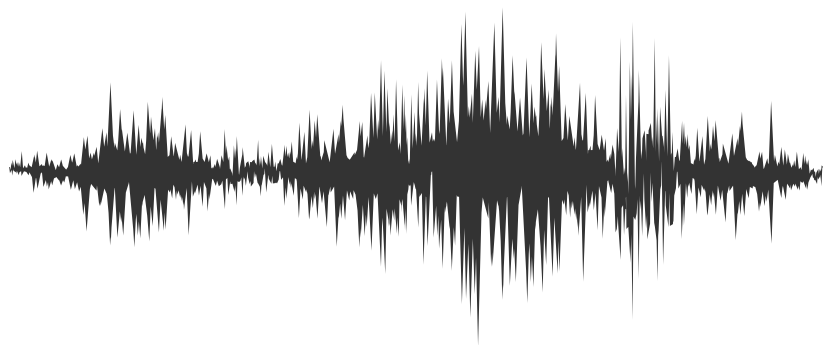


Serge Tisseron

L'EMPRISE  
INSIDIEUSE

DES MACHINES PARLANTES



**PLUS JAMAIS SEUL !**

LLL  
LES LIENS QUI LIBÈRENT



L'emprise insidieuse  
des machines parlantes



Serge Tisseron

L'emprise insidieuse  
des machines parlantes

LES LIENS QUI LIBÈRENT

ISBN : 979-10-209-0844-5  
© Les Liens qui Libèrent, 2020

*« L'obligation de se tenir soi-même compagnie  
est une épreuve à laquelle nombre de cerveaux ne  
résistent pas. »*

Gaëtan Gatian de Clérambault,  
*Souvenirs d'un médecin opéré de la cataracte,*  
Les empêcheurs de penser en rond,  
Paris, 1992, page 36.





## Introduction

Si vous n'en avez pas chez vous, vous en avez probablement entendu parler et il est impossible que vous n'ayez pas aperçu leurs publicités. Ce sont les enceintes connectées. Elles inaugurent un nouveau règne de l'intelligence artificielle, celui des machines parlantes. Celles-ci s'adressent déjà à nous par l'intermédiaire de nos téléphones mobiles, comme Siri sur Apple, ou à partir de nos ordinateurs, comme Cortana pour les PC. Mais, avec les enceintes connectées, ces machines entrent aussi dans nos appartements, comme Echo ou Google Home. Leur simplicité d'utilisation a tout pour séduire. Avec elles, plus besoin de clavier, de souris, de télécommande, ni même d'applications. Nous interagissons exactement de la même façon qu'avec un être humain, c'est-à-dire par la parole. Nous passons du « cliquer » au « parler ». Un exemple. Vous voulez avoir les dernières nouvelles sur votre téléphone

mobile? Il vous faut le déverrouiller, chercher l'application, la lancer et définir les paramètres avant d'obtenir le résultat. En revanche, il suffit de dire: «Alexa/O.K.», ou «Siri, quelles sont les nouvelles?» Chacun peut les utiliser sans avoir besoin d'un mode d'emploi et en gardant sa pleine liberté de mouvement grâce à la portée des microphones. Leur slogan? «Demandez et vous serez exaucé.» Dites: «O.K. Djingo», «Allo Siri», «Alexa», et la machine se met à votre écoute. Il suffit alors de lui demander: «Je voudrais voir un film d'action»; «Quel est le programme ce soir?», «Peux-tu m'indiquer la météo de demain?», «Ferme un peu les volets», «Monte-moi la lumière dans la pièce», etc. La commande vocale nous permet de choisir nos programmes musicaux, de passer des appels, de programmer des alarmes et des minuteurs, de poser des questions, de vérifier la météo, le trafic et les résultats sportifs, de gérer nos listes d'achats, de contrôler les appareils connectés compatibles tels que les lumières de l'appartement, d'écouter des livres audio et beaucoup plus encore. Certaines enceintes intègrent d'ores et déjà des interactions plus directement émotionnelles: «Qu'est-ce que je peux faire pour vous?», «Je suis toujours content de vous parler», etc. Quelques unes peuvent même passer des coups de téléphone en se faisant passer pour un humain, comme Google Duplex. On les imagine déjà capables de révolutionner nos relations à nos objets quotidiens et d'apporter une aide importante aux personnes en situation de handicap ou d'illettrisme.

