

Édition Bibliothèque IA en français

L'intelligence artificielle traduit le langage des animaux

L'histoire d'une IA qui
apprend à écouter les
baleines, les dauphins, les
oiseaux et les abeilles

Kim Kyung-jin, avocat

kimkj.com · Bibliothèque IA · 2026

L'intelligence artificielle traduit le langage des animaux

L'histoire d'une IA qui apprend à écouter les baleines, les dauphins, les oiseaux et les abeilles

Kim Kyung-jin, avocat

Tout commence avec trois histoires. Une chercheuse a vécu 40 ans avec les dauphins des Bahamas. Une linguiste a trouvé des motifs proches de voyelles dans les clics des cachalots. Des scientifiques ont échangé des sons pendant 20 minutes avec une baleine à bosse appelée Twain. Ce livre présente tout cela en douze chapitres. Il explique comment l'intelligence artificielle dessine les sons des animaux comme une carte pour trouver des règles. Il présente de nouveaux outils qui lisent les signaux des oiseaux, des abeilles et des singes. Il parle du respect du monde sensoriel des animaux. Il explique aussi le danger de tromper les animaux avec de faux sons, et les règles pour l'éviter. Il aborde enfin le débat sur l'idée de donner une personnalité juridique aux baleines. Les phrases sont courtes et faciles, afin que des élèves de fin d'école primaire puissent lire ce livre seuls. Chaque sous-chapitre contient des sources vérifiées. Cette version a été corrigée après la vérification de deux examinateurs utilisant l'intelligence artificielle.

Archives publiques des documents de recherche

Ce livre a été préparé et produit avec l'aide de l'intelligence artificielle. Il rassemble des documents écrits dans les principales langues du monde. NotebookLM ne limite pas la langue des questions selon la langue des documents. Le lecteur peut poser des questions dans sa propre langue aux mêmes archives de recherche que celles utilisées pour préparer ce livre. Cela reste vrai même si certains documents sont en coréen, en anglais ou dans d'autres langues. Dans NotebookLM, ouvrez les paramètres et choisissez la langue de sortie voulue. Les réponses et les contenus créés apparaîtront alors dans cette langue.

Archives publiques NotebookLM utilisées pour préparer ce livre (L'intelligence artificielle traduit la vie de la Terre) : <https://notebook.google.com/notebook/bb152971-e02c-4008-b56e-4076e1180e53>

La version originale coréenne de ce livre se trouve dans la bibliothèque IA de kimkj.com : <https://kimkj.com/ai-library/?mod=document&uid=12655>

Table des matières

Appuyez sur le titre de chaque chapitre pour aller au texte correspondant. Pour lire depuis le début, choisissez l'introduction.

Introduction: Ouvrir les voix de la communauté de la Terre ·	4
Chapitre 1. Le Wild Dolphin Project et le monde des dauphins tachetés ·	11
Chapitre 2. L'alphabet phonémique du cachalot et Project CETI ·	18
Chapitre 3. Rencontre avec la baleine à bosse Twain ·	25
Chapitre 4. L'orchestre des profondeurs ·	32
Chapitre 5. Le langage de la terre ·	39
Chapitre 6. La symphonie du ciel et la danse invisible ·	45
Chapitre 7. Du cerveau des animaux au canevas du deep learning ·	51
Chapitre 8. Le monde propre des animaux et les frontières sémiotiques ·	58
Chapitre 9. La croisée des chemins entre tromperie et protection ·	65
Chapitre 10. Personne non humaine et droits assistés par la machine ·	71
Épilogue: L'Internet interespèces et une Terre de coexistence ·	78

Introduction: Ouvrir les voix de la communauté de la Terre

0.1 Le nouveau regard ouvert par le télescope Hubble

En décembre 1995, des scientifiques ont tourné un télescope spatial vers le ciel. Ce télescope s'appelait le télescope spatial Hubble (Hubble Space Telescope). Le télescope a visé l'endroit le plus sombre et le plus vide du ciel. Beaucoup de savants pensaient qu'il n'y avait rien là-bas. Pendant plusieurs jours, le télescope a recueilli la lumière dans cette obscurité. Puis un résultat inattendu est apparu. Des milliers de galaxies étaient cachées dans le noir. Une galaxie est un immense groupe formé par des étoiles réunies. Cet endroit n'était pas vide. Il était plein. Cela a laissé une grande leçon à la science. Quand un nouvel outil éclaire un lieu, même ce qui semblait vide peut contenir beaucoup de choses [1].

Aujourd'hui, l'intelligence artificielle fait quelque chose de semblable. Cette fois, elle ne vise pas le ciel. Elle vise le monde des sons que les êtres vivants s'échangent. Ce nouvel outil s'appelle la bioacoustique numérique (Digital Bioacoustics). C'est une science qui enregistre les sons produits par les êtres vivants et les étudie avec des ordinateurs. Pendant longtemps, la science occidentale a jugé les sons des animaux comme peu importants. Elle les voyait comme de simples bruits produits par l'instinct. Mais quand l'intelligence artificielle a commencé à examiner ces sons, cette idée a changé. Les forêts et les mers étaient des scènes de dialogue pleines d'informations [2].

Les yeux et les oreilles des humains ne perçoivent qu'une toute petite partie du monde. Il est facile de croire que ce que nous ne percevons pas n'existe pas. Il existe un petit oiseau appelé junco ardoisé (dark-eyed junco). Pour l'oreille humaine, son chant paraît plat. Mais les oiseaux ont une perception du temps plus rapide que les humains dans l'oreille. Pour un oiseau, ce son est une mélodie fine, tissée en deux couches. Les éléphants parlent très loin avec des sons très graves. Ces sons se situent entre 14 et 35 hertz (Hz). Dans cette plage, les sons les plus graves échappent à l'oreille humaine. Ils se propagent sur plusieurs kilomètres. Si les conditions sont bonnes, ils vont encore plus loin. Les chauves-souris se disputent avec des sons aigus que les humains n'entendent pas. Même les jeunes coraux, qui n'ont pas d'oreilles, trouvent où aller grâce aux sons [3].

Les sons entendus par les humains ont un mur en haut et un mur en bas. Les sons trop aigus pour être entendus s'appellent ultrasons (ultrasonic). « Ultra » veut dire qui dépasse. Ce sont des sons qui dépassent le mur supérieur de l'oreille humaine. Les sons trop graves pour être entendus s'appellent infrasons (infrasonic). Ce sont des sons qui dépassent le mur inférieur de l'oreille humaine. Les sons des chauves-souris appartiennent aux ultrasons. La partie grave des sons des éléphants appartient aux infrasons. Pendant longtemps, les humains ont vécu sans entendre ces deux types de sons. Avec de meilleurs appareils d'enregistrement et de meilleures techniques d'analyse, les chercheurs ont pu étudier ces sons un par un. Les animaux qui ont des oreilles ne sont pas les seuls à entendre les sons. Des recherches montrent que même des plantes sans oreilles réagissent aux sons.

Les scientifiques captent aussi ces sons sous l'eau. Un micro que l'on place dans l'eau s'appelle un hydrophone (hydrophone). Les sons ainsi recueillis sont analysés par l'apprentissage profond (deep learning). L'apprentissage profond est une technique qui trouve elle-même des règles dans de grandes quantités de données. Ainsi, l'intelligence artificielle casse la boîte des sens humains. Une porte s'ouvre vers l'ancienne conversation de la Terre.

Sources liées

- [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 L'intelligence artificielle qui transforme les mots en carte

Comment l'intelligence artificielle traite-t-elle les mots humains ? Elle transforme les mots en une sorte de carte. Cette carte s'appelle l'espace latent (Latent Space). C'est une carte invisible qui place les mots de sens proche les uns près des autres. Il est facile de l'imaginer comme un ciel nocturne rempli d'étoiles. Chaque mot devient une étoile. Les mots aux sens proches se regroupent. Les mots aux sens éloignés se tiennent loin les uns des autres [1].

La force de cette carte ne vient pas seulement des distances. Les relations entre les mots apparaissent aussi par des directions et des distances. Par exemple, il existe une flèche qui relie « roi » et « homme ». Il existe aussi une flèche semblable qui relie « reine » et « femme ». Donc, si l'on prend « reine », que l'on enlève « femme » et que l'on ajoute « homme », on se rapproche de « roi ». Dans certaines cartes de mots, ce calcul fonctionne assez bien. Le mot « chien » n'est pas seul non plus. Il prend sa place en se reliant à des mots comme « humain », « aboiement » et « poil ».

Un ordinateur ne sent pas le sens caché des mots comme un humain. Il ne connaît pas non plus des émotions comme la tristesse ou la joie. Son travail consiste seulement à calculer des relations. Il examine comment les mots se relient les uns aux autres. Cela ressemble à un grand exercice où il faut remplir des blancs. Comme on choisit un mot pour chaque blanc, l'intelligence artificielle trouve peu à peu la place de chaque mot. Ainsi se forme la structure précise du langage [2].

Les sons peuvent aussi être transformés en carte de cette manière. L'intelligence artificielle découpe les sons en petits morceaux et les change en nombres. Elle place ces nombres comme des étoiles sur une carte. Les sons semblables se regroupent. Les sons différents s'éloignent. La méthode de cartographie est semblable pour la parole humaine et pour les sons animaux. Mais les chercheurs étudient encore si le sens peut vraiment passer d'une espèce à une autre [3].

Les mots ressemblent aux ombres dessinées par le monde. L'intelligence artificielle suit ces ombres en sens inverse. Elle trace ainsi une carte de la structure cachée du monde. Sans cette carte, l'histoire suivante ne peut pas commencer. Si l'on peut tracer une carte des mots humains, on peut aussi tracer une carte des sons animaux.

Sources liées

- [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 La traduction sans dictionnaire et l'Earth Species Project

Vers 2017, un grand événement s'est produit dans la traduction automatique. Il n'existait aucun dictionnaire reliant deux langues. Autrefois, il fallait une pierre de Rosette (Rosetta Stone). La pierre de Rosette est une ancienne stèle où le même texte est gravé en plusieurs langues. Grâce à cette stèle, les chercheurs ont pu déchiffrer l'écriture de l'Égypte ancienne. Pourtant, l'intelligence artificielle a associé les mots de deux langues sans ce type de dictionnaire. Elle y est arrivée grâce à l'apprentissage automatique, qui apprend par lui-même [1].

La méthode consistait à superposer les cartes des deux langues. L'intelligence artificielle a créé une carte d'étoiles pour l'anglais. Elle a aussi créé séparément une carte d'étoiles pour l'espagnol. Puis elle a posé les deux cartes l'une sur l'autre avec prudence. Beaucoup de savants pensaient que les deux cartes ne correspondraient pas. Chaque langue regarde le monde à sa façon. Pourtant, les deux cartes se sont superposées en douceur. L'étoile qui signifie « chien » est presque tombée au même endroit. Cette méthode, qui étire et rétrécit deux cartes de formes différentes pour les faire correspondre, s'appelle l'alignement de variétés (manifold alignment).

Ce phénomène ne s'est pas produit seulement dans les langues européennes. Des formes semblables sont apparues aussi dans des langues à grammaire très différente. Cela s'est même produit avec l'espéranto, une langue créée par des humains. Les nombreuses langues humaines semblaient partager une grande forme commune. On pense que c'est parce qu'une même espèce, l'être humain, ressent la même Terre [2].

Aza Raskin, de l'Earth Species Project, a alors posé une question. Si l'on peut superposer ainsi les langues humaines, ne pourrait-on pas employer la même méthode pour les sons des animaux ? Cette organisation est un institut de recherche à but non lucratif fondé en 2017. Elle rassemble les sons des bélugas, qui sont des baleines blanches, des orques et des orangs-outans. Elle crée une carte d'étoiles séparée pour chaque type de son. Puis elle pose ces cartes avec prudence sur la carte des mots humains. Les humains peuvent alors observer les sons des êtres vivants voisins avec un regard nouveau. Leur outil récent est NatureLM-audio. C'est une intelligence artificielle qui écoute les sons des animaux et devine l'espèce et le type de cri. En 2025, une démonstration ouverte à tous a été publiée. On téléverse un fichier sonore, on écrit une question, puis une réponse apparaît. En 2026, cette fonction s'est encore améliorée. En juillet 2026, un entrepôt de données appelé alp-data a aussi ouvert. Plusieurs laboratoires y partagent leurs données. Ce mouvement rejoint l'idée que des intelligences artificielles différentes finissent par arriver à des cartes de forme semblable. Cette idée s'appelle l'hypothèse de la représentation platonicienne (Platonic Representation Hypothesis) [3][4][5].

Sources liées

- [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 Trouver les règles cachées dans les sons

Les sons des animaux sont-ils des bruits sans sens ? Ou bien ont-ils des règles, comme les mots humains ? Les scientifiques examinent cela en mesurant les statistiques des sons. Une règle bien connue dit que les mots les plus utilisés deviennent plus courts. Cette règle s'appelle la loi d'abréviation de Zipf (Zipf's law of abbreviation). Elle porte le nom du linguiste George Kingsley Zipf. Les humains raccourcissent les mots qu'ils utilisent souvent pour économiser leurs efforts [1].

Le linguiste Stuart Semple et l'équipe de Ramon Ferrer i Cancho ont examiné cette règle chez plusieurs animaux. Ils ont rassemblé et comparé les sons de plusieurs mammifères, dont des singes. L'ordinateur a découpé les sons en petits morceaux et les a comptés. Le résultat était régulier. Plus un son était produit souvent, plus il était court. Une règle qui ressemble à celle de la parole humaine existait aussi dans les sons animaux. Une autre règle est apparue : plus un son est long, plus les morceaux qu'il contient sont courts. Cette règle s'appelle la loi de Menzerath-Altmann (Menzerath-Altmann law).

Il existe aussi une valeur qui mesure à quel point il est difficile de deviner le son suivant. Elle mesure la difficulté de savoir quel son viendra après un autre. Cette valeur s'appelle l'entropie de Shannon (Shannon entropy). Elle porte le nom de Claude Shannon, le créateur de la théorie de l'information. Si un son n'est qu'un cri sans sens, cette valeur penche fortement d'un côté. Si c'est un langage avec des règles, elle se situe entre les deux [4].

L'équipe de Whale-SETI a examiné les sons des baleines avec cette valeur. Brenda McCowan, Fred Sharpe et Laurence Doyle y ont participé. Ils ont étudié le chant de la baleine à bosse (humpback whale). Le chant des baleines se plaçait bien entre règle et liberté. Dans ce chant, les sons fréquents étaient courts et les sons rares étaient longs. Cette structure était visible. Ainsi, le même sens peut être envoyé avec moins d'effort. En informatique aussi, il existe une méthode de compression qui écrit court ce qui est fréquent et long ce qui est rare, afin de réduire la taille. Cette méthode s'appelle le codage de Huffman (Huffman coding) [2].

L'équipe a aussi mené une expérience d'échange sonore avec une baleine vivante. Dans la mer d'Alaska, elle a rencontré une baleine à bosse appelée Twain. Les chercheurs ont diffusé sous l'eau, avec un haut-parleur, un son de salutation de baleine enregistré à l'avance. L'intelligence artificielle n'a pas créé une conversation en temps réel. Ce son de salutation s'appelle un whup call. La baleine a répondu en respectant son tour, comme un humain. Elle attendait la fin du son, puis remplissait le silence. Cet échange a duré 20 minutes. C'était comme parler chacun son tour avec un ami à la cantine [3].

Sources liées

- [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 Les voyelles trouvées dans les clics du cachalot

L'endroit où cette méthode a le mieux brillé était la mer profonde. Le personnage principal était le cachalot. Le cachalot mesure jusqu'à 20 mètres de long. Son cerveau pèse environ 9 kilogrammes. C'est plus de 6 fois le poids du cerveau humain. Cette baleine chasse dans les profondeurs sans lumière. Elle voit donc le monde avec le son, à la place des yeux. Elle envoie de courts clics pour explorer ce qui l'entoure. Un groupe de ces clics s'appelle une coda [1].

Pendant longtemps, les savants ont vu les codas comme des signaux sans sens. Ils pensaient qu'il n'existait que quelques sons d'alerte fixes. En 2020, Project CETI a changé cette idée. C'est un groupe qui réunit des spécialistes de l'intelligence artificielle et des spécialistes des baleines venus de plusieurs universités. Ils ont rassemblé des enregistrements dans la mer de la Dominique. Plus de 8 700 codas avaient été recueillies pendant plus de 10 ans. Le nombre exact était 8 719. L'intelligence artificielle a analysé ces sons. Elle a trouvé beaucoup plus de combinaisons que les 21 types déjà connus. Elle a aussi découvert que les baleines changeaient la vitesse des sons et les intervalles entre eux. Elles ajoutaient parfois un clic à la fin, comme un point à la fin d'une phrase. Elles changeaient aussi très légèrement la longueur des codas. Cette petite variation de rythme s'appelle rubato (rubato). Cela ressemble à un chanteur qui ralentit ou avance un peu le rythme. C'est une manière de relier de petits morceaux de son pour fabriquer beaucoup de sens. Cette méthode existe aussi dans la parole humaine, et les chercheurs l'ont donc clairement observée chez un animal. Les cachalots employaient aussi des codas différentes selon la situation d'écoute. Cette capacité à choisir ses mots selon l'interlocuteur s'appelle la conception pour l'auditoire (audience design).

En 2025, une nouvelle étude a annoncé une découverte plus grande. C'était une étude menée par l'université de Berkeley, aux États-Unis, avec Project CETI. Elle était dirigée par le professeur Gašper Beguš. Les chercheurs ont réduit les silences entre les clics des baleines. Puis ils les ont écoutés à une vitesse plusieurs fois plus rapide que l'original. Alors, des sons séparés se sont reliés en un seul flux. Ce son ressemblait aux voyelles humaines. Les chercheurs ont distingué le son « a » et le son « i ». Ils ont aussi trouvé des diphtongues, où une voyelle glisse vers une autre. Les chercheurs ont aussi utilisé des outils pour regarder à l'intérieur de l'intelligence artificielle. Ils ont observé les parties auxquelles l'intelligence artificielle réagissait quand elle jugeait les sons des baleines. Ils ont ainsi montré que les pics de résonance du son étaient importants. On appelle ces pics de résonance des formants (formant). Ce sont des sommets sonores qui ressortent selon la forme de la bouche ou des cavités de résonance dans la tête. Les humains utilisent aussi ces valeurs pour distinguer les voyelles. En 2026, une étude de suivi a repris cette idée. Elle a examiné plus en profondeur la structure des sons des baleines en la comparant à la phonologie humaine [2][5][3].

En mars 2026, Project CETI a annoncé une autre étude. Cette étude avait enregistré une naissance de cachalot dans la mer de Dominique. Elle réunissait plus de six heures de sons et des images filmées depuis le ciel. D'autres baleines du groupe ont aidé pendant la naissance. C'était la première fois qu'une telle coopération était montrée avec des données chiffrées chez des animaux qui ne sont pas des primates. Cette étude a montré encore une fois que la société des cachalots est étroitement reliée par le son [4].

Le cachalot n'a pas de cordes vocales comme les humains. À la place, il utilise un grand organe situé à l'avant de sa tête et des tissus dans son nez. Il fait vibrer cette zone pour créer la résonance du son. Une forme de vie qui a suivi un chemin très différent de celui des humains a donc créé par elle-même des voyelles proches de celles des humains.

Sources liées

- [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 Écouter, et non dominer

La recherche en intelligence artificielle porte une vieille leçon. Le chercheur en informatique Richard Sutton l'a appelée la leçon amère (The Bitter Lesson). Cela veut dire que de grandes quantités de données et de gros calculs gagnent souvent contre une conception ingénieuse. Les outils qui déchiffrent les sons des animaux ont grandi de cette façon. L'un d'eux est DolphinGemma. Google, Georgia Tech et Wild Dolphin Project l'ont créé ensemble. Wild Dolphin Project est une recherche sur les dauphins qui dure depuis plus de 40 ans. DolphinGemma est une petite intelligence artificielle qui contient environ 400 millions de nombres internes pour juger les sons. Elle est assez petite pour fonctionner sur un téléphone Pixel tenu dans la main. DolphinGemma découpe les sons en petits morceaux avec une méthode appelée SoundStream. Elle regarde ces morceaux et devine à l'avance le son suivant. Sa méthode ressemble à celle d'une intelligence artificielle qui continue une parole humaine jusqu'au mot suivant. Aujourd'hui encore, cet outil est testé en mer en suivant les dauphins. Google veut ouvrir cet outil pour que tout le monde puisse l'utiliser [3][1].

Mais cette technologie rapide a aussi un côté inquiétant. Aza Raskin le répète souvent. L'intelligence artificielle peut imiter des sons avant même d'en comprendre le sens. Elle peut produire les sons des baleines à l'identique et continuer une conversation. Cela revient à répandre dans la mer de faux sons sans en connaître le sens. Les cachalots partagent une culture par les sons et par les comportements. Quand les anciens baleiniers sont apparus, les cachalots ont appris les uns des autres comment éviter le danger. Deux ans et demi après l'arrivée des bateaux, le nombre de baleines capturées a baissé de 58 %. On pense que c'est le résultat d'un avertissement transmis entre baleines. Leur communication s'étend plus loin que nous ne le pensons. Le chant des baleines à bosse est aussi une culture qui traverse les mers et se diffuse largement. De faux sons créés par les humains peuvent troubler ces liens. Créer des sons avant d'en comprendre tout le sens est dangereux [2][4][5].

La direction de ce voyage compte donc beaucoup. Cette technologie ne doit pas devenir un outil pour commander les animaux à notre guise. Elle doit servir à reconnaître et à écouter leur monde intérieur tel qu'il est. On appelle cela la compréhension entre espèces (Interspecies Understanding). Les humains changent non pas quand ils couvrent les autres de paroles, mais quand ils baissent la voix et accueillent l'autre.

Autrefois, les humains croyaient que la Terre était le centre de l'univers. Le télescope leur a montré que cette idée était fausse. Aujourd'hui, la carte des sons nous réveille encore une fois. Les humains ne sont pas le centre du monde vivant. Ils ne sont qu'une partie d'un grand cercle où chaque être possède sa voix. Cette prise de conscience est le premier pas pour ouvrir les voix de la Terre.

Sources liées

- [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.

Chapitre 1. Le Wild Dolphin Project et le monde des dauphins tachetés

1.1 La longue observation dans les Bahamas couleur émeraude

Dans la mer vivent des intelligences que les humains ne connaissent pas encore entièrement. Le dauphin en fait partie. Comment comprendre l'esprit d'un dauphin ? On peut enfermer un dauphin dans un bassin et l'observer. Mais une personne a choisi un autre chemin. C'est la docteure Denise Herzing. La docteure Herzing est allée elle-même dans la mer où vivent les dauphins [1].

La docteure Herzing a fixé un principe. C'est le principe « dans leur monde, à leur manière ». Cela veut dire qu'on ne fait pas agir les dauphins selon la volonté des humains. Cela veut aussi dire qu'on les observe sans leur faire de mal. Cette méthode s'appelle l'observation non invasive. Non invasive signifie qu'on ne perce pas le corps et qu'on ne capture pas l'animal.

Si on enferme des dauphins dans un bassin, il est difficile de voir leur vrai comportement. Un dauphin enfermé change sa façon de vivre. C'est pourquoi aller les observer dans la mer a tant de valeur.

La docteure Herzing dirige le Wild Dolphin Project. Cette recherche dure depuis près de 40 ans. 40 ans, c'est environ quatre fois l'âge d'un élève d'école primaire. Le lieu de recherche est la mer des Bahamas. L'eau y est claire et transparente comme l'eau d'une piscine. Elle est donc bonne pour enregistrer à la fois les gestes et les sons des dauphins [2][3].

Cette recherche a commencé en 1985. C'est la plus ancienne recherche au monde qui suit des dauphins sous l'eau. Chaque année, l'équipe est allée retrouver le même groupe. Grâce à cela, elle a pu observer toute la vie d'un dauphin, de sa naissance à sa croissance, puis à la naissance de ses petits. Ces dossiers accumulés deviennent ensuite une base pour enseigner à l'intelligence artificielle [4].

Dans cette mer vit le dauphin tacheté de l'Atlantique. Son nom scientifique est *Stenella frontalis*. Le nom scientifique est le nom officiel que les chercheurs utilisent partout dans le monde. La recherche se fait dans une mer peu profonde, sur des bancs de sable. L'eau est peu profonde, donc la lumière du soleil atteint le fond. Ainsi, on peut voir clairement les gestes des dauphins même depuis la surface.

Ce dauphin est plus petit que le grand dauphin. Mais c'est un animal intelligent. Il existe une valeur qui mesure la taille du cerveau par rapport à la taille du corps. Cette valeur s'appelle le quotient d'encéphalisation. Elle compare la taille du cerveau avec celle du corps. Les dauphins ont une valeur assez élevée. Chez les mammifères, ils font partie des plus élevés après les humains [5].

Les chercheurs ont appris à distinguer chaque dauphin. Les blessures sur leurs nageoires sont toutes différentes. Les motifs de taches sur leur corps sont aussi différents. Le dauphin tacheté gagne des taches en grandissant. Quand il est jeune, il n'a pas de taches. Puis elles deviennent plus nombreuses avec l'âge. On peut donc estimer son âge rien qu'en regardant ses taches. Les chercheurs ont mémorisé les dauphins comme on mémorise les visages des camarades de classe. Grâce à cela, ils ont pu suivre trois générations : grand-mère, mère et petit.

Une longue observation a une grande valeur pour une raison. Les habitudes des dauphins passent de la mère au petit. Le petit apprend auprès de sa mère comment trouver de la nourriture. Quand un comportement appris passe de la mère au petit, on appelle cela la transmission culturelle. Si l'on observe peu de temps, on ne peut pas voir ce passage. Il faut observer longtemps un même groupe pour enfin le voir.

La société des dauphins est aussi complexe que la société humaine. Les mâles nouent des amitiés pour la vie et se déplacent ensemble. Les femelles s'occupent longtemps de leurs petits et leur enseignent les règles de conduite. Pour former une société aussi complexe, il faut une façon de communiquer. C'est pour cela qu'il faut déchiffrer les sons des dauphins.

Sources liées

- [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/10.1159/000006540>

1.2 La mer rude et le système CHAT

Enregistrer les sons des dauphins n'est pas facile. Beaucoup de sons de dauphins sont trop aigus pour l'oreille humaine. L'oreille humaine peut entendre jusqu'à environ 20 kilohertz. Le kilohertz est une unité qui mesure la hauteur d'un son. Les dauphins produisent aussi des sons aigus proches de 150 kilohertz. C'est bien plus haut que les sons que les humains peuvent entendre [1].

Les dauphins utilisent plusieurs sortes de sons. Trois types sont représentatifs. Le premier est un sifflement qui monte et descend. Le deuxième est un clic. Le troisième est un son bref appelé burst pulse. Le burst pulse est une suite rapide de petits éclats sonores. Les dauphins l'utilisent souvent quand ils se disputent ou quand ils parlent de près. Le clic frappe un objet puis revient. Le dauphin sait ce qui se trouve devant lui grâce à cet écho. Cela s'appelle l'échololocation. L'échololocation est la capacité de connaître la position des objets avec le son.

Il y avait un problème. Mettre un seul micro dans l'eau ne servait pas assez. Parmi plusieurs dauphins, on ne savait pas lequel avait produit le son. La docteure Herzing est donc allée voir un ingénieur. C'est le docteur Thad Starner, du Georgia Tech. Le docteur Starner crée des technologies de petits ordinateurs que l'on porte sur le corps.

Le docteur Starner pensait que ce travail se terminerait vite. Mais la mer était un adversaire difficile. Il a fallu 4 ans seulement pour rendre la machine stable. Le plus grand ennemi était l'eau salée. L'eau salée entraînait par les fentes et brûlait la machine [2].

La boîte en aluminium qui contenait la machine posait aussi problème. Elle absorbait le soleil brûlant, et l'intérieur chauffait. Alors l'ordinateur s'arrêtait. L'aluminium bloquait aussi les signaux sans fil. Les chercheurs ont remplacé l'arrière de la boîte par du plastique transparent. Ils ont complètement enveloppé les câbles et les circuits dans une résine spéciale. Alors seulement, la machine a bien fonctionné sous l'eau.

Laisser longtemps la machine dans l'eau n'était pas facile non plus. Le sel ronge peu à peu le métal. Les chercheurs lavaient et examinaient la machine après chaque utilisation. Une seule petite fente pouvait rendre la machine inutilisable.

