



VOICE APP GOOGLE ET ALEXA : LES EFFETS DE LA VOIX DE SYNTHÈSE SUR LES REACTIONS DES CONSOMMATEURS

Nicolas Kusz

Université Paris 1 Panthéon-Sorbonne, PRISM Sorbonne

Jean-François Lemoine

Université Paris 1 Panthéon-Sorbonne, PRISM Sorbonne

ESSCA School of Management

Résumé : Les « voice app » intégrées aux systèmes Google et Alexa se multiplient, de plus en plus d'entreprises proposent ce nouveau canal d'interaction à leurs consommateurs en complément de leur site web et de leur application smartphone. Avec les progrès majeurs des technologies de reconnaissance vocale et de voix de synthèse, l'assistant vocal permet d'instaurer un véritable dialogue entre l'homme et le système. Si la voix n'est désormais plus le privilège des humains, notre étude cherche à explorer les effets du type de voix d'un assistant vocal (humain versus artificiel) sur les réactions des consommateurs. En nous appuyant sur une étude qualitative exploratoire menée auprès de 15 répondants, les résultats suggèrent que le type de voix de l'assistant influence la facilité d'utilisation perçue et l'anthropomorphisme perçu par les consommateurs. En outre, notre étude révèle que l'anthropomorphisme a un impact négatif sur la facilité d'utilisation perçue ; l'utilisateur semble oublier qu'il s'adresse à une machine et qu'il doit ajuster ses requêtes en conséquence.

Mots clefs : assistant vocal ; voix de synthèse ; facilité d'utilisation ; anthropomorphisme.

GOOGLE AND ALEXA VOICE APP: THE INFLUENCE OF THE VOICE ON CONSUMERS

Abstract : More and more companies develop “voice app” integrated into Google and Alexa systems, to offer a new channel of interaction to their consumers in addition to their website and their smartphone application. With spectacular progress in voice recognition and synthetic voice technologies, the voice assistant makes it possible to establish a real dialogue between the human and the system. If the voice is no longer the privilege of humans, our research aims to explore the influence of the type of voice of a voice assistant (human versus artificial) on the cognitive reactions of consumers. Based on an exploratory qualitative study conducted with 15 people, the results suggest that the voice type of the assistant influences ease of use and perceived anthropomorphism by users. Furthermore, our study reveals that anthropomorphism has a negative impact on usability; the user seems to forget that he's talking to a machine and forget adjusting his requests accordingly.

Keywords: voice assistant ; synthetic voice; trust ; usability ; anthropomorphism.

Résumé managérial

En complément de leur site web ou de leur application mobile, chaque entreprise peut développer une application vocale et venir se brancher sur le système de l'assistant vocal Google ou Alexa, pour offrir un canal d'interaction supplémentaire à leurs clients. Communément appelées « actions » pour Google et « skills » chez Amazon, ces applications dites « tierces » ou « de marque » permettent de créer une discussion via une commande vocale telle que « Ok Google, je veux parler à [nom de l'entreprise] », qui déclenche l'expérience vocale (avec des fonctionnalités et des services spécifiques) et le Dialogue Homme-Machine imaginés par chaque entreprise. En outre, lors de la conception d'une application vocale, les entreprises ont la possibilité d'opter pour une voix spécifique et unique plutôt que la voix du système Google ou Alexa, leur permettant de créer leur propre identité sonore. L'application vocale permet à l'utilisateur d'échanger directement avec une entreprise ou une marque comme il le ferait en venant naviguer sur un site web, de s'informer sur les produits ou les services, voire même acheter par la voix ou accomplir certaines actions spécifiques par commande vocale. Cette technologie ouvre ainsi de nouvelles perspectives inédites de relation client personnalisée, permettant une accessibilité immédiate, renforçant la satisfaction et l'expérience client. Les enjeux financiers sont considérables, le « voice-commerce » serait porté à 80 milliards de dollars dans le monde.

