ। इसमें बताया गया है कि स्पर्म व्हेल की स्वरों जैसी आवाज़ें इंसानों के स्वरों की तरह नियमों का पालन करती हैं। स्पर्म व्हेल इंसानों की तरह आवाज़ें बनाती हैं, इसका सबूत और मज़बूत हो गया।

अर्थ स्पीशीज प्रोजेक्ट (Earth Species Project) एक अलग रास्ता खोलता है। इस संगठन ने NatureLM-audio नाम का AI बनाया है। यह जानवरों की आवाज़ों के लिए पहला बड़ा ऑडियो भाषा मॉडल है। यह मॉडल केवल आवाज़ सुनकर पहचान लेता है कि वह कौन-सा जानवर है। रिकॉर्डिंग में कतिने जानवर हैं, यह उन्हें भी गनि लेता है। जसि जानवर की आवाज़ इसने पहले कभी नहीं सीखी, उसे भी यह कुछ हद तक पहचान लेता है [2]।

2026 में, इस संगठन ने अपने शोध के तरीके का दायरा बढ़ाया। केवल एक ही प्रजाति पर ध्यान देने के बजाय, वे कई प्रजातियों की आवाज़ों को एक साथ इकट्ठा करके देखते हैं। इस तरह वे कंप्यूटर से कई जानवरों की आवाज़ों का एक साथ विश्लेषण करने के तरीके का वसितार कर रहे हैं। जुलाई 2026 में, उन्होंने इसके लिए एक सार्वजनिक डेटा संग्रह alp-data जारी किया। यह एक ऐसा आधार है जहाँ दुनिया भर के शोधकर्ता एक ही डेटा साझा करते हैं। यह वैसा ही है जैसे पुस्तकालय से कोई भी एक ही किताब उधार लेकर पढ़ सकता है [3]।

मार्च 2026 में, Project CETI ने दो नए शोध प्रस्तुत किए। पहला शोध स्पर्म व्हेल के बच्चे के जन्म को वसितार से दर्ज करता है। समुद्र के भीतर जन्म को इतनी बारीकी से दर्ज करने वाला यह पहला रिकॉर्ड है। दूसरा शोध दिखाता है कि कैसे व्हेल का झुंड जन्म के समय मदद करता है। जब माँ बच्चे को जन्म देती है, तो दूसरी व्हेल उसके पास रहकर उसकी रक्षा करती है। यह प्राइमेट (primates) से अलग जानवरों में ऐसी सहयोगी देखभाल को वसितार से दिखाने वाला शोध है। इस व्यवहार को देखकर कुछ विद्वान व्हेलों के अधिकारों की बात करते हैं। उनका मानना है कि जटिल सामाजिकता और देखभाल का व्यवहार अधिकारों पर चर्चा का आधार बन सकता है। हालाँकि, एक-दूसरे की मदद करने के व्यवहार को सीधे कानूनी कर्तव्य मानना कठिन है। व्हेल का अपने साथी की मदद करना और इंसान का कानून का पालन करना दोनों अलग बातें हैं [4]।

कानून जानवरों के साथ इंसानों जैसा व्यवहार करे, इसमें अभी लंबा रास्ता बाकी है। अब तक कई देशों की अदालतों ने जानवरों के मुकदमा करने के अधिकार को स्वीकार नहीं किया है। 2022 में अमेरिकी न्यूयॉर्क राज्य की अदालत का फैसला इसका एक उदाहरण है। एक पशु अधिकार संगठन ने हैप्पी (Happy) नाम के हाथी के लिए मुकदमा दायर किया था। यह मुकदमा उसे एक तंग बाड़े से रहि कराने के लिए था। अदालत ने इसे स्वीकार नहीं किया। अदालत का कहना था कि इंसानों के लिए बनाई गई व्यवस्था को सीधे हाथी पर लागू नहीं किया जा सकता। इस प्रकार, सच्चाई को जानना और कानून को बदलना दोनों अलग-अलग बातें हैं [5]।

ये सभी सबूत एक ही दिशा की ओर इशारा करते हैं। वह यह सच है कि जानवरों की आवाज़ें केवल शोर नहीं हैं, बल्कि अर्थपूर्ण बातचीत हैं। जैसे-जैसे सबूत बढ़ते जाते हैं, कानून के लिए जानवरों को सिर्फ एक वस्तु मानना कठिन होता जाता है। विज्ञान आगे चलता है और कानून धीरे-धीरे उसके पीछे चलता है। AI जानवरों की आवाज़ को अदालत तक पहुँचाने वाला दुभाषिया बनता जा रहा है। धरती के जीवों की भाषा का अनुवाद करना उनके अधिकारों को स्थापित करने के काम से जुड़ता है।