Il y avait aussi beaucoup de dangers. Si une grosse batterie portée par un plongeur touchait l'eau, elle pouvait exploser. Un faible courant qui fuyait de la machine pouvait aussi attirer les requins. Les requins ont un sens qui leur permet de sentir les faibles signaux électriques. Pendant un essai, un requin est même apparu.

Les requins ont un organe sensoriel pour sentir les faibles courants électriques. Cet organe s'appelle les ampoules de Lorenzini. Il forme de petits trous autour du museau du requin. Ainsi, le requin sent même les faibles courants électriques produits par une proie. Les chercheurs craignaient que la machine sous l'eau ne déclenche ce sens.

L'équipement né de toutes ces difficultés est le système CHAT. Le nom CHAT a un sens. En coréen, on peut le traduire par appareil de télémétrie pour augmenter l'audition des cétacés. Plus simplement, c'est un ordinateur porté sur le corps, qui envoie, reçoit et reconnaît des sons sous l'eau. Les cétacés sont le groupe qui comprend les dauphins et les baleines [3].

CHAT est un ordinateur sous-marin qui écoute et enregistre les sons des dauphins. Cette machine enregistre le son de façon très serrée. Elle découpe et reçoit le son plus de 190 000 fois par seconde. Elle écoute donc de façon bien plus serrée que l'oreille humaine. Il faut écouter ainsi pour ne pas manquer les sons aigus. Mais l'espace de stockage se remplit vite. C'est pourquoi l'équipe utilise aussi un grand système de stockage.

Le plongeur porte cette machine sur le dos et un petit clavier au poignet. Il écoute les sons avec un casque à conduction osseuse. La conduction osseuse transmet le son en faisant vibrer les os du crâne, au lieu de passer par l'oreille. Quand le plongeur appuie sur le clavier, un son convenu part dans l'eau. Ce son n'est pas un son naturel des dauphins. C'est un son créé par les humains. Les chercheurs ont choisi exprès un nouveau son pour ne pas le confondre avec les sons des dauphins. La machine d'un autre plongeur reconnaît ce son. Si le son correspond à un son fixé à l'avance, elle indique le nom de l'objet attaché à ce son.

Quand les sons sont envoyés et reçus, une voix donne l'information. Quand un plongeur envoie un son, une voix d'homme se fait entendre. Du côté qui reçoit le son, une voix de femme lit le sens. Les deux voix sont séparées pour ne pas confondre qui a parlé.

Sources liées

- [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 Un jeu de communication appris en jouant ensemble

Même après la fabrication de la machine, un problème restait. Comment apprendre les sons aux dauphins ? Les chercheurs ont décidé de ne pas utiliser l'entraînement forcé. Ils ont choisi une méthode qui laisse les dauphins apprendre par eux-mêmes [1].

La méthode choisie par la docteure Herzing porte un nom. Elle s'appelle le « modelage par compétition sociale ». Elle consiste à montrer deux humains qui se posent des questions et se répondent. La docteure Irene Pepperberg a utilisé cette méthode pour enseigner à un perroquet. Quand deux humains

semblaient se mesurer, l'animal apprenait en les observant. Voici la méthode. Deux personnes montrent un jeu devant les dauphins. Une personne tient un objet. L'autre produit le son de cet objet. La personne qui a produit le son reçoit l'objet, puis elles jouent ensemble.

Dans le vrai jeu, cela se passe ainsi. Un plongeur tient une écharpe jaune. Un autre plongeur produit le son de l'écharpe avec le clavier à son poignet. Alors le plongeur qui tient l'écharpe la donne au plongeur qui a produit le son. Tous deux tirent l'écharpe dans l'eau et jouent avec plaisir.

Il existe plusieurs jouets pour jouer avec les dauphins. Ils jouent avec une écharpe jaune, mais aussi avec des algues sargasses qui flottent dans la mer. Les sargasses sont des algues brunes qui flottent à la surface. Ils jouent aussi en tirant une longue corde. Ils aiment également surfer sur la vague qui se forme à l'avant du bateau. Chaque jouet reçoit un son différent, que l'on envoie et reçoit.

Les dauphins ont observé ce jeu à côté des humains. Ils ont compris la règle : quand un son est produit, un jouet circule. Les dauphins ont vite commencé à imiter ces sons. Ils ont copié les montées et les descentes des sifflements créés par les humains. Le fait de reproduire un son de cette manière s'appelle l'imitation sonore. Les chercheurs l'appellent imitation acoustique. Cette observation par le jeu est devenue plus tard une base pour la recherche en intelligence artificielle [2][3].

Les dauphins ont reproduit ces sons plusieurs fois par eux-mêmes. Parfois, ils les imitaient tout de suite après les avoir entendus une seule fois. Parfois, ils les imitaient un peu plus tard. Cela a montré qu'ils pouvaient copier de nouveaux sons avec souplesse. Mais il faut préciser un point. On n'a pas encore confirmé qu'ils savaient que ce son désignait un objet. L'imitation sonore a réussi. Mais les chercheurs n'ont pas obtenu de preuve certaine que les dauphins utilisaient le son comme le nom de l'objet.

Pourquoi les dauphins ont-ils produit ces sons par eux-mêmes ? Ce n'était pas seulement pour obtenir de la nourriture. C'était parce qu'ils voulaient jouer avec les humains. Quand le jeu était agréable, les dauphins s'approchaient les premiers. Le plaisir a donc guidé la communication.

Cette expérience montre un fait important. Les dauphins peuvent apprendre et reproduire de nouveaux sons. Elle montre aussi qu'il n'est pas nécessaire de leur enseigner de force les mots humains. Les dauphins se sont approchés les premiers et ont échangé des sons avec les humains.

Sources liées

- [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., Starnes, 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 La carte des sons dessinée par DolphinGemma

Les règles sonores fixées par les humains avaient des limites. Les vrais sons que les dauphins échangent dans leur groupe sont bien plus complexes. Pour déchiffrer ces sons complexes, il fallait un outil puissant. Cet outil est l'intelligence artificielle [1].

Une grande nouvelle est arrivée en 2025. Google DeepMind et le Wild Dolphin Project ont présenté ensemble une intelligence artificielle. Google DeepMind est l'organisation de recherche en intelligence artificielle de Google. Cette intelligence artificielle s'appelle DolphinGemma. DolphinGemma repose sur

une petite intelligence artificielle de Google appelée Gemma. Les chercheurs l'ont entraînée avec des sons et des vidéos de dauphins recueillis pendant 40 ans [2].

L'ancien équipement était un lourd ordinateur porté comme un sac à dos. Le plongeur entraînait dans la mer avec cette lourde machine sur le dos. Aujourd'hui, un smartphone de la taille d'une paume fait ce travail. L'équipement est donc devenu beaucoup plus petit. DolphinGemma a été rendue petite pour une raison. Cette intelligence artificielle utilise environ 400 millions de valeurs de réglage quand elle apprend. Ces valeurs de réglage s'appellent des paramètres. Pour une grande intelligence artificielle, c'est plutôt petit. Elle peut donc fonctionner sur un smartphone tenu dans la main. Les chercheurs ont installé cette intelligence artificielle sur un téléphone Google Pixel. Ils peuvent ainsi l'utiliser directement en mer, sans laisser un lourd ordinateur sur le bateau [4].

Le nom DolphinGemma a aussi un sens. Il ajoute Gemma à Dolphin, qui signifie dauphin. Cela veut dire qu'elle a été créée en posant des sons de dauphins sur l'intelligence artificielle Gemma.

Le travail le plus difficile était de transformer les sons en nombres. Les sons sous l'eau coulent sans interruption. L'intelligence artificielle préfère les données faites de morceaux séparés. Les chercheurs ont donc découpé les sons en petits morceaux. C'est comme diviser une écriture en consonnes et en voyelles. Ces morceaux de sons s'appellent des jetons acoustiques. Cela signifie des morceaux obtenus en divisant finement le son.

Il existe aussi un outil utilisé pour diviser les sons. C'est SoundStream, une technologie de Google qui découpe les sons. Cette technologie divise les sons sous l'eau en petits morceaux. DolphinGemma apprend ensuite en reliant ces morceaux.

Avec ces morceaux de sons, DolphinGemma s'est entraînée à deviner le son suivant. Elle regarde le son précédent et essaie de prévoir le son qui vient après. Ainsi, DolphinGemma est devenue capable de bien deviner le son suivant. C'est comme une personne qui connaît les règles d'une phrase et peut prévoir les mots suivants.

Deviner le son suivant signifie connaître les règles. L'intelligence artificielle a trouvé ce type de règles dans les sons des dauphins.

DolphinGemma regroupe aussi tout seul les sons qui se ressemblent. L'intelligence artificielle transforme les sons en points et les place comme sur une carte. Les sons proches sont placés près les uns des autres. Les sons différents sont placés loin les uns des autres. Les chercheurs ont recueilli pendant longtemps des sons venus de nombreuses situations. Par exemple, il y a les sifflements-signatures par lesquels une mère et son petit s'appellent. Le sifflement-signature est un son différent pour chaque dauphin. Il ressemble à un nom. Il y a aussi les salves pulsées émises pendant les combats. Il y a aussi les clics émis pour chercher un partenaire ou poursuivre un requin. Associer ces sons à leurs situations a un avantage. Cela aide l'intelligence artificielle à trouver plus vite les indices cachés dans les sons.

La docteure Hejing a expliqué ce travail ainsi. Elle a dit qu'il faudrait 150 ans à une personne pour examiner toutes ces données à la main. Elle a ajouté que l'intelligence artificielle accomplit ce travail beaucoup plus vite. 150 ans, c'est plus long que deux vies humaines [3].

Sources liées

[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 Une vraie conversation née dans la mer

À l'été 2025, une nouvelle scène s'est ouverte. Le Wild Dolphin Project a commencé à utiliser DolphinGemma sur le terrain, en mer. Jusque-là, il avait appris seulement dans un ordinateur. Maintenant, un plongeur emporte cette intelligence artificielle sous l'eau. Les chercheurs ont utilisé un nouveau téléphone Google Pixel [1][9].

La méthode utilisée en mer est la suivante. Un plongeur emporte un téléphone Pixel placé dans un boîtier étanche. L'intelligence artificielle examine tout de suite les sons sous l'eau. Puis elle trouve les motifs sonores répétés et les signale aux chercheurs. Il n'est plus nécessaire de revenir au bateau pour examiner les données. La vérification se fait directement sous l'eau.

Cette étude en mer a un point particulier. DolphinGemma a créé tout seul de nouveaux sons qui ressemblent à ceux des dauphins. Ces sons étaient difficiles à produire avec les programmes utilisés par les humains. Les chercheurs prévoient de s'en servir dans de futures études d'échanges avec les dauphins. Ils veulent essayer de parler aux dauphins avec ces nouveaux sons [2].

Cette technologie est utilisée avec prudence. Si les humains émettent des sons n'importe comment, la société des dauphins peut être perturbée. Depuis très longtemps, les groupes de dauphins sont reliés par les sons. Les chercheurs essaient donc de ne pas déranger les sons naturels des dauphins. Les sons enseignés aux dauphins ont aussi été choisis parmi de nouveaux sifflements qui n'existent pas dans la nature. Le but est d'éviter toute confusion avec les sons déjà présents. Les chercheurs essaient aussi de ne pas interrompre le rythme des échanges sonores des dauphins. Ils observent les moments calmes où les dauphins cessent de parler. C'est une forme de politesse, comme attendre son tour dans une file pour recevoir de la nourriture.

Les dauphins ne sont pas les seuls à faire l'objet de ce type de recherche. Il existe aussi une organisation appelée Earth Species Project. Cette organisation examine avec l'intelligence artificielle les sons de plusieurs sortes d'animaux. Elle étudie des animaux différents, comme les oiseaux, les baleines et les éléphants. C'est parce que chaque animal produit des sons pour des raisons différentes. Quand on compare plusieurs animaux, les règles des sons apparaissent plus clairement. En 2026, elle a aussi ouvert un nouveau dépôt pour partager des données de sons animaux. Son nom est alp-data. Ce nom vient des premières lettres anglaises qui désignent le traitement du langage animal. Il a été créé pour que plusieurs chercheurs puissent partager leurs données [3][4].

Le partage des données a un avantage. Des chercheurs de plusieurs pays réunissent les sons animaux en un seul endroit. Plus les données sont nombreuses, mieux l'intelligence artificielle apprend. Ce dépôt devient une base commune pour plusieurs chercheurs qui étudient les sons des animaux.

Une seule intelligence artificielle peut aussi traiter plusieurs animaux à la fois. C'est parce que la manière de traiter les sons se ressemble d'un animal à l'autre. Une méthode apprise chez un animal est utilisée chez un autre. Ainsi, plusieurs recherches s'aident entre elles.

Les recherches pour comprendre les cachalots avancent aussi. Une organisation appelée Project CETI mène ce travail. Le cachalot produit plusieurs clics groupés. Ce groupe s'appelle une coda. Les chercheurs ont trouvé des règles sonores à l'intérieur des codas. Ils appellent ces règles l'alphabet

sonore des cachalots. Le mot alphabet est une comparaison. Il ne s'agit pas de lettres humaines. Il compare les règles d'assemblage des clics à un alphabet [5].

En 2026, un nouveau fait a été découvert. Des motifs ressemblant aux voyelles de la parole humaine ont été trouvés dans les codas des cachalots. Dans la parole humaine, il existe des voyelles comme a et i. Les codas des cachalots se divisaient aussi en deux groupes selon leurs propriétés sonores. La longueur des codas et l'ordre des sons suivaient aussi des règles. Les cachalots ont donc leur propre cadre sonore [6][7].

Cette recherche continuera avec prudence. Comprendre les paroles des animaux ne veut pas dire tout savoir d'eux. Les humains et les animaux vivent dans des mondes différents. Mais écouter reste un travail important. Quand on se connaît mieux, il devient plus facile de vivre ensemble [8].

Toutes ces recherches indiquent une même direction. Il ne s'agit pas de forcer les animaux à apprendre la langue des humains. Il s'agit d'écouter les sons des animaux tels qu'ils sont. Certaines personnes disent que cette recherche peut aussi aider les droits des animaux. Si l'on comprend ce que pensent les animaux, on peut mieux les protéger. Cette conversation née dans la mer devient un nouveau pont entre les humains et la nature.

Sources liées

- [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>

Chapitre 2. L'alphabet phonémique du cachalot et Project CETI

2.1 Celui qui déchiffre les codes rencontre le géant des grands fonds

L'animal qui possède le plus grand cerveau sur Terre n'est pas l'être humain. C'est le cachalot. Le nom scientifique du cachalot est *Physeter macrocephalus*. C'est une très grande baleine qui vit dans les eaux profondes. Son corps peut atteindre 20 mètres de long. C'est la longueur de deux grands bus mis bout à bout. Son poids peut atteindre 50 tonnes. Cela ressemble au poids de dix éléphants réunis [1][2].

Le cerveau de cette baleine pèse environ 8 à 9 kilogrammes. Le cerveau humain pèse environ 1,4 kilogramme. Le cerveau du cachalot est donc environ 6 fois plus lourd que celui de l'être humain. Ce grand cerveau sert à vivre en groupe et à gérer les émotions. Le cachalot passe 80 pour cent de sa vie dans la mer profonde et sombre. La lumière n'arrive pas jusque-là. Le cachalot voit donc le monde avec le son, au lieu de le voir avec les yeux.

Le cachalot envoie des sons et écoute l'écho qui revient. Le son frappe un objet et revient sous forme d'écho. En écoutant cet écho, il connaît la place et la distance de l'objet. Cette méthode s'appelle l'écholocalisation (echolocation). Cela veut dire regarder autour de soi avec le son. Les clics utilisés à ce moment servent à trouver son chemin et à évaluer les proies. Mais le cachalot utilise d'autres sons quand il dialogue. Il regroupe plusieurs clics courts selon un rythme fixé. Ce groupe de sons s'appelle une coda (coda). Une coda est un petit morceau sonore que les baleines utilisent pour se parler [3][5].

Pendant longtemps, les chercheurs ont peu prêté attention aux codas. Ils pensaient que ce n'étaient que de courts signaux. Mais en 2017, une nouvelle idée a germé au Radcliffe Institute de l'université Harvard, aux États-Unis. Le Dr David Gruber, qui étudiait les êtres vivants de la mer, avait enregistré des sons de baleines. Une personne a entendu ces sons par hasard. C'était la professeure Shafi Goldwasser, chercheuse en cryptographie au MIT, aux États-Unis. Un code secret est un signal avec des règles, créé pour cacher un secret. La professeure Goldwasser a senti des règles dans les sons des baleines. Ils ressemblaient à des codes créés par les humains. Elle a donc proposé d'utiliser l'ordinateur pour comprendre ces règles [4].

Cette idée a rejoint le travail d'un chercheur qui observait les baleines depuis longtemps. C'était le Dr Shane Gero. Le Dr Gero a suivi des familles de baleines pendant plus de 13 ans au large de la Dominique, dans la mer des Caraïbes. Puis, en 2020, Project CETI (Cetacean Translation Initiative) a commencé. En coréen, cela veut dire projet de traduction du langage des baleines. Des chercheurs de plusieurs pays se sont réunis. Ils ont installé une base de recherche dans la mer de la Dominique. Ils ont placé des microphones sous-marins jusqu'à 1 000 mètres de profondeur. Ils ont aussi fixé sur le corps des baleines de petits appareils qui mesurent le son et la position. Les ingénieurs de Project CETI ont fabriqué ces appareils [6][7][8].

Ce petit appareil est spécial. Il contient trois microphones. Comme il y a plusieurs microphones, il peut séparer les sons même dans la mer bruyante. Même si plusieurs baleines émettent des sons ensemble, il peut choisir le son d'une seule baleine. Cet appareil n'enregistre pas seulement le son. Il mesure aussi la profondeur où se trouve la baleine et ses mouvements. Il note aussi la température de l'eau et la position. On peut donc savoir quel son a été produit, quand et où. Les chercheurs ont analysé par ordinateur les codas ainsi recueillies. Ils ont analysé 8 719 codas. Ces groupes de sons sont devenus la clé pour comprendre le langage des baleines [9].

Sources liées

- [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-movin-g-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 Un son qui dépasse le feu de circulation : les quatre éléments de la coda

Les anciens manuels voyaient les sons des cachalots de façon trop étroite. Ils disaient que les baleines des Caraïbes utilisaient seulement 21 sortes de codas. Ils pensaient qu'une coda transmettait un seul sens. Ils les voyaient comme des feux de circulation. Rouge veut dire s'arrêter. Vert veut dire avancer. Mais quand l'intelligence artificielle de Project CETI a étudié les sons, cette idée s'est effondrée [1].

L'intelligence artificielle est un programme informatique créé par l'être humain. Elle trouve seule des règles cachées dans beaucoup de données. Cette intelligence artificielle a trouvé 156 sortes de codas dans les dialogues des baleines. C'est plus de 7 fois plus que les 21 sortes connues avant. En plus, plusieurs petits morceaux étaient cachés dans une seule coda. Les baleines changeaient quatre éléments pour produire des sons différents.

Le premier élément est le tempo (tempo). Il indique à quelle vitesse les clics sont produits. Les êtres humains parlent aussi plus vite quand ils sont pressés. Les baleines changent aussi la vitesse de leurs sons selon la situation. Le deuxième élément est le rythme (rhythm). Il indique l'intervalle entre les clics. Le signal change selon que les sons sont serrés ou espacés. Le troisième élément est l'ornementation (ornamentation). Il s'agit d'ajouter un clic à la fin d'un groupe de sons. Cela change un peu l'impression, comme un point d'interrogation à la fin d'une phrase. Le quatrième élément est le rubato (rubato). Comme en musique, quand on avance ou retarde un peu le temps, il change très finement la longueur du son. Les baleines règlent cette longueur selon le cours du dialogue [3][4].

Ces quatre éléments se combinent et créent 156 codas. Ces codas se suivent dans le temps et créent des signaux plus longs. Un petit morceau sonore n'a pas de sens tout seul. Mais si on assemble ces morceaux de plusieurs façons, de nombreux signaux apparaissent. Le langage humain fonctionne aussi ainsi. Les consonnes et les voyelles sont des sons sans sens. Elles se rassemblent pour former des mots et des phrases. Les chercheurs ont comparé ces petits morceaux sonores des codas aux consonnes et aux voyelles humaines. Ils les ont appelés alphabet phonémique. En linguistique, cette caractéristique s'appelle la dualité de structuration (duality of patterning). Jusqu'ici, les êtres humains croyaient que cette caractéristique existait seulement dans le langage humain. Mais on a découvert qu'une structure semblable pouvait aussi exister dans le langage des baleines [2].

Les baleines semblent aussi changer leurs sons selon la personne à qui elles parlent. Le son n'est pas le même pour un petit et pour un compagnon. Les chercheurs pensent que les baleines choisissent leurs sons en tenant compte de celle qui écoute. Ce n'est pas encore prouvé avec certitude. Les baleines attendent aussi que le son de l'autre finisse. Puis, quand un espace apparaît, elles ajoutent leur son. C'est comme attendre qu'un ami ait fini de parler en classe avant de parler. Cet échange de sons à tour de rôle s'appelle l'alternance vocale (turn-taking) [6].

Chaque groupe de baleines utilise des codas un peu différentes. C'est comme les façons de parler des humains, qui changent selon les régions. Les chercheurs appellent ces groupes des clans (clan). Un clan est un grand groupe qui partage les mêmes codas. Les baleines d'un même clan produisent des sons qui se ressemblent. Les petits apprennent les codas du groupe en grandissant. C'est comme un enfant qui apprend la langue de ses parents. Les codas des baleines ne sont donc pas innées. Elles s'apprennent [5].

Sources liées

- [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 Les voyelles trouvées dans les grands fonds : la découverte du professeur Beguš

Pendant longtemps, les recherches sur les sons des baleines ont mesuré une seule chose. C'était l'intervalle de temps entre les clics. Elles ne pouvaient pas étudier la hauteur des sons. Une personne a dépassé ce mur. C'est le professeur Gašper Beguš, de l'UC Berkeley, aux États-Unis. Le professeur Beguš a porté son attention sur la hauteur des sons. Un son est fait d'un mélange de plusieurs hauteurs. Ce mélange s'appelle le spectre de fréquences (spectrum). Quand la lumière du soleil passe dans du verre, elle se déploie en couleurs d'arc-en-ciel. Le son est aussi déployé ainsi selon ses hauteurs [1][3].

Le professeur Beguš a observé l'intelligence artificielle quand elle imitait les sons des baleines. L'intelligence artificielle traitait comme importantes non seulement les durées entre les clics, mais aussi les hauteurs des sons. Le professeur a donc ralenti les clics des baleines de plusieurs dizaines de fois. Il les a changés pour les adapter à la vitesse de l'oreille humaine. Alors, quelque chose de nouveau est apparu dans les clics. C'était un son qui ressemblait aux voyelles humaines [2][4].

Quand l'être humain produit une voyelle, certaines hauteurs résonnent fortement. Ces hauteurs qui résonnent fortement s'appellent des formants (formant). Cela ressemble au son grave qui apparaît quand on souffle dans une bouteille vide. Le professeur Beguš a séparé deux voyelles dans les sons des baleines. L'une est la a-coda (a-coda). Elle montre une structure de hauteurs semblable à la voyelle a des

humains. L'autre est la i-coda (i-coda). Elle montre une structure de hauteurs semblable à la voyelle i des humains.

Les baleines ne mélangeaient pas les deux voyelles de façon floue. Elles distinguaient clairement a et i. Il y avait aussi un son qui glissait d'une voyelle vers l'autre. Chez l'être humain, ce serait une diphtongue. C'est un son où deux sons se suivent, comme ai. Pour l'oreille humaine, cela ressemblait seulement au bruit d'un morceau de bois frappé. Mais quand on a ralenti le son, cela s'est changé en dialogue où a et i se répondaient.

Une intelligence artificielle spéciale se trouvait derrière cette découverte. C'était une intelligence artificielle dont les humains pouvaient examiner les raisons de ses réponses. Cette intelligence artificielle traitait les hauteurs des sons des baleines. Une partie à l'intérieur réagissait de façon très forte à certains sons seulement. C'est comme un interrupteur qui s'allume seulement pour un son précis. Il y avait une chose que les humains avaient longtemps laissée passer. Des hauteurs semblables à des voyelles étaient cachées dans les sons des baleines. L'intelligence artificielle l'a remarquée en premier.

L'intelligence artificielle a appris en imitant elle-même les sons des baleines. Elle a appris seule, même sans réponse donnée par les humains. Cette façon d'apprendre seule s'appelle l'apprentissage non supervisé. En produisant des sons, l'intelligence artificielle a choisi elle-même quelles parties étaient importantes. Ce qu'elle a choisi, c'était justement la hauteur des sons. En avril 2026, cette recherche a été publiée dans une revue de la Royal Society, au Royaume-Uni. Elle expliquait que la a-coda et la i-coda suivent des règles, comme les voyelles humaines.

Les voyelles des baleines avaient encore d'autres points communs avec le langage humain. Regardons le clic ajouté à la fin d'un groupe de sons. Ce son devenait semblable à la voyelle qui venait à côté de lui. Quand les humains parlent aussi, les sons voisins se colorent entre eux. Comme on l'apprend en cours de coréen, gukmul se prononce gungmul. C'est parce que les sons placés avant et après deviennent semblables. La même chose se produisait dans les sons des baleines.

Sources liées

- [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 Coda. *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 Le corps qui crée des voyelles dans les conduits du nez

Les êtres humains et les cachalots ont suivi des chemins différents. Ils ont évolué séparément pendant très longtemps. Pourtant, tous deux produisent des voyelles comme a et i. Ce phénomène s'appelle l'évolution convergente (convergent evolution). Cela veut dire que des êtres vivants différents trouvent la même réponse devant le même problème. C'est comme les oiseaux et les chauves-souris. Ils ont évolué séparément, mais tous deux ont des ailes. Les lois physiques du lieu où ils vivent sont les mêmes. Alors les réponses se ressemblent aussi [1][2].

Les êtres humains utilisent les cordes vocales de la gorge pour produire des voyelles. Les cordes vocales sont des plis dans la gorge qui créent le premier son. Pour produire des voyelles, les êtres humains ont fait descendre le larynx. Le larynx est une caisse de résonance au milieu de la gorge. Les cordes vocales

se trouvent à l'intérieur. Les êtres humains ont aussi évolué pour bouger finement les lèvres et la langue. Mais le cachalot ne produit pas de sons en faisant vibrer les cordes vocales de la gorge comme l'être humain. À la place, la baleine a développé son propre organe sonore dans les conduits du nez, à l'intérieur de sa tête. Elle devait produire des sons tout en supportant la forte pression de la grande profondeur.

L'organe sonore du cachalot réunit deux parties. La première est les lèvres phoniques (phonic lips). Ce sont des structures plissées à l'intérieur des conduits du nez. Elles font le même travail que les cordes vocales humaines. Quand l'air passe, elles vibrent et créent le premier son. Cela ressemble au bruit aigu qui sort quand on pince l'ouverture d'un ballon et qu'on laisse l'air sortir. La deuxième partie est faite de poches d'huile et de poches d'air. Ces organes remplissent le grand front de la baleine. Le premier son passe dans ces poches et rebondit plusieurs fois. Pendant ce passage, certaines hauteurs résonnent fortement. C'est comme la bouche humaine qui façonne le son [3].

Ces deux parties créent le son ensemble. L'endroit où le son naît d'abord s'appelle la source (source). Le passage qui façonne le son s'appelle le filtre (filter). Quand on pince une corde de guitare, l'endroit qui produit le son est la source. La caisse qui fait joliment résonner ce son est le filtre. Quand les deux se rencontrent, la texture du son se décide. Ce principe s'appelle la théorie source-filtre (source-filter theory). Le cachalot a réalisé ce principe par lui-même, sans utiliser de cordes vocales [4].

Cet organe sonore a encore un autre visage caché. C'est l'organe même qui produit les voyelles douces de la baleine. Mais le même organe produit aussi des sons très forts. Leur intensité est mesurée à environ 230 décibels. Cela fait partie des sons animaux les plus forts mesurés sous l'eau. Mais les décibels sous l'eau et les décibels dans l'air n'ont pas la même référence. Il ne faut donc pas comparer directement ces deux valeurs. L'idée que ce son pourrait blesser un être humain n'a pas encore été confirmée. La baleine ne dirige pas cette grande force contre les autres baleines. Elle les entoure plutôt d'une voix douce. C'est ainsi qu'elle a gardé l'amitié de la mer pendant très longtemps [5].

