

MARKETING/COMMUNICATION

LA RÉVOLUTION DES ASSISTANTS VOCAUX

Oxana
GOULIAÉVA

Éric
DOSQUET

Yvon
MOYSAN

Comprendre les enjeux et réussir
ses stratégies marketing

DUNOD

Crédits Iconographiques

Couverture : © emojiiez/Shutterstock ; figure 1.1. : © Ralf Schmitzer/© Caique Lima/© Very/© Ataur Rahman/© Luis Prado/© Danil Polshin – The Noun Project ; figure 2.4. : © Loïc Poivet – The Noun Project ; figure 2.5. : © anbilero adaleru/© Danil Polshin/© DAVIVONGSA PATHRPOL/© ProSymbols/© Trevor Dsouza – The Noun Project ; figure 2.6. : © DAVIVONGSA PATHRPOL/© MRK/© Manuela Barroso/© Danil Polshin/© Mavadee – The Noun Project ; figure 6.4. : © supalerk laipawat/© Lesha Petrick/© Smalllike/© nauracon – The Noun Project ; p. 226 : © Stuart O'Neil – The Noun Project

Composition : Nord Compo

Le pictogramme qui figure ci-contre mérite une explication. Son objet est d'alerter le lecteur sur la menace que représente pour l'avenir de l'écrit, particulièrement dans le domaine de l'édition technique et universitaire, le développement massif du photocopillage.

Le Code de la propriété intellectuelle du 1^{er} juillet 1992 interdit en effet expressément la photocopie à usage collectif sans autorisation des ayants droit. Or, cette pratique s'est généralisée dans les établissements

d'enseignement supérieur, provoquant une baisse brutale des achats de livres et de revues, au point que la possibilité même pour les auteurs de créer des œuvres

nouvelles et de les faire éditer correctement est aujourd'hui menacée.

Nous rappelons donc que toute reproduction, partielle ou totale, de la présente publication est interdite sans autorisation de l'auteur, de son éditeur ou du Centre français d'exploitation du

droit de copie (CFC, 20, rue des Grands-Augustins, 75006 Paris).



© Dunod, 2020

11 rue Paul Bert, 92240 Malakoff

www.dunod.com

ISBN 978-2-10-080571-6

Le Code de la propriété intellectuelle n'autorisant, aux termes de l'article L. 122-5, 2° et 3° a), d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L. 122-4).

Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L. 335-2 et suivants du Code de la propriété intellectuelle.

Sommaire



Remerciements	1
Préface	3
Introduction	5
Partie 1	
Contexte et enjeux des assistants vocaux	9
Chapitre 1 ■ Assistants vocaux : définitions et usages	11
Chapitre 2 ■ Contexte : marché, acteurs, écosystèmes	23
Chapitre 3 ■ Les grands verticaux impactés	59
Partie 2	
Marketing à l'ère des GAFA	
Enjeux, risques et opportunités pour les marques	75
Chapitre 4 ■ Intelligence Artificielle (IA), prédicatif et personnalisation	77
Chapitre 5 ■ Assistants vocaux et les 4P du marketing	91
Chapitre 6 ■ Accompagnement de nouveaux modèles économiques et de distribution	105

Partie 3

Développer une stratégie voix

Place des assistants vocaux dans le marketing mix 113

Chapitre 7 ■ Retour aux fondamentaux 115

Chapitre 8 ■ Identité sonore et vocale de la marque 125

Chapitre 9 ■ Podcasts : nouveau storytelling 149

Chapitre 10 ■ Recherche vocale & référencement vocal 169

Chapitre 11 ■ Applications vocales : nouvel outil marketing 191

Conclusion 231

Index 233

Remerciements



Penser, écrire, challenger ne se fait pas seul.