L'apport de notre recherche est double pour les managers : nous mettons en évidence le rôle majeur de la voix dans le cadre de la création d'un assistant vocal, puis nous renseignons les professionnels sur le type de voix à privilégier. Nos résultats contribuent également à une meilleure compréhension de l'effet des caractéristiques anthropomorphiques de la voix de l'assistant depuis une enceinte connectée. Le type de voix d'une application vocale tierce d'une entreprise induit un effet significatif sur la façon dont les utilisateurs souhaitent interagir avec la technologie, et comment ils réussissent finalement à le faire. La création d'une application vocale n'est pas une interface comme les autres telles qu'un chatbot ou un agent virtuel (et cela même s'ils sont dotés d'une voix de synthèse), le contexte d'utilisation sans écran apporte des attitudes et des comportements utilisateurs spécifiques. Tous les individus ne semblent pas à l'aise avec le fait de dialoguer avec un objet inerte telle qu'une enceinte connectée ; les utilisateurs ne comprennent pas exactement avec quoi et avec qui ils échangent, ils ne savent pas comment fonctionne une application vocale, ni le type de questions qu'ils peuvent poser. On observe finalement de meilleures réactions dans le cas où l'assistant vocal présente une voix de synthèse qui, bien que dotée de caractéristiques anthropomorphiques, soit bien identifiée par les utilisateurs comme une voix artificielle, et non pas comme une voix humaine. Nos résultats mettent en lumière qu'une voix très humaine induit un dialogue beaucoup plus riche de la part des utilisateurs, qui se mettent alors à interagir avec l'assistant de la même manière qu'ils le font avec un humain, avec toute la richesse et la complexité du langage, oubliant les limites techniques de l'appareil. Nous proposons ainsi aux entreprises et aux concepteurs d'intégrer des caractéristiques artificielles dans la voix de l'assistant pour aider l'utilisateur à garder en mémoire qu'il interagit avec une machine et non pas avec un humain, et lui permettre d'ajuster ses questions et ses attentes en conséquence. L'ensemble de ces résultats peut ainsi aider les spécialistes du marketing à comprendre les effets générés par la voix de l'assistant sur les réactions de leurs clients, afin de produire des efforts marketing ciblés pour créer une identité sonore qui puisse apporter de meilleures performances pour l'entreprise.

Introduction

Les progrès majeurs des technologies de reconnaissance vocale, du traitement du langage et de synthèse vocale ont donné naissance à des assistants vocaux (Moriuchi, 2020 ; Pitardi et Marriott, 2021). Si les premiers assistants vocaux sont apparus au XXe siècle, c'est avec l'arrivée des enceintes connectées Google Home et Amazon Echo en 2017 que cette technologie s'est faite connaître du grand public (2,5 milliards d'assistants vocaux dans le monde et 50 millions en France, source Hadopi). Si les premiers cas d'usage étaient assez basiques, les assistants vocaux proposent désormais de nombreuses fonctionnalités : rechercher une information, mener des activités de loisir et de divertissement (jeu, musique) ou contrôler à distance des objets connectés (Ammari et al., 2019 ; Gouliaéva et al., 2020). En complément, toutes les entreprises peuvent aussi créer leur propre « voice app » accessible depuis Google Assistant et Amazon Alexa, en complément de leur site web et de leur application smartphone, et ainsi proposer un nouveau canal d'interaction à leurs consommateurs (Vernuccio et al., 2022 ; Balakrishnan et Dwivedi, 2021). Les voice app (i.e. une application vocale) permettent aux consommateurs d'interagir vocalement avec une société ou une marque, de s'informer ou même d'acheter un produit ou un service à la voix (McLean et al. 2021). Tandis que l'ordinateur, le clavier et la souris apparaissent comme des dispositifs mécaniques et froids, le caractère anthropomorphique de l'assistant vocal rend le système plus humain (Moriuchi, 2019). L'acceptation de cette technologie est un domaine de recherche encore nouveau, en management des systèmes d'information la littérature apporte un éclairage sur les interactions entre l'homme et le système (McLean et Osei-Frimpong, 2019), alors qu'en marketing les chercheurs s'intéressent aux réactions des consommateurs à l'égard de la technologie (Moriuchi, 2019 ; Berriche et Benavent, 2021). Si l'interface vocale peut sembler plus simple d'utilisation, elle fait naître aussi de nouveaux défis, sans visuels ni aucun texte pour guider ou susciter l'intérêt, sans boutons sur lesquels cliquer pour naviguer (Devillers, Masure et Obin, 2020). En outre, le caractère humain de l'assistant est renforcé du fait de l'utilisation de la voix, le système s'humanise. Bien que plusieurs travaux menés sur les agents virtuels et des chatbots ont mis en évidence l'influence de la voix de synthèse sur les réactions des utilisateurs, à notre connaissance aucune recherche n'a étudié les effets des caractéristiques vocales d'un assistant vocal, sur les réactions des utilisateurs. La problématique d'étude que nous avons identifiée est la suivante : dans quelle mesure le type de voix de l'assistant vocal peut-il influencer sa facilité d'utilisation perçue ? D'un point de vue managérial, cette recherche a pour ambition de proposer des pistes de réflexions pour aider les entreprises à identifier la voix la plus adaptée à leur assistant vocal (i.e une voice app).