संबंधित संदर्भ सामग्री

- [1] Beguš, G., Sprous, C., Lebanidze, T., Gero, S., & Andreas, J. (2026). The phonology of sperm whale coda vowels. *Proceedings of the Royal Society B*, 293(2044). <https://doi.org/10.1098/rspb.2025.2994>
- [2] Robinson, D., Miron, M., Hagiwara, M., Weck, B., Keen, S., Alizadeh, M., Narula, G., Geist, M., & Pietquin, O. (2025). NatureLM-audio: An audio-language foundation model for bioacoustics. *International Conference on Learning Representations (ICLR) 2025*. <https://arxiv.org/abs/2411.07186>
- [3] Earth Species Project. (2026, July 22). Introducing alp-data: A shared data layer for animal language processing. Earth Species Project. <https://earthspecies.org/2026/07/22/introducing-alp-data-a-shared-data-layer-for-animal-language-processing/>
- [4] Project CETI. (2026, March). Studies documenting rare sperm whale birth and ancient cooperative care released. *EurekAlert!* <https://www.eurekalert.org/news-releases/1120823>
- [5] Nonhuman Rights Project, Inc. v. Breheny. (2022). 38 N.Y.3d 555, New York Court of Appeals. <https://www.nycourts.gov/ctapps/Decisions/2022/Jun22/520pn22-Decision.pdf>

उपसंहार: इंटरस्पीशीज इंटरनेट (Interspecies Internet) और सह-अस्तित्व की पृथ्वी

E.1 स्पर्म् व्हेल की क्लिक ध्वनियां और गहरे समुद्र की वर्णमाला

समुद्र की गहराइयां हमेशा अंधेरी होती हैं। जहां सूरज की रोशनी नहीं पहुंचती, वह हजारों मीटर नीचे की जगह काली रात जैसी होती है। उस अंधेरे में स्पर्म् व्हेल आवाज से दुनिया को देखती है। स्पर्म् व्हेल अपने माथे से 'टक-टक' जैसी आवाज निकालती है। इस एक छोटी सी आवाज को क्लिक (click) कहा जाता है [3]।

स्पर्म् व्हेल कई क्लिक लगातार निकालती है। आवाजों के ऐसे जुड़े हुए समूह को कोडा (coda) कहा जाता है [2, 3]। क्लिक एक अकेली आवाज है, और कोडा कई क्लिक से बना आवाजों का एक गुच्छा है। कोडा 3 से लेकर लगभग 40 क्लिक को जोड़कर बनता है। इंसानी कानों को कोडा बस एक जैसी 'टक-टक' आवाज ही सुनाई देती है। लंबे समय तक इंसानों ने इस आवाज को बर्ना मतलब का शोर माना [2]।

लेकिन AI ने इस आवाज को दोबारा सुना। प्रोजेक्ट सेटी (Project CETI) नाम की एक शोध टीम आगे आई। सेटी व्हेल की भाषा को समझने की कोशिश कर रहे कई देशों के वैज्ञानिकों का एक समूह है। इस टीम की शोधकर्ता प्रत्युषा शर्मा (Pratyusha Sharma) और प्रोफेसर जैकब एंड्रियास (Jacob Andreas) ने इन आवाजों का विश्लेषण किया। डोमिनिका की स्पर्म् व्हेल पर लंबे समय तक नज़र रखने वाले डॉ. शेन गेरो (Shane Gero) भी उनके साथ शामिल थे [1]।

शोध टीम ने बहुत सारी आवाजें इकट्ठा करके उनकी जांच की। उन्होंने स्पर्म् व्हेल के 8,719 कोडा (coda) को एक-एक करके ध्यान से देखा। उनमें उन्होंने कोडा बनाने वाले चार बुनियादी तत्वों को खोजा। ये तत्व आवाज की लय, गति, और जोड़े जाने वाले छोटे क्लिक जैसी चीजें हैं। इन तत्वों को अलग-अलग तरह से बदलने पर अलग-अलग कोडा बनते हैं।

AI ने एक जैसे दिखने वाले क्लिक के बीच छपि नयियों को ढूँढ निकाला। स्पर्म् व्हेल क्लिक के बीच के अंतराल को थोड़ा-थोड़ा बढ़ाती या घटाती है। वह आवाज की गति को भी बदलती है। वह क्लिक के समूह के आगे या पीछे छोटी आवाजें भी जोड़ती है। ये छोटे-छोटे अंतर मलिकर अलग-अलग ध्वनिसंरचनाएं बनाते हैं।

स्पर्म् व्हेल आवाज की गति को भी नियंत्रित करती है। वह एक ही कोडा (coda) को थोड़ा लंबा खींचकर भी निकालती है। वह उसे थोड़ा छोटा करके तेज़ी से भी निकालती है। यह गाना गाते समय ताल को लंबा या छोटा करने जैसा है। यह छोटा सा खिंचाव और बदलाव भी आवाज को बदलने की एक चाबी बनता है।

नए शोध ने यहाँ से एक कदम और आगे बढ़ाया। वैज्ञानिकों ने देखा कि कोडा (coda) के भीतर कसि पचि की गुंज तेज़ होती है। जब इंसान 'आ' और 'ई' स्वर बोलते हैं, तब भी तेज़ गुंजने वाले हस्से होते हैं। स्पर्म् व्हेल के कोडा में भी ऐसे हस्से दिखाई दिए। सेटी (CETI) टीम ने व्हेल के ऐसे दो स्वर खोजे। ये इंसानों द्वारा निकाली जाने वाली 'आ' और 'ई' की आवाजों से मलिते-जुलते हैं [4]।

