

FUTURA

Votre voix est-elle unique au monde ?

Podcast écrit et lu par Emma Hollen

(Liens de l'épisode en fin de document)

N.B. Les indications sonores sont notées de deux manières différentes : sans italique, pour indiquer que la partie entre crochets est lue par la narratrice et s'accompagne d'un bruitage correspondant, ou en italique, quand il s'agit d'un bruitage seul.

Quelle personne imaginez-vous derrière cette voix que vous entendez ? Une femme, un homme, ni l'un ni l'autre ? Quelqu'un de corpulent ou de mince, de grand ou de petit, de timide ou d'extraverti ? *[Avec une voix renfrognée :]* Saurez-vous deviner quelle expression je fais quand je parle comme ça ? *[Avec un sourire :]* Ou alors comme ça ? La voix est une partie essentielle de notre personnalité et elle peut révéler énormément de choses sur celui ou celle qui parle. Alors dans ce nouvel épisode d'INFRA, je vous propose que l'on commence à décrypter les grands fondamentaux du son, pour ensuite pouvoir plonger dans des sujets plus approfondis et variés. Et quoi de mieux, pour commencer, que de comprendre comment fonctionne la voix humaine ? Cette semaine, on va parler de confrontation de la voix, de théorème de Bernoulli, de jumeaux homozygotes, de castration et d'imitation. Avant de nous engager sur cette voie, pensez à vous abonner pour ne pas manquer les prochains rendez-vous et n'hésitez pas à laisser un commentaire pour dire ce que vous avez pensé de ce podcast à la fin de l'épisode.

[Générique : une musique percussive rythmée par des sons. Un réveil sonne. Une main tape dessus et un homme se met à bâiller en sortant de son lit. Il allume un robinet, se brosse les dents, et se rase. Puis il se verse des céréales tandis qu'un micro-onde sonne en arrière-plan. Une voiture démarre, un chien aboie et une sonnette de vélo résonne, puis la journée s'emballa et les sons se multiplient jusqu'à la dernière note, ponctuée par un feu d'artifice.]

Lorsque la photographie est apparue au début du XIXème siècle, il n'a pas fallu attendre plus d'une dizaine d'années pour que les selfies commencent à fleurir. On attribue le tout premier à Robert Cornelius, un Hollandais émigré aux États-Unis, qui en 1839, capture le premier autoportrait photographique dont nous ayons encore trace aujourd'hui *[un flash antique crépite]*. Depuis, inutile de préciser que la pratique s'est embrasée, devenant particulièrement répandue – et de plus en plus décomplexée – à partir des années 2000, avec l'ajout des premières caméras à l'avant des téléphones. Mais malgré l'apparition du phonautographe, le premier appareil d'enregistrement sonore, à peine 30 ans après la photographie, en 1857, *[une voix grésillante et toujours compréhensible d'enfant chante Au Clair de la Lune]* jamais, semble-t-il, notre voix n'a fait l'objet d'un culte similaire. Et pour cause. Si vous avez déjà entendu votre voix sur un portable ou un ordinateur, il y a plus