1. Le cadre conceptuel de la recherche

S'intéresser aux effets du type de voix d'un assistant vocal sur la facilité d'utilisation nous amène tout naturellement à étudier la littérature sur l'acceptation d'une technologie et la littérature sur l'anthropomorphisme. L'ensemble de ces deux domaines de recherche permettra d'obtenir un cadre théorique sur l'influence de la voix des assistants sur les consommateurs.

1.1. La facilité d'utilisation d'un assistant vocal

L'acceptation des assistants vocaux s'explique en grande partie par la facilité d'utilisation de son interface (Ammari et al., 2019 ; Rzepka et al., 2020), dans le cas présent la voix. Selon Cornut (2009), il s'agit du moyen le plus efficace pour communiquer entre les individus. Loin d'être anodine, la voix remet totalement en perspective la manière dont les individus consomment du contenu sur le web, recherchent une information et achètent des produits ou services (Ramadan et al., 2019). Pour comprendre les facteurs qui influencent l'acceptation

d'une nouvelle technologie, les chercheurs s'appuient sur les fondements théoriques de l'acceptation d'une technologie (Moriuchi, 2019 ; McLean et Osei-Frimpong, 2019), le modèle TAM révèle notamment que plus l'utilité perçue et la facilité d'utilisation perçue sont élevées, plus l'attitude de l'individu envers le système est positive. Appliquées aux assistants vocaux, les études ont été menées dans plusieurs contextes expérimentiels (i.e. avec les smartphones, les enceintes connectées ou encore les assistants embarqués dans les voitures) ou alors pour étudier différentes variables dépendantes (i.e. l'intention d'utilisation, l'intention d'achat d'un assistant vocal, l'intention d'achat d'un produit ou service par le biais de l'assistant). Les études suggèrent que les assistants vocaux sont pratiques dans une grande variété de situations (Ammari et al., 2019) ; ils sont faciles à déclencher, sans effort, sans avoir à ouvrir un ordinateur, saisir un mot de passe, cliquer et rechercher l'information (Porcheron et al., 2018 ; Rzepka et al., 2020), enfin l'appareil permet à l'utilisateur de poursuivre d'autres tâches en parallèle (McLean et al., 2019). Toutefois, si l'interface vocale peut sembler plus simple d'utilisation, dans la pratique l'introduction d'une nouvelle technologie constitue souvent un changement profond des usages (Rogers, 1962). L'acceptation semble encore balbutiante et de nombreux freins demeurent, les fantasmes technologiques des utilisateurs pensant pouvoir dialoguer de façon parfaitement fluide font face à des frustrations majeures qui nuisent à l'expérience globale (Poirier, 2017). Les études démontrent que l'interface vocale n'est finalement pas aussi aboutie que les promoteurs tentent de le faire croire, en réalité il existe encore un gouffre entre les attentes ou les croyances des utilisateurs et les performances encore limitées des assistants (Luger et Sellen, 2016 ; Velkovska et al., 2018). Bien qu'intelligents, en réalité les assistants vocaux n'en sont pas moins que de « simples » systèmes dotés de parole et capables de traiter uniquement des requêtes basiques (Natale, 2020). Enfin la découvrabilité des fonctionnalités disponibles, difficile à déceler au son de la voix, rend les usages encore méconnus par les utilisateurs et freine l'adoption des assistants à plus grande échelle (Velkovska et al., 2018 ; Lahoual et Fregus, 2018). Selon McLean et Osei-Frimpong (2019) les travaux manquent encore de recul pour juger de leur pertinence ou de leur robustesse, notamment sur le long terme. En outre, à notre connaissance, aucune étude ne s'est intéressé au rôle des caractéristiques vocales en tant que stimuli sur la facilité d'utilisation de cette technologie. Bien que des contributions scientifiques menées sur les agents virtuels et des chatbots ont mis en évidence l'influence de la voix de synthèse sur les réactions des consommateurs (Chérif et Lemoine, 2019), avec un assistant vocal le contexte expérimentiel diffère ; contrairement aux chatbots et aux agents virtuels qui viennent en complément d'une navigation sur un site web ou sur une application mobile, l'assistant vocal s'utilise en autonomie sans aucune autre interface, ni écran ni clavier (Natale, 2020).