Nous remercions l'ensemble des experts ayant apportés leur éclairage particulier sur cette nouvelle révolution en marche : Jacques Balian, Thomas Belmont, Raphaël Bijjaoui, Ulysse Bottello, Luc Bretones, Laurent Cochini, Maaïke Coppens, Antoine Crehalet, Niraj Dawar, Sabrina Delale, Violaine Di Meglio, Frédéric Dosquet, Mathieu Gallet, Florian Guichon, Nabeel Hussain, Luc Julia, Arthur Larrey, Jamal Lassiri, Benjamin Le Breton, Alastair Mackie, Laurent Mathis, Vincent Picot, Timothée Retailleau, Nicolas Rieul, Emmanuel Ruiz, Nick Schwab, Maître Charles Simons, Philippe Vinci et Paul Wauters.

Nous remercions enfin nos éditrices, Ludivine Le Gall et Chloé Schiltz, pour leur patience et leur conseils attentifs.

Préface



C'est un peu, beaucoup, de notre faute si en 2011, l'ancêtre de tous les assistants vocaux d'aujourd'hui, *Siri*, nous a fait croire qu'il pouvait être intelligent. Nous connaissons les limites de la technologie, nous savions que ça ne marchait pas très bien : un taux de succès assez faible, d'environ 80 %, pour la reconnaissance de la parole. Imaginez qu'il manque 20 % des mots dans chacune des pages de ce livre, ça le rendrait très difficile à comprendre !

D'où l'idée de rendre *Siri* un peu sourd, un peu stupide. Oui, le rendre imparfait était notre stratégie pour le rendre plus humain en somme. On lui faisait parfois raconter des blagues, grâce à des scénaristes d'Hollywood, ou dire des phrases complètement hors contexte. Il devenait comme nous de temps à autre, distrait et attachant. Il a ainsi séduit pas moins de 300 millions d'utilisateurs dans le monde, en quelques mois.

Comme nous nous croyons plus ou moins parfaits, cette humanisation a rendu notre assistant « intelligent » aux yeux de beaucoup. La popularité de *Siri* a été aidée par des films tels que *Her* ou par les paroles et les écrits de quelques amateurs se disant spécialistes de l'intelligence artificielle, qui prétendaient que HAL 9000, l'omnipotent assistant dépeint en 1968 dans le film de science-fiction *2001 : L'Odyssée de l'espace*, avait enfin trouvé son incarnation dans *Siri*.

Il est temps de revenir sur terre. Et c'est ce que nous aident à faire brillamment les auteurs de cet ouvrage très documenté, qui présente les

assistants vocaux sous tous les aspects nécessaires à leur compréhension et, surtout, à leur utilisation dans le contexte du marketing d'entreprise.

Les auteurs nous expliquent les avantages de la voix par rapport aux autres canaux de communication, les technologies qui en découlent, la place qu'y tiennent le *Cloud* ou l'IA, leur adoption. Ils font bien sûr un tour d'horizon des différents acteurs du marché, des différents segments et usages possibles. Enfin, ils n'oublient pas d'évoquer les potentielles dérives comme les conséquences possibles que peuvent avoir les assistants vocaux sur notre vie privée.

Pour plus de pédagogie, tout au long du livre, les auteurs font intervenir des experts venant de multiples disciplines qui apportent des éclairages différents sur les divers sujets abordés. Et pour le rendre pratique et encore plus exploitable, on y trouve des cas réels d'intégration de ces nouvelles technologies aux stratégies marketing de plusieurs entreprises.

Pour résumer, *La révolution des assistants vocaux* fera de vous, lecteurs, des initiés mieux armés pour expliquer et convaincre votre entreprise de la nécessité de déployer ces assistants digitaux. Il vous aidera également à les implémenter, afin d'améliorer les performances de vos techniques et méthodes de stratégie commerciale.

Dr. Luc JULIA

**Co-créateur de *Siri*
et auteur de *L'Intelligence artificielle n'existe pas*
(First Edition, 2019)**

Introduction



« Nous avons tendance à surestimer l'incidence d'une nouvelle technologie à court terme et à la sous-estimer à long terme. »

Roy Amara, co-fondateur de l'*Institute for the Future* à Palo Alto.