d'une chance sur deux que vous ayez ressenti... un profond malaise. Et je rassure tout de suite celles et ceux qui se sentent concerné·e·s, c'est un phénomène on ne peut plus normal chez les personnes pourvues de cordes vocales et d'une audition fonctionnelles. Souvent, quand on écoute un enregistrement de sa propre voix, elle nous semble... eh bien gênante, un peu trop aiguë, à la fois familière mais aussi source d'une inquiétante étrangeté. En fait, disons-le clairement : à part les professionnels de la rhétorique, les acteurs, doubleurs et autres influenceurs, il n'y a pas grand-monde sur Terre qui aime sa voix. Ce phénomène psychologique a même un nom : on appelle ça la « *voice confrontation* » en anglais, la confrontation à la voix. Il se définit tout simplement comme le fait qu'une personne [*avec une voix téléphonique* :] n'apprécie pas d'entendre le son de sa voix, généralement sur un enregistrement. Petite précision, pour les personnes dont le timbre ne correspond pas à leur genre, on parlera plutôt de dysphorie de voix ; c'est un autre sujet pour le coup. Il y a deux mécanismes qui sous-tendent la *voice confrontation*. D'abord, vous avez peut-être remarqué que votre voix vous semble plus aiguë quand vous entendez un enregistrement. C'est parce que lorsque vous parlez, [*la voix résonne comme en écho* :] le son émis par vos cordes vocales voyage non seulement dans l'air pour être perçu par vos oreilles, mais également [*une voix plus grave et étouffée* :] à travers votre crâne. Car oui, on en parlera sûrement dans un prochain épisode, l'être humain perçoit aussi les sons, et même les sons extérieurs, à travers les os de sa tête. C'est tout à fait négligeable quand le bruit vous parvient à plusieurs mètres de distance, mais quand ce sont vos cordes vocales qui s'expriment, votre crâne devient une vraie boîte de résonance. Vous percevez plus de nuances acoustiques et votre timbre vous semble plus profond que si vous perceviez votre voix depuis l'extérieur de votre corps, sans ces vibrations additionnelles. C'est pour ça qu'un enregistrement vous paraît plus aigu. La deuxième raison de cette *voice confrontation* est plus psychologique. Lorsque vous êtes engagé·e dans l'acte de la parole, votre cerveau est généralement trop occupé pour analyser en détails tout ce que vous dites et comment vous le dites. Ce n'est qu'avec le recul de l'enregistrement que votre oreille devient alors capable de percevoir l'insécurité, la tristesse ou encore la colère qui peuvent habiter votre voix. Vos hésitations deviennent plus remarquables, et cette personne timide derrière le micro peut clasher avec la représentation que vous vous êtes faite de vous-mêmes. Alors, vous vous dites peut-être : « *oui, non mais moi c'est bon, c'est pas pareil hein, j'ai vraiment une voix toute pourrie* ». Eh bien peut-être, je ne suis pas là pour vous contredire. Sauf qu'une étude de 2013 démontre que cette « confrontation » est plus piégeuse qu'il n'y paraît. Les chercheurs ont présenté plusieurs échantillons de voix à des participants en leur demandant de les noter. Ce qu'ils ne leur avaient pas dit, c'est que des enregistrements des participants eux-mêmes avaient été glissés dans le mix et en analysant les résultats, ils se sont rendu compte que les gens avaient tendance à mieux noter leur propre voix que celles des autres, même s'ils ne la reconnaissaient pas. Ainsi, même si vous êtes persuadé·e d'avoir la pire voix de la Terre, il n'est pas complètement impossible que votre cerveau ait un autre avis sur la question. Tant mieux d'ailleurs, il serait aussi dommage de détester sa voix qu'il est délétère de détester son corps. Alors la prochaine fois que vous vous entendrez, essayez d'apprécier la sympathie et le charme qui émergent de vos cordes vocales. Ça peut être un challenge les premières fois, mais je vous promets qu'avec un peu d'entraînement, on peut apprendre à apprécier sa propre voix. Croyez-en une podcastrice avec cinq ans de pratique.

Bon, venons-en maintenant au cœur du sujet : la voix, c'est quoi ?

On qualifie de voix l'ensemble des sons produits par l'être humain à l'aide de son canal vocal. Elle inclue [*la parole, la chanson, les cris, les rires, les pleurs*] et même l'éclat de voix