1.2.L'effet de l'anthropomorphisme (vocal) sur la facilité d'utilisation

Faire des machines performantes sur le plan technique, capables de comprendre et parler en langage naturel, ne suffit plus pour en assurer leur acceptation. Pour faciliter l'interaction entre l'individu et la technologie, l'enjeu est désormais de paraître humain en partageant certaines caractéristiques communes. Les travaux de Dubiel et al. (2018) révèlent notamment que les assistants vocaux dotés de capacités humaines conduisent à une acceptation plus forte des appareils et à une relation homme machine plus intuitive. Les voix de synthèse sont de nos jours modélisées sur la base de vraies voix humaines, qui imitent à la perfection une personne réelle. Doter une machine de caractéristiques anthropomorphiques, telles un corps, un visage ou des attitudes humaines, tend à améliorer la perception d'une technologie (Guthrie, 1997), cette discipline a ainsi reçu une attention particulière dans les recherches en systèmes d'information et en marketing (Guido et Peluso, 2015). En outre, l'homme a tendance à anthropomorphiser naturellement le monde qui l'entoure pour faciliter ses interactions (e.g. parler aux animaux, aux plantes ou à certains objets). Selon Natale (2020), la technologie de

synthèse vocale et ses caractéristiques paralinguistiques ne permettraient plus aux individus de discerner une voix de synthèse d'une voix humaine tant elle semble réaliste. Savoir dialoguer approximativement comme un humain, ou tout du moins en donner l'illusion, permet aux assistants de dissimuler les limites techniques des Intelligences Artificielles (Berriche et Benavent, 2021). Naturellement, l'utilisateur a tendance à s'adresser à un assistant vocal de la même manière qu'avec un autre individu (Sundar et al., 2017 ; Lopatovska, 2019 ; Canziani et MacSween, 2021). En outre, si la voix n'est qu'une onde, immatérielle et imperfectible au toucher, sans visage ni corps, les assistants vocaux viennent toutefois s'intégrer dans un objet physique matériel et tangible, comme une enceinte connectée ou un smartphone, ce qui participe à façonner la personnalité de la voix et permet à l'utilisateur de se faire une représentation mentale du « personnage » de l'assistant (Poirier, 2017). L'anthropomorphisme d'un assistant vocal peut se matérialiser de différentes façons ; le type de voix évidemment (artificiel ou humain), mais aussi à travers diverses modulations de la voix (i.e. la prosodie), c'est-à-dire l'intonation, le volume, le débit, les hésitations, les répétitions ou les corrections, voire les silences tels des moments de réflexion ou de respiration, voici autant de marqueurs qui donnent corps à la voix, délivrent des informations sur le locuteur (i.e. la marque associée à l'application vocale) et lui donnent une personnalité humaine (Stern et al., 2006 ; Zoghaib, 2019). La littérature suggère qu'en imitant les caractéristiques humaines, la voix renforce l'anthropomorphisme de l'assistant et facilite son utilisation (Patrizi et al., 2021 ; Balakrishnan et Dwivedi, 2021). Toutefois certains chercheurs évoquent la notion de piège anthropomorphique, en devenant la seule interface pour échanger avec un assistant vocal, la voix amène un changement de paradigme et un rapport ambivalent entre l'homme et la machine (Natale, 2020) ; d'un côté l'anthropomorphisme et l'attribution de caractéristiques humaines tendent à rapprocher l'utilisateur et l'assistant, d'un autre côté la voix met l'utilisateur à distance puisqu'il n'y a plus aucun contact physique avec l'appareil, plus de touches à presser, d'écrans à effleurer, de boutons à cliquer, la simplification apportée par la voix se traduit par une forme de complexité dans l'utilisation du système.