स्पर्म् व्हेल ऐसे स्वर कैसे निकालती है? स्पर्म् व्हेल अपने सरि के अंदर मौजूद 'ध्वनि होठों' से आवाज बनाती है। वह उस आवाज को अपने सरि के अंदर की हवा की धैलियों से तराशती है। यह इंसान के गले से आवाज निकालने और मुंह का आकार बदलकर स्वर बनाने जैसा है। यानी आवाज बनाने की जगह और उसे संवारने की जगह अलग-अलग बंटी हुई है।

शोध टीम ने बताया कि स्पर्म् व्हेल लगभग 156 अलग-अलग प्रकार के कोडा (coda) निकालती है। ये कोडा अकेले अक्षर नहीं हैं। ये लय और गति जैसे कुछ तत्वों को जोड़कर बनाए गए आवाजों के समूह हैं। कम चीजों से बहुत सारी आवाजें बनाने का हुनर गहरे समुद्र में भी मौजूद था। हालांकि, हर कोडा का क्या अर्थ है, यह अभी तक पता नहीं चला है। शोध टीम ने अभी केवल ध्वनियों की संरचना खोजी है।

एक ही कोडा (coda) भी हर झुंड में थोड़ा-थोड़ा अलग होता है। यह इंसानों की स्थानीय बोलियों जैसा है। एक झुंड के भीतर वे एक-दूसरे की आवाज पर ताल मलिते हुए एक साथ आवाजें निकालते हैं। जब एक व्हेल आवाज निकालती है, तो दूसरी व्हेल तुरंत अपनी आवाज उससे जोड़ देती है। इस तरह स्पर्म् व्हेल मलिकर बातचीत का प्रवाह बनाती है।

कुछ कोडा (coda) नेमटैंग जैसा काम करते हैं। स्पर्म् व्हेल बड़े पारिवारिक झुंडों में रहती है। हर झुंड का अपना पसंदीदा कोडा होता है। इस कोडा को सुनकर पता चल सकता है कि व्हेल कसि झुंड की है। यह इंसानों की बोली सुनकर उनके इलाके का अंदाज़ा लगाने जैसा है।

यह बात बहुत महत्वपूर्ण है कि ऐसे छोटे-छोटे अंतर मलिकर आवाजें बनाते हैं। स्पर्म् व्हेल ध्वनिके कुछ तत्वों को तरह-तरह से आपस में जोड़ती है। उन्हें जोड़ने के तरीके के आधार पर अलग-अलग कोडा (coda) बनते हैं। यह इंसानों द्वारा कुछ अक्षरों से अनगिनत शब्द बनाने जैसा है। गहरे समुद्र में भी ऐसी ही सुंदर व्यवस्था छपि हुई थी।

इन आवाजों को पकड़ने के लिए अच्छे रिकॉर्डिंग उपकरण की जरूरत होती है। सेटी टीम ने व्हेल के शरीर पर चपिकने वाला एक छोटा रिकॉर्डर बनाया। यह उपकरण व्हेल की आवाज के साथ-साथ पानी के अंदर की स्थिति भी रिकॉर्ड करता है। ऐसा इसलिए, क्योंकि किस स्थिति में कौन सी आवाज निकाली गई, यह देखकर ही उसका अर्थ समझा जा सकता है। इंसानों को AI जैसा नया उपकरण मलिन के बाद ही गहरे समुद्र की वर्णमाला दिखाई दी।

संबंधित संदर्भ सामग्री

- [1] Sharma, P., Gero, S., Payne, R., Gruber, D. F., Rus, D., Torralba, A., & Andreas, J. (2024). Contextual and combinatorial structure in sperm whale vocalisations. *Nature Communications*, 15, 3617.
<https://doi.org/10.1038/s41467-024-47221-8>
- [2] Oceanographic. (2026). Scientists listen in on sperm whale communication with new tech. *Oceanographic Magazine*.
<https://oceanographicmagazine.com/news/scientists-listen-in-on-sperm-whale-communication-with-new-tech/>
- [3] Project CETI. (2026). Research. Project CETI. <https://www.projectceti.org/research/index>
- [4] Kemeny, R. (2026, April 14). We're one step closer to understanding the sperm whale 'alphabet'. *National Geographic*.
<https://www.nationalgeographic.com/animals/article/sperm-whale-vowel-patterns-communication-discovery>

E.2 केवल इंसान ही बोलते हैं, इस सोच से परे

लंबे समय तक इंसानों ने खुद को सबसे खास समझा। वे मानते थे कि पृथ्वी पर केवल इंसान ही बात कर सकते हैं और वचिार साझा कर सकते हैं। इंसानी भाषा से अलग जानवरों की आवाजों को उन्होंने सिर्फ शोर समझा। केवल इंसानों को ही खास मानने की यह सोच बहुत लंबे समय तक चली [1]।