Sources liées

- [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.
- [5] Møhl, B., Wahlberg, M., Madsen, P. T., Heerfordt, A., & Lund, A. (2003). The monopulsed nature of sperm whale clicks. *Journal of the Acoustical Society of America*, 114(2), 1143-1154. <https://doi.org/10.1121/1.1586258>

2.5 Une bibliothèque marine de 30 millions d'années et les droits des baleines

Les recherches de Project CETI et du professeur Beguš nous laissent une grande question. Les cachalots ne sont pas un groupe qui crie au hasard par instinct. Ce sont des êtres reliés par les sons dans de nombreuses zones de la mer. Ils ont chacun leur culture et leurs liens familiaux [1].

Une étude montrant ce lien est sortie en mars 2026. Project CETI y a décrit en détail la naissance d'un petit cachalot. D'autres femelles étaient présentes au moment de la naissance. Elles sont restées près de la mère et l'ont aidée. Même des baleines sans lien de parenté ont participé. C'est un cas rare où l'aide à la naissance, chez un animal qui n'est pas un primate, a été mesurée avec des chiffres. À ce moment-là,

les sons des baleines ont aussi changé. À chaque moment important, la texture de leur voix changeait. On y entendait même des sons qui ressemblaient à des voyelles. Cette étude montre encore une fois que les baleines sont des êtres qui s'entraident [2].

Les baleines des Caraïbes utilisent souvent un son précis. Les chercheurs l'appellent la coda 1 plus 1 plus 3. Elle se compose de deux clics, d'une courte pause, puis de trois clics rapides. Les chercheurs essaient d'en comprendre le sens avec prudence. Ils pensent que cela peut vouloir dire : je suis ici. Ce son est un dialecte culturel partagé par le groupe des Caraïbes. Chaque groupe local utilise un dialecte différent. Les lignées de cachalots, elles, vivent depuis des dizaines de millions d'années en s'adaptant à la mer. Pendant très longtemps, dans la mer sombre, elles se sont reconnues par les sons [3].

Les sons des baleines étaient aussi un réseau pour signaler le danger. Dans les années 1800, des humains portaient en bateau pour chasser les baleines. À ce moment-là, certaines baleines survivantes ont transmis l'information. Elles ont indiqué aux autres, par des sons, comment éviter les bateaux. Résultat : le taux de réussite de la chasse à la baleine a beaucoup baissé. En deux ans et demi, il a chuté d'environ 58 pour cent. Ce récit montre que les sons des baleines forment un vaste réseau d'information [4].

Les paroles des baleines révélées par l'intelligence artificielle posent une nouvelle question. Comment les humains doivent-ils traiter les baleines comme des êtres qui vivent avec nous ? C'est pourquoi un mouvement grandit pour donner aux baleines une personnalité juridique. La personnalité juridique est un statut que la loi protège, comme pour les humains. Cela veut dire qu'on les voit non comme des choses, mais comme des êtres qui peuvent avoir des droits. En mars 2024, des dirigeants autochtones de Nouvelle-Zélande et d'États insulaires du Pacifique ont publié une déclaration. Elle reconnaît les baleines comme des personnes juridiques. Son nom est He Whakaputanga Moana. Mais cette déclaration n'est pas une règle imposée par la loi. Ce mouvement a travaillé avec Project CETI. Le programme MOTH de la faculté de droit de New York University y a aussi participé. MOTH est un nom qui désigne la vie au-delà de l'humain [5][6].

Pendant ce temps, le bruit produit par les humains bloque les sons des baleines. Il s'agit du bruit des hélices de bateaux et des sons liés à l'exploration de la mer. Ces bruits couvrent les sons que les baleines échangent entre elles. Alors les jeunes baleines ont plus de mal à apprendre la langue du groupe. La culture ne peut plus bien passer à la génération suivante. Earth Species Project aide aussi ce type de recherche. Ce groupe a créé une intelligence artificielle qui traite en même temps les sons de plusieurs animaux. Son nom est NatureLM-audio [7][8].

Ce que nous devons faire est clair. Nous ne devons pas envoyer aux baleines des sons fabriqués sans soin, au risque de troubler leur culture. Nous devons plutôt écouter les sons qu'elles ont accumulés dans la mer pendant très longtemps. C'est une bibliothèque de la mer, transmise depuis des dizaines de millions d'années. La première chose à faire est d'écouter doucement son rythme. C'est ainsi que les humains et les baleines peuvent vivre côte à côte.

Sources liées

[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>

Chapitre 3. Rencontre avec la baleine à bosse Twain

3.1 Un entraînement terrestre pour une intelligence extraterrestre

Imaginons qu'un signal arrive de très loin dans l'espace. Ce signal est-il juste du bruit, ou bien des mots envoyés par quelqu'un ? Le bruit est un son confus qui ne porte aucun sens. Nous n'avons pas de dictionnaire pour comprendre ces mots. C'est comme si nous n'avions aucune clé pour en connaître le sens.

Pour répondre à cette question, des personnes ont regardé non pas vers le ciel, mais vers la mer. Il existe un institut qui cherche des signaux intelligents venus de l'espace. Son nom est SETI Institute. Cet institut a travaillé avec des chercheurs de l'université de Californie à Davis (UC Davis), aux États-Unis. Des spécialistes qui étudient depuis longtemps les sons des baleines les ont rejoints. Ensemble, ils ont créé une équipe de recherche appelée Whale-SETI [1].

La personne qui a dirigé l'étude est la professeure Brenda McCowan. Le docteur Laurance Doyle, du SETI Institute, a aussi participé. Le docteur Fred Sharpe, spécialiste des baleines à bosse, a rejoint l'équipe. Ils partageaient une idée. Avant de rencontrer des êtres inconnus dans l'univers, il fallait apprendre à connaître nos voisins de la Terre.

Ce voisin est la baleine à bosse (*Megaptera novaeangliae*), une sorte de baleine à fanons. La baleine à bosse est un grand animal marin qui chante de longs chants. Ce chant est surtout produit par les mâles. Les mâles d'un même groupe apprennent leurs chants les uns des autres. Ainsi, le chant se transmet et change peu à peu. Les baleines à bosse traversent aussi de vastes mers chaque année, et elles restent en contact par le son [2].

Le plan des chercheurs était le suivant. Ils mettent un petit bateau à l'eau au large de l'Alaska. Ils coupent entièrement le moteur du bateau. Ils descendent un micro et un haut-parleur dans l'eau. Puis ils font entendre une fois un son de baleine. Il existe une méthode qui consiste à rejouer un son enregistré pour voir comment l'animal réagit. Cette méthode s'appelle une expérience de playback [3].

Le son diffusé était un « appel de contact ». C'est un son qu'une baleine produit pour indiquer sa position dans le groupe. On appelle ce son whup call. Ce n'est pas un son d'alerte. C'est un son calme. Si on le traduit en mots humains, il ressemble à « je suis ici ».

Ce son était la vraie voix d'une baleine enregistrée la veille dans cette mer. Dans ce groupe se trouvait une femelle baleine à bosse âgée de 38 ans. Elle s'appelait Twain. Les chercheurs observaient Twain depuis longtemps.

Pourquoi les chercheurs ont-ils choisi les baleines ? Les baleines vivent longtemps, comme les humains. Elles vivent en groupe. Elles communiquent entre elles par le son. Elles transmettent aussi des chants appris à d'autres baleines. Ces points ressemblent aux humains. Les chercheurs ont donc pensé qu'elles convenaient pour s'entraîner à comprendre une intelligence inconnue.

Comprendre un signal venu de l'espace pose une grande difficulté. Celui qui envoie le signal et celui qui le reçoit n'ont pas fixé de règles à l'avance. Quand on a un livre qui place deux langues côte à côte, la traduction est plus facile. Mais pour un signal extraterrestre, ce livre n'existe pas. Il faut donc trouver les règles depuis le début.

Sur le bateau, juste avant l'expérience, tout était calme. Les chercheurs retenaient leur souffle et regardaient la surface de l'eau. Personne ne savait si Twain allait répondre.

Sources liées

- [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 Vingt minutes avec Twain

Le premier son est sorti du haut-parleur. Peu après, Twain est apparue. Twain s'est approchée doucement du bateau. Elle a fait un grand tour autour du bateau. Puis elle a produit son propre son vers le haut-parleur.

Twain n'a pas produit un son avant de disparaître. Quand les humains diffusaient un son, Twain répondait par un son. Quand les humains s'arrêtaient, Twain attendait aussi. Quand le son reprenait, elle répondait à nouveau. Cet échange a duré 20 minutes. 20 minutes, c'est à peu près la durée d'un court dessin animé.

Pendant 20 minutes, les chercheurs ont diffusé 38 whup calls. Twain a produit 36 whup calls. Parmi eux, 33 étaient des réponses qui correspondaient tout de suite au son diffusé par les humains. Les 3 autres sons ont été produits après l'arrêt des sons humains. C'était la première fois qu'un humain et une baleine sauvage échangeaient des sons aussi longtemps.

Twain a même ajusté l'intervalle entre les sons. Quand les humains laissaient un long espace entre deux sons, Twain laissait aussi un long espace. Quand l'espace était court, Twain répondait aussi vite. Elle suivait donc la manière de l'autre.

La docteure McCowan, qui a dirigé l'étude, l'a expliqué ainsi. Selon elle, cet événement semble être un échange sonore important entre humains et baleine. Mais les humains n'ont pas traduit le sens des mots de la baleine. Les chercheurs ont compris cela comme un échange de sons où chacun respectait son tour.

Twain n'a pas produit des sons joyeux pendant toute la durée des 20 minutes. Au début, elle participait bien à l'échange. Mais quand les chercheurs ont répété le même whup call encore et encore, sa réaction a changé. Twain semblait devenir de plus en plus nerveuse. Elle respirait souvent et frappait aussi la surface de l'eau. Puis elle a fini par quitter l'endroit. Elle a donc montré trois étapes : participation, nervosité, départ.

Cette rencontre a une limite claire. Le sujet de l'expérience était une seule baleine, Twain. L'expérience a aussi été une seule occasion venue par hasard. On ne peut donc pas affirmer, avec ce seul résultat, que les baleines conversent. Les chercheurs l'ont dit avec prudence.

Pourtant, cette rencontre a montré la force de l'expérience de playback. On fait entendre un son enregistré, puis on observe la réponse. Grâce à cette méthode, on peut voir de près les réactions des animaux sauvages. Les chercheurs l'ont appelée un nouvel outil de recherche.

Sources liées

- [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 Une conversation qui respecte les tours

Quand les humains se parlent, ils suivent des règles. Ils attendent que l'autre ait fini de parler. Quand l'autre s'arrête, ils commencent à parler. Cette règle, où chacun parle à son tour, s'appelle le tour de parole vocal (turn-taking). Cela ressemble à une file d'attente où chacun garde sa place pour recevoir son repas.

Il existe un moment où l'autre finit de parler et où le silence dure un court instant. Ce petit intervalle s'appelle une pause conversationnelle (conversational pause). Cette pause permet aux paroles de ne pas se chevaucher. Les baleines à bosse suivent aussi cette règle dans la nature.

Pendant les 20 minutes d'échange, Twain n'a pas enfreint cette règle. Elle n'a pas coupé la parole pendant que le son humain était diffusé. Elle a attendu que le son soit complètement terminé. Elle a choisi le moment où le silence arrivait. Puis elle a placé son propre son dans cet espace vide.

Les chercheurs ont mesuré le temps de réponse de Twain. Ils ont regardé combien de secondes passaient après la fin d'un son avant sa réponse. Ils n'ont pas utilisé l'intelligence artificielle pour cela. Ils ont étudié un par un les enregistrements et les statistiques. Ce temps correspondait bien aux règles de la conversation humaine.

Ce résultat nous apprend quelque chose. Twain n'a pas crié parce qu'elle était irritée par un son de machine. Twain distinguait le tour de l'autre et son propre tour avant de produire un son. Les chercheurs ont vu cela comme une participation de Twain à l'échange sonore.

Les humains apprennent cette règle dès le bébé. Ils échangent des sons en regardant leur mère dans les yeux. Même avant d'apprendre à parler, ils apprennent d'abord cet ordre. C'est une règle qui sert de base à la conversation.

La pause conversationnelle est un moment très court. Elle est parfois plus courte qu'un clignement d'œil. Twain n'a pas manqué ce bref intervalle. Elle a placé son son au moment exact.

Une telle précision est difficile sans entraînement. Dans la nature, Twain échangeait toujours des sons de cette manière. Elle a utilisé cette méthode familière même avec les sons venus des humains. Même si l'autre n'était pas une vraie baleine, elle a respecté la règle.

Sources liées

- [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 La grammaire des baleines révélée par les mathématiques

Le docteur Doyle a essayé d'expliquer les chants des baleines avec des nombres. Il pensait que les chants des baleines n'étaient pas de simples sons sans ordre. Il a donc utilisé un outil appelé théorie de l'information (information theory). La théorie de l'information est une méthode qui mesure, avec des nombres, combien de règles se trouvent dans des sons ou dans des lettres.

Les chercheurs ont appliqué la loi de Zipf (Zipf's law) aux chants des baleines. La loi de Zipf montre une règle cachée dans le langage. Dans le langage humain, les mots très fréquents et les mots rares suivent une proportion régulière. Par exemple, en coréen, « eun, neun, i, ga » apparaissent très souvent. Les mots difficiles n'apparaissent que de temps en temps. Entre ce qui apparaît souvent et ce qui apparaît peu, il existe presque toujours une proportion semblable.

En 2025, une étude a trouvé cette règle dans les chants de baleines. L'équipe d'Inbal Arnon a étudié longtemps les chants des baleines à bosse. Ces chants avaient été recueillis pendant 8 ans dans le Pacifique Sud. Ils correspondaient bien à la loi de Zipf. C'était la première fois que cette règle était trouvée aussi clairement en dehors du langage humain.

Que signifie ce résultat ? Il peut indiquer qu'il existe des règles dans les chants des baleines. Mais cela ne prouve pas, à lui seul, que les chants des baleines ont un sens ou une grammaire. Les chercheurs n'ont pas dit non plus que les chants des baleines étaient identiques au langage humain. Ils ont plutôt pensé qu'ils pouvaient être plus proches de la musique humaine.

Les chercheurs ont aussi étudié l'entropie de Shannon (Shannon entropy). L'entropie de Shannon est une méthode qui mesure, avec des nombres, à quel point il est difficile de prévoir le son suivant. Si les sons sont totalement en désordre, cette valeur est élevée. Si le même son se répète sans cesse, cette valeur est proche de 0. Les chants des baleines se trouvaient entre ces deux extrêmes. Ils suivaient des règles, mais ils n'étaient pas toujours identiques.

Cette organisation apparaît aussi dans d'autres études sur les baleines. Prenons ici un autre exemple de baleine. Il existe une équipe appelée Project CETI. Elle étudie les cachalots. Le cachalot est une espèce différente de la baleine à bosse. Le cachalot produit de courts groupes de clics. Ces groupes de sons s'appellent des coda.

En 2024, les chercheurs de Project CETI ont expliqué l'organisation des coda. Les cachalots créent leurs sons en mélangeant quatre éléments : le rythme, la vitesse, le rubato et les ornements. Le rubato consiste à allonger ou à raccourcir un peu le tempo. Un ornement consiste à ajouter un clic de plus au son. En combinant ces quatre éléments de plusieurs façons, ils créent un très grand nombre de sons. On pensait que seuls les humains utilisaient ce type de combinaison, mais les cachalots le font aussi [4].

Sources liées

- [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 Une lettre envoyée vers l'espace

La recherche Whale-SETI a un dernier but. Elle ne s'arrête pas à la connaissance des baleines. Elle veut créer un outil capable, un jour, de reconnaître un signal inconnu venu de l'espace. Il faut un critère pour séparer le bruit d'un signal qui porte du sens.

Il existe des méthodes qui expliquent les chants des baleines avec des nombres. Ce sont, par exemple, la loi de Zipf et l'entropie de Shannon. On pourra essayer ces méthodes plus tard sur des signaux venus de l'espace. Mais elles n'ont pas encore été appliquées directement à un signal spatial. Les chercheurs les étudient comme des méthodes qui pourraient servir à l'avenir.

Beaucoup de chercheurs ont pointé l'orgueil humain. Ils ont rappelé que nous ne comprenons même pas les sons des baleines qui vivent sur la même Terre. Ils ont alors demandé comment nous pourrions comprendre un signal venu de l'autre côté de l'espace. Cette question rend les humains plus modestes.

Autrefois, les humains croyaient que la Terre était au centre de l'univers. Mais on a découvert que cette croyance était fausse. La croyance selon laquelle seuls les humains échangent du sens pourrait lui ressembler. Dans la mer aussi, il existe peut-être des voisins qui échangent des sons.

Pendant longtemps, les humains ont considéré les chants des baleines comme du bruit. Ils ont aussi chassé d'innombrables baleines. Avec les bateaux et les machines, ils ont rendu la mer bruyante. Pourtant, les baleines n'ont pas arrêté d'échanger des sons. Depuis des temps très anciens, elles se transmettent des sons.

Voilà pourquoi les 20 minutes avec Twain étaient un entraînement. C'était un moment pour apprendre à l'avance une méthode utile quand nous rencontrerons une intelligence inconnue. Avant d'envoyer nos mots vers l'espace, nous devons écouter les sons près de nous. Bien écouter est le début de la traduction.

Sources liées

- [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 Une nouvelle porte ouverte par l'intelligence artificielle

Après la rencontre avec Twain, la recherche progresse vite. L'intelligence artificielle permet d'étudier les sons des baleines plus en profondeur. Un travail qui prendrait des dizaines d'années à une personne seule peut avancer plus vite grâce à l'intelligence artificielle.

Il existe une organisation appelée Earth Species Project. Elle crée une intelligence artificielle qui comprend les sons des animaux. Cette intelligence artificielle s'appelle NatureLM-audio. Cet outil écoute seulement le son et reconnaît quel animal l'a produit. Il classe aussi les types de sons. Il réussit déjà bien ce genre de tâches.

Cette intelligence artificielle possède aussi des capacités encore expérimentales. Elle essaie de compter combien d'animaux se trouvent dans un enregistrement. Elle essaie aussi de reconnaître des sons d'animaux qu'elle n'a jamais appris. Ces tâches ne sont pas encore parfaites. Elle réussit seulement un peu mieux qu'une réponse au hasard.

L'intelligence artificielle peut reposer sur une grande intelligence artificielle de base. On appelle ce type de grande intelligence artificielle un modèle de fondation (foundation model). C'est une intelligence artificielle qu'on peut utiliser pour plusieurs tâches avec un même système. Perch, créé par Google, en est un exemple.

On peut raconter l'aide apportée par l'intelligence artificielle à la recherche sur les baleines en suivant l'ordre du temps. D'abord, il y a eu l'année 2018. Google et l'Agence américaine d'observation océanique et atmosphérique (NOAA) ont travaillé ensemble. La NOAA est un organisme américain qui étudie la mer et le temps qu'il fait. Les deux équipes ont examiné des sons sous-marins recueillis pendant 14 ans. Cela faisait plus de 187 000 heures. Même si une personne les écoutait sans arrêt, il lui faudrait environ 21 ans. L'intelligence artificielle a retrouvé les chants des baleines à bosse dans cet immense ensemble de données.

Ensuite, il y a eu l'année 2025. Google a retravaillé Perch et a présenté Perch 2.0. Perch est une intelligence artificielle qui a d'abord appris avec des chants d'oiseaux. Mais elle convenait aussi très bien aux sons sous-marins. Elle a obtenu de bons résultats pour reconnaître les sons des baleines. Une intelligence artificielle entraînée avec les oiseaux a donc aidé la recherche sur les baleines.

La grande force de l'intelligence artificielle tient à sa vitesse. Des sons qu'un humain devrait écouter pendant des mois, l'intelligence artificielle les parcourt en quelques heures. La recherche avance donc plus vite. Elle trouve aussi de petits sons qu'un humain manquait. Grâce à cela, de nouveaux faits apparaissent dans d'anciennes données.

Il faut pourtant rester prudent. Ce n'est pas parce que l'intelligence artificielle reconnaît les sons des baleines qu'elle a déchiffré toute leur langue. Elle a trouvé des règles dans les sons, mais elle n'en connaît pas encore tout le sens. Aucune preuve ne montre encore que le sens de la langue des baleines a été entièrement compris. Le dictionnaire entre les humains et les baleines vient seulement de tourner sa première page.

L'Earth Species Project n'étudie pas seulement les baleines. Il étudie aussi les cris des corbeaux. Il étudie aussi les sons des éléphants. Il étudie aussi les sons des bélugas. Il traite les sons de plusieurs animaux avec un même outil. À l'avenir, il prévoit aussi de faire écouter aux animaux des sons produits par des humains. Le but est d'observer comment les animaux répondent à ces sons. L'échange sonore avec Twain en a été le premier pas.

Des recherches récentes ont aussi éclairé la structure du chant des baleines. En 2026, une équipe de Julia Zeh, de Syracuse University aux États-Unis, a mené cette étude. Les chercheurs ont fixé de petits appareils d'enregistrement sur le corps de mâles baleines à bosse en train de chanter. Ce sont des appareils posés brièvement avec des ventouses. Ils enregistrent à la fois le son et les mouvements. Cette étude porte sur un autre sujet que l'échange sonore avec Twain. Elle montre que les mouvements des baleines qui montent et descendent dans l'eau sont liés à leur chant. Les chants étaient différents quand les baleines descendaient en profondeur et quand elles remontaient. Leur chant était donc organisé selon les mouvements de leur corps [5].

La recherche continue donc encore aujourd'hui. Les chercheurs remplissent une page après l'autre du dictionnaire entre les humains et les baleines. Les 20 minutes de Twain ont ouvert la première phrase de ce dictionnaire.

Sources liées

[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/>

Chapitre 4. L'orchestre des profondeurs

4.1 Isoler une voix dans une mer bruyante

La mer n'est pas un lieu silencieux. Sous l'eau, c'est un vaste espace rempli de sons. Le chant des baleines, les sifflements des dauphins et le bruit des vagues se mélangent sans arrêt. Sous l'eau, le son se propage beaucoup plus vite que dans l'air. Il va environ 4,5 fois plus vite. C'est pourquoi les baleines lisent le monde avec leurs oreilles plus qu'avec leurs yeux. Elles utilisent le son pour se retrouver au large. Elles utilisent aussi le son pour poursuivre leur nourriture et éviter le danger [1][2].

Sous l'eau, le son parcourt environ 1 500 mètres en 1 seconde. Il va beaucoup plus loin que sur la terre ferme. C'est pourquoi une baleine peut entendre la voix d'une amie très éloignée.

Le problème vient du bruit fabriqué par les humains. Le bruit des hélices des grands navires secoue l'eau. Le son des appareils qui explorent la mer avec des ondes sonores, les sonars, est aussi fort. Il y a aussi les explosions utilisées pour chercher des ressources sous le fond marin. Ces bruits se propagent jusqu'à des milliers de kilomètres. Aujourd'hui, le bruit des navires est le bruit humain qui se répand le plus largement dans la mer. Quand le bruit est fort, la voix des baleines s'y perd. Les chercheurs disent que le son est masqué. Pour les baleines, qui vivent par le son, c'est une grande menace.

Le bruit rend aussi le corps des baleines malade. En 2001, après un grand accident aux États-Unis, le nombre de navires a beaucoup baissé pendant un moment. La mer est alors devenue plus calme. L'équipe de Rosalind Rolland a observé cette période. Elle a recueilli et étudié les excréments de baleines noires vivant dans la baie de Fundy, au Canada. Quand la mer est devenue plus calme, les substances de stress dans le corps des baleines ont diminué. Cela ressemble à ce que ressent une personne dont le cœur bat vite près d'un chantier bruyant. C'était la première preuve que le bruit fatigue les baleines [3].

Il est difficile d'isoler la voix d'une seule baleine dans un lieu où beaucoup de sons se mélangent. Les chercheurs appellent cette difficulté le « problème du cocktail ». C'est comme écouter seulement les paroles d'un ami dans une salle pleine de monde. Cela ressemble aussi au fait d'entendre l'ami assis à côté de soi, même quand la cantine est bruyante. Les anciennes machines faisaient mal ce travail. Quand le bruit métallique d'un navire se mêlait au son d'une baleine, elles se trompaient souvent.

L'intelligence artificielle a résolu ce problème d'une nouvelle manière. C'est une technique qui isole le son voulu dans un mélange de sons. La machine transforme d'abord le son en carte de nombres. Sur cette carte, le bruit du navire et le son de la baleine occupent des zones différentes. C'est comme remettre par couleur des perles mélangées. La machine sépare ces zones et rassemble seulement le son de la baleine. Les chercheurs de l'Earth Species Project ont créé une méthode pour enlever le bruit même sans son propre servant de bonne réponse. Grâce à cela, les humains peuvent réécouter plus clairement la voix des baleines [4].

Sources liées

- [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 Le dialecte des familles d'orques

Une fois le bruit écarté, les chercheurs ont commencé à lire les secrets cachés dans les voix des baleines. Ce secret est un « dialecte » différent selon les familles. Sur la côte ouest du Canada, il existe une base de recherche sur les orques. Elle s'appelle OrcaLab. Helena Symonds et le docteur Paul Spong y enregistrent les sons des orques sauvages depuis plusieurs dizaines d'années. Ils ne capturent pas les baleines et ne les dérangent pas. Ils recueillent calmement leurs sons avec des microphones placés sous l'eau [1].

Les orques vivent autour de leur mère. Les petits ne quittent jamais le groupe de leur mère pendant toute leur vie. C'est pourquoi chaque famille possède des sons propres transmis de génération en génération. Les chercheurs appellent cela un dialecte partagé par la famille du côté maternel. C'est un peu comme les façons de parler de notre classe et de la classe d'à côté, qui changent un peu. Grâce à ce dialecte, les orques reconnaissent leur famille dans la vaste mer. Il sert aussi à éviter que des proches parents ne s'accouplent entre eux.

OrcaLab a recueilli des sons pendant très longtemps. Les enregistrements approchent à eux seuls 20 000 heures. Il est difficile pour un humain de tout écouter. Les chercheurs ont donc confié ce travail à l'intelligence artificielle. Le programme créé par des chercheurs allemands et canadiens s'appelle ORCA-SPOT. Il trouve automatiquement les sons d'orques dans de longs enregistrements. Comparé aux bonnes réponses données par des humains, il a eu raison 93 fois sur 100. La machine a terminé en huit jours un travail qui aurait pris plusieurs années [5].

Quand l'intelligence artificielle sépare les sons, elle met l'accent sur les parties importantes. C'est comme souligner les mots importants dans un long texte. Ainsi, la machine repère même de fines différences d'intonation. Le petit orque apprend le dialecte en écoutant les sons de sa mère et de son groupe. C'est comme un bébé humain qui apprend en imitant la parole de ses parents. Les chercheurs ont aussi découvert un fait nouveau en regardant la carte sonore dessinée par l'intelligence artificielle. Les sons d'une famille ont peu à peu changé au fil des générations. C'est comme une langue utilisée longtemps, qui change un peu selon les époques. Cela signifie que les sons des baleines ne viennent pas seulement de l'instinct. C'est la preuve d'une culture qui s'apprend, se transmet et change, comme la parole humaine.

Connaître les dialectes aide aussi à protéger les baleines. Le son permet de savoir quelle famille vit à quel endroit. On peut vérifier quelle famille se trouve dans une mer où les navires passent souvent. En comptant les familles par le son, on voit vite si leur nombre baisse. Ces recherches continuent encore. En 2026, un nouveau programme appelé orcaAI est apparu pour trouver et classer automatiquement les sons d'orques. Ce programme n'a pas appris les sons d'orques du Canada, mais ceux des eaux d'Islande. Là-bas, les orques vivent en chassant le hareng. Même si les humains n'écoutent pas pendant longtemps, la machine examine à leur place les longs enregistrements [6].

Sources liées

- [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 Faire revivre les sons des bélugas sous la glace

Si les orques cultivent leurs dialectes sur des côtes claires, d'autres baleines vivent dans des eaux froides et troubles. Ce sont les bélugas, ou Beluga Whale, qui vivent dans l'Arctique et dans le fleuve St. Lawrence River, au Canada. Comme leur corps est blanc, on les appelle les anges de l'Arctique. Comme ils produisent bien des sons hauts et clairs, semblables à ceux des oiseaux, on les surnomme aussi les canaris de la mer. Les bélugas du St. Lawrence River sont aujourd'hui en forte baisse. Le bruit des usines et la pollution de l'eau se sont accumulés. Même après la création d'une zone protégée, leur nombre augmente mal. Cela inquiète beaucoup [1][7].

