

Thomas Hueber,
chercheur au CNRS,
tente de rendre la
parole à ceux qui l'ont
perdue (ici à l'aide
d'un synthétiseur
connecté
au conduit vocal).



LA VOIX SACRÉE

L'organe vocal humain est si complexe qu'il est difficile de le reproduire. Mais les progrès technologiques récents, intelligence artificielle en tête, pourraient bien changer la donne. Rencontre avec des spécialistes de la voix, devenue un enjeu industriel majeur.

Par Lorraine Rossignol
 Photos Léa Crespi pour Télérama

Tous les jours, sur l'application Lullaby Experience, qu'il a lancée il y a peu, le compositeur de musique contemporaine Pascal Dusapin reçoit de véritables trésors vivants : des berceuses du monde entier (géolocalisées en Bolivie, Albanie, Arabie saoudite... et même en Mongolie !), chantées pour lui par des anonymes qui ont bien voulu participer à sa collecte – ils devaient interpréter la chansonnette ayant marqué leur enfance. Celle dont ils restent nostalgiques à jamais. Le résultat est frappant : « *D'une puissance inouïe ! Il y a dans ces berceuses une incroyable spiritualité, une mémoire, une essentialité, une minimalisation de la ligne mélodique...* » Bref, il y a quelque chose de métaphysique » s'enthousiasme Dusapin. Et même, quelque chose de chamanique. Destinés à apaiser, rassurer, à tenir à distance les mauvais esprits, ces condensés de douceur et d'intimité, transmis de génération en génération, ont un pouvoir magique : celui que peut exercer une voix aimante et aimée...

Par quelle prouesse les centaines de berceuses collectées seront-elles, au final, « interconnectées » pour former l'installation sonore « Lullaby Experience », soit « *un énorme nuage mélodique et chantant* » (à découvrir au Centquatre, à Paris, en juin prochain) ? Grâce à l'expertise de l'Ircam (Institut de recherche et coordination acoustique/musique), véritable « laboratoire du son » fondé en 1970 par Pierre Boulez, où, chaque jour, des dizaines d'ingénieurs-designers-scientifiques auscultent et sculptent la matière sonore à des fins artistiques (lire p. 22). Là, un logiciel harmonisera les chansons recueillies pour produire une synthèse archétypale de la berceuse humaine. « *Jamais un cerveau humain n'aurait pu faire ce travail à une telle échelle. Mais ne nous trompons pas : la vraie merveille, ici, c'est ce qu'un corps chantant est capable de donner. Aucun ordinateur ne peut rivaliser avec cela* », insiste le compositeur.

De l'ultra-anthropologique à l'ultra-technologique, le raccourci est vertigineux. Il résume pourtant l'extraordinaire potentiel du médium vocal, dont on commence à peine à entrevoir l'étendue. Non seulement en matière de création artistique – l'Ircam ne cesse d'être sollicité pour des projets liés à la voix, et s'attend à l'être toujours plus dans les années à venir – mais dans des domaines aussi variés que la médecine, l'éducation ou le judiciaire... « *La révolution de l'intelligence artificielle a permis la réalisation d'avancées spectaculaires dans la plupart des domaines scientifiques et technologiques. On peut envisager aujourd'hui ce qui semblait impossible il y a encore quelques années* », explique Nicolas Obin, enseignant-chercheur à l'Ircam. Les avancées technologiques réalisées sur la voix en ont fait le mode privilégié de communication entre l'humain et la machine, des assistants personnels (Siri, Google Home...) aux robots. De la capacité à reconnaître les paroles prononcées par un être humain (reconnaissance vocale) jusqu'à la création de voix de synthèse capables de prononcer n'importe quelle parole avec un réalisme encore jamais atteint, la voix s'est imposée comme un enjeu technologique et industriel majeur. »

Au départ, la voix serait ce que nous avons en nous de plus humain. Ce simple souffle d'air émis par les poumons, qui, en passant entre les cordes vocales puis en résonnant à l'intérieur du conduit bucco-nasal – propre à la physiologie de chacun –, acquiert un caractère unique, et devient ainsi l'expression de notre identité profonde. Les fameux >>>



» Gafam (les géants de l'industrie numérique Google, Amazon, Facebook, Apple et Microsoft) ne s'y sont pas trompés, chacun y allant de son « assistant » connecté censé devenir un alter ego numérique. Et ces gadgets ont montré à quel point il est difficile de créer ex nihilo une voix humaine.