de votre grand-mère quand elle éternue [*un éternuement tonitruant*]. Quand on parle de la voix, on pense bien sûr aux cordes vocales, mais en réalité c'est tout un orchestre qui participe à cette symphonie. Elle prend naissance au niveau du diaphragme, ce muscle en forme de parachute, responsable du hoquet [*hic !*] mais aussi et surtout de la respiration. Avec l'aide d'autres muscles situés au niveau des côtes, des abdominaux ou encore du cou, c'est lui qui dirige le déploiement et la contraction des poumons. Parce que les poumons ne sont pas des muscles. Ils sont plutôt à voir comme des espèces d'éponges qui dépendent du diaphragme et de ses acolytes pour fonctionner. Les muscles alentour se contractent et compressent l'intérieur de la cage thoracique à l'[expiration], puis ils se relâchent pour permettre aux poumons de reprendre leur volume initial [*une inspiration*]. Au moment de vocaliser, notre diaphragme et ses compères, donc, se contractent en pressant sur les poumons. Ces derniers [expulsent l'air] qu'ils contiennent le long de la carène trachéale et de la trachée, jusqu'aux portes des cordes vocales, qui sont alors fermement closes [*une porte qu'on claque avec un écho*]. Incapable d'aller plus loin, l'air s'accumule sur place, la pression augmente [*un ballon qui gonfle et se distend*] jusqu'à ce que la glotte, l'espace qui sépare les cordes vocales, n'y tienne plus et s'ouvre pour laisser passer l'expiration [*un souffle*]. À cause de la pression qui s'est accumulée, cette expiration prend de la vitesse au moment de sa libération en créant un appel d'air derrière elle. C'est ce qu'explique le théorème de Bernoulli : lorsqu'un fluide accélère, il entraîne une diminution de pression. Vous pouvez tester cette expérience chez vous. Suspendez deux ballons de baudruche avec des ficelles depuis votre plafond et [utilisez une paille pour aspirer l'air entre les deux]. Cet air accélère vers votre bouche, créant une baisse de pression derrière lui et les deux ballons se rapprochent, attirés par le « pseudo-vide » qui s'est créé entre eux. C'est exactement la même chose qui se passe avec les cordes vocales. Lorsque l'air est libéré par la glotte, il accélère vers votre bouche en créant une diminution de pression dans son sillage, et les cordes vocales se rapprochent, comme nos ballons de baudruche. Et comme vos poumons continuent d'expulser de l'air, le cycle recommence des centaines de fois par seconde : pression, libération, accélération, diminution de pression, glotte qui se referme, et cætera, et cætera. Quand on regarde une visualisation de ce phénomène, on s'aperçoit que l'air voyage comme une vague à travers la glotte. Et vous l'avez peut-être déjà deviné, c'est ça qui fait vibrer vos cordes vocales. Ça y est, vous produisez de la voix ! [*Une voix masculine produit un « aaah » continu.*] Toujours au niveau de votre larynx, vous pouvez étirer vos cordes vocales pour produire un son plus ou moins aigu, un peu comme avec une corde de guitare [*toujours avec un « aaah », la voix monte progressivement d'un son grave vers un son aigu*], et pour monter le volume, il faudra augmenter la pression de l'air au niveau de votre glotte [*la voix vocalise de plus en plus fort*]. Une fois qu'elle a pris naissance dans votre gorge, cette vibration sonore va remonter jusqu'à votre bouche où votre palais mou, vos dents, vos lèvres, votre mâchoire, et votre langue vont travailler de concert pour moduler les sons et les mots que vous souhaitez produire. On y consacrera tout un épisode séparé parce qu'il y a beaucoup de choses à dire à ce sujet. Ce sera l'occasion de découvrir l'immense répertoire des sons créés par l'humanité pour communiquer et de parler de quelques langues aux sonorités fascinantes. [*Un homme parle en Khoisan, une langue dite « à clics ».*]

Avant de poursuivre sur le sujet qui nous intéresse aujourd'hui, je vous propose de faire une courte pause pour digérer toutes ces informations anatomiques. On se retrouve dans trente secondes maximum pour essayer de comprendre pourquoi votre voix est unique au monde.

[Coupure pub.]

Avec désormais 8 milliards de personnes sur Terre, il ne serait pas complètement aberrant de se dire qu'on a sûrement un sosie vocal quelque part dans le monde. Et pourtant, votre voix est aussi unique que vos empreintes digitales. En fait, c'est assez logique, quand on y réfléchit. Parce que votre signature vocale ne dépend pas seulement de la forme de votre larynx. Si l'on regarde simplement du côté du corps, votre taille, votre posture, votre carrure, votre âge, votre façon de respirer, votre sexe biologique, l'aspect et la taille de vos poumons, de votre buste, de votre cou, de votre trachée, de vos cordes vocales, de votre langue, de votre cavité nasale, de votre palais, de vos dents, de votre mâchoire... bref, tout cela contribue à rendre votre voix unique. Et les chances que quelqu'un combine exactement les mêmes paramètres physiques que vous sont infinitésimales. En 2019, les chercheurs ont même réussi à créer un algorithme, Speech2Face, qui était capable de prédire votre apparence à partir d'un simple enregistrement de votre voix. Bon d'accord, je ne parle pas de portraits-robots dessinés au poil de moustache près. L'IA avait clairement des limitations et des angles morts, mais elle avait quand même réussi à établir, dans une bonne partie des cas, le sexe, l'âge, la forme du nez ou encore la corpulence des personnes qui lui étaient présentées. Et dans une certaine mesure, vous êtes sûrement capables de faire pareil. Par exemple, vous savez que de façon générale, une voix masculine a tendance à être plus grave qu'une voix féminine. C'est parce qu'en moyenne, les cordes vocales des personnes biologiquement mâles sont plus longues que chez celles de sexe féminin et peuvent donc vibrer à des fréquences plus basses. Vous pouvez faire un test simple chez vous avec deux élastiques de longueurs différentes. En tirant dessus pour faire varier la tension, vous verrez que l'élastique le plus court produit des sons plus aigus [*un élastique produit un chtong ! aigu*] et que le plus long résonne à des fréquences plus graves [*un autre élastique produit un chtong ! grave*]. C'est pour la même raison que votre voix est généralement corrélée à votre taille. Par exemple, avec mon mètre quatre-vingt-sept, j'ai une voix qui se situe en moyenne entre 165 et 200 Hz, la moyenne basse des fréquences vocales pour les personnes de mon sexe. Et contrairement à ce qu'on pourrait penser, la pomme d'Adam pour sa part, c'est-à-dire l'angle formé par le cartilage thyroïde qui protège nos cordes vocales, ne semble pas avoir d'impact sur la profondeur de la voix, même si elle est plus proéminente chez les hommes biologiques que chez les femmes. Concernant les personnes intersexes, dont les chromosomes, les gonades ou les taux d'hormones ne permettent pas de leur assigner un sexe binaire, homme ou femme, je n'ai malheureusement trouvé aucune étude sur l'anatomie des cordes vocales, ni même sur la hauteur de la voix. Alors si des personnes intersexes écoutent ce podcast et souhaitent s'exprimer sur ce sujet, n'hésitez pas à laisser un commentaire, à nous écrire via le site, ou sur les réseaux sociaux.