2. Méthodologie

Des questions demeurent sur les effets de la voix de d'un assistant vocal, il nous apparaît nécessaire de s'interroger sur les avantages et les freins à l'utilisation d'un assistant vocal, la perception de la technologie et les réactions des consommateurs. Cette étude qualitative exploratoire repose sur quinze entretiens semi-directifs menés auprès de consommateurs en face à face durant une cinquantaine de minutes. Afin de bénéficier d'un échantillon diversifié, nous avons sélectionné un public varié (cf. annexe) en termes de variables socio-démographiques et d'expertise. Les entretiens ont été menés auprès des consommateurs à l'aide d'un guide d'entretien scindé en trois grandes parties majeures ; (i) les usages des répondants avec leur assistant, (ii) leur opinion de la technologie vocale avec les avantages et les freins relevés dans le cadre de leur utilisation, (iii) leurs réactions vis-à-vis de la voix des assistants.

Application vocale	Fonctionnalité	Genre de la voix	Type de voix	Tonalité
Guerlain	Conception d'une fragrance	femme	Robotisée	sérieuse
Hello Bank	Accès au compte	femme	Humaine	Sympathique
Boulangier	Trouver un magasin, un produit etc.	homme	Robotisée	Sérieuse
Oui.scnf	Trouver un voyage et un billet	femme	Robotisée	Sérieuse
BNP Paribas	Accès au compte	homme	Humaine	Sympathique

Dans un second temps, nous avons utilisé la méthode des protocoles, nous avons invité les répondants à manipuler et interagir avec les enceintes Google Home ou Amazon Echo, puis à déclencher cinq applications vocales (connectées aux systèmes Google ou Amazon) sélectionnées parmi un benchmark d'une cinquantaine d'applications, conçues par diverses marques. Nous avons sélectionné des assistants vocaux tous secteurs confondus afin offrir une plus grande variété de situations.

En outre, pour cette étude nous avons retenu des applications ayant une grande diversité de caractéristiques vocales, parmi lesquelles nous pouvons citer ; le genre de la voix, la vitesse d'élocution, la tonalité, le timbre de la voix et la prosodie, et enfin le type de voix (synthèse vocale ressemblant fortement à une voix humaine ou au contraire plus robotique).

3. Résultats et discussion

L'analyse de notre étude qualitative menée auprès des consommateurs révèle que le type de voix des assistants semble avoir un effet sur la facilité d'utilisation perçue et l'anthropomorphisme perçu. Pour analyser le contenu des interviews auprès des consommateurs, nous avons regroupé et classé manuellement l'ensemble des réponses des répondants en thèmes et sous thèmes dans une grille d'analyse (Bardin, 2003).

3.1 L'effet de la voix sur la facilité d'utilisation du système

La facilité d'utilisation apportée par l'interface vocale ressort comme l'un des principaux facteurs de succès des assistants vocaux et confirme ainsi la littérature (Kowalczyk, 2018). La rapidité de réponse participe à la facilité d'utilisation perçue, « *Un assistant vocal c'est quelque chose qui est censé me faciliter mon quotidien. C'est censé nous éviter d'aller sur internet, sur notre téléphone, sur notre PC ou chercher dans un livre. Voilà, j'ai une question, là, du tac au tac j'ai besoin d'une réponse* » (F, 39), un répondant ajoute « *Je n'ai pas besoin de taper, vous voyez... Je dis juste « OK Google » et il me répond quoi... C'est rapide.* » (F, 40). L'interface vocale amène une facilité d'utilisation évidente, « *Sans notice, c'est quand même d'une utilisation simplissime... Simplissime* » (H, 65), une opinion partagée par la plupart des répondants « *Même ma grand-mère pourrait s'en servir, vous avez juste à parler et l'appareil se met en route* » (H, 41), ou encore « *« Mon fils qui a 6 ans, va demander à jouer sa playlist sur Spotify. Il demande sa playlist et, du coup, je trouve ça assez bien. Il s'est approprié l'usage de la Google Home* » (F, 38). Les répondants mettent en évidence la possibilité de poursuivre d'autres tâches en parallèle, « *Je suis en train de cuisiner et que je veux connaître la météo je n'ai pas besoin d'utiliser mes mains qui sont déjà occupées* » (F, 24), un répondant nous illustre une autre situation « *Tu es entrain de petit-déjeuner et tu as cette information qui te tombe un peu comme ça, et ça peut faire gagner du temps sur la préparation du matin.* » (F, 38). Paradoxalement, certains répondants évoquent une utilisation presque trop simple qui aurait tendance à accentuer la fainéantise « *Même s'il apporte des facilités de réponses aux questions, s'il est rapide, il est aussi un peu lénifiant parce qu'il t'enlève la nécessité de te lever et d'aller chercher le renseignement ailleurs, sur un dictionnaire, sur un autre support, dans la presse [...] ça engendre un peu de la paresse intellectuelle* » (H, 65), un sentiment partagé par un autre répondant « *Je suis souvent tentée de dire « OK Google, qu'est-ce que tu peux nous raconter là-dessus » mais je ne le fais pas parce que je n'ai pas envie d'habituer mon fils à cet automatisme-là. Oui je n'ai pas envie de lui donner l'impression que tout est facile et que l'information se trouve comme ça, en claquant des doigts* » (F, 38). Les répondants pointent du doigt une facilité parfois mitigée, « *très rapidement on revient à la réalité avec des « je n'ai pas compris la question, une erreur est survenue etc.* » (H, 16). Malgré une facilité d'utilisation évidente, les répondants évoquent le syndrome de la page blanche, « *utiliser un assistant vocal c'est quand tu sais ce que tu peux faire avec [...] comme tu le sais pas tu ne l'utilises pas* » (F, 32), un inconvénient relevé par de nombreux répondants « *quand on achète ça on n'a même pas de manuel d'utilisation, donc tu sais pas vraiment comment les utiliser à 100%* » (F,40).