यह विश्वास एक पुरानी दीवार की तरह पक्का था। उस दीवार के कारण जानवरों की आवाजें अंधेरे में दबी रही। इंसानों ने अपने पड़ोसी जीवों के वचिारों को सुनने की कोशिश नहीं की। इस बीच कतिनी सारी कहानियां यूँ ही बीत गई, यह हम नहीं जानते।

यह दीवार इंसानी तरीके से जानवरों को नापने की आदत से खड़ी हुई थी। इंसान मानते थे कि केवल शब्द और वाक्य होने पर ही कोई भाषा बनती है। उस ढाँचे में न बैठने वाली आवाजों को उन्होंने बस अर्थ का कहकर नकार दिया। लेकिन जानवरों की आवाजों को इंसानी भाषा के ढाँचे में ढलाने की कोई जरूरत नहीं है। ऐसा इसलिए, क्योंकि वे आवाजें उनके अपने अलग शरीर और अलग जीवन से निकलती हैं।

AI इस दीवार को धीरे-धीरे तोड़ रहा है [3, 4]। AI जानवरों की आवाजों को एक वैचारिक नक्शे पर व्यवस्थित करता है। वैज्ञानिक इस नक्शे को लैटेंट स्पेस (latent space) कहते हैं। इस नक्शे पर, अगर आवाजें मलित-जुलती हों, तो बड़ि भी पास-पास इकट्ठा होते हैं। अगर आवाजें अलग हों, तो बड़ि भी दूर चले जाते हैं [4]।

वैज्ञानिकों ने इंसानी भाषा के नक्शे और जानवरों की आवाजों के नक्शे को साथ रखकर देखा। ये दोनों नक्शे एक-दूसरे से पूरी तरह अजनबी नहीं थे। नक्शों के आकार मिलाकर जब उन्हें एक-दूसरे पर रखा गया, तो मलित-जुलते हिस्से दिखाई दिए। इंसानों का कोई शब्द और जानवरों का कोई संकेत एक ही जगह पर भी आए।

इसके पीछे एक दलिचस्प तरीका छपा है। दोनों नक्शों को मलिते समय पहले से बनी अनुवाद सूची की जरूरत नहीं पड़ती। वैज्ञानिक दोनों नक्शों की बनावट देखकर ही आपस में मेल खाने वाली जगहें ढूँढ लेते हैं। यह खड़िकी के कांच पर दो कागजों को रखकर उनके डिजाइन मलाने जैसा है। वैज्ञानिक अभी इस प्रकार आकार मलाकर देखने के तरीके का परीक्षण कर रहे हैं।

इस तरह के शोध का नेतृत्व करने वाले कुछ वैज्ञानिक हैं। अमेरिका के भाषाविद् गाशपेर बेगुश (Gašper Beguš) AI की मदद से जानवरों की आवाज के नयिम खोजते हैं। वे इंसानी भाषा के नयिमों को जानवरों पर जबरदस्ती नहीं थोपते। वे आवाज की अपनी बनावट से नयिमों के खुद सामने आने की प्रतीक्षा करते हैं।

यहाँ से एक नया वचिार जन्म लेता है। जानवर इंसानी भाषा का इंतजार करने वाले जीव नहीं हैं। वे अपनी-अपनी दुनिया को महसूस करने और उसे आवाजों में ढालने वाले हमारे साथी हैं। जानवर अपनी इंद्रियों से जसि दुनिया को अनुभव करते हैं, उसे उम्बवेल्ट (Umwelt) कहा जाता है। चमगादड़ आवाज से दुनिया को देखते हैं, और कुत्ते सूँघकर दुनिया को समझते हैं। व्हेल की अपनी दुनिया है, और पक्षियों की अपनी दुनिया है।

नक्शों को मलाने का यह तरीका कठनि परस्थितियों में बहुत मददगार साबित होता है। इंसान और जानवरों के बीच पहले से बनी कोई डिक्शनरी नहीं है। शब्दों के जोड़े बनाने वाली कोई सूची भी नहीं है। फरि भी AI केवल दोनों नक्शों के आकार देखकर जोड़े ढूँढने की कोशिश करता है। यह दो अलग-अलग दुनियाओं को जोड़ने वाला एक नया पुल बन सकता है।

इसलिए हमें जानवरों की आवाजों को इंसानी भाषा से कमतर नहीं समझना चाहिए। हमें जानवरों के वचारों और उनकी लय का बराबरी से सम्मान करना चाहिए। यह हर जीव की आवाज को समान रूप से अनमोल समझने की सोच है। अगर इंसान को पृथ्वी के एक सच्चे परिवार के रूप में बड़ा होना है, तो पहले यह सोच अपनानी होगी [2]।