Il y a un autre facteur qui peut vous aider à émettre des hypothèses sur la voix d'une personne : son âge. Vous l'aurez deviné, le timbre d'un enfant est plus aigu tout simplement parce que ses cordes vocales sont plus courtes, en proportion avec son gabarit. [*Une petite fille dit : « Tu sais que la vie, elle est belle. »*] Au moment de la puberté, douce période durant laquelle le corps décide de pousser dans tous les sens, de se couvrir de poils et de boutons – une pensée pour vous si vous êtes dans cette situation – l'ensemble du système vocal (votre larynx, mais aussi vos os crâniens, vos sinus ou l'arrière de votre gorge) traverse sa propre poussée de croissance. Elle est visible chez les deux sexes, mais elle est généralement beaucoup plus marquée chez les personnes de sexe masculin, dont la voix devient notablement plus grave et peut aussi muer plus fréquemment, signe que le cerveau a encore besoin de s'habituer à ces nouveaux changements. Quand la personne a du mal à s'adapter à ces transformations que ce soit physiquement ou psychologiquement, elle peut

alors être atteinte de puberphonie, c'est-à-dire qu'elle garde une voix d'enfant. Voici par exemple un enregistrement d'un jeune garçon après des séances avec un orthophoniste [*une voix grave*] et... avant [*une voix frêle et peu assurée*].

Ce n'est d'ailleurs pas la seule chose qui peut amener un individu de sexe masculin à garder une voix d'enfant à l'âge adulte. On peut aussi tout simplement... lui couper les testicules [*un bruit de métal, comme une épée qu'on sort de son fourreau*]. Alors aux trois personnes qui sont déjà en train de sortir la scie et le scalpel, stop ! Cette pratique est aujourd'hui strictement illégale et de toute façon, je ne vous la recommanderais pas. Mais, entre le XVI^{ème} siècle et jusqu'à la fin du XIX^{ème}, le monde occidental s'est bel et bien pris de passion pour le phénomène des castrats, des chanteurs ayant été châtrés avant la puberté. Cette technique – d'autant plus barbare qu'elle était pratiquée sur des enfants – avait pour avantage de leur permettre de conserver leur voix aiguë tout au long de leur vie, tout en bénéficiant d'une cage thoracique pleinement développée une fois arrivés à l'âge adulte. Et si vous êtes curieux·se d'entendre ce que ça donne et êtes quand même bien tenté·e de tester l'expérience sur votre cousin de 8 ans, je vous renvoie plutôt vers les enregistrements d'Alessandro Moreschi, le dernier castrat de l'histoire, qui meurt en 1922. [*Un extrait de Preghiera de Francesco Paolo Tosti, chanté par Moreschi.*]