D'abord une pulsation, ensuite un souffle, puis un cri et enfin une voix. D'abord des babillages, ensuite quelques mots, puis des phrases et enfin une pensée, partagée ou polarisante mais force élémentaire d'une société dans toutes ses composantes. La voix est un marqueur de maturité de notre humanité comme elle le devient pour les objets.

Plus encore, la voix est considérée par la psychanalyse comme une expression du Moi intérieur qui permet de « prendre possession de l'espace et de nous situer par rapport à l'autre ». C'est ainsi que la voix devient aussi l'instrument du dire et du faire, un outil de la socialisation et de l'existence, un outil de transition entre soi et l'extérieur constitutif, du Moi dans un monde social tout comme la peau dans un monde physique. Prendre la parole ou avoir voix au chapitre, donner sa voix ou voter, écouter sa voix intérieure ou convoquer la voix de la raison, depuis les premiers mois de la vie utérine jusqu'au dernier souffle, la voix est un puissant levier de notre vie. La tonalité, la respiration, l'intensité, l'articulation, la vitesse signent de manière unique une personne, comme un second visage pouvant exprimer une émotion, séduire ou irriter, convaincre ou polariser, en tout état de cause un levier de l'action et de l'existence.

Quelle est alors cette voix sortant de ces assistants, nouveaux objets de notre quotidien ? Quels pouvoirs inconscients projetons-nous sur *Siri* et *Alexa* quand nous leur posons des questions sur le monde, comme l'enfant interrogeant sa mère, quand ces nouvelles voix artificielles ont, elles aussi, réponse à tout ? Quelles régressions infantiles et quelles dépendances construisons-nous quand, contrairement au surmoi sonore paternel, la machine satisfait toujours à la demande comme si le savoir ne nécessitait aucun effort ? Sur quelles règles éthiques et légales acceptons-nous de laisser l'objet collecter tout de nos interrogations, de nos croyances, de nos constructions mentales à celui (ou celle) que nous croyons intime alors qu'il (ou elle) nous est en tout point étranger ? Voire ce grand frère qui saurait tout de la vie des autres ?

« Pour qui sont ces serpents qui sifflent sur vos têtes ?

À qui destinez-vous l'appareil qui vous suit ? »

Jean Racine, *Andromaque*, 1667.

Et pourtant, ces assistants vocaux constituent une nouvelle évolution d'objets ayant déjà changé notre vie. Voilà une bonne vingtaine d'années, les ordinateurs personnels connectés à un Internet balbutiant nous ouvraient déjà une somme inédite de connaissances, immédiates et (presque) gratuites. Voilà une bonne dizaine d'années, les smartphones propulsés par les « apps » résolvait nombre de nos tracas quotidiens peu importe l'endroit où nous nous trouvions. Aujourd'hui, les assistants vocaux représentent la nouvelle génération d'objets intelligents du quotidien. Leur qualité première est singulière : leur charge cognitive est quasi-nulle. Ainsi, plus besoin de savoir naviguer dans un système d'exploitation, plus besoin même de savoir lire et écrire, il suffit de savoir parler. Ce sont ainsi trois grandes populations nouvelles, et probablement ultimes, qui se voient ouvrir l'accès la nouvelle Alexandrie numérique.

D'abord la population illettrée, majoritairement dans les pays les moins avancés, ceux dont la démographie est la plus dynamique et celle des pays les plus peuplés. Ensuite, tous les enfants de moins de six ans n'ayant pas encore eu accès à la lecture et à l'écriture. Enfin, les plus âgés qui auront raté les vagues des ordinateurs et des smartphones dont

la charge cognitive aura été trop complexe à maîtriser, devenus citoyens de seconde zone d'une société numérique.