Jaclyn Aubin, chercheuse à l'University of Windsor au Canada, a participé à ce travail. Elle a décidé de décrypter les sons des bélugas. Ce travail n'était pas facile. Les bélugas sont des animaux sociaux qui vivent en groupe. Des dizaines d'individus se rassemblent et produisent des sons. Ils émettent en même temps des sifflements et des clics rapides. Les sons se mélangeaient, et il était difficile pour un humain de dire où ils commençaient et où ils finissaient.

L'équipe de recherche a confié ces sons à l'intelligence artificielle. D'abord, une machine qui repère les sons a découpé seulement les sons de bélugas dans de longs enregistrements. Ensuite, une autre machine a classé ces sons par type. À ce moment-là, l'équipe a aussi utilisé le programme AVES créé par l'Earth Species Project. AVES est un modèle de base qui apprend les sons d'animaux. Un modèle de base est une intelligence artificielle qui apprend d'abord sur beaucoup de données, puis sert à de nombreuses tâches. AVES apprend seul la structure des sons, même sans bonnes réponses données par des humains [8].

Quand les chercheurs ont comparé les sons classés par l'intelligence artificielle avec ceux classés par des humains, les deux résultats étaient presque identiques. Même sans intervention humaine, la machine a classé les sons de bélugas aussi bien qu'un spécialiste. Cette méthode a un avantage. La machine examine vite de longs enregistrements qu'un humain devrait écouter pendant plusieurs jours. Elle peut enregistrer sans interruption quand et où les bélugas du fleuve chantent. Ces données deviennent une base pour protéger les bélugas en déclin.

Le bruit de la glace qui se brise et se heurte était un gros obstacle. Pour enlever ce bruit, les chercheurs ont aussi utilisé une intelligence artificielle spécialisée dans la réduction du bruit. Cette machine fait revivre seulement la voix des bélugas au milieu des sons bruyants. Grâce à elle, il est devenu possible de réentendre les sons de bélugas qui risquaient d'être perdus. Ainsi, l'intelligence artificielle a protégé des voix enfouies sous la glace.

Sources liées

- [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 Une IA formée avec les chants d'oiseaux écoute la mer

Pourquoi parler soudain d'oiseaux dans une histoire sur la mer ? Il y a une raison. Une intelligence artificielle formée avec des chants d'oiseaux a beaucoup aidé à décrypter les sons des animaux marins. Tout a commencé chez Google. Il s'agit du projet « Perch », mené par l'équipe du chercheur en informatique Tom Denton. Le rêve de cette équipe est grand. Elle veut organiser ensemble tous les sons naturels de la Terre [10].

L'équipe Perch a enseigné les chants d'oiseaux à l'intelligence artificielle. Elle lui a fait apprendre les chants de plus de 10 000 espèces d'oiseaux vivant dans le monde. Le temps appris atteint plusieurs centaines de milliers d'heures. L'intelligence artificielle ainsi créée transforme les sons en coordonnées numériques et les range. On peut imaginer qu'elle dessine une grande carte où chaque son reçoit sa place.

Puis un fait inattendu est arrivé. L'intelligence artificielle qui avait appris les chants d'oiseaux a aussi bien classé les sons d'animaux sous l'eau. Une règle apprise chez un animal fonctionnait donc aussi pour les sons d'autres animaux. Les oiseaux chantent dans l'air, et les baleines chantent dans l'eau. Leurs lieux de vie et leurs sons sont différents. Et pourtant, la carte apprise avec les chants d'oiseaux convenait aussi aux sons des baleines.

Les chercheurs ont vérifié cette force avec des grenouilles. Les données utilisées alors s'appellent AnuraSet. C'est un ensemble de sons recueillis chez 42 espèces de grenouilles du Brésil. Ils n'ont pas été enregistrés en Amazonie, mais dans des forêts appelées Cerrado et forêt atlantique. Ces données n'ont pas été créées par Google, mais par l'équipe de Cañas. Google a emprunté ces données publiques pour tester son intelligence artificielle. Même après avoir vu seulement quelques sons de grenouilles, l'intelligence artificielle a reconnu les sons avec plus de 95 % de réussite. C'est grâce à cette méthode, qui apprend une nouvelle tâche avec seulement quelques exemples. Cette force ne s'est pas arrêtée aux grenouilles. Elle a aussi fonctionné pour les clics des baleines et les sifflements des dauphins [9].

Cette recherche continue de grandir. En août 2025, Google a présenté une nouvelle version de Perch, appelée Perch 2.0. Cette intelligence artificielle a appris non seulement avec des oiseaux, mais aussi avec les sons de plusieurs animaux terrestres. Elle a appris les sons de plus de 14 000 espèces animales. Avec les oiseaux, elle a aussi appris les sons de mammifères, de grenouilles et d'insectes. Mais elle a très peu appris de sons sous-marins. Pourtant, en février 2026, Google a annoncé que cette intelligence artificielle avait obtenu de bons résultats dans la recherche sur les sons des baleines. Dans plusieurs tests sur les sons de baleines, elle a été la meilleure ou la deuxième meilleure. Google a aussi créé DolphinGemma, une intelligence artificielle pour les dauphins. Ce programme est petit, donc il fonctionne même dans un smartphone. Il a appris avec les sons recueillis depuis longtemps par le Wild Dolphin Project. Il n'a pas encore révélé le sens du langage des dauphins. C'est un outil de recherche qui cherche la structure et les règles des sons [11].

DolphinGemma ne fait pas que classer les sons. Il imite les sons des dauphins et crée aussi de nouveaux sons par lui-même. Des expériences continuent pour faire écouter ces sons aux dauphins dans la mer. Les sons appris par DolphinGemma ont été recueillis pendant longtemps. Le Wild Dolphin Project recueille des sons depuis 1985. Ce sont les sons des dauphins tachetés de l'Atlantique vivant dans les eaux des Bahamas. Le projet a enregistré ensemble les sons et les comportements, comme l'accouplement ou les disputes. Google a ouvert ce programme pour que tout le monde puisse l'utiliser.

Sources liées

[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 Une petite IA protège les dauphins roses dans l'eau boueuse

L'intelligence artificielle sous l'eau travaille non seulement dans la vaste mer, mais aussi dans les fleuves. Elle travaille dans les eaux boueuses de l'Amazonie, où les yeux ne voient rien. Là vit un dauphin de rivière au corps rose, appelé boto. Au Brésil, le Mamirauá Institute et l'INPA mènent le travail pour protéger le boto [12].

Le boto diminue vite aujourd'hui. Il se prend dans les filets jetés dans le fleuve par les humains. Il est contaminé par le mercure rejeté pendant l'extraction de l'or. Il souffre aussi du bruit du développement. Le fleuve Amazone contient beaucoup de terre et de feuilles, donc l'eau est toujours trouble. C'est pourquoi les caméras et les drones ne peuvent pas trouver les botos. Seul le son est la clé pour les protéger [13].

Les chercheurs ont placé dans l'eau du fleuve des microphones qui écoutent les sons. Ces microphones enregistrent sans repos pendant longtemps. Il est difficile pour un humain d'écouter tous ces longs enregistrements. L'intelligence artificielle les examine donc à sa place. L'intelligence artificielle utilisée est une machine qui juge en regardant des images de sons. Elle filtre le bruit de la pluie, du vent et des branches dans la jungle. Puis elle trouve les clics produits par les botos.

Cette intelligence artificielle a trouvé les sons de botos avec une précision de 95 %. Elle a aussi reconnu le bruit des bateaux avec une précision de 92 %. Sous l'eau boueuse où les yeux ne voient plus rien, la machine protège la place des botos seulement grâce au son. Elle a aussi reconnu les sons du tucuxi, un autre dauphin qui vit dans le fleuve. Ainsi, l'intelligence artificielle trouve par le son des dauphins invisibles.

Ce relevé sert directement à protéger les botos. Les chercheurs ont dessiné, avec le son, les passages que les botos empruntent souvent. Ils ont aussi découvert à quelle saison ils se rassemblent et à quels endroits. Quand vient la saison où l'eau monte, les botos quittent les grands fleuves et entrent dans les forêts inondées. En étudiant les sons, ce mouvement est apparu clairement. Cette carte aide à changer les routes des bateaux et les endroits où l'on pose les filets. Le son est devenu un panneau qui guide la protection de la vie des botos. Même dans les régions reculées de la jungle, sans gros ordinateurs, les recherches continuent pour trouver comment écouter longtemps les sons et protéger ces animaux [2].

Sources liées

[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 Prêter l'oreille à la mer

Les recherches sous l'eau avec l'intelligence artificielle, vues au chapitre 4, changent une vieille idée. Pendant longtemps, les humains ont pensé que les autres animaux suivaient seulement leur instinct [1]. Mais les recherches récentes montrent autre chose [1]. Les dialectes des orques montrent une culture transmise de génération en génération [1]. Les bélugas qui échangent des sons dans les fleuves montrent qu'ils forment une société faite de relations [1].

De nouveaux faits sur le langage des baleines sont aussi apparus. Ils viennent de Project CETI et de chercheurs de plusieurs universités. En 2024, ils ont montré la structure des sons des cachalots. Les cachalots créent des signaux en enchaînant plusieurs clics. Ces groupes de sons s'appellent des coda. Les chercheurs ont trouvé dans les coda des règles, comme la vitesse et le rythme. Quand on mélange ces règles, beaucoup de sons différents apparaissent. Cela veut dire que ces signaux sont bien plus organisés qu'on ne le pensait. Les chercheurs ont aussi présenté un « alphabet sonore » qui classe ces petits morceaux de son. En 2026, ils sont allés plus loin. Dans les sons des cachalots, ils ont trouvé deux formes sonores qui ressemblent aux voyelles de la parole humaine. Elles se séparent un peu comme le son « a » et le son « i » chez les humains. Mais cela ne veut pas dire que nous avons compris le sens du langage des cachalots [14][15].

L'intelligence artificielle qui aide à poser des questions sur les sons des animaux, et à y répondre avec des mots humains, progresse aussi. C'est NatureLM-audio, créé par Earth Species Project. C'est le premier modèle de base son-langage pour les sons des animaux. Quand une personne pose une question avec des mots ordinaires, l'intelligence artificielle étudie le son et répond. Elle classe les espèces, trouve les sons, et compte même combien d'animaux crient [16].

La technologie qui lit le langage des baleines pose aussi de nouvelles questions au droit. En octobre 2025, des juristes et des chercheurs de Project CETI ont publié ensemble un article. Ils ont vu les baleines comme des êtres qui partagent une culture et communiquent entre eux. Ils ont donc proposé de donner aussi aux baleines le droit de ne pas subir de traitements cruels. Le groupe « More-Than-Human Life (MOTH) » de la faculté de droit de l'Université de New York mène ce travail. La nouvelle oreille ouverte par l'intelligence artificielle mène ainsi à des questions de droit et d'éthique [17].

Pendant longtemps, nous avons versé du bruit dans la mer. Avec le bruit des grands navires et les explosions de prospection, nous avons coupé les chemins de communication des baleines. L'intelligence artificielle est maintenant devenue un stéthoscope qui lit avec prudence les sons de la mer. Cet outil dit une chose aux humains. Il nous demande de baisser notre voix et d'écouter les sons de la nature tels qu'ils sont. Quand nous abandonnons l'envie de dominer la mer, nous pouvons vraiment entendre le chant ancien des baleines. L'orchestre des grands fonds n'a pas encore cessé de jouer. Plus nous prêtons l'oreille, plus ce son nous revient [2][1].

Sources liées

- [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/>

Chapitre 5. Le langage de la terre

5.1 Les sons des chimpanzés ont des règles

La parole humaine se forme en attachant quelques sons les uns aux autres. Avec peu de sons, elle crée un nombre infini de sens. Pendant longtemps, les humains ont cru que ce talent appartenait seulement à l'être humain. Mais des chercheurs qui ont observé longtemps les sons des animaux ont commencé à revoir cette idée.

La scène se trouve dans le parc national de Taï, en Côte d'Ivoire, à l'ouest de l'Afrique. Des groupes de chimpanzés sauvages y vivent. Des chercheurs de l'Institut Max Planck d'anthropologie évolutionniste, en Allemagne, et du Centre national de la recherche scientifique, en France, les ont observés longtemps. Le docteur Cédric Girard-Buttoz et le docteur Roman Wittig ont dirigé ce travail [1].

Cette étude ne s'est pas arrêtée après un an ou deux. Elle vient de plusieurs années d'observation du même groupe. Les chercheurs ont recueilli les sons de 46 chimpanzés adultes. Ils ont enregistré 4 826 sons. Autrefois, on pensait que ces sons étaient seulement des gémissements qui montraient des émotions. Mais en les comptant un par un, les chercheurs ont trouvé un ordre à l'intérieur.

Les chercheurs ont longtemps écouté ces nombreux sons et les ont classés. Ils ont séparé la hauteur et la durée des sons comme avec une règle. Ils ont ainsi distingué 12 sons de base chez les chimpanzés. Ce sont des sons comme des grognements, des aboiements et des cris. Ce sont, en quelque sorte, les plus petits morceaux de son, comme les consonnes et les voyelles dans la parole humaine.

Les chimpanzés ne produisaient pas ces 12 sons au hasard. Ils les attachaient selon un ordre précis. Cela formait plus de 390 groupes de sons. Certains ordres revenaient souvent. D'autres apparaissaient presque jamais. Cela veut dire qu'il existe une forme pour relier les sons.

Aujourd'hui, quand il faut examiner beaucoup de sons, une intelligence artificielle appelée apprentissage automatique aide les chercheurs. L'apprentissage automatique trouve vite les structures qui se répètent dans un grand nombre de sons. Il fait en quelques jours des calculs qui prendraient plusieurs années à une personne seule. Ce sont les chercheurs qui trouvent les règles. L'intelligence artificielle est un outil qui les aide.

Les groupes de sons formaient des niveaux. Un groupe de deux sons pouvait entrer dans un groupe de trois sons. La parole humaine forme aussi des niveaux comme cela. C'est comme quand des mots forment des groupes, puis que ces groupes forment des phrases.

Les chercheurs citent un grand trait de la parole humaine. C'est la capacité de relier de petits morceaux de son sans sens pour créer un nouveau sens. Un seul morceau de son n'a pas de sens. L'ordre dans lequel on les attache crée enfin le sens. Cette capacité s'appelle la double articulation des sons. Mais les 12 sons des chimpanzés ne sont pas exactement comme les consonnes et les voyelles humaines. Les chercheurs les étudient encore avec prudence [2].

Avec le temps, de nouvelles études ont suivi. En 2025, des chercheurs ont montré que les chimpanzés associent deux sons pour changer le sens. En examinant ces groupes de deux sons, ils en ont trouvé 16 sortes. Certains groupes ajoutaient simplement le sens des deux sons. D'autres groupes prenaient un sens tout nouveau.

Par exemple, quand un son était suivi d'un autre, le sens changeait. Cela pouvait devenir un signal pour proposer de faire un nid dans un arbre. Ils pouvaient aussi relier un son d'alerte et un son d'appel à l'aide pour dire de faire attention à un serpent. Un seul son et deux sons n'avaient pas le même sens [3].

Ces découvertes sont des traces qui nous montrent comment la parole humaine a pu naître. Avec ces groupes de sons, les chimpanzés signalent un danger et coordonnent leurs actions. Cela veut dire que le talent de relier des sons n'appartient pas seulement aux humains. Les règles sonores des chimpanzés montrent des traces de communication venues d'un passé très lointain [4].

Sources liées

- [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 Le rythme de parole presque humain des geladas

Si les chimpanzés montrent des règles sonores dans la forêt, un autre singe produit des sons qui ressemblent encore plus à ceux des humains. C'est le gelada (*Theropithecus gelada*), qui vit dans les hautes montagnes d'Éthiopie. Ces singes vivent en grands groupes sur des falaises rocheuses difficiles. Les sons qu'ils échangent ressemblent à des humains qui bavardent doucement [1][2].

Les sons des geladas semblent familiers à l'oreille humaine. On dirait plusieurs personnes qui marmonnent à voix basse. Ces sons tournent toute la journée autour des falaises. C'est pour cette raison que les chercheurs se sont longtemps laissés captiver par ces sons.

Le professeur Thore Bergman, de l'Université du Michigan aux États-Unis, a longtemps étudié ces sons. Quand les geladas vivent en bonne entente, ils font vibrer rapidement leurs lèvres. On appelle ce son un claquement des lèvres. Ils y superposent un son grave qui fait vibrer la gorge. On appelle ce son un warble.

Le professeur Bergman a mesuré le rythme de ce son. Quand les humains parlent, ils ouvrent et ferment la bouche environ trois à huit fois par seconde. Le warble des geladas vibre six à neuf fois par seconde. Les deux rythmes se recouvrent. L'unité qui compte combien de fois un mouvement se répète en une seconde s'appelle le hertz (Hz).

Cette découverte a été publiée en 2013. Les humains et les geladas vivent très loin les uns des autres. Pourtant, le rythme de leurs sons était proche. C'est comme s'ils avaient pris rendez-vous.

Les geladas ne s'arrêtent pas au seul claquement des lèvres. Ils superposent deux sons pour porter plus de sens. Les chercheurs ont comparé les sons des geladas avec ceux des babouins chacma. Plus les groupes qui vivent ensemble sont grands, plus les sortes de sons sont nombreuses. Quand on a beaucoup de voisins avec qui partager, on a aussi beaucoup de mots à partager.

Aujourd'hui, ces données sonores se sont accumulées en très grand nombre. Les chercheurs ont donc commencé à utiliser l'intelligence artificielle comme outil. L'intelligence artificielle classe vite une grande quantité de sons. Elle montre clairement en chiffres un rythme que l'oreille seule devinait. Ce sont les chercheurs qui découvrent. L'intelligence artificielle est une main qui les aide [3].

Une question apparaît ici. Les humains et les geladas n'ont pas le même ancêtre proche. Alors pourquoi leurs rythmes se ressemblent-ils ? Quand des animaux différents avancent dans la même direction, on appelle cela l'évolution convergente. C'est comme deux classes différentes qui inventent chacune un cri

d'encouragement, mais avec un rythme proche. Quand des êtres vivent de façon proche, ils peuvent arriver à des talents proches.

Sources liées

- [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 Le dialecte souterrain des rats-taupes nus

Si les singes échangent des sons sur la terre lumineuse, il existe aussi une société des sons sous la terre sombre. C'est dans des tunnels de terre sans la moindre lumière. Le rat-taupe nu (*Heterocephalus glaber*) y vit. Cet animal forme des groupes comme les fourmis ou les abeilles. Une seule reine met au monde les petits, et les autres se partagent le travail. Une société où les rôles sont partagés comme entre reine et ouvriers s'appelle une société eusociale [1].

Des chercheurs du Centre Max Delbrück, à Berlin en Allemagne, ont recueilli leurs sons. La docteure Alison Barker a dirigé ce travail. Quand les rats-taupes nus se rencontrent, ils gazouillent doucement. Les chercheurs ont recueilli plus de 30 000 de ces sons et les ont examinés un par un. Ils ont extrait plusieurs traits, comme la hauteur et la durée des sons, comme s'ils les mesuraient avec une règle [2].

Voici ce qu'ils ont découvert. Chaque groupe avait ses propres habitudes sonores. Chez les humains, ce serait comme les dialectes différents selon les régions. On appelle cela un dialecte de groupe souterrain. Dans les tunnels sombres, les rats-taupes nus ne peuvent pas voir avec leurs yeux. Ils distinguent les leurs et les étrangers seulement par le son.

Dans les tunnels, il fait toujours noir. Sans lumière, il n'y a aucun moyen de reconnaître les visages. Le son sert donc d'étiquette de nom. Si un individu utilise le même dialecte, il appartient au même groupe. Si un dialecte inconnu se fait entendre, il est vu comme venant d'un autre groupe. Un seul son joue donc le rôle de gardien du groupe.

Ce dialecte n'était pas inné. Les petits nouveau-nés imitent sans cesse les sons des adultes. C'est ainsi qu'ils apprennent les habitudes sonores du groupe. Il en allait de même pour les rats arrivés depuis un autre groupe. Ils imitaient les sons du nouveau groupe et apprenaient son dialecte. Le dialecte n'était pas fixé par le sang. Il était fixé par l'apprentissage.

Ce chemin d'apprentissage ressemble à celui du bébé humain qui apprend à parler. Comme un bébé écoute les paroles des adultes et les imite peu à peu, les petits rats apprenaient les sons de cette manière [3].

Le point à regarder de près est le pouvoir de la reine. Quand la reine est vivante, le dialecte du groupe reste bien uni. Mais si la reine meurt soudain, le dialecte se dérègle. Les règles sonores tremblent toutes ensemble. La reine était donc comme le pilier qui tenait ensemble les sons du groupe.

L'intelligence artificielle a aidé à classer cette grande quantité de sons. Elle a vite distingué les habitudes sonores propres à chaque groupe. Mais c'est une longue observation qui a montré que les dialectes s'apprennent, et que la reine les protège. L'intelligence artificielle était un outil qui aidait cette observation.

Le fait que les sons s'effondrent quand la reine s'effondre a un sens profond. Pour tenir les sons ensemble, il fallait un centre. Cela ressemble aux règles décidées ensemble dans notre classe. Si le centre qui fait respecter les règles vacille, les règles vacillent aussi.

Sources liées

- [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 Les cris d'alerte des suricates et le rire des hyènes

Beaucoup d'animaux terrestres mettent aussi de nombreux sens dans leurs sons. Le suricate (*Suricata suricatta*) en est un bon exemple. Par ses sons, le suricate signale si l'ennemi vient du ciel ou du sol. Il ne fait pas le même son selon qu'un aigle ou un serpent arrive. Il met même dans le son le degré d'urgence du danger [1].

Ce fait a été découvert il y a longtemps par la professeure Marta Manser, en Suisse. La professeure Manser a longtemps observé les suricates dans le désert du Kalahari, en Afrique du Sud. Elle a écouté les cris d'alerte un par un et les a classés. Plus le danger était urgent, plus le son devenait rude et bruyant. Il y avait des sons pour dire le type de danger et d'autres pour dire l'urgence. Avec un seul son, le suricate faisait bouger tout le groupe ensemble.

Aujourd'hui, l'intelligence artificielle classe parfois ces cris d'alerte. Elle distingue vite de petits détails sonores difficiles à entendre pour les humains. Mais c'est une longue observation sur le terrain qui a d'abord montré que les sons portent le type de danger et son urgence.

Il en va de même pour la hyène tachetée (*Crocuta crocuta*), qui règne sur la nuit des savanes africaines. La hyène produit un son qui ressemble à un rire. Ce ricanement contient plusieurs informations. L'équipe de recherche française de Nicolas Mathevon a étudié ce son. Il ressemble à un bruit désordonné, mais ce n'était pas le cas. À l'intérieur, il y avait l'âge et le rang dans le groupe. Il y avait aussi des marques qui indiquaient chaque individu différent. C'est comme une empreinte vocale humaine [2].

L'équipe de Mathevon a enregistré les sons de 17 hyènes à Berkeley, aux États-Unis. Puis elle a extrait deux traits du rire. Le premier est le motif sonore créé par la forme de la bouche et de la gorge. On l'appelle formant. Le second est la hauteur du son. On l'appelle pitch. Chez les hyènes de haut rang et de bas rang, ces deux traits étaient différents. Il était donc possible de savoir qui était puissant seulement en écoutant le son.

Le chien (*Canis lupus familiaris*), proche des humains, suit la même logique. Des chercheurs de l'Université du Michigan, aux États-Unis, ont étudié les sons des chiens. Ils ont recueilli des aboiements, des grognements et des gémissements. Puis ils ont essayé sur ces sons une intelligence artificielle qui servait à analyser la parole humaine. Cette intelligence artificielle s'appelle Wav2Vec2. Elle a alors distingué si le son était joueur ou agressif. Elle a aussi deviné, dans une certaine mesure, la race du chien, son sexe et son âge. Le taux de réussite a atteint 70 %. Un outil conçu pour la parole humaine a donc aussi fonctionné avec les sons des chiens [3].

Une question apparaît ici. Plus haut, nous avons dit que le rythme proche des geladas et des humains venait de l'évolution convergente. Leurs ancêtres étaient différents, mais leurs vies proches les ont

rendus proches. Mais la base sonore des humains et des chiens se ressemble pour une autre raison. Les humains et les chiens sont tous deux des mammifères. Il y a très longtemps, ils se sont séparés d'un même ancêtre. Dès le début, la structure du corps qui produit les sons avec la gorge et les cordes vocales se ressemblait donc. Quand les racines se ressemblent parce que l'ancêtre est le même, on appelle cela l'homologie. C'est comme la main et la patte avant, qui viennent des mêmes os. La ressemblance suit donc deux chemins. L'un vient de vies séparées qui deviennent proches. L'autre vient d'une séparation à partir d'une même racine.

Sources liées

- [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 Une seule intelligence artificielle apprend les sons de plusieurs animaux

Jusqu'ici, les chercheurs étudiaient les sons animal par animal. Ils étudiaient les chimpanzés comme des chimpanzés, et les geladas comme des geladas. Mais aujourd'hui, une seule intelligence artificielle apprend en même temps les sons de plusieurs animaux [1].

L'Earth Species Project est en première ligne. Cette organisation cherche à comprendre les sons des animaux grâce à l'intelligence artificielle. Elle a créé un modèle appelé NatureLM-audio. C'est une intelligence artificielle qui écoute les sons des animaux et répond aux questions humaines. Elle a été présentée pour la première fois en novembre 2024, puis ouverte à tous en 2025.

Ce modèle a appris les sons de milliers d'espèces. Il reconnaît l'animal seulement en écoutant le son. Il distingue aussi si c'est une femelle ou un mâle, un adulte ou un petit. Il peut même compter combien d'animaux se trouvent dans un enregistrement. Il a parfois montré tout seul des capacités qu'on ne lui avait pas enseignées.

Cette organisation a reçu une grande aide en octobre 2024. Elle a obtenu 17 millions de dollars de financement de plusieurs donateurs. C'est une somme de plus de 23 milliards de wons. Avec cet argent, elle a beaucoup agrandi son équipe de recherche. L'organisation pense pouvoir beaucoup éclairer les sons des animaux d'ici 2030 [2].

À l'été 2026, plusieurs équipes de recherche ont aussi ouvert un dépôt de données à utiliser ensemble. Ce dépôt s'appelle Animal Language Processing Data (alp-data). Il rassemble en un seul lieu des sons d'animaux qui étaient dispersés. Aujourd'hui, on y étudie ensemble les sons des corbeaux, des bélugas et des éléphants. Plus les données s'accumulent, plus l'intelligence artificielle apprend vite [5].

Un prix est aussi né pour soutenir ce mouvement. C'est le Coller-Dolittle Prize. Il donne de la force aux recherches qui cherchent à comprendre les sons des animaux. La lauréate de 2026 est la docteure Julie Elie, de Berkeley, aux États-Unis. La docteure Elie a déchiffré les sons d'un petit oiseau appelé zebra finch. Elle a trouvé 11 sons de base utilisés par cet oiseau. Ils portaient des sens comme la faim, la peur ou le salut. La docteure Elie a elle aussi reçu l'aide de l'intelligence artificielle appelée apprentissage automatique [3].

L'intelligence artificielle a aussi trouvé, dans les sons des animaux, une règle d'économie semblable à celle du langage humain. Les humains raccourcissent les mots qu'ils utilisent souvent. Plus un mot est

utilisé, plus son son devient court. En 2025, un chercheur a étudié les sons de 16 espèces de baleines. Puis il les a comparés à 51 langues humaines de plusieurs pays. On a vu la même économie dans les chants des baleines à bosse et des baleines bleues. Plus un son était chanté souvent, plus il était court. La sagesse qui consiste à économiser les sons n'appartenait pas seulement aux humains. Mais toutes les baleines ne suivent pas cette règle. Les chercheurs continuent donc leurs études [4].

Pensons aux règles sonores des chimpanzés, au rythme de parole des géladas et aux dialectes des rats-taupes nus. Les sons de ces voisins nous donnent matière à réfléchir. Le langage humain n'est pas un miracle tombé du ciel. Il a fleuri sur un large sol commun, où de nombreuses formes de vie se tiennent ensemble.