संबंधित संदर्भ सामग्री

- [1] Rodríguez-Garavito, C. (2026, February 19). AI-enabled decoding of whale communication could bolster animal rights. NYU School of Law Ideas. <https://www.law.nyu.edu/news/ideas/cesar-rodriguez-garavito-animal-communication-ai>
- [2] MOTH - More Than Human Life. (2026). MOTH x CETI: Exploring how understanding sperm whale communications can be a force for good. mothrights.org. <https://mothrights.org/news/moth-x-ceti-exploring-how-understanding-sperm-whale-communications-can-be-a-force-for-good/>
- [3] Bioneers. (2026). The quest to decode whale-speak. Bioneers. <https://bioneers.org/the-quest-to-decode-whale-speak/>
- [4] Rutz, C., Bronstein, M., Raskin, A., Vernes, S. C., Zacarian, K., & Blasi, D. E. (2023). Using machine learning to decode animal communication. Science, 381(6654), 152-155. <https://doi.org/10.1126/science.adg7314>

E.3 पृथ्वी के जीवों को जोड़ने वाला इंटरस्पीशीज इंटरनेट

इंसानों के बीच इंटरनेट मौजूद है। इंटरनेट एक ऐसा जाल है जो दूर बैठे लोगों को संदेश भेजने और प्राप्त करने के लिए जोड़ता है। लेकिन पृथ्वी पर केवल इंसान ही नहीं रहते। यहाँ वृहल और हाथी, पक्षी और मधुमक्खियाँ, और पेड़-पौधे भी साथ रहते हैं। इन सभी जीवों को एक सूत्र में परिने का एक सपना है। इस सपने को इंटरस्पीशीज इंटरनेट (Interspecies Internet) कहा जाता है [1]।

यह सपना पहली बार 2013 में सामने आया था। इंटरनेट बनाने वाले वॉटि सर्फ (Vint Cerf) और संगीतकार पीटर गेब्रियल (Peter Gabriel) इसमें सबसे आगे रहे। वैज्ञानिक नील गर्शेनफेल्ड (Neil Gershenfeld) और डायना रीस (Diana Reiss) भी उनके साथ जुड़े। उन्होंने कहा कि इंटरनेट को केवल इंसानों तक सीमिति नहीं रखना चाहिए। उनका सोचना था कि यह जाल जानवरों के लिए भी खोला जाना चाहिए।

इंटरस्पीशीज इंटरनेट केवल किसी ऐप से जानवरों की आवाज का अनुवाद करने तक सीमिति नहीं है। इसका मतलब एक ऐसा विशाल नेटवर्क है जहाँ पृथ्वी के सभी जीव एक-दूसरे के समाचार और खतों की जानकारी साझा करते हैं। जैसे इंसानी नसे शरीर के हर हिस्से को जोड़ती हैं, वैसे ही यह जाल पूरी पृथ्वी के जीवन को जोड़ता है।

असल में, इस जुड़ाव के टुकड़े प्रकृति में पहले से मौजूद हैं। कई जानवर दूसरी प्रजातियों की चेतावनी भरी आवाजें सुनते हैं। जब मीरकैट खतरे का संकेत देता है, तो पास के दूसरे जानवर उस आवाज को समझकर साथ भाग जाते हैं। वैज्ञानिकों ने 70 से अधिक प्रजातियों के जानवरों में इस तरह की बातें सुनने की पुष्टि की है। यानी अलग-अलग जानवर खतरे की एक जैसी जानकारी आपस में बांटते हैं [2]।

डॉल्फिन भी ऐसा साझापन दिखाती हैं। जब समुद्र में डॉल्फिन के झुंड एक-दूसरे से मिलते हैं, तो वे सीटी की आवाजें भेजते हैं। हर डॉल्फिन की अपनी अलग सीटी नेमटैग जैसा काम करती है। वे आवाज से पहचानती हैं कि सामने वाला कौन है। आँखों से न देखने वाला यह आवाजों का रास्ता पहले से ही एक जाल बनाता है [3]।

यह जाल सिर्फ आवाजों को ही नहीं जोड़ता। पेड़-पौधे भी संकेत भेजते हैं। पानी सूखने पर टमाटर और तंबाकू जैसे पौधे ऐसी आवाजें निकालते हैं जो इंसानों को सुनाई नहीं देती। तनाव में आया एक पौधा एक घंटे में 30 से 50 बार 'टक-टक' की आवाज निकालता है। AI ऐसी आवाजों को इकट्ठा करके पौधों की स्थिति को पढ़ सकता है [4]।

AI इन टुकड़ों को आपस में जोड़ने में मदद करता है। आने वाले नए संचार नेटवर्क भी इसमें हाथ बंटा सकते हैं। 6G नाम का नेटवर्क रेडियो तरंगों को चीजों से टकराकर आसपास की हलचल को समझ सकता है। नेटवर्क को सेंसर की तरह उपयोग करने के इस विचार को 'नेटवर्क एज अ सेंसर' (network as a sensor) कहा जाता है। हालांकि, इस तकनीक को जानवरों के संचार नेटवर्क से जोड़ना अभी केवल एक संभावना है [5]।