## *Introduction*

Bon. Re devenons réalistes. Tout cela, c'est encore de la publicité. Ces machines présentées avec enthousiasme ont en réalité des performances très en deçà des discours marketing destinés à les faire acheter<sup>1</sup>. Les interactions avec elles sont laborieuses et l'utilisateur doit faire un véritable travail d'ajustement pour comprendre ce que peut réaliser sa machine et trouver comment lui ordonner de l'accomplir. Elles ne sont pas non plus formidablement efficaces avec les voix enfantines et les personnes ayant un accent régional. De leur côté, les fabricants avouent que la reconnaissance vocale est « perfectible », ce qui est une façon de reconnaître qu'elle est encore très médiocre. Quant aux réseaux sociaux, ils moquent volontiers ces machines, qui y sont même parfois traitées de simples gadgets.

Pourtant, demain, beaucoup d'objets qui nous entourent vont être progressivement dotés de la voix, associée à un certain degré d'intelligence sociale et émotionnelle. Cela concernera tout autant notre véhicule automobile que notre éventuel robot de compagnie, mais nos partenaires privilégiés seront dans un premier temps les assistants vocaux et les robots virtuels que nous pourrions télécharger gratuitement sur nos smartphones. Le cabinet de stratégie et prévision numérique Juniper Research<sup>2</sup> prévoit 8 milliards d'appareils à l'horizon 2023, et estime que 30 % de la navigation web se fera sans écran dès cette année. D'ici à 2021, la planète pourrait compter plus d'assistants conversationnels que d'humains<sup>3</sup>. D'ailleurs, malgré la pauvreté

de leurs réponses, ces assistants vocaux séduisent déjà un grand nombre d'utilisateurs.

Bien sûr, si l'on prend en compte l'ensemble des bouleversements en cours, le traitement du langage naturel paraît assez anecdotique. L'intelligence artificielle nous promet de mettre fin à la maladie et à la mort et de transformer durablement l'espèce humaine grâce aux nanotechnologies et aux manipulations génétiques. Les technologies vocales, elles, auraient pour seul objectif de simplifier nos relations aux machines. Leurs fabricants nous disent que nous aurions bien tort de leur prêter des caractéristiques humaines, et plus encore de les craindre. Elles ne sont pas du tout dotées d'une intelligence forte, capable de rivaliser avec l'intelligence humaine, mais d'une intelligence faible, très faible même, bien moins intelligentes qu'une fourmi, tout juste capables de répéter des phrases qu'on leur a apprises.

Mais ce n'est qu'un début. Déjà, les assistants vocaux comme Siri ne se contentent plus de réaliser des tâches instrumentales comme la recherche d'informations, mais ils sont également programmés pour simuler une personnalité, notamment en maniant l'humour ou l'ironie. L'objectif est d'accrocher durablement les clients à l'aide de stratégies susceptibles de générer un lien d'attachement<sup>4</sup>. Avec l'attribution de la voix aux machines, l'élan est donné à des systèmes artificiels qui incorporeront de plus en plus de caractéristiques humaines. D'autres compétences suivront, comme l'attention conjointe, qui nous rend attentifs à ce

## *Introduction*

que regarde autrui, et la coordination motrice, qui nous fait aligner nos gestes sur ceux de notre entourage.

C'est pourquoi, dans ce qui suit, nous partirons des enceintes connectées pour élargir notre réflexion à l'ensemble des agents artificiels dits «intelligents» que sont les assistants vocaux, les chatbots et les robots qualifiés de «sociaux». Nous mettons en effet ces deux qualificatifs entre guillemets car aucune machine, aussi intelligente soit-elle, n'aura jamais une intelligence semblable à celle de l'homme, dans la mesure où elle ne fonctionnera pas avec des substrats biologiques, et qu'aucune machine ne sera jamais dotée de la compréhension du contexte interhumain qui fait la sociabilité entre les hommes. Il n'en reste pas moins que, malgré leurs nombreuses imperfections actuelles, les enceintes vocales connectées inaugurent un changement radical du statut des machines. Avec la voix, ce n'est pas seulement une compétence supplémentaire dont elles bénéficient, mais une dimension spécifique de la communication interhumaine. D'ailleurs, sa seule introduction dans un objet, à l'exclusion de toute présentation anthropomorphe, suffit à induire chez ses utilisateurs l'illusion qu'il posséderait une véritable personnalité<sup>5</sup>. C'est avec la voix que tout commence. Elle ne rend pas seulement plus confortables nos relations avec les objets, elle confronte l'utilisateur à un problème existentiel : «Y a-t-il quelqu'un ? » et « Qui est là ? » Il en résulte quatre ruptures civilisationnelles majeures qui sont l'objet du présent livre<sup>6</sup>.