« La voix est un médium éminemment plastique et évolutif, elle change en fonction des émotions, de l'environnement... De cette réactivité au monde intérieur et extérieur, un robot ne sera jamais capable », explique la phoniatre Elizabeth Fresnel, qui, dans son Laboratoire de la voix, à Paris, soigne du matin au soir, depuis plus de trente ans, toutes sortes de voix, malades ou cassées, dont celles des stars de l'opéra, du théâtre ou du cinéma. Unique, plurielle, la voix d'un individu n'est donc pas près de se laisser réduire à un modèle informatique qui la figerait à la façon d'un instantané. Sans compter les mystères qui entourent toujours son fonctionnement, directement lié à celui du cerveau – lui-même encore si mal connu –, et dont elle est la porte-parole.

Parce que l'on arrive de mieux en mieux à se le figurer, grâce aux progrès de l'imagerie médicale, on sait en revanche à quel point une voix peut, en retour, avoir d'impact sur le cerveau – le sien ou celui d'autrui. Rien de bien nouveau, certes, depuis Homère, Ulysse et leurs sirènes, en passant par le charisme des grands orateurs qui galvanisent les foules. Tels Adolf Hitler ou Martin Luther King. Sans parler, bien sûr, des berceuses de Dusapin, et du souvenir indélébile qu'elles ont laissé. « La grande avancée consiste en ce que l'on pourra bientôt dire à quelles caractéristiques vocales précises réagit telle ou telle zone du cerveau », assure Jean-Julien Aucouturier, chercheur du CNRS en sciences cognitives à l'Ircam.

Et ainsi pouvoir « injecter » de façon ciblée des voix-remèdes dans des cerveaux cliniquement malades, grâce aux technologies qui permettent aujourd'hui de « colorer » une voix en temps réel. Et donc de lui faire prendre des intona-

« Les avancées technologiques réalisées sur la voix en ont fait le mode privilégié de communication entre l'humain et la machine. »

Nicolas Obin, enseignant-chercheur à l'Ircam

tions tristes, joyeuses ou détendues, et même de le faire changer de genre, d'âge... à la manière de filtres de Photoshop ou Instagram. « Le son est un moyen fabuleux d'entrer en contact avec les grands combattants ou les victimes de troubles post-traumatiques... Il a vraiment ici une carte à jouer pour l'innovation thérapeutique. Et ce n'est que justice que la science se serve de cet outil primordial qu'est la voix pour sonder ce qu'il y a de plus sophistiqué en nous. »

On n'en frémît pas moins à l'idée de l'usage industriel, commercial, médiatique ou politique, voire intime (via les objets connectés) qui pourrait être fait de voix persuasives, menaçantes ou culpabilisantes, pour manipuler des individus... « De même qu'il faut se méfier aujourd'hui des fake news et des photomontages, les gens vont devoir apprendre à faire preuve d'esprit critique vis-à-vis des "fake voices" », n'importe qui pouvant se doter d'un logiciel modifiant, maquillant ou contrefaisant les voix de manière ultra réaliste. Tout comme il vaut mieux aujourd'hui éviter de laisser des photos de soi sur le Net, il faudra éviter d'y laisser traîner sa voix », poursuit Nicolas Obin. Pour Jean-François Bonastre, qui, à Avignon, dirige le LIA, laboratoire dédié à la recherche sur le langage, « personne n'est préparé à ces nouvelles réalités, alors qu'elles sont en train d'explorer ». Lui-même, depuis des années, intervient auprès de tribunaux, pour tâcher d'identifier ou d'authentifier des voix, dans le cadre d'enquêtes judiciaires. « La demande est énorme, et elle risque de croître encore du fait de l'augmentation des situations où la voix sera en jeu. Par exemple, dans le cadre bientôt banal de paiements »