इस जाल का सही उपयोग करके प्रकृति की रक्षा की जा सकती है। मान लीजिए कि वर्षावनों में पेड़ों को काटना शुरू किया जाता है। तितर-बितर होते पक्षियों की हलचल और पौधों के असामान्य संकेत इस नेटवर्क के जरिए केंद्र में इकट्ठे होंगे। इंसान तुरंत वह खबर पाकर उसे रोक सकते हैं। इंटरस्पीशीज इंटरनेट नगरानी का नहीं, बल्कि साथ मिलकर जीने का साधन बन सकता है। यह एक ऐसा बड़ा तंत्रिका तंत्र बन जाएगा, जहाँ पृथ्वी के जीव एक-दूसरे के खतरे साझा करेंगे और एक-दूसरे की देखभाल करेंगे।

संबंधित संदर्भ सामग्री

- [1] Interspecies Internet. (n.d.). About. Interspecies Internet. <https://www.interspecies.io/about>
- [2] Magrath, R. D., Haff, T. M., Fallow, P. M., & Radford, A. N. (2015). Eavesdropping on heterospecific alarm calls: from mechanisms to consequences. *Biological Reviews*, 90(2), 560-586. <https://doi.org/10.1111/brv.12122>
- [3] Quick, N. J., & Janik, V. M. (2012). Bottlenose dolphins exchange signature whistles when meeting at sea. *Proceedings of the Royal Society B*, 279(1738), 2539-2545. <https://doi.org/10.1098/rspb.2011.2537>
- [4] Khait, I., Lewin-Epstein, O., Sharon, R., et al. (2023). Sounds emitted by plants under stress are airborne and informative. *Cell*, 186(7), 1328-1336. <https://doi.org/10.1016/j.cell.2023.03.009>
- [5] Nokia Bell Labs. (n.d.). Network as a sensor. Nokia. <https://www.nokia.com/bell-labs/research/6g-networks/6g-technologies/network-as-a-sensor/>

E.4 नए उपकरणों ने क्या उजागर किया: प्रकृति की भाषा सीखता AI

AI जानवरों की आवाजें सीखने की दशा में भी आगे बढ़ा। अर्थ स्पीशीज प्रोजेक्ट (Earth Species Project) ने एक नया AI पेश किया। इसका नाम नेचुरलएम-ऑडियो (NatureLM-audio) है। यह जानवरों की आवाजों के लिए बनाया गया पहला बड़ा ध्वनि AI है। इसे पहली बार नवंबर 2024 में सार्वजनिक किया गया, और 2025 में अंतरराष्ट्रीय सम्मेलन (ICLR) में औपचारिक रूप से पेश किया गया [1]।

यह AI एक साथ कई काम करता है। यह आवाज सुनकर पहचान लेता है कि वह कौन सा जानवर है। वह जानवर बच्चा है या बड़ा, यह भी पता लगा लेता है। यह आवाजों के प्रकारों को बांटता है और उन पर विवरण जोड़ता है। यह रिकॉर्डिंग में कितने जानवर हैं, उन्हें गिनता भी है।

सबसे खास बात यह है कि इसने वह काम भी कर दिखाया जो इंसानों ने इसे नहीं सिखाया था। आमतौर पर AI केवल वही कर सकता है जो उसे सिखाया गया हो। लेकिन इस AI ने बहुत सारी आवाजें सीखते हुए अपने आप नए कौशल हासिल कर लिए। जानवरों की गिनती करना भी इसने शुरू से नहीं सीखा था। बहुत सारी आवाजें सुनते-सुनते यह अपने आप ऐसा करने में सक्षम हो गया।

अर्थ स्पीशीज प्रोजेक्ट (Earth Species Project) कई जानवरों की भाषाओं की एक साथ जांच करता है। अभी यह कौबो, बेलुगा व्हेल और हाथियों की आवाजों का गहरा अध्ययन कर रहा है। यह दुनिया भर के जीववैज्ञानिकों के साथ मिलकर डेटा इकट्ठा करता है। केवल एक प्रजाति को नहीं, बल्कि कई प्रजातियों को साथ देखना इस टीम का तरीका है।

जुलाई 2026 में एक और नया उपकरण सामने आया। इसका नाम एएलपी-डेटा (alp-data) है। यह कई जानवरों की आवाज के डेटा को एक जगह इकट्ठा करने वाला उपकरण है। इसमें पक्षियों, व्हेल और कीड़ों की आवाजों के 35 डेटासेट शामिल हैं। अब वैज्ञानिक कई जानवरों की आवाजों को एक ही जगह से निकालकर इस्तेमाल कर सकते हैं [2]।

अर्थ स्पीशीज प्रोजेक्ट इस नई धारा को एनमिल लैंग्वेज प्रोसेसिंग (Animal Language Processing) कहता है। यह AI से जानवरों के संवाद का अध्ययन करने का तरीका है। यह इंसानी भाषा के नियमों को जानवरों पर वैसा का वैसा नहीं थोपता। यह पहले इस बात को स्वीकार करता है कि हर जानवर के बातचीत करने का तरीका अलग हो सकता है। इसके आधार पर AI डेटा से नियमों को खोजता है [3]।