La beauté de l'avenir se cache dans l'inconnu. Dans un contexte d'accélération et de ruptures technologiques, aucune prédiction ne peut avoir lieu avec certitude, seulement des hypothèses dont la plupart seront oubliées quand sonnera l'heure de leur réalisation. Quand il s'agit d'une technologie aussi naturelle mais disruptive que la voix (omniprésente, omnisciente, presque magique, mais avec des apparences, ou du moins des sonorités humaines), les prévisions ne manquent pas d'audace. Tous les acteurs majeurs, sur l'ensemble des verticaux, explorent, élaborent et expérimentent avec les nouvelles opportunités offertes par les assistants vocaux. C'est à l'aube de ces ruptures que peuvent se conforter ou se défaire des empires. L'enjeu est d'importance et c'est pourquoi nous nous efforcerons de comprendre ce qui peut se cacher derrière l'écran de fumée des prévisions.

Les assistants vocaux associent deux aspects : des assistants virtuels et une interface vocale, d'où le terme complet « d'assistants virtuels à commande vocale ». Ils conjuguent des algorithmes d'intelligence artificielle (IA), des systèmes de reconnaissance et de générations vocales voire, plus important encore, une nouvelle approche dans l'accès à l'information. Ils s'inscrivent dans les grandes **tendances technologiques** de fond (montée des plateformes, IA, données, personnalisation, multiplication d'appareils, etc.) et dans la lignée des transformations **d'interfaces Homme-Machine**, ce qui ne laisse aucun doute sur leur potentiel à modifier profondément **nos comportements et nos usages**. Ils apportent encore un nouveau spectre d'interrogations sur le respect de la vie privée, de la protection des données et de la confiance, principaux freins à l'adoption et garde-fous fondamentaux envers des tentations orwelliennes.

Les marques s'interrogent désormais sur la place des assistants dans leur marketing mix. Sont-ils un médium publicitaire, une interface de vente, un nouveau canal de distribution, un point d'achat ou un point de contact après-vente, concurrents en marque privée ou d'une intimité inédite avec leur client ? Nous essayerons de comprendre le

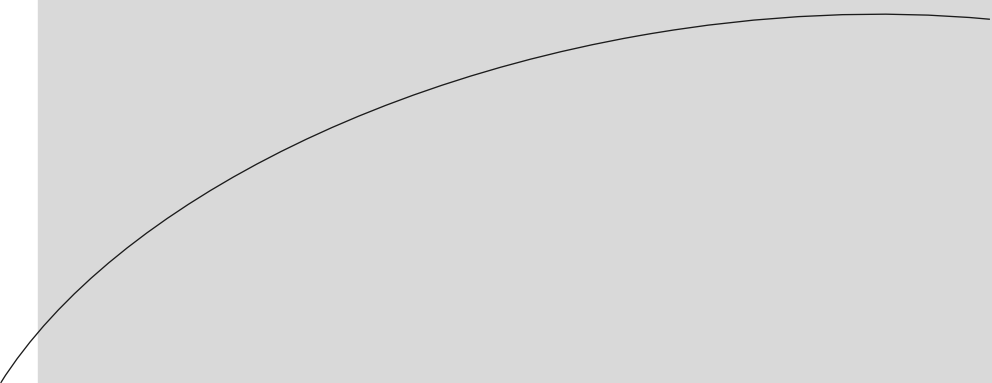
potentiel impact des assistants sur les marques, notamment de grande consommation, et comment celles-ci peuvent intégrer cette nouvelle technologie dans leur relation client.

Pour se faire, il conviendra d'étudier :

- l'offre mondiale en matière d'assistants et d'enceintes connectées. Notre attention ira avant tout à l'*Assistant Google* et *Alexa d'Amazon*, deux leaders actuels sur le marché occidental. Quelques particularités, différences et éléments de stratégie de ces deux acteurs seront analysés pour permettre l'identification des potentielles implications de l'intelligence artificielle (IA) et de la voix sur le marketing des marques ;
- l'aspect « IA » des assistants, comme les principes de leur fonctionnement ;
- l'aspect « voix » des assistants (interface et médium) ;
- les principaux usages ;
- la protection des données personnelles et de la vie privée.
- les questions que les marques doivent se poser et les actions qu'elles doivent entreprendre.