L'intelligence artificielle trouve à notre place les règles cachées dans les sons. Elle fait en quelques jours un travail qui prendrait plusieurs années à une personne seule. Grâce à elle, les chercheurs peuvent étudier beaucoup plus de sons. Mais l'intelligence artificielle ne traduit pas les paroles des animaux telles quelles. Elle aide seulement à classer les sons et à trouver des règles.

Comprendre les sons des animaux aide aussi à les protéger. Si l'on sait quel son signifie la peur, on peut reconnaître le moment où un animal souffre. On peut aussi deviner, grâce aux sons, ce qui les dérange. La capacité de lire les sons devient alors la capacité de prendre soin de nos voisins.

Pendant longtemps, nous avons pris ces sons pour du bruit dans les broussailles. Nous avons parfois aussi bouleversé sans respect les terres où ils vivent. Pourtant, depuis très longtemps, ils transmettent l'histoire de leur groupe par les sons. Ce que nous devons faire maintenant, ce n'est pas leur faire entendre d'abord des sons artificiels. Nous devons commencer par écouter humblement leurs voix.

L'intelligence artificielle devient une bonne oreille qui aide cette écoute. Elle entend à notre place les petites nuances que les humains manquaient. Mais ce qu'il faut écouter, et pourquoi l'écouter, ce sont les humains qui le décident. Il faut tenir un bon outil dans la main, et garder en même temps un cœur humble qui écoute. Alors nous pourrions vivre longtemps aux côtés des voisins de la terre.

Sources liées

- [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 Collier Foundation. (2026, July 3). Scientists win US\$100,000 in 2026 Collier 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/>

Chapitre 6. La symphonie du ciel et la danse invisible

6.1 La conversation entre les oiseaux et l'intelligence artificielle

Une conversation humaine n'est pas un texte écrit à l'avance. Nous écoutons les mots de l'autre personne et sa façon de parler. Nous observons même les petits silences. Puis nous ajoutons notre réponse au bon moment. Le chant des oiseaux contient aussi ces échanges.

Pendant longtemps, les scientifiques ont étudié les cris d'oiseaux en leur faisant écouter des enregistrements. Cette méthode s'appelle la réécoute sonore, ou playback. Le playback classique utilisait des sons courts choisis à l'avance ou des enregistrements fixes. Il captait mal le fil d'une longue conversation entre oiseaux. C'est pourquoi les gens ont longtemps pensé que le chant des oiseaux était un son fixé dès la naissance [1].

Deux équipes de recherche ont travaillé ensemble pour dépasser cette limite. Il s'agit de Earth Species Project, qui veut comprendre la communication animale avec l'intelligence artificielle, et de McGill University au Canada. Le travail a été mené par le Dr Logan James et la professeure Sarah Woolley. Ils ont recueilli les sons de femelles zebra finch, de petits oiseaux de la taille d'un moineau. Les sons échangés dépassaient 1 000 heures [2].

L'équipe a créé un modèle d'intelligence artificielle avec ces sons. Il s'appelle ZF-AIM. Quand un oiseau chante une petite phrase, ce modèle crée aussitôt la phrase suivante pour répondre. Cela ressemble à un smartphone qui propose le mot suivant quand on écrit. Cette méthode, qui écoute le son de l'autre et répond tout de suite, s'appelle l'échange conversationnel.

L'équipe a préparé l'expérience avec soin. Elle a placé chaque femelle zebra finch seule dans une boîte insonorisée. Dans la boîte, l'oiseau échangeait des sons avec un microphone et un haut-parleur. Quand l'oiseau produisait un son, ZF-AIM créait le son suivant adapté et le faisait entendre. L'oiseau et l'intelligence artificielle échangeaient donc des sons en temps réel.

L'équipe a comparé trois situations. Une conversation avec un oiseau vivant. La lecture d'un simple enregistrement. Un échange avec une intelligence artificielle en temps réel. Les traits naturels sont apparus seulement quand l'oiseau échangeait avec un oiseau vivant ou avec l'intelligence artificielle. Ils n'apparaissaient pas avec un simple enregistrement.

Le résultat a renversé une vieille idée. Les manuels disaient que les appels des femelles zebra finch étaient fixés par les gènes et ne changeaient pas. Mais les oiseaux qui échangeaient des sons avec l'intelligence artificielle étaient différents. Ils modifiaient leurs propres sons selon les sons de l'autre. Cela a montré que la conversation des oiseaux n'est pas une répétition de sons fixés.

Pourquoi l'équipe a-t-elle choisi le zebra finch ? Le zebra finch ne vit pas seul. Plusieurs individus forment un groupe et échangent des sons sans arrêt. Ils s'appellent, se répondent et changent de place. Cette conversation active convenait bien à l'étude. L'intelligence artificielle a imité ces échanges rapides.

L'équipe a trouvé trois traits dans les conversations des oiseaux. Premier point, les sons des deux oiseaux devenaient plus semblables. Deuxième point, ils changeaient vite leurs sons. Troisième point, ils ne répondaient pas à n'importe quel son. Ils choisissaient le son auquel répondre. Cela ressemble à une salle de classe bruyante où l'on se retourne seulement quand on entend son nom.

L'équipe a arrêté une partie du modèle à la fois. Quand elle a arrêté la partie qui réglait le rythme, la conversation s'est décalée. Quand elle a arrêté la partie qui fabriquait la forme du son, autre chose s'est décalé. Les deux parties jouaient donc des rôles différents.

L'équipe a estimé que cette méthode pouvait aussi aider à comprendre la conversation humaine. Quand les humains parlent, ils devinent aussi les prochains mots. Cette forme de prédiction est apparue dans les échanges entre l'oiseau et l'intelligence artificielle. Les règles de l'échange sonore peuvent peut-être fonctionner chez plusieurs animaux.

L'équipe a placé deux versions de ZF-AIM face à face et les a fait dialoguer. Sans intervention humaine, les deux intelligences artificielles ont échangé des sons. Cette conversation a recréé les traits des vraies conversations entre oiseaux. Cela signifie que le modèle avait bien appris les règles.

Cette étude mène à un but plus grand. Il s'agit de parler aux animaux avec des sons créés par l'intelligence artificielle. Aujourd'hui, nous sommes arrivés au stade où l'oiseau répond aux sons de l'intelligence artificielle. L'étape suivante sera d'envoyer des sons qui portent une intention humaine. Quand ce jour viendra, les humains et les oiseaux auront peut-être une vraie conversation [3].

Sources liées

- [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 Trouver les règles cachées dans le chant

Pendant que le zebra finch échange des sons avec l'intelligence artificielle, son cerveau apprend aussi. Comme les humains, les oiseaux imitent les sons des adultes quand ils sont jeunes. Cette période s'appelle période sensible (sensitive period). C'est le bon moment pour apprendre les sons. Cet apprentissage n'est pas fixé seulement par les gènes. L'ambiance du groupe dans lequel l'oiseau vit a aussi un grand effet [1].

Un jeune zebra finch grandit en écoutant le chant des oiseaux adultes. Au début, il imite maladroitement. Peu à peu, il se rapproche du chant adulte. Cela ressemble à un bébé humain qui apprend à parler. Plus il entend souvent les sons dans son groupe, mieux il apprend.

Dans la tête de l'oiseau, une zone guide l'ordre du chant. Elle s'appelle HVC. On peut l'appeler le centre de commandement du chant. Cette zone décide quel son produire et à quel moment. Cela ressemble à un chef d'orchestre qui guide le rythme.

L'équipe du professeur Timothy Gentner, à UC San Diego aux États-Unis, a étudié l'European starling. Cet oiseau chante en reliant des centaines de petits morceaux sonores. Il est difficile pour une personne de trouver ces règles seulement à l'oreille. L'équipe a donné à l'intelligence artificielle des chants sans étiquette [2].

L'intelligence artificielle a regroupé les morceaux sonores semblables. Elle a déployé le chant comme une carte. Sur cette carte, les sons formaient des groupes. Une structure que les humains ne voyaient pas est devenue visible. Cette méthode a permis de comparer les sons de plusieurs animaux.

L'équipe a utilisé cette méthode avec les sons de plusieurs espèces animales. Elle a placé sur une même carte les sons d'oiseaux, de souris, de singes, d'humains et de baleines. Même quand les espèces étaient différentes, elle pouvait les comparer de la même manière. Elle a aussi révélé des traits qui indiquent qui a produit chaque son.

Même en écoutant longtemps le chant des oiseaux, les humains ne trouvent pas toutes les règles. Les sons sont rapides et formés de nombreux morceaux. L'intelligence artificielle examine des milliers d'heures de chant sans se fatiguer. Elle repère un à un les ordres que les humains ont manqués.

Le chant du canary a aussi révélé des règles cachées. Une équipe de Boston University, aux États-Unis, a analysé les longs chants du canary. Avant, on utilisait une méthode qui regardait seulement le son juste avant pour prédire le son suivant. Cette méthode assez courte s'appelle un modèle de Markov. Mais le chant du canary n'était pas aussi court [3].

L'équipe a découvert que le début du chant décidait des sons beaucoup plus loin. Un son chanté au départ guidait même des sons situés loin après. Cette propriété s'appelle la dépendance à long terme. Cela ressemble au premier mot et au dernier mot d'une longue phrase qui vont ensemble. Même si le chant semblait libre, une structure y était cachée.

À mesure que ces recherches s'accumulent, une grande image apparaît. Earth Species Project a créé une intelligence artificielle qui apprend ensemble les sons de plusieurs animaux. Ce modèle s'appelle NatureLM-audio. Il a appris la parole humaine, la musique et les sons de la nature. Il reconnaît l'animal seulement en écoutant le son. Il devine aussi les sons d'animaux qu'il n'a jamais appris. Quand une personne écrit une question, il écoute le son et répond. Si l'on écrit quel animal a produit ce son, il choisit l'espèce. Il compte aussi combien d'individus crient dans l'enregistrement [4].

En 2026, Earth Species Project a donné un nom à ce mouvement. C'est l'étude qui traite les sons des animaux comme des données. Elle s'appelle le traitement du langage animal (Animal Language Processing). Ce nom a été utilisé pour la première fois en 2024. Il vient du traitement automatique du langage naturel, qui étudie les paroles humaines. Le travail qui déchiffre les règles du chant des oiseaux grandit aussi dans ce domaine [5].

Sources liées

- [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 La danse en 8 dans une ruche noire

L'ordre du ciel n'existe pas seulement chez les grands oiseaux. Il se trouve aussi dans le battement d'ailes des petites abeilles. La première communication animale précise découverte par les humains est la danse des abeilles. Le savant autrichien Karl von Frisch l'a découverte et a reçu le prix Nobel [3].

Quand une abeille trouve du nectar, elle retourne à la ruche. Puis elle secoue son corps en dessinant la forme du chiffre 8. Cette danse s'appelle la danse en 8 (waggle dance). L'intérieur de la ruche est complètement noir, sans une seule lumière. Sur la paroi de la ruche, l'abeille prend le haut comme repère. Le haut est la direction opposée à la gravité.

L'abeille traverse tout droit le milieu de la danse en 8. À ce moment, l'angle de son corps indique la direction. L'inclinaison par rapport au haut donne la direction du nectar. Le temps pendant lequel elle secoue son corps indique la distance. Plus elle secoue longtemps, plus l'endroit est loin. Avec cette seule danse, elle transmet une information jusqu'à plusieurs kilomètres.

Si le nectar est tout près, l'abeille danse en rond au lieu de faire un 8. Dans ce cas, elle ne donne pas la direction en détail. Comme c'est proche, il y a moins de choses à préciser. Plus la distance augmente, plus elle passe à la danse en 8.

Le signal de l'abeille n'est pas seulement un geste du corps. Les petites vibrations de son thorax et le son de ses ailes agissent aussi. Il y a aussi des vibrations qui passent par la paroi de la ruche. Cette méthode qui emploie plusieurs signaux ensemble s'appelle multimodale. Elle utilise en même temps le geste, le son et la vibration. Même dans une ruche noire, les autres abeilles sentent ce signal avec leur corps.

Ces signaux ressemblent aux différents sens des humains. Quand nous parlons avec un ami, nous n'utilisons pas seulement la voix. Nous utilisons aussi le regard et les gestes des mains. Les abeilles utilisent aussi les gestes, les vibrations et les sons ensemble.

Von Frisch a observé cette danse pendant longtemps et en a compris les règles. Il a marqué chaque abeille avec une couleur. Il a vérifié une par une quelle abeille était allée à quel endroit. Les humains avaient mis longtemps à faire ce travail à la main. L'intelligence artificielle l'a fait beaucoup plus vite [1].

L'équipe du professeur Tim Landgraf, à la Free University of Berlin en Allemagne, a créé une méthode pour lire cette danse automatiquement. L'équipe a installé une caméra dans la ruche. Cette méthode ne blessait pas les abeilles. L'intelligence artificielle a trouvé toute seule les danses des abeilles dans les vidéos.

Cette intelligence artificielle a lu rapidement des mouvements que l'œil humain manquait. Elle a lu la direction dans laquelle l'abeille avançait tout droit. Elle a aussi mesuré le temps de secousse du corps. En combinant les deux, elle a deviné l'endroit où se trouvait le nectar. La machine a fait le travail que les humains faisaient à la main.

L'équipe ne s'est pas arrêtée là. Elle a mis la danse lue dans une abeille robot. Ce robot est une machine qui imite la danse en 8 dans la ruche. Il est différent du petit robot volant de Harvard University. Quand le robot a dansé au milieu de la ruche, de vraies ouvrières ont suivi cette danse [2].

Les ouvrières ont changé leur direction de vol vers le côté indiqué par le robot. Pendant plusieurs années, l'équipe a amélioré les mouvements et les vibrations du robot. Elle a aussi diffusé une odeur de fleur. Ainsi, les abeilles ont mieux suivi la danse du robot.

Cette expérience montre deux choses. La première est que les humains ont bien lu la danse des abeilles. Si le robot avait envoyé un mauvais signal, les abeilles ne l'auraient pas suivi. La seconde est que les abeilles font confiance à la danse des autres et se déplacent selon elle. Un petit geste d'insecte contenait une direction et une distance.

Sources liées

[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 La conversation en trois dimensions que les corbeaux dessinent dans l'air

Parmi les intelligences qui volent dans le ciel, les corbeaux sont différents. Les corbeaux sont connus pour leur intelligence. Le cerveau des oiseaux est plus petit que le cerveau humain. Mais ses neurones sont très serrés. C'est pourquoi on appelle parfois les corbeaux des singes à plumes [1].

Des chercheurs de University of León en Espagne et de Earth Species Project ont étudié les corbeaux. Ils ont observé le carrion crow, qui vit dans le nord de l'Espagne. L'équipe a fixé un tout petit capteur sur les plumes de la queue des oiseaux. Il pesait seulement 12.5g. Il était conçu pour tomber tout seul après un certain temps. C'était un appareil qui ne blessait pas l'oiseau.

Ce capteur contenait un petit microphone et un détecteur de mouvement. Le microphone enregistrait les sons de l'oiseau. Le détecteur enregistrait les mouvements des ailes et du corps. Les chercheurs ont donc rassemblé les sons produits, les moments et les mouvements associés. Cela ressemble à une montre qui compte les pas d'une personne.

L'équipe a posé ces capteurs pendant trois ans. Elle a recueilli les sons quand les petits avaient environ dix à quinze jours. L'intelligence artificielle a divisé les sons recueillis en trois groupes. Les sons des adultes portant un capteur, les sons des autres adultes proches, et les sons des petits. Elle a donc distingué qui criait et à quel moment.

L'intelligence artificielle a trié les sons recueillis. Earth Species Project a créé un outil appelé Voxaboxen. Cet outil trouve chaque cri dans de longs enregistrements. Il a trouvé plus de 120 000 sons de corbeaux. Même un chercheur qui étudiait les corbeaux depuis plus de 25 ans a entendu des sons pour la première fois. Ces sons étaient trop faibles pour l'oreille humaine. Cet outil séparait aussi les sons quand ils se chevauchaient. Il faisait son travail même dans des enregistrements où plusieurs oiseaux criaient ensemble [4].

L'équipe de recherche a étudié séparément les données de sons et de mouvements. Les mettre au même endroit et les regarder ensemble reste un travail pour plus tard. Pour l'instant, elle a créé des données qui pourront être vues ainsi. Grâce à ces données, une voie s'est ouverte pour observer plus en profondeur les conversations des corbeaux.

L'intelligence artificielle n'a pas révélé tout le sens des corbeaux. Elle a aidé à classer de petits sons par groupes. Ce sont les zoologues d'autrefois qui ont éclairé l'esprit des corbeaux [2].

L'équipe du professeur John Marzluff, à l'University of Washington, aux États-Unis, a montré que les corbeaux se souviennent des visages. Les chercheurs ont porté des masques en caoutchouc et ont capturé des corbeaux. Ensuite, les corbeaux ont crié et attaqué les personnes qui portaient le même masque. Ils avaient mémorisé le visage d'une personne dangereuse. Ils ont aussi transmis ce souvenir à d'autres corbeaux.

Dans la même université, Kaeli Swift et le professeur Marzluff ont étudié un autre comportement. Les corbeaux se rassemblent près d'un compagnon mort et crient. Les gens ont appelé cela des funérailles. Mais ce n'était pas seulement de la tristesse. C'était un moment pour apprendre ce qui avait tué leur compagnon. Ils signalaient ce danger au groupe et s'en souvenaient longtemps [3].

Autrefois, on pensait que les cris des corbeaux n'étaient que des bruits forts. Quand l'intelligence artificielle les a classés par groupes, ce n'était pas le cas. Chaque son semblait porter un sens différent. Il semblait y avoir des sons pour dire de se méfier de quelqu'un, et des sons pour dire où aller. Le chant des oiseaux ressemblait presque à la parole humaine.

Une autre étude sur les oiseaux a fait un pas de plus. En 2026, Earth Species Project a présenté une intelligence artificielle qui reconnaît chaque animal seulement à partir du son. Cette étude ne portait

pas sur les corbeaux, mais sur les diamants mandarins. Une voie s'est ouverte pour distinguer la voix de chaque oiseau dans un groupe. Maintenant, les humains peuvent retrouver quel oiseau a dit quoi, et quand. Ils peuvent voir qui crie souvent et qui reste silencieux. Ces outils peuvent aussi servir à l'étude des corbeaux [5].

La recherche avance plus vite quand les données sont partagées. En juillet 2026, Earth Species Project a présenté un nouvel outil. Son nom est alp-data. C'est un outil qui rassemble des données de sons animaux dispersées ici et là. Il importe plus de 35 collections de sons de la même manière. Avant, chaque équipe de recherche préparait ses données séparément. Maintenant, elles les utilisent ensemble avec un seul outil. L'étude des sons des corbeaux avance aussi plus vite sur ces données communes. Derrière le ciel dégagé par l'intelligence artificielle, les oiseaux échangeaient chaque jour des conversations que les humains ne connaissaient pas [6].

Sources liées

- [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/>

Chapitre 7. Du cerveau des animaux au canevas du deep learning

7.1 Une intelligence artificielle qui apprend les sons par elle-même

Autrefois, les scientifiques écoutaient les sons des animaux avec leurs oreilles et les mesuraient à la main. Ils notaient un par un la hauteur et la durée des sons. Cette méthode pouvait servir dans une pièce calme. Mais elle marchait mal au bord de la mer ou dans la forêt. Le vent et les vagues se mêlaient aux sons. Beaucoup de sons d'animaux n'avaient encore aucun sens connu. Le travail était trop grand pour être fait seulement à la main [1].

Aujourd'hui, une intelligence artificielle appelée deep learning s'occupe de ce travail. Le deep learning a plusieurs méthodes. Dans certaines, une personne donne les bonnes réponses. Dans d'autres, l'IA trouve les règles par elle-même, sans bonnes réponses. Pour les sons d'animaux, l'IA utilise souvent cette deuxième méthode. Elle écoute d'immenses quantités de sons d'animaux. Puis elle trouve seule les règles cachées dedans.

Il existe des IA qui apprennent ainsi sur un vaste ensemble, puis servent à beaucoup de tâches. On appelle ce type d'IA un modèle de base pour les sons d'animaux. En anglais, on dit Bioacoustic Foundation Model. Foundation veut dire base. Si la base est solide, on peut l'utiliser vite pour de nouveaux problèmes [2, 4].

Un modèle de base partage ce qu'il a appris entre plusieurs tâches. Il sert à séparer les sons. Il sert aussi à enlever le bruit. Quand un nouveau problème apparaît, il ne recommence pas tout depuis le début. Il suffit d'ajouter un peu d'apprentissage sur la base déjà prête. Ces petites tâches faites sur la base s'appellent des tâches d'application. En anglais, on dit downstream task [2, 4].

L'une des premières IA importantes dans ce domaine s'appelle AVES. AVES a appris sans tableau de réponses écrit par des humains. Elle a écouté environ 5 000 heures de sons venus de plusieurs endroits. Même en écoutant 24 heures par jour, cela fait plus de 200 jours.

La méthode d'apprentissage d'AVES ressemble à un jeu de texte à trous. AVES reçoit directement la forme d'onde du son. Elle découpe cette onde en petits morceaux. Puis elle efface exprès quelques parties. Elle devine les parties effacées en regardant les sons d'avant et d'après. Elle répète ce jeu un très grand nombre de fois. Elle apprend à partir du son lui-même, sans le transformer en image.

Ce qu'AVES apprend, c'est la texture du son. Elle regarde si le son se répète de façon régulière. Elle observe comment le son monte et descend. Elle regarde la forme que le son dessine. Quand elle connaît cette texture, elle peut traiter le son de n'importe quel animal. Même si les espèces sont différentes, la texture du son peut se comparer.

Avant AVES, les tableaux de réponses bloquaient le travail. Pour apprendre, il fallait qu'une personne donne un nom à chaque son. Les sons sans étiquette ne pouvaient pas être appris. AVES a ouvert une voie pour apprendre sans étiquette. Elle a donc pu utiliser toutes sortes de sons du monde comme matière d'apprentissage.

Après ce processus, l'IA transforme les caractéristiques du son en nombres. Ce groupe de nombres s'appelle un embedding. Le son devient ainsi une coordonnée numérique que l'ordinateur peut reconnaître. Grâce aux embeddings, on peut traiter les sons de tous les animaux de la même manière.

Ces coordonnées ressemblent à un plan de places dans une classe. Avec ce plan, on sait qui est assis où. Avec un embedding, on sait où se place un son. Les sons semblables se regroupent près les uns des autres. Les sons éloignés sont des sons différents.

Cette méthode a un grand avantage. Les humains n'ont pas à donner une bonne réponse à chaque son. Donner des réponses prend beaucoup de temps. Pour nommer un son, un expert doit l'écouter lui-même. L'IA trouve les règles des sons sans bonnes réponses. Elle peut donc traiter même des sons dont personne ne connaît encore le sens. C'est une grande différence avec les anciennes méthodes.

Sources liées

- [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 Les sons de plusieurs animaux à la fois

Pour entraîner un modèle de base, il faut beaucoup de données sonores. Mais chaque jeu de données avait son propre format. Cela demandait beaucoup de travail. En juillet 2026, Earth Species Project a publié un assistant. Son nom est alp-data. Il rassemble plus de 35 collections dispersées de sons d'animaux. Il met la vitesse des sons et les étiquettes dans un même format. Les chercheurs passent moins de temps à préparer les données. Ils peuvent se concentrer sur l'apprentissage [10].

Earth Species Project a aussi publié un outil pour voir les données. Cet outil permet d'observer d'un coup d'œil les sons de plusieurs espèces. Rassembler et choisir les données est déjà la première étape pour entraîner une IA. Plus les bonnes données sont nombreuses, plus l'IA devient intelligente.

Quand les données ont été prêtes, une IA plus grande est apparue. Elle s'appelle BioLingual. Elle a été créée par l'équipe de David Robinson. Cette IA a appris les sons de centaines d'espèces terrestres et marines. Les données d'apprentissage associaient des sons à des textes descriptifs. Il y avait plus d'un million de paires [2].

BioLingual relie les sons et les mots. Elle associe les descriptions d'animaux écrites dans les encyclopédies avec de vrais enregistrements. La méthode utilisée s'appelle apprentissage contrastif (contrastive learning). Elle rapproche, sur une carte interne de l'ordinateur, les sons et les textes qui ont le même sens. Elle éloigne les sons et les textes qui ont des sens différents.

Cette idée de relier sons et textes vient d'une IA appelée CLAP. Des chercheurs de Carnegie Mellon University et de Microsoft l'ont créée ensemble. CLAP a associé toutes sortes de sons humains et de textes qui les décrivent. BioLingual a transporté cette idée vers les sons d'animaux [3].

Grâce à cela, une personne peut chercher un son avec des mots. Elle écrit par exemple : "le chant aigu d'un rouge-gorge qui appelle un partenaire au printemps". Alors l'IA retrouve ce son parmi d'innombrables enregistrements. Il n'est pas nécessaire de l'avoir entraînée séparément pour cela. Réussir directement une tâche jamais apprise s'appelle le zéro-shot (zero-shot). C'est la capacité de trouver la bonne réponse du premier coup, même sans apprentissage séparé.

BioLingual fait plusieurs tâches avec un seul corps. Elle reconnaît l'espèce de l'animal qui a produit le son. Elle trouve les endroits où il y a des sons d'animaux dans un long enregistrement. Si on pose une question par écrit, elle trouve le son qui correspond. Si elle apprend un peu plus avec de nouvelles

données, ses capacités augmentent encore. Une seule IA fait un travail que plusieurs personnes devraient se partager.

Les sons appris par BioLingual viennent de plusieurs grands dépôts de données. Il y a Xeno-canto, qui rassemble des sons d'oiseaux. Il y a aussi iNaturalist, où des personnes mettent en ligne des observations de la nature. Pendant de longues années, les chercheurs ont enregistré ces sons un par un. L'IA a appris en utilisant ces dépôts comme nourriture de base.

Ces IA aident beaucoup les chercheurs. Avant, il fallait écouter des enregistrements pendant des jours pour trouver un son. Maintenant, une courte ligne de texte suffit. La recherche sur les sons va beaucoup plus vite [2, 3].

Sources liées

- [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://earthspecies.org/2026/07/22/introducing-alp-data-a-shared-data-layer-for-animal-language-processing/>

7.3 Une intelligence artificielle qui écoute les sons et les explique avec des mots

Earth Species Project est un groupe de recherche qui veut comprendre le langage des animaux avec l'IA. L'IA créée par ce groupe s'appelle NatureLM-audio. Cette IA traite les sons des oiseaux, des animaux terrestres et des animaux marins [4].

NatureLM-audio écoute un son et répond avec des mots. Quand une personne lui fait écouter un son d'animal, elle explique par écrit ce que fait cet animal. Elle devine l'espèce. Elle compte aussi combien d'animaux se trouvent dans l'enregistrement. Elle distingue si l'animal qui produit le son est jeune ou adulte. Elle a fait ces tâches toute seule, même sans apprentissage séparé.

Cette IA écoute les sons et écrit du texte. Ce n'est pas une IA qui crée de nouveaux sons. Les IA savaient déjà comprendre la parole humaine et répondre par écrit. Maintenant, cette capacité est passée aux sons d'animaux. C'est le premier grand modèle de langage sonore créé pour les animaux.

Les capacités de cette IA se mesurent avec un test. Ce test s'appelle BEANS-Zero. Il ne contient que des problèmes que l'IA n'a pas appris. Dans ce test, l'IA a réussi à reconnaître les espèces. Elle a aussi réussi à classer les types de cris. Elle a encore réussi à écouter un son et à expliquer par écrit ce qui se passe.

Plus une IA avait appris largement, plus elle était forte devant les tâches inconnues. Elle obtenait de meilleurs résultats que les IA entraînées seulement sur les sons d'animaux. C'est parce qu'elle avait aussi entendu la parole humaine et toutes sortes de sons de l'environnement. Les petites IA entraînées sur un domaine étroit se perdaient devant les nouvelles tâches. Il faut une base large pour résoudre de nouveaux problèmes.

Earth Species Project utilise cette IA dans plusieurs recherches sur les animaux. Elle sert à étudier les sons des corbeaux. Elle examine aussi les sons des bélugas. Elle traite encore les sons des éléphants. Une seule IA permet d'étudier plusieurs animaux ensemble. Cette idée, qui traite les sons de plusieurs espèces avec une même méthode, s'appelle le traitement du langage animal (animal language processing) [11].