इस टीम ने लोगों की राय भी पूछी। उन्होंने 2025 में दुनिया के 67 देशों के 1,057 लोगों से सवाल किए। बहुत से लोगों ने जानवरों की भाषा को समझने की बात का स्वागत किया। साथ ही उन्होंने यह चिंता भी जताई कि जानवरों के साथ गलत बर्ताव नहीं होना चाहिए। जैसे-जैसे तकनीक बढ़ रही है, वैसे-वैसे सावधानी की भावना भी बढ़ रही है [4]।

ये नए उपकरण वैज्ञानिकों के काम करने का तरीका बदल रहे हैं। पहले एक वैज्ञानिक लंबे समय तक केवल एक ही जानवर का गहरा अध्ययन करता था। डेटा भी अलग-अलग रूपों में बखिरा हुआ था। अब कई जानवरों के डेटा को एक ही जगह पर एक साथ देखा जा सकता है। अलग-अलग बंटे हुए शोध अब एक बड़े प्रवाह में एक साथ आ रहे हैं।

अगला लक्ष्य ध्वनि-प्लेबैक परीक्षण है। यह AI द्वारा बनाई गई आवाजों से जानवरों से बात करने की कोशिश करने का परीक्षण है। इस योजना की घोषणा मुख्य रूप से Project CETI ने की थी। Earth Species Project बनी सोचे-समझे आवाजें सुनाने के बजाय डेटा इकट्ठा करने और उसका अध्ययन करने पर ध्यान लगाता है। इंसानों द्वारा बनाई गई आवाज को जानवर समझ पाते हैं या नहीं, यह अभी तक साबित नहीं हुआ है [5]।

संबंधित संदर्भ सामग्री

- [1] Earth Species Project. (2024, November 11). Introducing NatureLM-audio: An audio-language foundation model for bioacoustics. Earth Species Project. <https://earthspecies.org/2024/11/11/introducing-naturelm-audio-an-audio-language-foundation-model-for-bioacoustics/>

- [2] Earth Species Project. (2026, July 22). Introducing alp-data: A shared data layer for animal language processing. Earth Species Project. <https://earthspecies.org/2026/07/22/introducing-alp-data-a-shared-data-layer-for-animal-language-processing/>
- [3] Earth Species Project. (2026, April 2). Animal language processing: An AI convergence in animal communication. Earth Species Project. <https://earthspecies.org/2026/04/02/animal-language-processing-an-ai-convergence-in-animal-communication/>
- [4] Earth Species Project. (2025, September 29). What the world thinks about AI and animal communication: Findings from our first global survey. Earth Species Project. <https://earthspecies.org/2025/09/29/what-the-world-thinks-about-ai-and-animal-communication-findings-from-our-first-global-survey/>
- [5] Project CETI. (2026). Research. Project CETI. <https://www.projectceti.org/research/index>

E.5 सावधानी से सुनना: ज़िम्मेदार उपकरण और सह-अस्तित्व का वादा

नए उपकरण बहुत शक्तिशाली होते हैं। शक्तिशाली उपकरण का गलत इस्तेमाल करने पर नुकसान हो सकता है। AI अब जानवरों की बिल्कुल वैसी ही आवाज़ें बना सकता है। अगर इस ताकत का गलत इस्तेमाल किया जाए, तो जानवरों को धोखा भी दिया जा सकता है। नकली आवाज़ों से जंगली जानवरों को फुसलाने की बात सोचिए। शिकारी भी शिकार करने के लिए इन नकली आवाज़ों का सहारा ले सकते हैं [1][2]।

इसलिए हमें तीन वादों का पालन करना होगा।

पहला, बना सोचे-समझे दखल न देने का वादा है। जानवरों के समूहों ने लंबे समय में अपनी आवाज़ की अनूठी संस्कृति विकसित की है। इंसानों को व्यापार या नयितरण के लिए नकली आवाज़ें डालकर उस व्यवस्था को बिगाड़ना नहीं चाहिए। AI जानवरों की आवाज़ छीनने का हथियार नहीं है। इसे जानवरों की संस्कृति को उनकी मर्जी के अनुसार सुरक्षित रखने का उपकरण होना चाहिए।

ऐसे वादे की ज़रूरत क्यों है? इंसान लंबे समय से सिर्फ़ प्रकृति को जीतने की कोशिश करता रहा है। आवाज़ को समझने की शक्ति भी आसानी से दूसरों को नयितरित करने की ताकत में बदल सकती है। उनके समूह में बना सोचे-समझे दखल देने से जानवरों का जीवन अस्त-व्यस्त हो जाता है। इसलिए जितना अधिक हम सीखते हैं, हमें उतना ही पीछे हटना भी आना चाहिए।

दूसरा, आवाज़ में नशान छोड़ने का वादा है। वैज्ञानिक सुझाव देते हैं कि AI द्वारा बनाई गई जानवरों की आवाज़ों में एक अदृश्य नशान लगाया जाए। इस नशान को ऑडियो वॉटरमार्क (audio watermark) कहा जाता है। नकली नोटों में छपि चित्र की तरह, इंसान के कान इसे नहीं सुन सकते, लेकिन मशीनें इसे पहचान लेती हैं। तब यह पहचाना जा सकता है कि कौन सी आवाज़ असली है और कौन सी नकली [3]।