3.2 Les effets de la voix sur l'anthropomorphisme

L'anthropomorphisme motive l'utilisation d'un assistant vocal, un répondant indique « *Je trouve ça très très impressionnant, parce que je n'ai pas l'impression que c'est une machine*

qui parle, c'est presque effrayant. Le ton est aussi très impressionnant, le langage non verbal, et c'est extrêmement naturel » (H, 45), un répondant ajoute « La voix était vraiment très réelle. Vraiment, il y avait de l'intonation, la ponctuation... c'était très naturel. On avait l'impression de parler à un vrai humain » (F, 39). La personnification des assistants renforce également la perception de l'anthropomorphisme « ils ont essayé d'humaniser ça en lui donnant un petit prénom. Je me rappelle qu'il s'appelle Gustave, je trouve ça assez drôle » (F, 40). Toutefois notre étude révèle qu'une voix de synthèse anthropomorphe très réaliste a une influence négative sur la facilité d'utilisation, les utilisateurs ayant alors tendance à exprimer leur requête de façon naturelle avec toute la richesse et la complexité du langage humain en omettant les limites techniques de l'appareil « Pourriez-vous me conseiller pour un placement d'un montant de 15 000 euros ? » (H, 45), amenant à une erreur du système qui ne comprend pas la requête. Contrairement aux études menées sur les chatbots et agents virtuels (Lee et Nass, 2003 ; Wang et al., 2007 ; Chérif et Lemoine 2019), nos résultats indiquent qu'une voix artificielle génère de meilleures réactions sur la facilité d'utilisation « Je parle à une machine ou je parle à un humain ? Il faut que ce soit humain oui, mais avec quelques indices qui permettent de bien identifier le caractère artificiel de la voix et du système avec lequel je parle » (H, 34), un répondant ajoute « Si je lui pose des questions comme à un humain, je vais avoir envie de lui poser des questions comme à un humain, mais je vais être déçu, alors que là j'ai conscience des limitations, que j'échange avec un algorithme, je vais pouvoir structurer ma pensée en ce sens. Une voix un peu mécanique me permet de ne pas oublier que j'échange avec une machine et que je dois adapter mon discours » (H, 45). Enfin les répondants concluent « Je n'aime pas trop les voix humaines, je préfère lorsqu'elle est plus robotisée, un peu spatiale parce que c'est plus moderne, c'est plus futuriste ... c'est une machine, je m'attends à une voix de machine » (H, 16), ou encore « J'ai pas envie que le robot imite l'humain, c'est pas un humain » (H, 45), un phénomène qui n'est pas sans rappeler la vallée de l'étrange (MacDorman, 2019).