Bien. Si la personne n'est pas atteinte de puberphonie, et pour peu qu'on laisse ses gonades tranquilles, la voix d'adulte se stabilise en moyenne dans les deux ou trois ans qui suivent la puberté. Elle peut devenir un peu plus profonde au fil des années, à mesure que la gravité entraîne votre larynx un peu plus en arrière dans votre gorge, mais dans l'ensemble, elle peut demeurer relativement inchangée jusqu'aux premiers signes de ce qu'on appellera arbitrairement « la vieillesse ». Selon votre hygiène de vie, votre historique de santé et votre génétique, elle commencera à s'affaiblir, à trembler ou à s'érailler autour de 50 ans... ou de 80 ans selon les individus. Quoi qu'il en soit, dans tous les cas, les causes de ces changements sont relativement similaires. Chez les personnes atteintes de ménopause, les cordes vocales deviennent plus raides, plus sèches et plus épaisses, menant à une voix plus grave. [*Une voix de femme âgée : « on n'a pas le même corps à 50 ans qu'à 80 ans. »*] Tandis que chez les personnes de sexe masculin, à l'opposé, le timbre a de fortes chances de devenir plus aigu. Les cordes vocales se rigidifient aussi, mais au lieu de s'épaissir, elles s'affinent. La glotte a pour sa part plus de mal à se refermer et laisse passer plus de souffle dans la voix. [*Une voix d'homme âgé : « J'en sais rien, j'ai attendu. »*] Des troubles neurologiques peuvent d'autre part s'ajouter à ces changements physiques et rendre la vocalisation plus chevrotante et difficile. Et bien sûr, il va sans dire que tout au cours de la vie, la voix peut aussi subir des modifications liées à un changement de sexe, un accident, des problèmes de santé, du tabagisme, de la pollution et bien d'autres facteurs. Ainsi, même les jumeaux ont des voix différentes. Ce n'est peut-être pas tout à fait évident aux oreilles d'une personne extérieure, surtout si elle les connaît peu ; *mais* les technologies de reconnaissance vocale modernes, elles, sont capables de distinguer les voix de jumeaux homozygotes, d'après une étude de 2021. Ainsi, ce n'est pas juste une affaire de génétique ou d'apparence : de nombreux facteurs impactent votre voix et contribuent à la rendre unique. J'en profite pour signaler que les scientifiques ignorent encore si l'origine ethnique a vraiment une influence sur les dimensions de vos cordes vocales. Les études à ce sujet restent rares et le débat demeure donc ouvert.

Mais on pourrait aussi parler, au-delà de l'anatomie, de la façon dont vous utilisez votre voix, et de l'impact sur la parole de votre entourage, de votre psychologie, des normes sociales au milieu desquelles vous évoluez ou encore de votre langue, maternelle ou secondaire.

Ce sont toutes ces variables, tous ces petits détails, que les imitateurs et imitatrices identifient, décortiquent et s'approprient pour recréer la voix d'une autre personne. Tandis que leur texte reproduit les tournures et les tics de langage les plus caractéristiques de leur victime, leur corps entier et leur cerveau participent à une véritable transformation physique qui les aide à parachever l'illusion. Ils vont jouer sur la position de leur larynx et la tension de leurs cordes vocales, ou encore modifier l'angle du cartilage thyroïde et l'ouverture de leur pharynx pour créer une voix plus ou moins riche et profonde. Mais tout ne se passe pas seulement au niveau des cordes vocales. Vous verrez souvent ces artistes reprendre les expressions faciales des personnes qu'ils imitent. Au-delà de l'effet comique, c'est aussi parce qu'en étirant la bouche, en fronçant le nez, ou en crispant la mâchoire par exemple, ils tentent de reproduire au plus près l'anatomie de leur sujet. Ça ne se limite d'ailleurs pas au visage, puisque ces acteurs auront aussi tendance à modifier leur posture, à rentrer les épaules ou à courber le dos. Une étude de 2008 démontre d'ailleurs que l'imitation active autant les régions cérébrales impliquées dans la parole que celles impliquées dans la visualisation mentale et la perception de son propre corps. En imitant la configuration corporelle de leurs sujets, ces artistes se garantissent donc la meilleure chance de reproduire une voix la plus identique possible. Ah, et bien entendu, être bon imitateur·ice, c'est aussi avoir une bonne oreille. Et justement, on va voir que ce n'est pas le cas chez tout le monde. Après une dernière pause, on parlera de phonagnosie, de dyslexie et de deepfakes.

[Coupure pub.]