यह नशान डेटा की सुरक्षा के काम भी आता है। वैज्ञानिक समुद्र और जंगलों में लंबे समय तक रिकॉर्डर लगाकर छोड़ देते हैं। इंसान के दखल दिए बिना केवल आवाज़ें इकट्ठा करने के इस काम को नैसर्गिक ध्वनिक निगरानी (passive acoustic monitoring) कहा जाता है। अगर इसमें नकली आवाज़ें मलि जाएँ, तो यह डेटा दूषित हो जाता है। आवाज़ में नशान लगाकर नकली आवाज़ों को अलग किया जा सकता है और डेटा को शुद्ध रखा जा सकता है।

तीसरा, साथ मलिकर तय करने का वादा है। न्यूयॉर्क यूनिवर्सिटी स्कूल ऑफ़ लॉ का MOTH कार्यक्रम इस काम का नेतृत्व कर रहा है। MOTH गैर-मानव जीवों के अधिकारों का अध्ययन करने वाला एक समूह है। मशीनों द्वारा खोजे गए डेटा को समुद्र की रक्षा करने वाले मूल नविसायियों के पुराने ज्ञान से मलिना चाहिए। मूल नविसायियों द्वारा लंबे समय के अवलोकन से मलि प्राकृतिक ज्ञान को पारंपरिक पारस्थितिक ज्ञान (traditional ecological knowledge) कहा जाता है। नियम केवल इंसानों के फायदे को देखकर नहीं बनाए जाने चाहिए।

ये तीनों वादे एक ही सोच से निकलते हैं। वह सोच यह है कि नई ताकत मलिन पर भी दूसरों के साथ मनमानी न की जाए। इंसान लंबे समय से प्रकृति को सिर्फ़ एक औज़ार मानता आया है। अब उस आदत पर विचार करने का समय है। AI जैसी नई ताकत घमंड करने की चीज़ नहीं, बल्कि एक भारी ज़िम्मेदारी है।

जब हम स्पर्म् व्हेल की भाषा समझ लेंगे, तो उनके नए अधिकारों के बारे में भी सोच सकते हैं। वैज्ञानिक कहते हैं कि व्हेल को परेशान न किए जाने का अधिकार होना चाहिए। वे उनकी अपनी संस्कृति को सुरक्षित रखने के अधिकार की भी बात करते हैं। 2024 में, न्यूजीलैंड और कुक आइलैंड्स के मूल नविसायी नेताओं ने घोषणा की कि व्हेल को इंसानों जैसे समान अधिकारों वाला जीव माना जाना चाहिए। इसका मतलब है कि व्हेल कोई ऐसी वस्तु नहीं है जिसे इंसान अपनी मर्जी से इस्तेमाल करे।

यह किताब पृथ्वी के जीवों की भाषा का अनुवाद करने की कहानी बताती है। AI ने स्पर्म् व्हेल के क्लिक (clicks), पक्षियों के गीतों और हाथियों की धीमी गूँज को एक-एक करके सुलझाया है। अनुवाद दूसरों को जीतने के लिए नहीं, बल्कि उन्हें जानने और समझने के लिए है। AI पृथ्वी के जीवों का अनुवाद करने वाला एक पुल है। उस पुल को सावधानी से पार करके ही हम अपने पड़ोसी जीवों के साथ मलिकर रहने वाली पृथ्वी बना सकते हैं। अनुवाद का काम यहाँ समाप्त नहीं होता, यह तो बस अभी शुरू हुआ है।

संबंधित संदर्भ सामग्री

- [1] Inside Climate News. (2025, October 29). AI is decoding whales' communications. Could that be a turning point in the push for their rights? Inside Climate News. <https://insideclimatenews.org/news/29102025/ai-sperm-whale-communications-legal-rights/>
- [2] MOTH - More Than Human Life. (2026). MOTH x CETI: Exploring how understanding sperm whale communications can be a force for good. mothrights.org. <https://mothrights.org/news/moth-x-ceti-exploring-how-understanding-sperm-whale-communications-can-be-a-force-for-good/>
- [3] Rutz, C., Bronstein, M., Raskin, A., Vernes, S. C., Zacarian, K., & Blasi, D. E. (2023). Using machine learning to decode animal communication. Science, 381(6654), 152-155. <https://doi.org/10.1126/science.adg7314>

इस AI पुस्तकालय पुस्तक का समर्थन करें

यदि यह पुस्तक उपयोगी रही, तो आप PayPal के माध्यम से आगे के अनुवाद, सार्वजनिक संस्करण और खुली कृत्रिम बुद्धिमत्ता ज्ञान परियोजनाओं का समर्थन कर सकते हैं।

PayPal

<https://paypal.me/kimkjkj>



QR कोड स्कैन करें या ऊपर दिया गया लिंक खोलें।

अधिवक्ता कमि क्यूंग-जनि