4. Apport de la recherche et conclusion

L'intérêt théorique de cette recherche est d'étudier les effets de la voix de l'assistant vocal sur sa facilité d'utilisation et sur les réactions des consommateurs, ainsi qu'enrichir la littérature en marketing digital. Malgré les études portant sur l'acceptation des assistants vocaux ou encore les travaux sur l'anthropomorphisme des chatbots et agents virtuels, nous n'avons trouvé aucune étude qui s'est intéressée spécifiquement à l'anthropomorphisme des assistants vocaux. Si la voix donne corps à une machine, et plus particulièrement à un assistant vocal dépourvu de représentation visuelle, nous sommes face à deux choix possibles : doter l'assistant d'une voix de synthèse qui ressemble en tout point à une voix humaine ou opter pour une voix de synthèse qui, bien que basée sur des caractéristiques humaines, permet à l'utilisateur d'identifier qu'il échange avec une machine, et non pas avec un humain. Notre étude qualitative menée auprès des consommateurs confirme la littérature sur la facilité d'utilisation apportée par l'interface vocale, toutefois nos résultats s'opposent à la littérature menée sur les agents virtuels, nous mettons en évidence qu'un assistant vocal doté d'une voix de synthèse artificielle amène de meilleures réactions qu'une voix humaine réaliste. D'un point de vue managérial, cette recherche a pour ambition de proposer des pistes de réflexions pour les concepteurs des assistants vocaux ainsi que les marques souhaitant développer leur assistant vocal et améliorer la relation avec leurs consommateurs (Viot et Bressolles, 2012). Compte tenu du caractère exploratoire de notre recherche, nos conclusions ont tout de même certaines limites ; les marques associées aux applications vocales manipulées dans le cadre de l'étude qualitative ont pu venir biaiser les perceptions des répondants. Il conviendrait, d'entreprendre une expérimentation en créant un assistant vocal dépourvu de marque, puis de moduler l'anthropomorphisme de la voix pour tester les effets sur la facilité d'utilisation de l'appareil.

Références

- Ammari, T ; Kaye, J. Tsai, J.Y ; Bentley, F. (2019). « Music, Search, and IoT: How People (Really) Use Voice Assistants », *ACM Transactions on Computer-Human Interaction* 26(3), pp. 1-28. DOI:10.1145/3311956
- Balakrishnan, J., Dwivedi, Y.K. (2021) Conversational commerce: entering the next stage of AI-powered digital assistants. *Ann Oper Res* (2021). <https://doi.org/10.1007/s10479-021-04049-5>
- Bardin, L. (2003), « L'analyse de contenu », Presses Universitaires de France, Paris.
- Berriche, A. et Benavent, C. (2021). « Comprendre les profils des voicenautes et leur intention d'achat par les assistants vocaux : Les apports de la théorie du conflit décisionnel », 20th International Marketing Trends Conference, Digital Conference.
- Canziani, B. et MacSween, S., 2021. Consumer Acceptance of Voice-Activated Smart Home Devices for Product Information Seeking and Online Ordering. *Computers in Human Behavior*, p.106714. <https://doi.org/10.1016/j.chb.2021.106714>
- Cornut, G. (2009). « La voix », Presses Universitaires de France, « Que sais-je ? ». DOI:<https://www.cairn.info/la-voix--9782130576747.htm>
- Chérif, E. et Lemoine, J.F. (2019). « Les conseillers virtuels anthropomorphes et les réactions des internautes : une expérimentation portant sur la voix du conseiller », *Recherche et Applications en Marketing*, Janvier 2019, vol. 34, no. 1, pp. 29-49.
- Devillers, L. ; Masure, A. ; Obin, N. (2020). « Assistants connectés : des ordres dans la voix », interview dans le cadre l'émission de radio Magnétique, Lausanne, Radio Télévision Suisse.
- Dubiel, M., Halvey, M., & Azzopardi, L. (2018). « A Survey Investigating Usage of Virtual Personal Assistants », *Computer Science ArXiv*, abs/1807.04606.
- Gouliaéva, O ; Dosquet, E ; Moysan, Y (2020). « La révolution des assistants vocaux : comprendre les enjeux et réussir ses stratégies marketing ». Collection Marketing/Communication, Dunod
- Guido G. er Peluso A.M. (2015), “Brand anthropomorphism: Conceptualization, measurement, and impact on brand personality and loyalty”, *Journal of brand Management*, vol. 22, n. 1, pp. 1-19.
- Guthrie, S.E. (1997), « Anthropomorphism: A definition and a theory », in Mitchell R., Thompson N. et Miles, H. (eds) *Anthropomorphism, anecdotes, and animals*. State University of New York Press, pp. 50-58.
- Kowalczyk, P. (2018). « Consumer acceptance of smart speakers: a mixed methods approach: An International Journal », *Journal of Research in Interactive Marketing* 12, 418–431. <https://doi.org/10.1108/JRIM-01-2018-0022>

Lahoual, D. et Fréjus, M. (2018). « De l'utilisabilité à l'appropriabilité des assistants vocaux. Étudier les interactions vocales en situation domestique à partir d'une démarche centrée utilisateurs », Conférence internationale ERGO'IA.