Cette IA n'est pas difficile à utiliser. On lui fait écouter un son enregistré et on pose une question. On lui demande quel animal a produit ce son. Alors elle répond par écrit. C'est comme poser une question avec des mots à un outil qui connaît les sons [4, 12].

Quand une IA écoute un son et répond par écrit, il faut rester prudent. Parfois, elle invente des faits qui semblent vrais. Elle peut donner un mauvais nom d'espèce. Elle peut aussi écrire qu'un animal a fait une action qu'il n'a pas faite. Produire ainsi une réponse fausse s'appelle une hallucination (hallucination).

Il existe une méthode pour réduire ces erreurs. On demande à l'IA de chercher d'abord de vrais enregistrements avant de répondre. Elle doit écrire seulement ce qu'elle a vérifié avec de vraies données animales. Cette méthode s'appelle une méthode augmentée par récupération (retrieval-augmented). C'est comme parler après avoir vérifié dans un livre à la bibliothèque.

Earth Species Project veut protéger la nature en comprenant le langage des animaux. Si nous connaissons les sons, nous pouvons repérer plus tôt les difficultés des animaux. Les sons peuvent montrer qu'un groupe diminue. L'IA trouve ces signaux plus vite que les humains [4, 11].

Sources liées

[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://earthspecies.org/2026/04/02/animal-language-processing-an-ai-convergence-in-animal-communication/>

[12] Earth Species Project. (2024-11-06). Introducing NatureLM-audio: An audio-language foundation model for bioacoustics. earthspecies.org. <https://www.earthspecies.org/blog/introducing-naturelm-audio-an-audio-language-foundation-model-for-bioacoustics>

7.4 Comment traduire sans dictionnaire

Pour passer d'une langue à une autre, il faut d'habitude un dictionnaire. Quelqu'un doit nous dire que le mot anglais dog veut dire 개 en coréen. Mais il n'existe pas de dictionnaire pour le langage des animaux. Personne n'a noté quel son de baleine signifie quoi. Comment traduire le sens sans dictionnaire [5] ?

En 2017, des informaticiens ont trouvé une méthode. Elle permet de traduire même sans aucune phrase d'exemple associée entre deux langues. On l'appelle traduction automatique non supervisée (unsupervised machine translation). Cela veut dire que l'IA traduit par elle-même, même sans bonnes réponses données à l'avance.

Pour les baleines, il est presque impossible de créer des exemples associés. Personne ne sait quel son de baleine veut dire quoi. Pour faire des paires, il faudrait d'abord connaître le sens. Mais pour connaître le sens, il faudrait déjà une traduction. La traduction non supervisée donne une piste pour sortir de cette impasse.

La clé de la méthode se trouve dans la forme que créent les mots. Quand on met beaucoup de mots d'une langue dans une IA, elle dessine une carte. Les mots aux sens proches se regroupent. Des règles apparaissent dans les distances et les directions entre les mots. Sur la carte, roi et reine sont séparés comme homme et femme. Cette carte créée par les mots s'appelle un espace latent (latent space). C'est une carte qui montre le sens des mots par leur position.

Là, quelque chose d'étonnant se produit. Quand on dessine les cartes de deux langues différentes, leurs formes sont presque les mêmes. La carte de l'anglais ressemble à la carte du finnois. C'est parce que le cerveau humain découpe le monde de façons proches. Si l'on tourne une carte pour la poser sur l'autre,

les mots se superposent. Le mot finnois *koira* vient se placer sur la position du mot anglais *dog*. Un dictionnaire de traduction apparaît donc sans dictionnaire.

Les chercheurs essaient de transporter cette méthode vers les baleines. Project CETI mène ce travail. Il dessine séparément la carte sonore des cachalots et celle des grands dauphins. Puis il étudie comment tourner les deux cartes pour les superposer. Il utilise des mathématiques qui font s'ajuster doucement les deux cartes. Le sens n'est pas encore entièrement expliqué. Les chercheurs en sont au stade où ils cherchent la structure des sons et les endroits où les cartes pourraient se recouvrir.

Project CETI a trouvé, dans les sons des cachalots, des motifs qui ressemblent aux voyelles humaines. Le chercheur qui les a mis en lumière est le linguiste Gašper Beguš. Il a travaillé à University of California, Berkeley. Dans les codas, qui sont des morceaux de sons de cachalots, il y avait deux branches. Une coda est un court groupe de clics que les baleines produisent quand elles dialoguent. L'une est la coda *a*, l'autre est la coda *i*. Comme les humains distinguent *a* et *i* avec la forme de la bouche, les baleines les distinguent par la résonance du son. Cette étude a été publiée en 2026 dans une revue de la Royal Society [6].

Project CETI avait déjà montré la structure des sons des baleines. Il a étudié 8 700 codas de cachalots avec l'intelligence artificielle. Le résultat a montré que les codas s'assemblent comme des briques Lego. Quelques traits, comme la vitesse et le rythme, se combinent de plusieurs façons. Les chercheurs ont comparé cette structure aux lettres et sons de la parole humaine. Ils l'ont appelée alphabet sonore. Mais ils n'ont pas encore montré ce que chaque son veut dire. Ils ont montré les règles de la structure [7].

En mars de la même année, Project CETI a annoncé une autre découverte sur les cachalots. Il a enregistré en détail le moment de la naissance d'un petit. À ce moment-là, plusieurs baleines s'occupaient ensemble du petit. Cette étude montre les soins coopératifs chez les baleines. C'est une étude différente de celle qui a montré les motifs sonores regroupés [13].

Chercher la structure des sons aide la traduction sans dictionnaire. Si les sons des baleines ont une structure proche de la parole humaine, il devient plus facile de trouver où superposer les deux cartes. Plus on divise les sons en petites parties, plus la carte devient précise. Des cartes précises se superposent mieux [5, 6, 7].

Sources liées

[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 L'intelligence artificielle qui dépasse les espèces et les murs qu'elle ne dépasse pas

Une intelligence artificielle qui a appris les sons d'un animal peut parfois reconnaître les sons d'autres animaux. On appelle cela la capacité d'adaptation au-delà des espèces. En anglais, c'est la généralisation entre espèces (cross-species generalization). Par exemple, une intelligence artificielle qui a appris les chants d'oiseaux peut aussi classer les sons de grenouilles [8, 9].

L'exemple le plus clair est Perch, de Google. Perch 2.0, sorti en 2025, a surtout appris avec des chants d'oiseaux. Il a entendu un très grand nombre de sons d'oiseaux et d'autres animaux terrestres. Les animaux présents dans ces données dépassent 14 500 espèces. Cette intelligence artificielle a très peu appris les sons sous-marins. Pourtant, elle a bien reconnu les sons de baleines. Elle a obtenu de bons résultats avec un classificateur créé à partir de peu de données. Une intelligence artificielle entraînée sur des oiseaux a reconnu des baleines [8, 9] [8].

Le mur que Perch a dépassé ne concerne pas seulement les baleines. Perch, entraîné avec des chants d'oiseaux, a aussi traité des sons de grenouilles. Un ensemble de sons de grenouilles de l'Amazonie brésilienne s'appelle AnuraSet. Perch a classé ces sons de grenouilles par type. Il y est arrivé sans avoir appris les sons de grenouilles [8, 9].

Ces résultats montrent un côté prometteur. Mais il existe aussi de grands murs que l'intelligence artificielle ne peut pas dépasser.

Premier point : les données sont trop rares. Les intelligences artificielles entraînées sur la parole humaine ont grandi avec des centaines de milliers d'heures d'enregistrements. Tous les sons d'animaux sauvages réunis restent bien moins nombreux. En plus, les anciens enregistreurs captaient surtout les sons forts. Ils ont manqué les petits chuchotements que les baleines échangent de près. Ces petits sons s'appellent des soft clicks. Des sons importants ont donc manqué dans les données [6].

Les soft clicks sont de petits sons que les baleines échangent de près. Chez les humains, cela ressemble à un chuchotement à l'oreille. Ces petits sons peuvent porter des messages liés aux émotions. Comme les anciens enregistreurs les ont manqués, des conversations importantes ont disparu.

Deuxième point : leurs mondes de vie sont trop différents. Le cachalot vit dans les profondeurs de la mer, sans lumière. Il lance des ultrasons avec la poche d'huile de son front et voit son environnement par le son. L'être humain voit le monde avec ses yeux. Une baleine qui voit par le son et un humain qui voit par les yeux ne sentent pas le monde de la même manière. Les deux cartes peuvent donc ne pas se superposer complètement.

Le monde que chaque être sent à sa façon s'appelle Umwelt. Ce mot veut dire monde environnant. Le biologiste Jakob von Uexküll a créé ce mot. Il veut dire que chaque animal perçoit un monde différent avec ses propres sens. Le philosophe Ludwig Wittgenstein a dit quelque chose de proche. Il a dit que même si un lion pouvait parler, les humains ne pourraient pas le comprendre. C'est parce que le lion et l'humain vivent dans des mondes différents.

Troisième point : mal utilisée, cette technologie est dangereuse. L'intelligence artificielle pourrait imiter le son d'un oiseau femelle pour appeler un mâle. Cela peut troubler l'ordre des oiseaux qui cherchent un partenaire et élèvent leurs petits. Tromper les animaux sauvages avec de faux sons s'appelle tromperie écologique (ecological deception) [14].

Le danger va plus loin que l'imitation. L'intelligence artificielle peut regarder la réaction d'un animal et modifier le son. Si l'animal répond, elle ajuste le son selon cette réponse. Une expérience qui change les sons en observant la réaction de l'animal en temps réel s'appelle une expérience sonore réactive. Bien utilisée, elle peut devenir une clé de communication. Mal utilisée, elle devient un outil pour tromper les animaux.

Les baleines ont gardé leurs propres cultures sonores pendant des dizaines de millions d'années. Les faux sons peuvent salir cette longue culture. Une fois l'ordre sonore dérangé, il est difficile de le rétablir. Il faut donc être prudent avant d'imiter leurs sons.

Quand on apporte de grandes intelligences artificielles dans les forêts et les mers, un autre problème apparaît. Les grandes intelligences artificielles sont lourdes et consomment beaucoup d'énergie. Il est difficile de les placer dans des enregistreurs au milieu de la forêt. Les chercheurs déplacent donc le cœur d'une grande intelligence artificielle vers une plus petite. C'est comme verser seulement le contenu d'un grand récipient dans un petit récipient.

C'est pourquoi les chercheurs disent qu'il faut établir des principes à respecter. Le principe est de ne pas intervenir sans prudence dans les sons des animaux. Ils l'appellent le principe à respecter en premier. Avant d'imiter les sons des animaux, nous devons d'abord apprendre à écouter en silence.

Si nous comprenons la parole des baleines, le droit peut aussi changer. En 2026, une discussion a proposé de donner des droits juridiques aux baleines. Si les sons nous apprennent ce que les animaux sentent et vivent, les humains auront plus de mal à les traiter sans respect. Décrypter les sons est lié à la protection des animaux [15].

L'intelligence artificielle a ouvert un nouveau chemin vers la parole des animaux. Elle a réduit à quelques années un travail qui aurait pris des centaines d'années aux humains. Mais ouvrir un chemin ne veut pas dire être arrivé. Les données restent rares, et le monde des animaux est différent du monde des humains. L'intelligence artificielle n'est qu'un outil. C'est aux humains de décider comment utiliser cet outil. Le cœur qui écoute avec humilité doit passer avant la technologie.

Sources liées

- [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>

Chapitre 8. Le monde propre des animaux et les frontières sémiotiques

8.1 Des mondes différents pour chacun, l'Umwelt

Le monde n'est pas un seul monde. Chaque être vivant vit dans son propre monde. Une personne a exprimé cette idée clairement pour la première fois. C'est le biologiste Jakob von Uexküll, né en Estonie. C'est un chercheur qui pensait que chaque animal sent le monde autrement [1].

Uexküll a appelé ce monde l'Umwelt. En allemand, on dit Umwelt. L'Umwelt est le monde propre qu'un être vivant sent et construit avec ses propres sens. Les humains voient les couleurs avec les yeux et entendent les sons avec les oreilles. Mais tous les animaux ne sentent pas comme les humains.

Le mot Umwelt n'est pas difficile. Même sur le même terrain de sport, un chien et un humain ne sentent pas les mêmes choses. Le chien lit le monde par les odeurs. L'humain le lit par les yeux. Chaque être vivant construit le monde avec les sens qui lui conviennent.

Regardons le cachalot (*Physeter macrocephalus*). Le cachalot descend à plusieurs centaines de mètres sous la mer. Là, c'est une obscurité sans aucune lumière. Dans cette obscurité, les couleurs n'ont aucun sens. À la place, le cachalot envoie des sons puissants depuis son front. Il écoute le retour de ces sons et lit le monde. Le monde du cachalot n'est pas dessiné par les couleurs, mais par les sons [2].

Pendant longtemps, les humains se sont trompés sur un point. Ils ont pensé que ce qu'ils ne sentaient pas n'existait pas. Ils ont pris leurs propres sens comme seule mesure. Cette attitude s'appelle l'anthropocentrisme. L'anthropocentrisme consiste à prendre les sens et les idées des humains comme norme du monde. À cause de cette attitude, les humains ont vu les sons des animaux comme de simples cris ou réactions.

Il existe une science qui étudie les sons et les gestes des animaux. C'est la zoosémiotique (Zoosemiotics). La zoosémiotique étudie le sens des signaux que les animaux échangent. Un signal est un son, une odeur ou un geste qui porte un sens. Le chercheur américain Thomas Sebeok a donné son nom à cette science. Cette science repose sur l'idée d'Umwelt de Uexküll. Les signaux porteurs de sens n'existent pas seulement dans la parole humaine. Tous les êtres vivants échangent des signaux dans leur propre monde [3].

L'idée de Uexküll mène à une nouvelle question. Le chercheur italien Nicola Zengiaro a proposé l'idée de zoosémiotique générative. Il dit que l'intelligence artificielle ne peut pas comprendre le sens des signaux animaux si elle se contente de les imiter avec une machine. Il faut aussi regarder l'environnement où vit l'animal et ses gestes. Alors seulement, on peut lire le sens correctement [4].

Aujourd'hui, l'intelligence artificielle ouvre ce problème d'une façon nouvelle. L'intelligence artificielle lit par le calcul une zone sonore plus large que les humains. Des lieux que les humains prenaient pour un silence vide étaient pleins de sons. Dans la mer, la forêt et le sol, des signaux circulaient. L'intelligence artificielle aide à trouver ces règles cachées.

Ce problème rejoint aussi la société humaine. Aza Raskin, qui a fondé Earth Species Project, dit que la technique sépare les humains les uns des autres. Les humains se sont aussi éloignés de la nature. Écouter les sons des animaux avec l'intelligence artificielle, c'est renouer ce lien rompu [5].

Sources liées

[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.ctv2hnhkcc3>
- [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>
- [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/>

8.2 Sons et vibrations au-delà des sens humains

L'oreille humaine a des limites. Les humains n'entendent que les sons entre 20 hertz et 20 000 hertz. Le hertz est l'unité qui compte combien de fois un son vibre en 1 seconde. Les sons plus hauts ou plus bas ne s'entendent pas pour les humains. Mais ces sons n'ont pas disparu. L'intelligence artificielle les capte et les montre aux humains [1].

Regardons d'abord les sons des plantes. Pendant longtemps, on a pensé que les plantes n'avaient aucun lien avec les sons. Mais les équipes du professeur Yossi Yovel et de la professeure Lilach Hadany, de l'université de Tel Aviv en Israël, ont changé cette idée. Des plants de tomate et de tabac qui manquaient d'eau ou dont la tige était blessée émettaient des ultrasons. Les ultrasons sont des sons aigus que les humains ne peuvent pas entendre. Ces sons se propageaient dans l'air avec une intensité proche de celle d'une conversation humaine.

Pourquoi les plantes produisent-elles des sons ? Quand l'eau manque, de petites bulles d'air apparaissent dans les fins canaux d'eau de la tige. Quand ces bulles éclatent, elles font un clic. L'oreille humaine ne l'entend pas, mais les machines le captent.

On a donné ces sons à une intelligence artificielle d'apprentissage profond. L'apprentissage profond est une intelligence artificielle qui trouve elle-même les caractéristiques des sons. Alors les sons de manque d'eau et les sons de blessure se sont séparés en groupes différents. Les deux sons avaient des formes d'onde bien différentes. La forme d'onde est la forme du son qui ondule comme une vague. Les animaux qui entendent les ultrasons, comme les insectes, peuvent entendre ces sons.

Les plantes entendent aussi les sons. Les équipes du professeur Yovel et de la professeure Hadany ont étudié l'onagre. Elles ont diffusé près de la fleur le son produit par le vol d'une abeille. Alors, en moins de 3 minutes, la fleur a rendu son nectar plus sucré. Elle n'a pas réagi aux autres bruits. Les pétales ont donc joué le rôle d'oreilles qui rassemblent le son [2].

Les équipes de la professeure Heidi Appel et du professeur Rex Cocroft, de l'université du Missouri aux États-Unis, ont fait une expérience proche. Elles ont fait entendre plusieurs sons à une petite plante appelée *Arabidopsis thaliana*. La plante restait calme face aux bruits de pluie et de vent. Mais c'était différent quand elle entendait les vibrations d'une chenille qui ronge une feuille. La plante a aussitôt produit des substances qui repoussent les insectes. Elle avait distingué les vibrations d'une feuille rongée des autres vibrations [3].

Une étude publiée en 2026 a poussé ce fait un pas plus loin. Les équipes du professeur Yovel et de la professeure Hadany ont observé des papillons de nuit. Quand une femelle choisissait un endroit pour pondre ses œufs, elle écoutait les sons des plantes. Le papillon évitait les sons des plantes desséchées et choisissait les plantes saines. C'était la première preuve qu'un animal réagit aux sons des plantes [4].

Il en va de même dans le sol et dans le monde des petits êtres vivants. Les abeilles dansent en forme de 8 dans la ruche sombre. Cette danse est un signal qui indique la direction et la distance des fleurs. Mais ce signal n'était pas seulement une danse vue par les yeux. Les vibrations des ailes et les vibrations qui

passent dans la ruche s'y mêlaient aussi. Les équipes du professeur Tim Landgraf ont lu cette danse avec l'intelligence artificielle. Des robots et des caméras ont observé la ruche pendant plus d'un mois. L'intelligence artificielle a automatiquement décodé les danses de la vidéo, une par une [5].

Les souris parlent aussi avec des sons que les humains n'entendent pas. Quand elles sont inquiètes ou de bonne humeur, elles émettent des ultrasons. L'équipe de Kevin Coffey, à l'université de Washington, a créé une intelligence artificielle appelée DeepSqueak. Cette intelligence artificielle a découpé et lu les sons des souris en petits morceaux. Résultat : on pouvait lire l'état émotionnel des souris grâce aux sons [6].

Les sons des souris ne sont pas captés par l'oreille humaine. Quand une souris a peur, elle émet des sons autour de 22 kilohertz. Quand elle est de bonne humeur, elle émet des sons aigus dans la zone des 50 kilohertz. Un kilohertz veut dire 1000 hertz. DeepSqueak a transformé ces sons en images et a dessiné une carte des émotions.

Ces découvertes disent toutes la même chose. Ce n'est pas parce que les humains n'entendent pas un son qu'il n'existe pas. La terre, les plantes et les petits animaux échangeaient sans cesse des signaux. L'intelligence artificielle déploie ces sons cachés devant les humains.

Sources liées

- [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 Une intelligence artificielle qui lit plusieurs sens à la fois

Les animaux ne parlent pas avec un seul son. Les sons, les gestes et les odeurs se mêlent pour créer du sens. Une communication qui utilise plusieurs canaux en même temps s'appelle multimodale (multimodal). C'est une communication qui mélange non seulement les sons, mais aussi le regard, les gestes et les odeurs [1].

Si l'on regarde seulement un son, on peut facilement manquer le sens. Le même son peut changer de sens selon le geste. Chez les humains aussi, les mêmes mots ne veulent pas dire la même chose quand on les dit en souriant ou en colère. Il faut donc regarder plusieurs signaux ensemble pour comprendre le sens.

L'étude des dauphins tachetés dans la mer des Bahamas en donne un bon exemple. La docteure Denise Herzing et le professeur Thad Starner ont créé une machine appelée CHAT. Cette machine aide les dauphins et les humains à échanger des sons. Mais le dialogue ne pouvait pas être compris avec les sons seuls. Il fallait aussi enregistrer le moment où le dauphin croisait le regard de l'humain et l'angle dessiné par son corps. Il fallait regarder plusieurs signaux ensemble pour comprendre le sens.

Les recherches sur les corbeaux montrent aussi que plusieurs signaux sont importants. Les équipes du professeur John Marzluff, à l'université de Washington aux États-Unis, ont étudié les corbeaux. Les chercheurs ont porté un masque spécial, puis ont capturé des corbeaux avant de les relâcher. Plus tard, quand une personne portant le même masque est apparue, les corbeaux ont crié fort et se sont rassemblés autour d'elle. Les corbeaux avaient mémorisé le visage de la personne qui les avait dérangés. Quand cette personne arrivait, ils lançaient des cris d'alerte pour prévenir leurs compagnons du danger [2].

Cette mémoire a duré longtemps. Les corbeaux n'ont pas oublié ce visage pendant plus de 2 ans. Il est difficile de découvrir cela en écoutant seulement les sons. Le sens apparaît quand on regarde ensemble les sons et les comportements des corbeaux.

Earth Species Project s'est lancé pour élargir ce type de recherche. Cette organisation est un groupe de recherche qui veut décoder les langues animales avec l'intelligence artificielle. Elle dit qu'il faut sortir des méthodes qui regardent seulement les sons. C'est parce qu'il existe des données recueillies par des capteurs fixés sur le corps des animaux. Les capteurs mesurent des choses comme les sons, les mouvements et la profondeur dans l'eau [3][4].

Les capteurs fixés sur le corps des animaux mesurent beaucoup de choses. Ils recueillent la distance nagée, la profondeur dans l'eau et le rythme du cœur. Quand on ajoute ces données aux sons, le sens devient plus clair. On peut distinguer un son de faim d'un son de peur. L'intelligence artificielle lit ensemble ces différentes données dans l'ordre du temps.

Earth Species Project a présenté un nouveau nom dans un texte de 2026. C'est le traitement du langage animal (Animal Language Processing). Ce terme a été utilisé pour la première fois dans une publication de 2024. Cette méthode ne force pas la grammaire de la langue humaine sur les animaux. Elle analyse les sons, les gestes et d'autres signaux comme des données qui ont des règles. C'est une attitude qui ne prend pas la parole humaine comme mesure, mais qui essaie de lire le monde animal tel qu'il est.

Un dépôt de données qui aide cette recherche est aussi apparu. En 2026, Earth Species Project a ouvert un outil appelé alp-data. Il rassemble en un seul lieu plus de 35 ensembles de données sonores d'oiseaux, de baleines et d'insectes. Pour l'instant, les données sonores sont au centre. Les données de mouvement et de vidéo seront ajoutées plus tard. Le dépôt est ouvert pour que chacun puisse travailler sur les mêmes données. Ainsi, la recherche peut s'accumuler vite.

Sources liées

- [1] Herzing, D. L., Starnes, 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 L'alphabet sonore du cachalot et l'intelligence artificielle

Le cachalot parle avec des clics. Un petit groupe de plusieurs clics s'appelle un coda. Un coda est un court paquet de clics que les cachalots échangent entre eux. Project CETI étudie ces codas. CETI est un projet de recherche qui veut décoder la langue des baleines avec l'intelligence artificielle [1].

En 2024, CETI et des chercheurs du Massachusetts Institute of Technology (MIT), aux États-Unis, ont fait une grande découverte. Ils ont analysé avec l'intelligence artificielle plus de 8700 morceaux de sons de baleines. Alors, les codas se sont révélés être un système sonore avec des règles. Les chercheurs ont appelé ces règles l'alphabet vocal du cachalot [2].

Cet alphabet contenait quatre éléments. Le premier est le rythme, c'est-à-dire une pulsation régulière entre les clics. Le deuxième est le tempo, c'est-à-dire la vitesse générale. Le troisième est le rubato, qui accélère ou ralentit selon le fil de la conversation. Le quatrième est un ornement, qui ajoute un clic à la fin. Quand on mélange ces éléments, on crée plusieurs codas. Les chercheurs ont trouvé plus de 143 combinaisons fréquentes. La structure ressemble un peu aux consonnes et aux voyelles dans la parole humaine.

Une étude de 2025 est allée un pas plus loin. Les équipes de Gašper Beguš, de l'université de Berkeley aux États-Unis et de CETI, ont étudié les codas en détail. Dans les clics, il y avait des sons proches des voyelles humaines. Les chercheurs ont trouvé deux voyelles de baleines différentes. Elles étaient différentes l'une de l'autre, comme le son a et le son i chez les humains. Cela veut dire que les cachalots divisent les sons avec finesse, comme les humains [3].

La forme des codas changeait selon la situation de conversation. La baleine suivante modifiait ses sons en fonction des sons produits par la baleine précédente. Cela ressemble aux humains qui se répondent et se relancent en parlant. Ce déroulement est apparu plus clairement grâce à beaucoup de données et à l'analyse par calcul.

Project CETI a aussi annoncé une nouvelle information en 2026. Les chercheurs ont enregistré en détail la naissance d'un cachalot. Ils ont capté les sons sous l'eau pendant plus de 6 heures et ont aussi recueilli des images filmées depuis le ciel. Ces données ont montré le groupe aidant ensemble à la naissance du petit. C'était le premier cas où l'on démontrait par des chiffres ce type de soin coopératif chez des animaux qui ne sont pas des humains ni des singes [4].

Transformer la langue des baleines en langue humaine n'est pas facile. C'est parce qu'il n'existe pas de dictionnaire où les deux langues seraient écrites côte à côte. Les équipes de Beguš cherchent une nouvelle méthode. Cette méthode trouve seule les règles des sons sans fiche de réponses, puis devine le sens [5].

Ce type de recherche était difficile sans intelligence artificielle. Avec l'oreille humaine, il est difficile de trouver les règles dans tant de clics. L'intelligence artificielle compare très vite de nombreux sons et trouve les règles cachées.

Il ne faut pas oublier une chose. Trouver les règles des codas ne veut pas dire que l'on connaît tout leur sens. Connaître la forme des sons et connaître leur sens, ce n'est pas la même chose. On n'a pas encore confirmé ce que disent les cachalots. La recherche vient seulement d'ouvrir la porte.

Sources liées

- [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>

8.5 Frontières sémiotiques et diplomatie entre espèces

L'intelligence artificielle ouvre aux humains un monde de sens plus large. Alors, avec quoi recueille-t-on les sons de la vie sous la mer ? Les humains utilisent des hydrophones. Un hydrophone est un micro qui capte les sons dans l'eau. On fixe aussi de petits enregistreurs sur le corps des baleines ou des dauphins. Dans différentes zones de la mer, on met à flot des réseaux de capteurs sur bouées pour capter les sons. Puis l'intelligence artificielle lit les sons ainsi recueillis [1].

Il existe aussi une idée au sujet d'une technologie à venir. Elle propose d'utiliser le futur réseau de communication de 6e génération (6G) comme capteur. La 6G est un réseau sans fil beaucoup plus rapide que ceux d'aujourd'hui. L'idée est que l'on peut connaître les mouvements en lisant la façon dont les ondes de ce réseau frappent les objets. Mais ce n'est encore qu'une proposition. En plus, les ondes radio traversent très peu l'eau de mer. Cette méthode serait donc utilisée pour observer la terre ferme et l'air. Elle ne convient pas pour recueillir les sons sous la mer [2].

Ici apparaît une frontière importante. Une frontière sémiotique est une ligne qui dit jusqu'où l'on peut lire les signaux et leur sens. Imaginons que l'intelligence artificielle ait recueilli tous les signaux des animaux. Peut-on alors forcer ces signaux à entrer dans le cadre du langage humain ? Par exemple, peut-on associer chaque son de baleine à une partie de la musique humaine ? Ce genre d'association impose la pensée humaine aux animaux [3].

Prenons un exemple de mauvaise association. Quelqu'un pourrait dire que 3 dauphins sur 10 utilisent des noms. Il pourrait aussi faire correspondre chaque clic de baleine à une note de musique humaine. Mais cela impose de force le cadre humain. Les sons des animaux doivent être lus dans le monde des animaux.