Vous avez peut-être déjà entendu parler de prosopagnosie. Ce trouble neurologique touche 2 à 3 % de la population et il a fait les gros titres l'année dernière lorsque Brad Pitt s'en est avoué atteint. Les patients concernés ont de grandes difficultés voire sont incapables de reconnaître les visages autour d'eux, même lorsqu'ils appartiennent à leurs propres enfants ou à leur partenaire. Le plus souvent, ces personnes sont obligées de s'appuyer sur une coupe de cheveux, un accessoire récurrent comme une paire de lunettes, ou la voix de leur interlocuteur pour parvenir à le reconnaître. Mais que se passe-t-il lorsque même sa voix vous échappe ? C'est ce qui arrive aux gens atteints de phonagnosie, un trouble a priori tout aussi fréquent qui vous empêche, cette fois, d'identifier la voix de vos proches et encore plus, celle de votre banquière ou de votre boulanger. Si ça vous paraît beaucoup moins gênant ou en tout cas plus anodin que la prosopagnosie, imaginez regarder un film et ne plus savoir qui parle dès que la caméra s'éloigne des acteurs, ou sursauter en entendant votre moitié vous parler depuis l'autre pièce. Et si je cite ces deux troubles en parallèle, eh bien, c'est parce que la mémoire des visages pourrait bien être étroitement corrélée avec celle des voix. Une étude de 2021 démontre que les super reconnaisseurs de visages (un groupe de personnes avec une appellation maladroitement traduite de l'anglais) seraient aussi des super reconnaisseurs de voix. Des profils hautement recherchés dans les milieux de la sécurité. Une autre expérience, menée en janvier 2023, démontre pour sa part que le gyrus fusiforme, impliqué dans le traitement d'indices visuels, s'active lorsque l'on essaye de reconnaître quelqu'un uniquement grâce à sa voix. La vision et l'audition semblent donc intimement liées dans la reconnaissance vocale, même si ce n'est, bien sûr, pas le seul critère déterminant. Il va sans dire que les personnes non-voyantes et entendantes, qui s'appuient beaucoup plus sur leurs oreilles pour interagir avec le monde, battent les sujets voyants à tous les coups lorsqu'il s'agit de reconnaître une voix.

Un autre critère déterminant, mis en évidence par les scientifiques, est la compétence linguistique : la capacité à comprendre les règles implicites de sa propre langue. C'est comme ça que des chercheurs du MIT ont démontré que les personnes dyslexiques avaient plus de mal à reconnaître les voix. Et si vous apprenez une nouvelle langue, il n'est pas impossible que vous soyez confronté·e au même problème.

Malgré toutes ces limitations, pendant longtemps, les humains ont été bien meilleurs que les IA pour identifier la personne derrière la voix, en particulier si cette voix appartenait à un proche. Mais avec l'explosion du *deep learning*, en seulement quelques années, les machines ont non seulement rattrapé leur retard, mais elles ont tout simplement pulvérisé toute concurrence. Comme je l'ai mentionné plus tôt, en 2021, des technologies de reconnaissance vocale étaient déjà capables de distinguer des vrais jumeaux. Mais on peut même remonter plus loin, en 2013, alors que des chercheurs utilisaient l'IA pour distinguer les cris d'individus au sein d'une meute de loups à Yellowstone. [*Des hurlements de loups.*] En dépit de ces progrès, il est important de noter que les technologies de reconnaissance vocale, comme beaucoup d'autres, sont porteuses de biais raciaux importants. Comme les modèles sont entraînés avec une majorité d'enregistrements produits par des individus blancs, la précision de leur analyse est meilleure pour les caucasiens que pour les personnes d'origine latine ou africaine par exemple.

Mais bon... généralement mis de côté, ces problèmes n'ont pas pour autant mis un frein à la multiplication des innovations dans ce domaine. Dix ans après l'expérience de Yellowstone, la biométrie vocale est en passe de devenir un business florissant. Web, banque, aéroports, télécommunications, sécurité, santé, elle prolifère au milieu de nombreux secteurs et plusieurs rapports indiquent que son marché pourrait atteindre près de 5 milliards de dollars d'ici 4 ou 5 ans, autant que le marché de la reconnaissance faciale en 2021.