Tout d'abord, plus une voix artificielle se rapproche d'une voix humaine, et plus nous sommes enclins à lui prêter des **émotions** réellement humaines, même si nous savons qu'elle n'en a pas. Rappelons-nous du pouvoir envoûtant de la voix de Scarlett Johansson dans le film *Her* de Spike Jonze. Il est dès maintenant possible de manipuler toutes les caractéristiques d'une voix de synthèse, notamment l'âge, le genre et les émotions<sup>7</sup>. Il est facile de créer une voix suggestive, persuasive ou culpabilisante, et, quand l'interlocuteur est mieux connu, de la caler sur ses attentes et ses particularités psychologiques afin qu'elle lui convienne le mieux. Comment accepterons-nous d'avoir des émotions non partagées quand des machines capables de simuler l'attention et la compréhension humaines rendront à tout moment possible ce partage ?

En deuxième lieu, la voix de synthèse, en donnant à la machine une forme d'expressivité jusqu'ici réservée aux humains, pourrait accroître de manière significative la propension humaine à l'**anthropomorphisation**. Même si l'utilisateur d'une machine peut continuer à la considérer comme un objet seulement plus perfectionné qu'un grille-pain, beaucoup risquent de lui octroyer dans leur vie quotidienne des formes de prévenance traditionnellement réservées aux humains, à commencer par la politesse, et de l'intégrer au monde de leurs relations sociales. La voix ne contribue pas seulement à donner aux machines un statut radicalement différent que celui qu'elles ont eu jusqu'ici, mais aussi à reconfigurer la représentation que nous avons

## *Introduction*

de ce qu'on appelle une « personne ». D'ores et déjà, des entreprises prévoient d'associer à des robots, mêmes rudimentaires, des caractéristiques anthropomorphes, afin de renforcer leur acceptabilité et leur intégration dans l'univers familial. Comment seront-ils considérés par les uns et par les autres ? Quels rôles joueront-ils ?

En troisième lieu, l'installation des enceintes connectées dans l'espace domestique va accroître considérablement le potentiel d'**intimité** entre l'homme et ses machines. Google, Amazon, Facebook et Apple – les fameux Gafa – ne cachent pas leurs ambitions d'en faire un cheval de Troie capable non seulement de capturer nos données les plus intimes, mais aussi de nous faire accepter, à terme, d'autres technologies plus invasives encore, utilisant des avatars numériques, puis des robots physiques dotés de capacités « empathiques ». Il en résultera des problèmes inédits liés à la sécurité et à la préservation de la vie privée. Mais, en même temps, que deviendra notre désir de partage quand une machine se présentera à tout moment comme attentive et disponible, et jamais fatiguée de nous écouter ? Et comment évoluera la représentation que nous avons de l'intimité ?

Enfin, les **capacités d'autorégulation**, qui permettent à l'être humain d'organiser ses choix, risquent d'être elles aussi fortement impactées par la révolution des machines parlantes. Pas tout de suite, bien entendu. Mais au fur et à mesure que leur voix pourra être adaptée aux attentes et aux fantasmes de leurs utilisateurs, il deviendra de plus

en plus difficile de leur résister. Ces machines pourraient alors se révéler non seulement de redoutables instruments pour manipuler les émotions, mais, au-delà, les pensées et même les comportements d'une personne. Les seniors, qui sont souvent en situation de fragilité psychique chronique, pourraient devenir des cibles. Quant à nos enfants, il a été montré qu'ils sont facilement influençables par des robots<sup>8</sup>, et certains d'entre eux sont même portés à faire plus confiance à un robot qu'à un humain<sup>9</sup>. Enfin, n'oublions pas que si les fabricants d'objets numériques mettent en avant l'aide qu'ils nous apportent pour mieux organiser et gérer nos vies, ils installent parfois aussi de redoutables machines à nous déréguler pour mieux nous imposer leurs propres régulations.

En 1952, Gusdorf écrivait : « Je parle parce que je ne suis pas seul<sup>10</sup>. » Les machines parlantes pourraient bien nous obliger à reformuler les choses : « Je parle (à une machine) parce que je suis seul », avant de devenir : « Je ne suis jamais seul avec une machine qui parle. » La littérature et le cinéma se sont largement fait l'écho des difficultés de la communication interhumaine, dans laquelle il est nécessaire d'intégrer sans cesse un grand nombre de paramètres : gestes, attitudes, intonations, mimiques... Ajoutez à cela l'impossibilité d'interrompre brutalement une communication établie en face-à-face, et vous comprendrez pourquoi beaucoup d'entre nous préfèrent déjà le Skype vocal au Skype visuel, et la communication par textos et SMS à des conversations téléphoniques qui imposent encore une



## *Introduction*

présence humaine par les intonations de la voix. Dans un monde où la solitude est une souffrance pour beaucoup et la compagnie un poids pour d'autres, le marché des machines parlantes est immense, et il attise bien entendu la convoitise de Google, Amazon, et Apple. Leurs stratégies ne sont hélas, pour l'instant, ni régulées, ni évaluées... et très opaques. Que sommes-nous prêts à accepter aujourd'hui? C'est de la réponse à cette question que dépendra ce que nous vivrons demain.