Déchiffrer les sons pose aussi des questions éthiques. Les chercheurs disent que les humains ne doivent pas imiter les paroles des animaux sans précaution. Mal utilisée, cette imitation peut servir à tromper ou à déranger les animaux. C'est pourquoi ils disent qu'il faut d'abord établir des règles [4].

L'idée d'Umwelt de Uexküll sert encore de guide aujourd'hui. Les animaux et les plantes ne vivent pas pour devenir des mots traduits dans un dictionnaire humain. Leurs signaux sont leur propre culture, transmise depuis très longtemps. Les chercheurs d'aujourd'hui reprennent l'idée de Uexküll et disent qu'il ne faut pas utiliser l'intelligence artificielle pour imiter les animaux sans prudence. Le but de l'intelligence artificielle n'est pas de fabriquer un jouet animal qui parle.

Il existe des expériences où l'on enregistre les sons des animaux, puis on les leur fait réentendre. On appelle cela la diffusion sonore, ou playback. On fait entendre un son enregistré avec un haut-parleur, puis on observe la réaction de l'animal. Si l'on utilise cet outil sans prudence, on peut troubler le monde des animaux. Un animal surpris par un son diffusé peut confondre un vrai signal et un faux signal. C'est pourquoi les chercheurs disent qu'il faut être prudent.

Le chemin à suivre n'est pas la conquête, mais l'écoute. Les humains doivent quitter la place de seul interprète de la Terre. Ils doivent protéger les sons des animaux, pour qu'ils continuent longtemps sur leurs terres et dans leurs mers. Les chercheurs appellent cela la diplomatie entre espèces (interspecies diplomacy). La diplomatie entre espèces est un accord pour que les humains et les autres formes de vie coexistent sans se nuire [5].

L'écoute ne s'arrête pas au fait d'écouter en silence. Elle mène aussi à la protection des mers et des forêts où vivent les animaux. Pour protéger les sons, il faut aussi protéger les lieux où ces sons naissent. Voilà le vrai sens de la diplomatie entre espèces.

Sources liées

- [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.ctv2hnhkcc3>
- [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/>

Chapitre 9. La croisée des chemins entre tromperie et protection

9.1 La confusion appelée par les faux sons

L'intelligence artificielle peut maintenant imiter les sons des animaux de façon très proche. Elle crée aussi des chants d'oiseaux, des chants de baleines et des cris d'éléphants. Comme une personne apprend à parler, l'intelligence artificielle écoute des sons et apprend leurs règles. Après cet apprentissage, elle invente elle-même des sons d'oiseaux. On peut employer cette technologie pour aider les animaux. Mais si on l'emploie mal, elle déranger beaucoup la nature [1].

Dans la recherche sur les animaux, il existe une méthode appelée « playback ». C'est une expérience où l'on fait réécouter un son enregistré, puis où l'on observe la réaction de l'animal. Autrefois, on employait seulement de vrais sons enregistrés par des personnes. Les anciens enregistrements contenaient du bruit, et les animaux pouvaient parfois s'en rendre compte. Maintenant, l'intelligence artificielle crée de faux sons très lisses. Les animaux peuvent donc être trompés plus facilement. Ces faux sons peuvent les embrouiller pour l'accouplement et les soins aux petits [7].

Des chercheurs de plusieurs pays ont averti de ce danger. Le professeur Christian Rutz et ses collègues en sont un bon exemple. Ils ont estimé que les faux sons changent le comportement des animaux sauvages. Tromper les animaux avec des sons de cette manière s'appelle la « tromperie écologique ». Cela signifie tromper des animaux dans la nature pour changer leur conduite. Un oiseau qui entend un faux son d'accouplement cherche un partenaire pour rien. Il perd alors de l'énergie et du temps, et il a moins de temps pour s'occuper de ses petits.

La professeure Karen Bakker a aussi signalé ce danger pendant longtemps. Elle a dirigé des recherches qui écoutent les sons des animaux grâce aux technologies numériques. Elle a expliqué cette idée dans son livre de 2022, « The Sounds of Life ». La professeure Bakker est morte en 2023. L'avertissement qu'elle a laissé continue encore aujourd'hui [2].

Les humains ajoutent souvent leurs propres sentiments aux sons des animaux. Il est facile de penser qu'un chant de baleine est une chanson d'amour. Pourtant, personne ne connaît encore le sens de ce son. Si l'on devine trop vite, on se trompe. Les chercheurs disent qu'il faut aussi faire attention à cette habitude.

Il existe une histoire racontée par la docteure Joyce Poole, qui a étudié les éléphants pendant longtemps. Cela s'est passé à Amboseli, au Kenya. L'équipe de recherche a fait écouter à un groupe le cri d'une mère éléphant déjà morte. Les éléphants ont entendu ce son qui leur manquait, puis ils ont couru partout pour la chercher. Parmi eux, la fille de la mère morte l'a appelée plusieurs fois et a attendu une réponse. Mais la propriétaire de la voix n'était nulle part. Les éléphants se souviennent même de la voix d'un membre de leur famille mort. Ils parlent au loin avec des sons graves que l'oreille humaine entend mal. Grâce à cela, ils se souviennent séparément des voix de dizaines d'éléphants. Un seul son a fortement secoué le cœur du groupe [3][4].

Il faut aussi faire attention avec les baleines. La baleine à bosse est une baleine qui chante. Quand un nouveau chant apparaît dans une mer, il se répand comme une chanson à la mode. Un chant venu des eaux australiennes peut même traverser le Pacifique Sud. Ce chant se répand lentement, sur plusieurs mois ou plusieurs années. Il traverse peu à peu environ 6 000 kilomètres. C'est une très longue distance, comme aller de Séoul jusqu'à un pays très lointain. Les baleineaux écoutent aussi le chant des adultes et apprennent en les imitant. Cela ressemble à la façon dont une chanson à la mode se répand avec les smartphones. Apprendre des chants les uns des autres est rare chez les êtres vivants autres que les

humains. Si l'on diffuse en mer un faux chant de baleine créé par l'intelligence artificielle, un gros problème peut apparaître. Une longue tradition de chants peut être embrouillée. Ce serait comme détruire leur culture avant même d'avoir compris leur langage [5].

C'est pourquoi les chercheurs mènent aussi ces expériences avec prudence. Ils les font seulement quand elles sont indispensables, et pendant peu de temps. Quand l'expérience se termine, ils vérifient si l'animal va bien. Même technologie, mais tout dépend de la façon dont on l'emploie : elle peut aider ou nuire.

Sources liées

- [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 Les réseaux de surveillance partout et la vie privée des animaux

Autrefois, les forêts profondes et les mers lointaines étaient des refuges pour les animaux. Les yeux et les oreilles des humains n'y arrivaient pas. Maintenant, c'est différent. De petits capteurs sonores intelligents sont placés dans beaucoup d'endroits des forêts et des mers. Ces appareils écoutent les sons 24 heures sur 24. L'intelligence artificielle analyse aussitôt ces sons. Ces appareils automatiques d'enregistrement, placés dans les forêts et les mers, s'appellent des « réseaux d'observation acoustique passive » [6].

Les animaux sauvages se trouvent maintenant devant des yeux qui les observent partout. C'est comme si une tour de surveillance invisible restait toujours allumée. Aza Raskin, qui a cofondé Earth Species Project, a décrit ce changement. Il dit que l'intelligence artificielle permet maintenant d'observer sans arrêt les forêts et les mers. Mais il avertit aussi qu'il faut créer des règles pour que cette force ne devienne pas une surveillance qui serre les animaux de trop près [8].

Ces appareils servent aussi à de bonnes choses. Beaucoup de baleines meurent après avoir heurté de grands navires. Avec des capteurs sonores, on trouve où se trouvent les baleines et on change la route des bateaux. Alors les grands navires ne les heurtent pas. Dans les forêts, ces appareils détectent aussi le bruit des scies qui coupent des arbres en secret. Ils aident donc à protéger les animaux qui disparaissent et à empêcher le braconnage.

Mais entre de mauvaises mains, ils deviennent des armes. C'est comme un couteau, qui peut servir à cuisiner ou à blesser. On peut aussi imiter de faux sons d'accouplement ou de nourriture pour attirer des animaux. Cela n'arrive pas encore partout, mais les chercheurs s'inquiètent de ce danger. Une technologie créée pour aider les animaux peut servir à les capturer [7].

La vie privée des animaux pose aussi problème. La vie privée, c'est l'espace personnel que l'on ne veut pas montrer aux autres. La professeure Karen Bakker a dit que les baleines et les dauphins devaient être vus comme des êtres qui ont des droits. Il arrive que des humains écoutent en secret même les sons calmes échangés entre une mère et son petit. Ils rassemblent parfois ces sons comme ils veulent, puis les

rendent disponibles pour tout le monde. Les chercheurs pensent qu'il faut traiter cela comme une atteinte à la vie privée. C'est un peu comme si quelqu'un lisait mon journal intime en secret et le mettait sur Internet [2].

Il existe aussi les droits des peuples autochtones, qui protègent depuis longtemps les forêts et les mers. De grandes entreprises prennent parfois en secret des sons venus de leurs terres. Elles les utilisent pour entraîner l'intelligence artificielle. Les données sonores sont aussi une richesse de cette terre. Il ne faut pas les prendre sans l'accord de leurs propriétaires. Le droit des habitants d'une région à décider eux-mêmes de l'usage des données venues de leur terre s'appelle la « souveraineté autochtone des données ».

Sur Internet, on raconte aussi que des animaux résistent à la surveillance. Une histoire dit qu'un groupe de cachalots aurait arrêté de faire du bruit tous ensemble. Une autre histoire dit que des corbeaux auraient créé des sons d'oiseaux pour embrouiller l'intelligence artificielle. Mais la science n'a pas encore confirmé ces faits. Ce ne sont que des histoires vues dans des vidéos sur Internet, donc il ne faut pas les croire comme des faits.

Les bons usages et les mauvais usages viennent de la même technologie. C'est pourquoi il est important de savoir qui l'utilise, et avec quelle intention. Les animaux aussi ont un espace personnel qu'il faut protéger.

Sources liées

- [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 Une marque humaine cachée dans les sons : le tatouage audio

Il faut une méthode pour reconnaître les faux sons. Les chercheurs et les ingénieurs ont proposé une solution. Elle consiste à cacher une « marque » dans les sons créés par l'intelligence artificielle. Cette méthode s'appelle le « tatouage audio (audio watermarking) » [9].

Le tatouage audio est une technologie qui grave une marque invisible dans un fichier sonore. Cette marque est un signal informatique caché dans le fichier. Les humains ne remarquent pas bien ce signal. Mais un programme de vérification le lit tout de suite. SynthID, créé par Google, est un exemple de ce type de marque. Cette marque est dispersée à plusieurs endroits du son. Il est donc difficile de la couper ou de l'effacer. Même si l'on baisse le son ou si l'on change sa vitesse, la marque reste.

Cette marque indique l'origine du son. Une machine peut vérifier si un son est un faux créé par des humains. On peut alors traiter séparément les vrais sons d'animaux et les faux sons. Séparer le vrai du faux est un premier pas pour protéger la nature. Sans marque, il est difficile de savoir ce qui est vrai [1].

Cette marque peut aussi servir à protéger les données de recherche. Dans les forêts et les mers, il existe des milliers d'appareils qui collectent des sons. Si de faux sons se mélangent à ces données, les résultats de recherche deviennent mauvais. Si l'on compte ensemble de vrais chants d'oiseaux et des chants d'oiseaux créés par une machine, la réponse sera fausse. Avec une marque, on peut filtrer les faux sons. Alors les chercheurs peuvent faire confiance aux données et les utiliser [7].

Certains proposent aussi de l'employer contre le braconnage. L'idée est de mettre une marque dans les sons de recherche pour vérifier d'où ils viennent. Cela permet de distinguer les sons de recherche autorisés et les faux sons dont l'origine est inconnue. Mais cette méthode a des limites. Si une mauvaise personne crée des sons sans marque, on ne peut pas les filtrer. Si elle utilise un son dont la marque a été effacée, il est difficile de le retrouver. Une seule technologie ne peut donc pas arrêter tout le braconnage [8].

Cette méthode est encore au stade de proposition. C'est une bonne idée, mais elle ne fonctionnera pas toute seule. Elle aura de la force seulement si tous les pays et toutes les entreprises promettent d'ajouter ces marques. Si un seul endroit laisse circuler de faux sons sans marque, un trou apparaît. C'est comme un filet : un seul trou suffit pour que les poissons s'échappent. La technologie seule ne suffit donc pas. Il faut des règles que tous respectent ensemble.

Cette technologie peut aussi servir pour d'autres sons créés par des humains. On peut mettre une marque dans une voix ou une chanson créée par l'intelligence artificielle. Cela peut réduire les tromperies avec de fausses voix. Une méthode créée pour protéger les animaux protège donc aussi les humains. Aujourd'hui, Internet est rempli de chansons et de voix créées par l'intelligence artificielle. Ces marques seront donc encore plus utiles à l'avenir.

Sources liées

- [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>
- [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 Créer des règles entre les humains et les animaux

Il faut des règles pour traiter les sons des animaux avec l'intelligence artificielle. Une personne puissante ne doit pas bouleverser la nature comme elle veut. Le son traverse les frontières et voyage loin. Un seul pays ne suffit donc pas. Les chercheurs disent que plusieurs pays doivent créer des règles à respecter ensemble. Cela signifie des promesses à tenir entre les humains et les animaux [10].

À la faculté de droit de New York University, aux États-Unis, il existe un groupe de recherche appelé « More Than Human Life ». On l'appelle MOTH en abrégé. Ce groupe a été créé en 2022. Des spécialistes du droit et des scientifiques s'y sont réunis. Ce groupe dit qu'il faut voir les animaux non comme des données, mais comme des êtres qui ont chacun leur vie. Il a proposé des règles avec quatre promesses : préparer, discuter ensemble, empêcher les dégâts et protéger.

La première règle consiste à examiner les choses à l'avance. Certaines expériences font échanger des sons entre l'intelligence artificielle et des animaux sauvages. Avant de faire ces expériences, il faut d'abord vérifier les risques. Il faut s'assurer qu'elles ne nuisent pas à l'accouplement, aux soins des petits et à la culture des animaux. Après cela seulement, l'expérience peut être autorisée. C'est comme recevoir un contrôle de sécurité avant de commencer des travaux.

La deuxième règle consiste à empêcher les usages sans contrôle. La technologie qui imite les sons des animaux ne doit pas être accessible à n'importe qui. Le groupe dit qu'il faudrait interdire complètement les technologies utilisées pour la chasse. Il faudrait les autoriser seulement avec prudence, dans des

lieux qui ont passé une vérification. C'est comme ne pas laisser un outil dangereux n'importe où. Il propose aussi d'empêcher l'usage sans contrôle des technologies clés pour gagner de l'argent.

La troisième règle consiste à empêcher les annonces exagérées. Il existe des nouvelles gonflées qui disent : « L'intelligence artificielle a créé un traducteur capable de parler librement avec les animaux. » Les chercheurs pensent que ce genre de nouvelle fait du mal. Elle pousse les gens à traiter les animaux comme des jouets. Ils disent donc qu'il faut expliquer honnêtement ce que la technologie ne peut pas encore faire [11].

Les règles ne cherchent pas à bloquer la recherche. Elles sont une clôture qui aide la recherche à se faire correctement. C'est comme sur un terrain de jeu : il faut des lignes au sol pour courir et jouer en confiance. Pour créer ces règles, il faut écouter plusieurs voix. Pas seulement celle des scientifiques. Les personnes qui connaissent le droit, les peuples autochtones qui vivent dans les forêts et près des mers, et les personnes qui protègent les animaux doivent décider ensemble. Quand plusieurs regards observent la même chose, on oublie moins d'endroits. Le but des règles est clair. Il ne s'agit pas de dominer les animaux, mais de les respecter [7][8].

Sources liées

- [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-rodriquez-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 Les nouvelles récentes et la croisée des chemins où nous nous trouvons

En 2026 aussi, la recherche sur le langage des animaux a avancé vite. Earth Species Project a présenté ce domaine sous le nom de « traitement du langage animal (Animal Language Processing) ». C'est une méthode qui rassemble beaucoup de sons d'animaux et les analyse, comme un ordinateur comprend la parole humaine. Elle s'éloigne de l'ancienne façon de regarder une espèce séparément. Elle observe les sons de nombreux animaux en même temps. Ce nom n'a pas été inventé pour la première fois ici ; il rassemble des recherches dispersées en un seul domaine. Cette annonce est sortie en avril 2026 [12].

En juillet de la même année, un nouveau dépôt de données a aussi ouvert. Son nom est « alp-data ». C'est un lieu qui rassemble des données sonores d'animaux pour que plusieurs chercheurs puissent les utiliser ensemble. Il contient plus de 35 ensembles de données. Quand on partage les données, la recherche avance plus vite. C'est comme des élèves d'une classe qui partagent leur matériel pour finir le travail plus vite [13].

À la base de cette recherche, il y a NatureLM-audio. C'est une grande intelligence artificielle créée pour les sons des animaux. Cette intelligence artificielle écoute seulement un son et devine de quel animal il vient. Elle compte aussi combien d'animaux se trouvent dans le son. Elle peut même reconnaître le stade de croissance, par exemple si c'est un jeune oiseau ou un oiseau adulte. Elle a montré qu'elle savait aussi repérer les sons d'animaux qu'elle n'avait jamais appris. Une telle intelligence artificielle fait en quelques minutes un travail qui prendrait plusieurs jours à une personne. Les chercheurs peuvent donc examiner les sons beaucoup plus vite [6].

La recherche sur les baleines a aussi avancé. Project CETI a rassemblé des codas, des morceaux de sons produits par les cachalots. Une coda est un groupe de clics qui se suivent, comme « clic, clic, clic ». L'équipe de recherche a analysé plus de 8 700 codas. Elle a ainsi trouvé les éléments de base qui composent ces sons. Les chercheurs ont classé ces éléments de base en quatre groupes. Ils les ont appelés « lettres sonores du cachalot ». Cela ressemble un peu à la séparation des consonnes et des voyelles dans l'écriture humaine. Les baleines changent le sens en modifiant la vitesse et le rythme des sons. Elles changent aussi le sens en ajoutant des clics. Les scientifiques ont même trouvé des règles pour assembler les sons, un peu comme les voyelles chez les humains. Cela ne veut pas dire que les baleines produisent des sons comme « a » ou « i » à la manière des humains. C'est une comparaison. De même que les humains relient consonnes et voyelles pour fabriquer beaucoup de mots, les baleines relient aussi des sons selon des règles [14][15][16].

L'intelligence artificielle sert aussi à protéger les animaux qui disparaissent. Le corbeau d'Hawaï en est un exemple. Cet oiseau a longtemps vécu en captivité, puis il est revenu dans la nature. Les chercheurs vérifient si, pendant ce temps, il n'a pas oublié certains sons de son groupe. Avec l'intelligence artificielle, ils comparent les anciens sons et les sons actuels. Leur but est d'aider l'oiseau à retrouver les sons oubliés. Earth Species Project étudie aussi les sons des corbeaux, des bélugas et des éléphants. Quand on observe plusieurs animaux ensemble, on voit des règles communes dans leurs sons.

Nous sommes maintenant à un croisement. Un chemin mène à la tromperie. C'est le chemin où l'on utilise de faux sons pour tromper les animaux et les perturber. L'autre chemin mène à la protection. C'est le chemin où l'on écoute en silence pour protéger les animaux. Ces deux chemins reposent sur la même technologie. Le chemin que nous prendrons dépend du choix des humains. L'intelligence artificielle n'est qu'un outil. Plus notre pouvoir d'écouter les sons grandit, plus les règles qui protègent ce pouvoir doivent grandir aussi. Quand un pouvoir devient fort, nous devons apprendre en même temps à le manier. Quand nous abandonnons notre arrogance et que nous ouvrons nos oreilles, les humains et les animaux peuvent vivre côte à côte [10][1].

Sources liées

- [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>

Chapitre 10. Personne non humaine et droits assistés par la machine

10.1 Du silence au témoignage

Pendant longtemps, le droit des humains a donné des droits seulement aux humains. Il en a aussi donné à des organisations créées par des groupes de personnes, comme les entreprises. On appelle ces organisations des personnes morales. Une personne morale est une fausse personne créée par le droit. Les entreprises et les écoles en font partie. Une personne morale n'est pas un humain, mais devant la loi, elle est traitée comme une personne [3].

L'ancien droit croyait une chose. Seul celui qui connaît ses devoirs et peut conclure un accord possède des droits. Les animaux ne peuvent pas se défendre avec des mots. Le droit ne les a donc pas vus comme des détenteurs de droits. Il les a vus comme des choses et des biens que les humains peuvent acheter et vendre.

Ce vieux mur tremble aujourd'hui. L'intelligence artificielle et la bioacoustique travaillent ensemble. La bioacoustique est la science qui étudie les sons produits par les animaux. Quand ces deux techniques se rencontrent, elles commencent à montrer en chiffres ce qui se passe dans la vie intérieure des animaux [2].

Autrefois, les humains écoutaient les sons des animaux seulement avec leurs oreilles. Ils manquaient donc beaucoup de sons. Aujourd'hui, on place beaucoup d'enregistreurs sonores sous la mer. Project CETI (Cetacean Translation Initiative) fait ce travail. En coréen, on pourrait l'appeler le projet de traduction du langage des baleines. Ce projet collecte depuis longtemps les sons des cachalots. L'intelligence artificielle examine un par un ces sons collectés. L'ordinateur trouve même des règles que les humains n'avaient pas vues [4].

Les cachalots échangent des morceaux de sons appelés coda. Une coda est une suite de plusieurs clics courts. Elle contient une coda d'identité (identity coda), qui indique le groupe. C'est comme un son-signé partagé par une même famille. Les cachalots reconnaissent leur groupe grâce à ce son. Il est difficile de dire que c'est exactement comme un prénom humain. Cela ressemble plutôt à un cri de classe partagé par toute une classe.

Des règles sont cachées dans les sons des cachalots. Ils assemblent de plusieurs façons quelques unités sonores pour créer de nombreuses formes. La parole humaine fonctionne aussi ainsi. Nous assemblons quelques consonnes et voyelles pour créer un grand nombre de mots. Mais les chercheurs étudient encore le sens exact de chaque coda. La structure des sons est devenue plus claire, mais leur sens n'est pas encore entièrement expliqué.

Les chercheurs qui étudient le cerveau humain expliquent le langage d'une nouvelle manière. Il s'agit de l'équipe de la professeure Evelina Fedorenko, du Massachusetts Institute of Technology (MIT). En 2024, cette équipe a publié une étude dans la revue Nature. Elle considère que le langage n'est pas une condition obligatoire pour avoir des pensées profondes. Cela veut dire que le langage est seulement un outil pour transmettre sa vie intérieure aux autres. Penser et parler sont deux choses différentes. On ne peut pas dire qu'un être ne pense pas parce qu'il ne parle pas [1].

Cette idée ébranle une vieille obstination du droit. On ne peut pas ignorer l'intelligence des animaux parce qu'ils ne parlent pas comme les humains. Les animaux aussi échangeaient des sons précis à leur manière. L'intelligence artificielle aide à trouver les règles sonores que les humains ne voyaient pas.

Il faut ici séparer deux choses. Savoir qu'il existe des règles dans les sons n'est pas la même chose que changer la loi. Même si la science révèle un fait nouveau, la loi ne change pas tout de suite. Il faut un accord social, de longues discussions et de nouvelles décisions de justice pour faire bouger le droit. Mais si les sons des animaux se révèlent être une communication porteuse de sens, le droit change aussi. Il devient difficile de laisser les animaux au rang de simples choses.

Sources liées

- [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 Le programme MOTH de NYU et les droits assistés par la machine

Un centre cherche à transformer ce changement en droit. C'est le programme sur la vie au-delà de l'humain de la faculté de droit de New York University (NYU). En anglais, il s'appelle MOTH (More Than Human Life). Ce programme est dirigé par le professeur César Rodríguez-Garavito, spécialiste du droit de l'environnement [1].

Le professeur Garavito travaille avec plusieurs chercheurs. Le docteur Gašper Beguš, de l'University of California, Berkeley (UC Berkeley), y participe aussi. Ils ont proposé une idée nouvelle. Ils veulent utiliser devant les tribunaux les sons d'animaux déchiffrés par l'intelligence artificielle. Ils appellent cela les droits assistés par la machine (Machine-Assisted Rights). Ce terme n'est pas encore beaucoup utilisé. C'est une proposition nouvelle de ces chercheurs.

On peut expliquer simplement les droits assistés par la machine ainsi. L'intelligence artificielle lit la communication des animaux et leur milieu de vie. Elle mesure en chiffres les dommages que les humains causent aux animaux. L'idée est de présenter ces chiffres comme preuves devant un tribunal. La machine devient donc une interprète et une représentante pour les animaux qui ne parlent pas. Cela ressemble à un adulte qui aide un enfant à déposer des documents quand l'enfant ne peut pas le faire seul.

Cette méthode cherche à contourner une vieille condition du droit. L'ancien droit donnait seulement aux humains la capacité d'engager un procès. Cette capacité s'appelle la qualité pour agir. Les animaux ne l'ont pas obtenue. Les chercheurs veulent donc utiliser comme preuves les dommages mesurés par l'intelligence artificielle. Ainsi, ils veulent frapper à la porte du tribunal pour les animaux.

L'équipe du professeur Beguš a fait une grande découverte. Elle a trouvé une structure de formants (formant) dans les sons des cachalots. Un formant est une zone d'un son où l'énergie se concentre et résonne avec une clarté particulière. Cela ressemble à ce qui se produit quand les humains changent la forme de leur bouche pour produire des voyelles. Dans une publication de 2025, l'équipe a montré qu'il existe dans les codas des cachalots une coda a (a) et une coda i (i). Ce sont des sons différents, comme les voyelles humaines [2].

Mais la carte sonore recréée par l'intelligence artificielle a aussi révélé un fait triste. Les bruits d'hélices des grands navires et les explosions utilisées pour chercher des ressources remplissent la mer. On

appelle ces sons le bruit marin créé par les humains. Une étude publiée en 2021 dans la revue Science a traité ce sujet. Elle expliquait que la mer est devenue bien plus bruyante depuis le développement de l'industrie. Certaines recherches montrent aussi que le bruit des navires peut influencer la structure des codas des cachalots [3].

Le programme MOTH affirme que ces dommages peuvent servir de preuves devant un tribunal. Il n'est pas nécessaire de traduire parfaitement le sens des sons. Il suffit de montrer en chiffres à quel point la communication d'un groupe de cachalots est abîmée. L'idée est de prendre comme preuve le graphique même des sons déformés. Alors un tribunal pourrait ordonner de détourner les routes des navires ou bloquer des dispositifs sonores militaires. Cette thèse parle de deux droits pour les cachalots. Le droit d'échapper à la souffrance et le droit de protéger leur culture [4].

Cette idée rend aussi plus fort le droit qui existe déjà. Aux États-Unis, une loi protège les animaux menacés de disparition. Une autre loi protège les mammifères qui vivent dans la mer. Les chiffres des dommages mesurés par l'intelligence artificielle donnent de la force pour faire respecter ces lois dans la réalité. Cela veut dire qu'avant même de créer de nouveaux droits, on peut utiliser plus fortement les lois qui existent déjà.

Ici aussi, il faut séparer deux choses. Savoir qu'un son a été abîmé n'est pas la même chose que changer la loi. On ne sait pas encore avec certitude s'il existe déjà un procès où ces preuves ont été réellement présentées au tribunal. Les chercheurs en sont au stade où ils proposent cet usage.

Sources liées

- [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., Sprou, 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 Le témoignage des Caraïbes et la sagesse autochtone

Le débat sur la personne non humaine ne reste pas enfermé dans les ordinateurs des chercheurs occidentaux. Une personne non humaine désigne un être qui n'est pas humain, mais que le droit peut protéger. Ce débat a grandi en rencontrant la sagesse des peuples autochtones, qui protègent la mer depuis longtemps [2].

Le 28 mars 2024, un événement historique a eu lieu à Rarotonga, dans les Cook Islands du Pacifique. Le roi maori et les dirigeants des Cook Islands se sont réunis. Des dirigeants de Tahiti, de Tonga et d'Hawaï étaient aussi présents. Ils ont reconnu les baleines comme des sujets indépendants qui possèdent des droits. Ils les ont déclarées êtres protégés par le droit, même si elles ne sont pas humaines. Cette déclaration commune s'appelle He Whakaputanga Moana, c'est-à-dire la Déclaration de l'océan. Cette déclaration a de nouveau été largement connue en février 2026 à Waitangi, en Nouvelle-Zélande [1][4].