Bon, ceci dit, elle devra quand même faire face à un nouveau concurrent, parce qu'en parallèle, les deepfakes vocaux ont eux aussi pris de la vitesse. Si vous suivez notre podcast Vitamine Tech, vous vous souviendrez peut-être de VALL-E, une IA capable d'imiter votre voix à partir de seulement... 3 secondes d'enregistrement. Après s'être entraîné sur 60 000 heures de fichiers audio en langue anglaise, produits par plus de 7 000 voix différentes et accompagnés de leur texte, l'algorithme de *deep learning* est parvenu à extrapoler les blocs de production du langage, d'une part, mais aussi le comportement typique d'une voix lorsqu'elle lit une phrase. Par exemple, voici un enregistrement qui lui a été fourni : [*une voix masculine humaine* : « He descended the ladder and found himself soon upon firm rock »]. Et ça, c'est la ~~voix~~ phrase qu'il a produite en imitant cette voix et son intonation : [*une voix identique à la précédente* : « They moved thereafter cautiously about the hut »]. Lorsqu'on met les deux morceaux bout à bout [« He descended the ladder and found himself soon upon firm rock. They moved thereafter cautiously about the hut »], eh bien, l'illusion est parfaite. Ou en tout cas suffisamment poussée pour tromper un paquet de monde et peut-être même... une autre IA. À peine sommes-nous entrés dans l'ère du deepfake vocal que déjà les médias font état de toutes sortes d'arnaques, d'usurpations d'identité ou même de propagande employant ces technologies. On pense par exemple à cette fausse vidéo du président ukrainien Volodymyr Zelenskyy invitant ses compatriotes à baisser les armes face à l'armée russe en 2022. [*Un extrait de cette vidéo.*] Ou plus récemment, la prolifération d'escroqueries au kidnapping virtuel. Un parent reçoit un appel d'un numéro inconnu et entend la voix de son enfant à l'autre bout de la ligne, implorant de lui venir en aide. En avril dernier, une mère américaine a ainsi été contactée par sa fille, en larmes, qui lui a expliqué qu'elle avait été enlevée. L'arnaqueur, bien réel pour sa part, a repris la ligne pour réclamer une rançon de 50 000 dollars, sans quoi il droguerait

l'adolescente, se servirait d'elle puis l'abandonnerait au Mexique. Si je donne ces détails, ce n'est pas pour verser dans le macabre mais pour souligner à quel point ce type d'arnaque peut prendre aux tripes. Même quand on est averti·e de la situation, il est difficile de ne pas être envahi·e par la panique et d'agir dans la précipitation.

Et les humains ne sont pas les seules victimes de ces tours de passe-passe. En mars de cette année, en Australie, une IA est parvenue à tromper le système de reconnaissance vocale des impôts en se faisant passer pour l'un de ses millions d'utilisateurs. Ainsi, aussi vite qu'elle s'est fait remarquer, la biométrie vocale pourrait donc être finalement tuée dans l'œuf par sa jumelle maléfique.

[La musique percussive du générique apparaît discrètement en fondu tout au long du paragraphe.]

Alors, qu'est-ce qu'il faut retenir de tout ça ? Eh bien déjà, que votre voix est unique, et qu'au fond, vous l'aimez sûrement plus que vous ne le pensez. Prenez cette occasion pour partir à sa rencontre, apprenez à apprécier son charme et explorez tous ces petits détails que vous avez absorbés au contact des gens qui ont marqué votre vie. Votre voix fait décidément autant partie de votre histoire et de votre identité que votre visage, vos tatouages ou vos choix vestimentaires. Et ça, aucune IA ne peut vous le dérober. Bon, ceci étant dit, si vous voulez absolument éviter de finir en deepfake – et je vous comprends –, ne faites pas comme moi avec ce podcast. Évitez de disséminer votre image et votre voix aux quatre vents sur les réseaux sociaux ou ailleurs sur le web, protégez vos comptes et quoi qu'il arrive, exercez une bonne dose de prudence si jamais vous recevez un appel suspect, ce que je ne vous souhaite pas. Si ce nouvel épisode vous a plu ou si vous avez des remarques ou des demandes, n'hésitez pas à laisser un commentaire, ou à nous écrire sur Twitter @futurasciences ou @Emma_Hollen. Si vous connaissez une personne sourde ou malentendante à qui ce podcast pourrait plaire, n'hésitez pas à le lui recommander. Comme toujours, des transcriptions détaillées sont fournies en description pour que tout le monde puisse en profiter. Surtout, pensez bien à vous abonner pour ne pas manquer le prochain épisode. Vous avez été des dizaines de milliers à écouter celui sur le Hum mais peu de gens se sont souvenus de cliquer sur « suivre » pour être tenus informés des prochaines sorties. Je compte donc sur vous ; on se retrouve dans deux semaines et d'ici là, écoutez le monde autrement.

Si vous avez suivi cet épisode jusqu'ici, voici un dernier fun fact pour vous remercier. En 2012, le chanteur Tim Storms a établi le record de la note la plus grave jamais produite par une voix humaine. À seulement 0,189 Hz, ce son est tellement bas que ni lui ni vous ne pourriez l'entendre. Je ne vais donc pas vous le jouer mais voici un extrait d'une de ses chansons. *[Tim Storms interprète une de ses chansons – que je n'ai pas reconnue –, descendant à une note tellement basse qu'elle pourrait visiblement faire vibrer une enceinte.]*

LIENS ET SOURCES :

[Front-facing camera - Wikipedia](#)

[Selfie Waves](#)

[Qu'entend-on lorsque l'on s'entend ? | Cairn.info](#)