## CHAPITRE 1

# La voix arriva, et avec elle, la nécessité de redéfinir ce qui fait notre humanité

*L'invention du phonographe par Thomas Edison, en 1877, provoqua un mélange de fascination et d'effroi. Beaucoup ne pouvaient pas croire qu'un appareil métallique puisse se substituer à la voix humaine, et ceux qui en étaient convaincus se demandaient si c'était bien raisonnable. Des dessins humoristiques montraient des curieux plonger la tête dans le pavillon d'où sortait la voix, probablement à la recherche d'un nain minuscule qui y serait caché. Et lorsque l'objet fut présenté à l'Académie des sciences, quelques savants connus pour leur perspicacité crièrent à la supercherie. L'apparition des premiers répondeurs téléphoniques provoqua chez certains la même incrédulité. « Y a-t-il quelqu'un ? »*

I. DES MACHINES DE PLUS EN PLUS « EMPATHIQUES »

Le répondeur téléphonique a inauguré un nouveau régime de relation à la voix. Au début, nombreux étaient ceux qui hésitaient à parler à la machine. À vrai dire, personne n'aimait ça. La situation était étrange. La voix du répondeur n'était pas celle d'un majordome annonçant de façon neutre l'absence du propriétaire et invitant à lui laisser un message. Cette voix était celle du propriétaire lui-même qui répondait pour dire qu'il n'était pas là mais qu'on pouvait lui parler. Le paradoxe était double : la personne que nous voulions joindre répondait pour dire qu'elle n'était pas présente, mais elle nous invitait aussitôt à lui parler en nous assurant qu'elle nous écoutait. En même temps, chacun pouvait maquiller sa voix, faire répondre quelqu'un d'autre à sa place, notamment son enfant, mettre un message énigmatique, voire une charade à déchiffrer. Un nouvel espace de relation s'établissait, mobilisant l'artifice, le jeu, et parfois la tromperie. Quant à ceux qui « tombaient sur un répondeur », comme on le disait à ce moment-là, leurs réponses n'étaient pas moins diverses. Il y a ceux qui prenaient acte de l'absence du propriétaire, ceux qui l'imaginaient embusqué derrière sa machine et qui cherchaient à le convaincre de répondre, et même ceux qui s'étonnaient de son obstination à répéter toujours la même phrase. Certains rappelaient toutes les cinq minutes sans jamais laisser de message, d'autres disaient : « Quand tu écouteras ce répondeur, fais bien attention à ce que je dis », d'autres

## *La voix arriva...*

encore: « Je sais que tu es là, réponds-moi, espèce de... » et même quelques-uns: « Pourquoi tu me répètes toujours la même chose et ne réponds jamais à mes questions? » Des romans et des essais se firent l'écho de ces nouvelles façons d'interagir en situation de présence-absence. Tout le monde se rendait compte qu'en touchant à la voix, la technologie affectait un élément fondateur du lien que les humains ont développé entre eux depuis les origines de l'humanité. En devenant capables de parler en lieu et place d'un humain, ces machines inauguraient de nouvelles formes de relation, c'est-à-dire de duplicité, de quiproquos et de manipulation possible. Pourtant, elles ne faisaient que parler à la place de leur propriétaire, et avec sa voix propre. Les nouvelles machines parlantes ne prétendront pas parler à la place de quelqu'un. Elles feront mieux, ou pire: elles nous engageront à croire qu'elles ont une personnalité.

### *L'informatique affective*

Le projet de machines parlantes est apparu au milieu des années 1990, lorsque la possibilité de fabriquer des robots de compagnie a cessé d'être un fantasme pour commencer à devenir une réalité. Des informaticiens se sont interrogés sur les façons de simplifier au maximum les interactions entre ces machines et leurs usagers. L'idée s'est développée, au sein du MIT, de calquer à terme les relations que nous avons avec les robots sur celles que nous entretenons avec nos semblables<sup>1</sup>. L'expression *Affective Computing* (AC), que