À VOIR

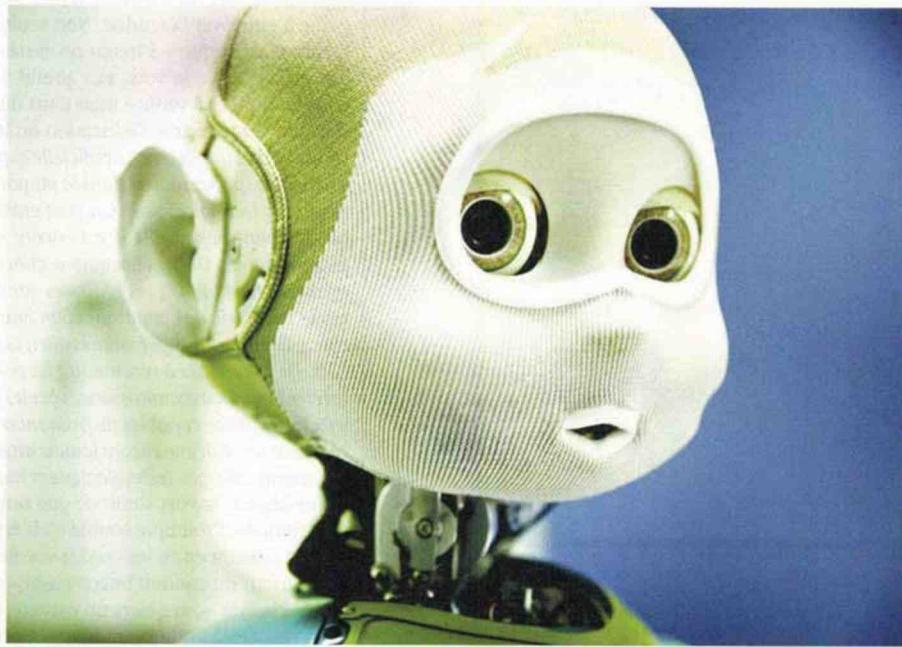
Vino Voce, festival dédié à la voix dans tous ses états. La prochaine édition aura lieu du 6 au 8 septembre, à Saint-Emilion (33).

À LIRE

Vox Confidential, de Christophe Haag, éd. Michel Lafon, 215 p., 17€.

À CONSULTER

Lullaby Experience, l'application de Pascal Dusapin, à laquelle tout le monde peut encore participer. www.lullaby-experience.eu



Le robot Nina est doté de capacités de communication verbale (développées au GIPSA-lab de Grenoble).



» bancaires sur Internet effectués uniquement par la voix. On imagine déjà à quelles escroqueries cela va donner lieu. »

Pourtant, à Grenoble, au département Parole et cognition du Gipsa-lab (l'un des plus grands laboratoires de France dédiés à l'étude de la voix et de la parole), l'équipe d'une trentaine de chercheurs permanents (en linguistique, phonétique, sciences cognitives et traitement de l'information) veut croire au côté progressiste et humaniste d'un « retour à la voix » — là où les réseaux sociaux et l'emprise de l'informatique sur nos vies ont eu tendance à la mettre en sourdine, ces trente dernières années. A cet égard, les travaux de Thomas Hueber, chercheur au CNRS, prennent une tournure emblématique, lui qui s'efforce, via des systèmes de suppléance vocale basés sur l'intelligence artificielle, de rendre la parole à ceux qui l'ont perdue après un accident de la vie ou une maladie. Car, sans aller jusqu'à l'ablation du larynx, nombre de gens perdent régulièrement leur voix. A commencer par les enseignants, dont les cordes vocales sont le principal outil de travail. Pour y remédier, Maëva Gar-

nier, elle aussi chercheuse au CNRS et spécialiste du « forçage vocal », a mis au point un micro-casque doté d'un petit capteur de vibrations, qui, placé sous le menton, permet aux instituteurs et professeurs de surveiller en temps réel le niveau d'intensité de leur voix. Mais aussi de mieux repérer les effets de celle-ci. Et donc de prendre conscience de la façon dont ils doivent la moduler pour mieux faire passer leur message pédagogique auprès de leurs élèves... Une petite révolution, que l'on pourrait envisager de transposer dans toutes sortes de situations du quotidien, « notamment dans le cadre familial, où le niveau sonore peut monter très vite ! »