[Voice confrontation - Wikipedia](#)

[Voice Confrontation in Individuals with Normal and Defective Speech Patterns](#)
[Do You Like Your Voice? A Population-Based Survey of Voice Satisfaction and Voice Enhancement - Naunheim - The Laryngoscope - Wiley Online Library](#)
[Scientists explain in more detail how we hear via bones in the skull | Imperial News](#)
[I like my voice better: self-enhancement bias in perceptions of voice attractiveness](#)
[The real reason the sound of your own voice makes you cringe | Psychology | The Guardian](#)
[Normal Voice Function](#)
[How the Lungs Work - How Your Body Controls Breathing | NHLBI, NIH](#)
[Mechanics of human voice production and control - PMC](#)
[Idéologies de genre et construction des savoirs en sciences phonétiques](#)
[Puberphonia, A Voice Problem of Puberty Before-After Voice Therapy](#)
[Bernoulli's principle - Wikipedia](#)
[Why your voice is like a fingerprint](#)
[A Morphometric Study of the Larynx - ScienceDirect](#)
[Does Size Matter in Laryngology? Relation Between Body Height and Laryngeal Morphometry](#)
[Voice change - Wikipedia](#)
[What Is an Adam's Apple?](#)
[Une petite fille qui parle sur la vie - YouTube](#)
[Puberphonia - Wikipedia](#)
[Castrat — Wikipédia](#)
[Castrato Singers Today: Their History and Modern Equivalent](#)
[Aging Voice | Condition | UT Southwestern Medical Center](#)
[👴 Vivre jusqu'à 100 ANS ? - Allo Docteurs](#)
[Marcel Vincent, le centenaire tranquille et heureux de la maison de retraite de Clairvaux-les-Lacs](#)
[The Voice & Ageing](#)
[Speech2Face Sees Voices and Hears Faces: Dreams Come True with AI - Belitsoft](#)
[University researchers say voice biometrics can tell identical twins apart](#)
[Normative standards for vocal tract dimensions by race as measured by acoustic pharyngometry](#)
[Do people of different races have different voices?](#)
[Effect of Ethnicity on the Speaking Fundamental Frequency Values of Indians vis-a-vis Caucasians, Afro-Americans and Mongoloids:](#)
[Study of vocal impersonations reveals how we manipulate our voices](#)
[Reading about embodiment \(Heidemann on timbre\)](#)
[Imagination is the key to vocal mimicry | New Scientist](#)
[Human voices are unique – but our study shows we're not that good at recognising them](#)
[How Common Is Face Blindness? | Harvard Medical School](#)
[Prosopagnosia - StatPearls - NCBI Bookshelf](#)
[Three per cent of people are unaware they cannot recognise familiar voices, study suggests | The Independent](#)
[Are super-face-recognisers also super-voice-recognisers? Evidence from cross-modal identification tasks](#)
[Super-recognisers: some people excel at both face and voice recognition](#)
["Are There Differences in Vocal Parameters Across Various Ethnic Groups" by Mason P. Meek](#)

[Effects of Race and Sex on Acoustic Features of Voice Analysis - Steve An Xue, Donald Fucci, 2000](#)

[Gender Differences in Familiar Voice Identification. | Request PDF](#)

[Familiarity and Voice Representation: From Acoustic-Based Representation to Voice Averages](#)

[We Can Identify A Friend's Voice After Hearing Just Two Words | HuffPost Impact](#)

[To identify a voice, brains rely on sight | ScienceDaily](#)

[Superiority of blind over sighted listeners in voice recognition | The Journal of the Acoustical Society of America | AIP Publishing](#)

[Recognizing voices depends on language ability | MIT News](#)

[New research is helping scientists to understand how humans recognize voices better than computers - ABC News](#)

[Could a New "Voice Recognition" Software Program Help Protect Wolves?](#)

[Voice Biometrics Market Size USD 4.90 Bn by 2027 | CAGR of 23.6%](#)

[Voice Biometrics Market Share to Reach \\$4.82Bn by 2028 at](#)

[Exploring racial and gender disparities in voice biometrics | Scientific Reports](#)

[AI scam calls imitating familiar voices are a growing problem – here's how they work](#)

[Deepfake video of Zelenskyy could be 'tip of the iceberg' in info war, experts warn](#)

[AI scam calls: This mom believes fake kidnappers cloned her daughter's voice | CNN](#)

[AI can fool voice recognition used to verify identity by Centrelink and Australian tax office |](#)

[Artificial intelligence \(AI\) | The Guardian](#)