Partie 1

Contexte et enjeux des assistants vocaux



Chapitre 1

Assistants vocaux : définitions et usages



Executive summary |

- ▶▶ **Comprendre les technologies derrière les assistants vocaux**, ainsi que les grandes familles d'usages.

Depuis quelques années, nous assistons à la convergence de plusieurs technologies émergentes. Cependant, pour comprendre le monde et préparer l'avenir, les organisations doivent cesser de suivre des tendances isolées, mais s'attacher à comprendre les liens entre elles. Nous avons identifié trois tendances majeures dans lesquelles s'inscrivent les assistants virtuels à commande vocale, afin de mettre en perspective et mieux comprendre leur potentiel impact sur le marketing des marques.

1. Vers la fin du smartphone, la multiplicité des objets et une nouvelle génération d'interfaces¹

La multiplication d'objets technologiques semble marquer le début de la fin des smartphones tels que nous les concevons aujourd'hui. Dans ses prévisions pour 2018, le *Future Today Institute* indiquait que la prochaine décennie serait une transition vers une nouvelle ère du *computing* ambiant, c'est-à-dire vers des objets connectés et des interfaces invisibles que nous porterons sur nous et qui nous accompagneront dans notre mobilité (*wearables*, lunettes intelligentes, casques de réalité virtuelle et de musique, écrans flexibles, pliables et enroulables, etc.). Il sera possible de les contrôler par le toucher, les gestes, la voix et plus tard la pensée, grâce à l'intelligence artificielle.

2. Le silence est d'or²

Le temps passé sur nos écrans d'ordinateurs et smartphones atteint un seuil de saturation. Les taux des désabonnements, de déconnexions, de « *opt out* » (action de se désengager d'une option cochée par défaut) sont en nette hausse pour faire barrage au contenu venant polluer notre quotidien. Il est presque question de santé publique, notamment pour les plus jeunes. Les organisations et les marques doivent apprendre à offrir de la valeur aux utilisateurs qui cherchent désormais du calme, de la sérénité et de la concentration dans un monde devenu trop bruyant.

¹ Futur Today Institute Trends, 2018.

² Fjord Trends, 2019.

3. La montée en puissance du canal audio

Les interactions peuvent faire appel à nos cinq sens, mais le canal visuel reste le plus usité. On constate aujourd'hui que le canal auditif, jusqu'ici sous-exploité, gagne en popularité. Il permet de continuer de s'informer et de se divertir sans écran et en situation de mobilité, voire d'augmenter la productivité dans un cadre de travail.

On observe concomitamment une montée significative d'intérêt pour les nouvelles formes de consommation radio et audio (web radio, services de musique en streaming, podcasts, livres audio, dont les *Audible*[®] d'*Amazon*) sur de nouveaux supports d'écoute (comme les enceintes connectées) avec une augmentation de la durée d'écoute et de volumes d'impressions servies (Semestre 1 2018 : +15 % vs Semestre 1 2017¹).

Ces observations se confirment avec l'émergence d'une nouvelle gamme de produits, les *hearables* : mini-écouteurs (*earbuds*), casques, appareils auditifs et même accessoires et bijoux. Selon Statista, ce marché est d'ailleurs en forte croissance (10,5 milliards de dollars en 2016 contre 22 milliards de dollars estimés en 2020).

Toutes ces évolutions, combinées avec la miniaturisation des composants et la multiplication des canaux de communication, participent à une intégration de plus en plus étroite d'Internet et de l'IA dans notre quotidien. Ces technologies deviendraient, à l'instar de l'électricité, omniprésentes mais invisibles, pour aboutir à une « intelligence ambiante », qui se distinguera par ces trois composantes principales : ubiquité, contextualisation et interaction naturelle.

Contrairement aux interfaces informatiques traditionnelles, leur utilisation se fera de façon naturelle et intuitive, par la voix, le regard et les gestes. Par son caractère omniprésent et universel, l'intelligence ambiante deviendra une niche d'innovation et de croissance économique, capable de favoriser des modes de vie moins consommateurs

1 Baromètre de l'Audio Digital, 1^{er} semestre 2018, Kantar Media.

d'énergie, favorisant les gains de productivité en entreprise et améliorant le confort de vie.