Cette déclaration est la première déclaration menée par des peuples autochtones qui reconnaît les baleines comme des êtres sensibles. Elle affirme que les baleines ont le droit de nager librement et le droit de vivre dans une mer saine.

Cette déclaration s'enracine dans le savoir écologique traditionnel. Le savoir écologique traditionnel est un ancien savoir qui voit la nature comme une famille avec laquelle on vit. Les peuples autochtones ont transmis ce savoir pendant très longtemps. Ils savaient que le chant des baleines n'était pas un bruit dû au hasard. Ils croyaient que ce chant portait la mémoire de la mer et l'histoire des familles.

L'intelligence artificielle moderne va dans le même sens que cette vieille intuition. Project CETI a longtemps observé les baleines au large de la Dominique, dans les Caraïbes. L'équipe de recherche a étudié les codas de ces baleines avec l'intelligence artificielle. En combinant quelques caractéristiques, le nombre de codas reconnaissables a beaucoup augmenté. Avant, on pensait qu'il existait environ vingt et un types de codas. Mais quand la méthode de combinaison est devenue claire, il y a eu plus de cent formes reconnaissables. Cela ressemble aux humains, qui créent beaucoup de mots avec peu de lettres. Cela rejoint aussi la croyance autochtone selon laquelle les chants portent du sens [3].

Depuis longtemps, les dirigeants des Cook Islands considèrent les baleines comme des parentes. Ils les traitent comme de grandes aînées de la mer. Ce sentiment existait bien avant l'arrivée de la science. Aujourd'hui, l'intelligence artificielle ajoute à ce sentiment une preuve en chiffres.

Les peuples autochtones et les scientifiques ont donc créé ensemble une nouvelle méthode. Elle place côte à côte les calculs de l'intelligence artificielle et le savoir marin des peuples autochtones. Ici, les peuples autochtones ne sont pas des assistants de recherche. Ils sont des acteurs principaux qui décident dès la conception de l'intelligence artificielle. Les scientifiques et les peuples autochtones protègent la mer sur un pied d'égalité.

Un tel accord ne reste pas dans un seul pays. L'océan traverse les mers de plusieurs pays. Les baleines ne connaissent pas les frontières et se déplacent dans les vastes mers. La loi d'un seul pays ne peut donc pas protéger toutes les baleines. La Déclaration de l'océan est un premier pas où plusieurs pays et tribus agissent ensemble.

Sources liées

- [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>
- [4] Inside Climate News. (2026, February 5). How a 'powerful vision' on whales and oceans could change worldviews. Inside Climate News. <https://insideclimatenews.org/news/05022026/maori-scientist-daniel-hikuroa-legal-rights-of-whales/>

10.4 Le droit d'une espèce à évoluer

Au bout du chemin ouvert par l'intelligence artificielle se trouve une idée plus grande. Elle dépasse l'ordre de détourner une route maritime. C'est le droit d'une intelligence non humaine à avancer d'elle-même vers l'avenir. Certains chercheurs appellent cela le droit d'une espèce à évoluer. Ce n'est pas encore un droit fixé par la loi. C'est une proposition de chercheurs [1].

La parole humaine a grandi pendant très longtemps. Elle a commencé par quelques sons, puis elle est devenue une grammaire complexe. La culture sonore des animaux n'est pas non plus une chose morte et figée. C'est un patrimoine vivant qui grandit au fil des générations, en apprenant, en imitant et en

changeant. C'est aussi pour cela que les chants des baleines diffèrent un peu selon les régions. Cela ressemble aux langues humaines, qui se divisent en dialectes selon les régions [3].

On peut le comprendre en observant les mères cachalots. Une mère enseigne les sons du groupe à son petit nouveau-né. Le petit imite longtemps avec des sons maladroits, comme un babillage. Cela ressemble à un bébé qui apprend à parler en imitant maladroitement. Cela signifie que les sons des cachalots ne sont pas un instinct gravé à la naissance. Ils sont un héritage culturel transmis de génération en génération. Comme les humains transmettent des livres et des chansons, les baleines transmettent des sons.

Mais le bruit que les humains versent dans la mer bloque cet apprentissage. Les bruits industriels et les explosions militaires jour et nuit remplissent la mer. Les jeunes baleines n'entendent pas correctement les sons de leurs parents. Les bandes de hauteur importantes pour apprendre les sons se mélangent. C'est comme dans une classe bruyante, où l'on n'entend pas la voix du professeur. Cela ne blesse pas seulement un individu. Cela coupe le fil de la culture et de la mémoire d'un groupe [2].

C'est pourquoi certains philosophes de l'écologie et juristes défendent une idée. Ils veulent protéger le droit des animaux à apprendre leurs sons en sécurité et à les transmettre. Ici, le monde où vit chaque animal est appelé un monde propre, perçu par ses propres sens. Le monde perçu par une chauve-souris et celui perçu par une baleine sont différents. La proposition consiste à établir un droit à grandir par soi-même, sans bruit humain ni sons artificiels. Ce n'est pas encore du droit. C'est encore une discussion entre chercheurs.

Il faut aussi faire attention aux faux sons de baleines créés par l'intelligence artificielle. Des faux sons diffusés sans prudence par les humains peuvent troubler l'apprentissage des petits. Cette proposition inclut donc aussi un rôle pour empêcher ces faux sons. Si on l'emploie mal, une bonne technique peut devenir un nouveau bruit. Le traducteur doit être un outil pour écouter discrètement le monde des baleines. Il ne doit pas devenir un outil qui dérange ce monde.

Quand on réduit le bruit dans la mer, il diminue vite. L'étude de 2021 publiée dans Science l'a montré. Le bruit naît sur place. Donc, si l'on réduit sa cause, la mer redevient vite plus calme. Si nous réduisons le bruit commercial, les baleines peuvent de nouveau entendre clairement les sons des autres.

Sources liées

- [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 Nouvelles preuves et suite du récit

Ce domaine grandit encore vite aujourd'hui. Chaque année, de nouvelles études paraissent et de nouveaux outils apparaissent. Il y a quelques années, ces choses étaient difficiles à imaginer [1].

Les recherches de Berkeley et de Project CETI en sont un bon exemple. Comme nous l'avons vu, l'équipe a trouvé, dans les codas de cachalots, des sons qui ressemblent à des voyelles. En 2026, une étude suivante a paru dans *Proceedings of the Royal Society B*. Elle dit que ces sons de cachalots suivent des

règles, comme les voyelles humaines. Les preuves sont devenues plus solides. Les cachalots fabriquent des sons d'une manière qui ressemble davantage à celle des humains.

Earth Species Project ouvre une autre voie. Cette organisation a créé une intelligence artificielle appelée NatureLM-audio. C'est le premier grand modèle de langage sonore pour les sons d'animaux. Ce modèle écoute seulement un son et devine de quel animal il vient. Il compte aussi combien d'animaux se trouvent dans un enregistrement. Il reconnaît même, jusqu'à un certain point, les sons d'animaux qu'il n'a jamais appris [2].

En 2026, cette organisation a élargi sa méthode de recherche. Au lieu d'étudier une seule espèce de près, elle rassemble les sons de plusieurs espèces en même temps. Elle développe ainsi une méthode qui analyse ensemble, par ordinateur, les sons de nombreux animaux. En juillet 2026, elle a publié alp-data, un dépôt commun de données pour cela. C'est une base qui permet aux chercheurs du monde entier de partager les mêmes données. Cela ressemble à une bibliothèque où chacun peut emprunter les mêmes livres [3].

En mars 2026, Project CETI a publié deux nouvelles études. L'une décrit en détail la naissance d'un cachalot. C'est la première fois qu'une naissance sous la mer est enregistrée avec autant de précision. L'autre étude montre comment le groupe aide pendant la naissance. Quand la mère met bas, d'autres baleines restent près d'elle et la protègent. Cette étude montre en détail ce type de soin collectif chez un animal qui n'est pas un primate. Certains chercheurs parlent des droits des baleines en observant ces comportements. Selon eux, une vie sociale complexe et des soins entre individus peuvent servir de référence dans le débat sur les droits. Mais il est difficile de transformer directement l'entraide en obligation juridique. Une baleine qui aide une amie et une personne qui respecte la loi, ce n'est pas la même chose [4].

La loi a encore beaucoup de chemin à faire avant de traiter les animaux comme des personnes. Jusqu'ici, les tribunaux de plusieurs pays ont rarement reconnu aux animaux le droit d'agir en justice. Une décision rendue en 2022 par un tribunal de l'État de New York, aux États-Unis, en est un exemple. Une association de défense des animaux a intenté une action pour l'éléphante Happy. Elle demandait qu'on la libère de son enclos étroit. Le tribunal a refusé. Il a expliqué qu'un système créé pour les humains ne pouvait pas être appliqué tel quel à un éléphant. Ainsi, connaître les faits et changer la loi sont deux choses différentes [5].

Toutes ces preuves montrent la même direction. Les sons des animaux ne sont pas du bruit. Ce sont des communications qui portent du sens. Plus les preuves s'accumulent, plus il devient difficile pour la loi de laisser les animaux au rang de choses. La science avance la première. La loi la suit lentement. L'intelligence artificielle devient peu à peu l'interprète qui porte la voix des animaux devant les tribunaux. Traduire la vie de la Terre rejoint alors le travail qui établit leurs droits.

Sources liées

- [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/Jun22/52opn22-Decision.pdf>

Épilogue: L'Internet interespèces et une Terre de coexistence

E.1 Les clics du cachalot et l'alphabet des grands fonds

Au fond de la mer, il fait toujours noir. À plusieurs milliers de mètres sous la surface, là où la lumière du soleil n'arrive pas, c'est comme une nuit noire. Dans cette obscurité, le cachalot voit le monde avec le son. Il envoie depuis son front de petits sons secs. On appelle chacun de ces sons courts un clic (click) [3].

Le cachalot produit plusieurs clics à la suite. On appelle ce groupe de sons enchaînés une coda (coda) [2, 3]. Un clic est un seul son. Une coda est un paquet de sons formé de plusieurs clics. Une coda contient environ 3 à 40 clics. Pour l'oreille humaine, une coda ressemble seulement à une suite de petits sons secs tous pareils. Pendant longtemps, les humains ont pris ces sons pour du bruit sans sens [2].

Mais l'intelligence artificielle a réécouté ces sons. Une équipe appelée Project CETI s'en est chargée. CETI réunit des chercheurs de plusieurs pays qui veulent comprendre la parole des baleines. La chercheuse Pratyusha Sharma et le professeur Jacob Andreas ont analysé les sons. Le docteur Shane Gero, qui observe depuis longtemps les cachalots de Dominique, a travaillé avec eux [1].

L'équipe a rassemblé et étudié une très grande quantité de sons. Elle a examiné une par une 8 719 codas de cachalots. Elle y a trouvé quatre éléments de base qui composent les codas. Il y a le rythme, la vitesse, et de petits clics ajoutés. Quand on change ces éléments de différentes façons, on obtient des codas différentes.

L'intelligence artificielle a trouvé des règles dans des clics qui semblaient identiques. Le cachalot allonge ou raccourcit un peu l'intervalle entre les clics. Il change aussi la vitesse du son. Il ajoute parfois de petits sons avant ou après un groupe de clics. Ces petites différences forment ensemble des structures sonores différentes.

Le cachalot travaille aussi la vitesse du son. Il peut produire la même coda en l'étirant un peu. Il peut aussi la resserrer et la produire plus vite. Cela ressemble au fait d'allonger ou de raccourcir le tempo quand on chante. Ces petits étirements et raccourcissements deviennent aussi une clé qui change le son.

La nouvelle étude est allée plus loin. Les chercheurs ont regardé quelles hauteurs résonnaient fortement dans les codas. Quand un humain prononce les voyelles « a » et « i », certaines zones résonnent aussi plus fort. On retrouve ce genre de zone dans les codas des cachalots. L'équipe CETI a trouvé deux voyelles de baleine de ce type. Elles ressemblent aux sons humains « a » et « i » [4].

Comment le cachalot produit-il ces voyelles ? Le cachalot fabrique les sons avec des lèvres sonores situées à l'intérieur de sa tête. Puis il façonne ces sons avec des poches d'air dans sa tête. Cela ressemble à l'humain qui produit un son avec sa gorge puis change la forme de sa bouche pour faire des voyelles. L'endroit qui produit le son et l'endroit qui le façonne sont donc séparés.

L'équipe a expliqué que le cachalot produit environ 156 codas différentes. Ces codas ne sont pas des lettres isolées. Ce sont des groupes de sons faits avec quelques éléments, comme le rythme et la vitesse. Même dans les grands fonds, il existe donc une manière de créer beaucoup de sons avec peu de matériaux. Mais on ne sait pas encore ce que chaque coda veut dire. L'équipe a seulement trouvé la structure des sons.

Une même coda change un peu selon les groupes. Cela ressemble aux accents régionaux chez les humains. Dans un même groupe, les cachalots superposent leurs sons et se mettent au même rythme.

Quand une baleine produit un son, une autre ajoute aussitôt le sien. Ainsi, les cachalots construisent ensemble le fil de la conversation.

Certaines codas servent comme des étiquettes de nom. Les cachalots vivent en grands groupes familiaux. Chaque groupe a des codas qu'il utilise souvent. Quand on entend ces codas, on peut savoir de quel groupe vient la baleine. Cela ressemble au fait de deviner la région d'une personne grâce à son accent.

Ce qui compte, c'est que de petites différences s'accumulent pour former le son. Le cachalot combine quelques éléments sonores de plusieurs façons. Selon la manière de les combiner, des codas différentes apparaissent. Cela ressemble aux humains qui fabriquent beaucoup de mots avec quelques lettres. Même dans les grands fonds, cette organisation était cachée.

Pour capter ces sons, il faut un bon appareil d'enregistrement. L'équipe CETI a fabriqué un petit enregistreur que l'on fixe sur le corps des baleines. Cet appareil enregistre les sons des baleines et aussi ce qui se passe sous l'eau. Il faut regarder dans quelle situation un son est produit pour en comprendre le sens. C'est seulement après avoir obtenu ce nouvel outil, l'intelligence artificielle, que les humains ont vu l'alphabet des grands fonds.

Sources liées

- [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 Dépassez l'idée que seuls les humains parlent

Pendant longtemps, les humains se sont crus à part. Ils pensaient être les seuls sur Terre à parler et à partager du sens. Les sons des animaux qui ne ressemblaient pas à la parole humaine étaient vus comme du bruit. Cette idée, selon laquelle seuls les humains seraient spéciaux, a duré longtemps [1].

Cette croyance était solide comme un vieux mur. À cause de ce mur, les sons des animaux sont restés enterrés dans l'obscurité. Les humains n'ont pas voulu écouter ce que pensaient les autres êtres vivants près d'eux. Nous ne savons pas combien d'histoires sont passées sans être entendues.

Ce mur a grandi à partir d'une habitude : mesurer les animaux avec la façon de parler des humains. Les humains pensaient qu'il fallait des mots et des phrases pour parler. Ils repoussaient les sons qui n'entraient pas dans ce cadre, comme s'ils n'avaient pas de sens. Mais les sons des animaux n'ont aucune raison d'entrer dans le cadre de la parole humaine. Ils viennent de corps différents et de vies différentes.

L'intelligence artificielle abat peu à peu ce mur [3, 4]. Elle range les sons des animaux sur une carte de pensée. Les chercheurs appellent cette carte un espace latent (latent space). Sur cette carte, les sons semblables se placent près les uns des autres. Les sons différents se placent loin les uns des autres [4].

Les chercheurs ont placé côte à côte la carte de la parole humaine et la carte des sons animaux. Ces deux cartes n'étaient pas étrangères l'une à l'autre. Quand on les superposait en faisant correspondre leurs

formes, des zones semblables apparaissaient. Certains mots humains et certains signaux animaux se retrouvaient parfois au même endroit.

Une méthode intéressante se cache ici. Pour superposer les deux cartes, on n'a pas besoin d'un tableau de traduction préparé à l'avance. Les chercheurs regardent seulement la forme des deux cartes et cherchent les endroits qui correspondent. Cela ressemble à deux feuilles de papier posées l'une sur l'autre contre une fenêtre pour faire coïncider les motifs. Les chercheurs testent maintenant cette superposition par les formes.

Certains chercheurs dirigent ce type d'étude. Le linguiste américain Gašper Beguš cherche les règles des sons animaux avec l'intelligence artificielle. Ces chercheurs ne forcent pas les règles de la parole humaine sur les animaux. Ils attendent que les règles apparaissent d'elles-mêmes dans la forme des sons.

Ici, une nouvelle idée grandit. Les animaux ne sont pas des êtres qui attendent la parole humaine. Chaque animal sent son propre monde et le dessine par le son. Le monde qu'un animal sent avec ses propres sens s'appelle l'Umwelt (Umwelt). La chauve-souris voit le monde par le son. Le chien lit le monde par l'odeur. La baleine a son monde de baleine. L'oiseau a son monde d'oiseau.

Cette méthode de superposition est utile dans les situations difficiles. Entre les humains et les animaux, il n'existe pas de dictionnaire préparé à l'avance. Il n'existe pas non plus de tableau qui associe les mots deux par deux. Pourtant, l'intelligence artificielle essaie de trouver des correspondances seulement avec la forme des deux cartes. Elle peut devenir un nouveau pont entre des mondes différents.

Nous ne devons donc pas placer les sons des animaux sous la parole humaine. Nous avons besoin de respecter leurs pensées et leurs rythmes au même niveau. C'est une attitude qui considère les sons de toute vie comme précieux. Pour devenir une famille de la Terre, les humains doivent d'abord avoir cette attitude [2].

Sources liées

- [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 L'Internet entre espèces qui relie la vie sur Terre

Les humains ont Internet. Internet est un réseau qui permet aux personnes éloignées d'échanger des nouvelles. Mais les humains ne sont pas les seuls à vivre sur Terre. Les baleines, les éléphants, les oiseaux, les abeilles et même les arbres vivent avec nous. Il existe un rêve : relier tous ces êtres vivants. On appelle ce rêve l'Internet interespèces (Interspecies Internet) [1].

Ce rêve est apparu en 2013. Vint Cerf, qui a participé à la création d'Internet, et le musicien Peter Gabriel l'ont porté les premiers. Les chercheurs Neil Gershenfeld et Diana Reiss les ont rejoints. Ils ont dit qu'Internet ne devait pas rester réservé aux humains. Leur idée était d'ouvrir aussi ce réseau aux animaux.

L'Internet interespèces ne consiste pas seulement à traduire les sons des animaux avec une application. Il désigne un grand réseau où toute la vie sur Terre partage des nouvelles et des dangers. Comme les nerfs humains relient toutes les parties du corps, ce réseau relie toute la vie de la planète.

En réalité, des morceaux de ce lien existent déjà dans la nature. Beaucoup d'animaux écoutent les cris d'alerte d'autres espèces. Quand un suricate annonce un danger, d'autres animaux à côté comprennent ce son et fuient avec lui. Les scientifiques ont observé cette écoute chez plus de 70 espèces animales. Des animaux différents partagent donc la même information sur un danger [2].

Les dauphins montrent aussi ce partage. Quand deux groupes de dauphins se rencontrent dans la mer, ils échangent des sifflements. Ces sifflements, différents pour chacun, servent comme des étiquettes de nom. Les dauphins vérifient par le son qui est qui. Ce chemin sonore invisible forme déjà un réseau [3].

Ce réseau ne relie pas seulement les sons. Les arbres et les herbes envoient aussi des signaux. Quand l'eau manque, des plantes comme la tomate et le tabac produisent des sons que l'oreille humaine n'entend pas. Une plante stressée produit environ 30 à 50 petits sons secs par heure. L'intelligence artificielle peut rassembler ces sons et lire l'état des plantes [4].

L'intelligence artificielle aide à relier ces morceaux. Les nouveaux réseaux de communication peuvent aussi aider. Un réseau appelé 6G peut faire rebondir les ondes radio sur les objets pour lire les mouvements autour de lui. Cette idée, qui transforme un réseau de communication en capteur, s'appelle réseau comme capteur (network as a sensor). Mais relier cette technique à un réseau de communication animale reste seulement une possibilité [5].

Si nous utilisons bien ce réseau, nous pouvons protéger la nature. Imaginons que des arbres commencent à être coupés sans autorisation dans une forêt tropicale. Les mouvements des oiseaux qui se dispersent et les signaux anormaux des plantes passent par le réseau et arrivent au centre. Les humains reçoivent vite cette nouvelle et peuvent arrêter cela. L'Internet interespèces peut devenir non pas un outil de surveillance, mais un outil pour vivre ensemble. Il peut devenir un grand réseau nerveux où les êtres vivants de la Terre partagent leurs dangers et prennent soin les uns des autres.

Sources liées

- [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 Ce que les nouveaux outils ont révélé : l'intelligence artificielle apprend la parole de la nature

L'intelligence artificielle a aussi avancé dans l'apprentissage des sons animaux. Earth Species Project a présenté une nouvelle intelligence artificielle. Son nom est NatureLM-audio. C'est la première grande intelligence artificielle sonore créée pour les sons animaux. Elle a été publiée pour la première fois en novembre 2024, puis présentée officiellement à la conférence internationale ICLR en 2025 [1].

Cette intelligence artificielle fait plusieurs tâches en même temps. Elle écoute un son et devine de quel animal il vient. Elle distingue aussi si l'animal est un petit ou un adulte. Elle classe les types de sons et ajoute des descriptions. Elle compte aussi combien d'animaux se trouvent dans un enregistrement.

Ce qui frappe, c'est qu'elle réussit des tâches que les humains ne lui ont pas enseignées. En général, une intelligence artificielle ne peut faire que ce qu'elle a appris. Mais celle-ci a acquis de nouvelles capacités par elle-même en apprenant beaucoup de sons. Compter le nombre d'animaux n'était pas une tâche apprise dès le départ. À force d'écouter beaucoup de sons, elle a fini par savoir le faire seule.

Earth Species Project étudie ensemble la parole de plusieurs animaux. Aujourd'hui, il examine de près les sons des corbeaux, des bélugas et des éléphants. Il rassemble des données avec des biologistes du monde entier. La méthode de cette équipe est de ne pas regarder une seule espèce, mais plusieurs espèces côte à côte.

En juillet 2026, un nouvel outil est apparu. Son nom est alp-data. C'est un outil qui rassemble au même endroit des données sonores de plusieurs animaux. Il contient 35 jeux de données sur les chants d'oiseaux, les sons de baleines et les sons d'insectes. Désormais, les chercheurs peuvent obtenir les sons de plusieurs animaux depuis une seule porte d'entrée [2].

Earth Species Project appelle ce mouvement le traitement du langage animal (Animal Language Processing). C'est une méthode pour étudier la communication animale avec l'intelligence artificielle. Elle ne force pas les règles de la parole humaine sur les animaux. Elle reconnaît d'abord que chaque animal peut avoir sa propre manière de communiquer. Ensuite, l'intelligence artificielle cherche les règles dans les données [3].

Cette équipe a aussi interrogé les gens sur leurs pensées. En 2025, elle a posé des questions à 1 057 personnes dans 67 pays. Beaucoup de personnes ont accueilli favorablement l'idée de comprendre la parole des animaux. Mais elles ont aussi exprimé une inquiétude : il ne faut pas traiter les animaux n'importe comment. Plus la technique grandit, plus la prudence grandit aussi [4].

Ces nouveaux outils changent la façon de travailler des chercheurs. Avant, un chercheur étudiait longtemps un seul animal. Les données étaient aussi dispersées, chacune dans un format différent. Maintenant, on voit les données de plusieurs animaux ensemble, par une seule porte d'entrée. Les recherches séparées se rejoignent dans un grand mouvement commun.

Le prochain objectif est un essai de lecture de sons. C'est un essai où l'on parle aux animaux avec des sons créés par l'intelligence artificielle. Ce projet a surtout été présenté par Project CETI. Earth Species Project met plutôt ses efforts dans la collecte et l'étude des données, au lieu de faire écouter des sons sans prudence. On n'a pas encore vérifié si les animaux comprennent les sons créés par les humains [5].

Sources liées

- [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/>

E.5 Écouter avec prudence : des outils responsables et une promesse de coexistence

Un nouvel outil est puissant. Un outil puissant peut faire du mal s'il est mal utilisé. L'intelligence artificielle peut maintenant reproduire les sons des animaux presque à l'identique. Si l'on utilise mal cette force, on peut tromper les animaux. Imaginons que de faux sons servent à attirer des animaux sauvages. Des braconniers pourraient partir chasser avec de faux sons [1][2].

Nous devons donc respecter trois promesses.

Premièrement, la promesse de ne pas intervenir sans raison. Les groupes d'animaux ont construit leur propre culture sonore pendant très longtemps. Les humains ne doivent pas troubler cet ordre avec de faux sons, pour faire du commerce ou pour dominer. L'intelligence artificielle n'est pas une arme pour voler les sons des animaux. Elle doit être un outil pour protéger leur culture selon leur propre volonté.

Pourquoi cette promesse est-elle nécessaire ? Pendant longtemps, les humains ont seulement voulu vaincre la nature. Le pouvoir de comprendre les sons peut lui aussi devenir un pouvoir pour contrôler l'autre. Si l'on entre sans prudence dans un groupe, la vie des animaux se dérègle. Donc, plus on sait, plus on doit aussi savoir reculer.

Deuxièmement, la promesse de laisser une marque dans les sons. Des chercheurs proposent de placer une marque invisible dans les sons d'animaux créés par l'intelligence artificielle. Cette marque s'appelle un filigrane audio (audio watermark). Comme l'image cachée dans un faux billet, l'oreille humaine ne l'entend pas, mais une machine la reconnaît. On peut alors distinguer les sons vrais des sons faux [3].

Cette marque sert aussi à protéger les données. Les chercheurs laissent longtemps des enregistreurs dans la mer et dans les forêts. Collecter seulement les sons, sans intervention humaine, s'appelle l'observation acoustique passive. Si de faux sons se mélangent aux enregistrements, ces données sont salies. Si l'on place une marque dans les sons, on peut retirer les faux sons et garder les données propres.

Troisièmement, la promesse de décider ensemble. Le programme MOTH de la faculté de droit de New York University mène ce travail. MOTH est un groupe qui étudie les droits des êtres vivants non humains. Les données trouvées par les machines doivent rencontrer l'ancienne sagesse des peuples autochtones qui protègent la mer. Le savoir sur la nature obtenu par une longue observation autochtone s'appelle le savoir écologique traditionnel. Les règles ne doivent pas être fixées seulement selon l'intérêt humain.

Ces trois promesses viennent du même esprit. C'est l'esprit de ne pas traiter l'autre comme on veut, même quand on a reçu une nouvelle force. Les humains ont longtemps vu la nature seulement comme un outil. Il est temps maintenant de regarder cette habitude en face. La nouvelle force appelée intelligence artificielle n'est pas un sujet de fierté. C'est une lourde responsabilité.

Si nous comprenons le langage des cachalots, nous pouvons aussi penser à de nouveaux droits. Des chercheurs parlent du droit des baleines à ne pas être harcelées. Ils parlent aussi de leur droit à protéger leur propre culture. En 2024, des chefs autochtones de Nouvelle-Zélande et des îles Cook ont déclaré que les baleines devaient être vues comme des êtres ayant des droits semblables à ceux des humains. Cela signifie qu'une baleine n'est pas un objet que les humains peuvent utiliser comme ils veulent.

Ce livre raconte comment on traduit la vie de la Terre. L'intelligence artificielle a commencé à déchiffrer, un par un, les clics des cachalots, le chant des oiseaux et les basses vibrations des éléphants.

Traduire ne veut pas dire vaincre l'autre. Cela veut dire apprendre à le connaître. L'intelligence artificielle est un pont qui traduit la vie de la Terre. Si nous traversons ce pont avec prudence, nous pouvons créer une Terre où nous vivons avec les autres formes de vie voisines. La traduction ne s'arrête pas ici. Elle vient seulement de commencer.

Sources liées

- [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>

Soutenir ce livre de la Bibliothèque IA

Si ce livre vous a été utile, vous pouvez soutenir les prochaines traductions, les éditions publiques et les projets ouverts de connaissance sur l'intelligence artificielle via PayPal.

PayPal

<https://paypal.me/kimkjkj>



Scannez le code QR ou ouvrez le lien ci-dessus.

Kim Kyung-jin, avocat