« La voix est un cadeau de la vie, conclut Nathalie Henrich-Bernardoni, qui, au Gipsa-Lab de Grenoble, est spécialiste de la voix chantée. Et pourtant, nous n'utilisons qu'une toute petite partie de son potentiel, tout comme de notre cerveau. Face aux paillettes du tout-numérique et à l'accélération de nos existences, elle va nous permettre de réagir, en réapprenant à communiquer entre nous comme à nous reconnecter avec nous-mêmes. Nous devons à notre voix de rester humains. » ●

L'IRCAM, ROYAUME DES SONS

Dans les sous-sols de l'Ircam, à l'architecture essentiellement souterraine (pour se protéger de la pollution sonore urbaine), sous la place Igor-Stravinsky, à Paris, on trouve de curieuses choses. Découpeuse laser, fraiseuse... Nous sommes dans l'Atelier mécanique de l'Institut, auquel on accède par un long couloir qui mène, dit-on, au Centre Pompidou (les édifices furent associés dès leur création, au début des années 1970). « Ici, explique Chalabi Djellal, ingénieur en conception mécanique, on revient au concret, à la matière première et aux lois de la physique. On tâche de comprendre le fonctionnement subtil du larynx, que l'on ne maîtrise pas encore. Si l'on arrive à saisir quels mouvements produisent quels sons, on pourra "robotiser l'ensemble", en concevant un modèle de larynx morphoréalistique en silicone, comme on l'a déjà fait pour la bouche et comme il reste à le faire pour le conduit buccal. Pour cela, il n'est pas rare que l'Ircam collabore avec des chirurgiens! »

Percer les mystères de la voix humaine... L'Ircam, incroyable laboratoire interdisciplinaire où cohabitent quelque deux cent cinquante personnes, y parviendra-t-il un jour ? Juste au-dessus de l'Atelier mécanique, des mathématiciens couvrent de signes des tableaux entiers, établissant des modèles pour la synthèse sonore. A côté, assis devant leur ordinateur,

ingénieurs du son et designers sonores travaillent sur cette matière du matin au soir, grâce aux logiciels les plus performants (Angus, Omax, Audiosculpt, IrcamTools Trax...). « Nous ne sommes pas encore des magiciens, juste des sorciers confirmés ! », plaisante Axel Roebel, directeur de recherche, spécialiste en analyse et synthèse sonores. Pourtant, de véritables exploits ont bel et bien lieu entre ces murs. Depuis la reconstitution stupéfiante de la voix du castrat Farinelli, pour le film de Gérard Corbiau, en 1994, jusqu'à celle, plus récemment, du maréchal Pétain, pour les besoins d'un documentaire de France 5. Sans oublier celle de Marilyn Monroe telle qu'elle devait être dans l'intimité, pour une installation sonore de Philippe Parreno au Palais de Tokyo. L'Ircam a même réussi à faire s'enchaîner un vibrato grave et un staccato aigu — « ce qui n'est physiologiquement pas possible pour un être humain, mais qu'un ordinateur parvient à faire faire à la version numérique d'une voix réelle, tout en gardant le côté organique de cette dernière », raconte Jérôme Nika, chercheur et réalisateur en informatique musicale (c'est lui qui, au côté de Thierry Coduys, spécialiste en dispositifs électroacoustiques, met en œuvre le projet Lullaby de Dusapin), qui, chaque jour, assis derrière une console de l'un des huit studios d'enregistrement de l'Ircam,

« démultiplie les possibles de la voix humaine ». Frank Madlener, le directeur de l'institution depuis 2006, s'en félicite : « Le secteur est en pleine expansion, il est évident que le design sonore va devenir l'un des domaines les plus répandus dans les prochaines années, que ce soit pour le cinéma, les jeux vidéo ou le spectacle vivant. L'opéra voit un monde s'ouvrir à lui. Il peut tout inventer en la matière. »

Nicolas Obin,
enseignant-
chercheur à l'Ircam,
dans la chambre
sourde de l'institut.