Il est opportun de rappeler ici la célèbre phrase de Nicholas Negroponte du MIT Lab : « *L'informatique ne concerne plus seulement les ordinateurs, mais notre vie tout entière* ». Les technologies vocales n'en sont que la première hirondelle d'un nouveau printemps numérique...

Définition : assistant virtuel

Un assistant virtuel est un programme informatique qui s'appuie sur plusieurs technologies d'intelligence artificielle pour effectuer des tâches et/ou rendre des services à l'utilisateur. Pour cela, il interprète ses requêtes (sous forme de texte, voix ou tactile) et interagit avec lui en mode conversationnel. Il serait plus précis de dire qu'un assistant vocal est une combinaison d'agents intelligents (dont chacun exécute une tâche spécifique), d'applications (dont les applications vocales, voiceapps, appelées *skills* sur *Alexa* et *actions* sur l'*Assistant Google*¹) et de connexions entre les données. Grâce au *machine learning*, l'assistant capitalise sur toutes les interactions pour s'améliorer et apporter des réponses plus personnalisées et plus efficaces. Le service conversationnel est assuré par les chatbots (aussi appelés « agents conversationnels » ou « bots conversationnels ») qui, généralement, fonctionnent à travers trois types d'interfaces :

- à travers le clavier, par exemple sur les messageries instantanées ;
- à travers des canaux non verbaux, par exemple par une sélection parmi les images proposées ;
- à travers la voix, grâce au micro du smartphone ou aux objets statiques tels que les enceintes connectées (ou *smart speakers*) du type *Écho* d'*Amazon*, *Google Home* de *Google* ou *HomePod* d'*Apple*. Une version vocale d'un chatbot s'appellera voicebot.

¹ En français, on rencontre parfois les termes « aptitudes » et « compétences ».

Dans cet ouvrage, nous nous intéresserons à ce troisième type d'assistants : les assistants virtuels à commande vocale, ou tout simplement, les assistants vocaux.

Les enceintes connectées sont parfois, à tort, assimilées aux assistants vocaux, mais ne sont en réalité que leur partie matérielle, visible. Les assistants vocaux revêtent des formes diverses : applications installées sur les smartphones, tablettes, ou directement embarqués sur les ordinateurs, terminaux mobiles, montres connectées et autres *wearables*, dans des voitures ou sur des équipements de la maison (*smart home*).

Nouvelle interface Homme-Machine

La voix est une nouvelle interface qui répond à une vieille question de l'interface universelle, celle dont la charge cognitive pour accéder aux machines sera des plus faibles. Au CES 2016, la start-up française *Sevenhugs* présentait sa télécommande *Smart Remote* qui se voulait un game changer : elle avait l'ambition de s'adapter à tout type d'équipement (TV, musique, stores, thermostats, etc.) et d'afficher les boutons de contrôle adaptés pour déclencher les actions. Deux ans ont été nécessaires pour industrialiser l'objet qui n'a finalement pas tenu toutes ses promesses et surtout s'est avéré trop coûteux (près de 250 euros). Les assistants vocaux, quant à eux, permettent de contrôler tous les objets connectés de la maison, voire bien plus, pour un coût d'acquisition de quelques dizaines d'euros. Selon Brian Roemmele, à la maturité de la technologie, « *la voix ne va pas nous aider à naviguer dans nos objets technologiques, mais les remplacera dans la plupart des cas*¹ ».

La raison principale de ce succès est l'universalité naturelle de cette interface. La voix est aussi une technologie qui enlève de la complexité à la technologie. Tous les dix ans, depuis soixante-dix ans, la technologie enlève une couche de complexité pour atteindre une

¹ Brian Roemmele, « There Is A Revolution Ahead », 12 avril 2016, techpinions.com