Lee K. M. et Nass C. (2003). « Designing social presence of social actors in human computer interaction », Proceedings of the Conference on Human Factors in Computing Systems - CHI '03, (5), 289.

Lopatovska, I. ; Rink, K. ; Knight, I. ; Raines, K. ; Cosenza, K. ; Williams, H. ; Sorsche, P. ; Hirsch, D. ; Li, Q. ; Martinez, A. (2018). « Talk to me: Exploring user interactions with the Amazon Alexa », Journal of Librarianship and Information Science. DOI:10.1177/0961000618759414

Luger, E. et Sellen, A. (2016). « Like having a really bad PA : The gulf between user expectation and experience of conversational agents », 2016 CHI Conference on Human Factors in Computing Systems, ACM, pp.5286–5297. DOI:10.1145/2858036.2858288

MacDorman, K.F. (2019). « Masahiro Mori's The Uncanny Valley: Significance and Impact on Aesthetics and Robot Design », Journal of the history of technology. DOI.org/10.4000/ephaistos.5333

McLean, G. et Osei-Frimpong, K. (2019). « Hey Alexa... Examine the variables influencing the use of Artificial Intelligent In-home Voice Assistants », Computers in Human Behavior (99), pp. 28-37. DOI:10.1016/j.chb.2019.05.009

McLean, G., Osei-Frimpong, K. and Barhorst, J. (2021). « Alexa, do voice assistants influence consumer brand engagement?—Examining the role of AI powered voice assistants in influencing consumer brand engagement », Journal of Business Research, 124, pp. 312-328.

Moriuchi, E., 2019. « Okay, Google!: An empirical study on voice assistants on consumer engagement and loyalty », Psychology & Marketing. <https://doi.org/10.1002/mar.21192>

Moriuchi, E. (2020). « An empirical study on anthropomorphism and engagement with disembodied AIs and consumers' re-use behavior », Psychology & Marketing. <https://doi.org/10.1002/mar.21407>

Natale, S. (2020). « To believe in Siri: A critical analysis of AI voice assistants », Communicative Figurations Working Papers, 2020, Vol. 32, pp. 130–146.

Patrizi, M. ; Vernuccio, M. ; Pastore, A. (2021). « Developing voice-based branding: Insights from the Mercedes case », Journal of Product & Brand Management, 30(5), 726– 739. <https://doi.org/10.1108/JPBM-08-2019-2490>

Pitardi, V, Marriott, HR (2021). « Alexa, she's not human but... Unveiling the drivers of consumers' trust in voice-based artificial intelligence », .Psychologie et Marketing. 2021; 38: 626– 642. <https://doi.org/10.1002/mar.21457>

Poirier, F. (2017). « Quelle modalité pour l'interaction avec les petits appareils mobiles et vestimentaires : texte ou vocal ? Comment choisir entre clavier et assistant personnel ? », 29ème

conférence francophone sur l'Interaction Homme-Machine, Poitiers, France. pp. A-67. hal-01577674

Porcheron, M. ; Fischer, J.E. , Reeves, S. ; Sharples, S. (2018). « Voice Interfaces in Everyday Life », 2018 CHI Conference on Human Factors in Computing Systems, ACM, paper 640, pp. 1-12. DOI :10.1145/3173574.3174214

Ramadan, Z. ; Farah, M. ; Audi, H. (2019). « The Advent of the Voice Moment of Truth: The Case of Amazon's Alexa », International Conference on Advances in National Brand.

Rogers, E.M. (1962) Diffusion of Innovations. Free Press, New York.

Rzepka C., Berger, B. et Hess, T. (2020). « Why Another Customer Channel ? Consumers' Perceived Benefits and Costs of Voice Commerce », 53rd Hawaii International Conference on System Sciences.

Stern S. E., Mullennix J. W. et Yaroslavsky I. (2006), Persuasion and social perception of human vs synthetic voice across person as source and computer as source conditions, International Journal of Human-Computer Studies, 64(1), pp. 43–52.

Sundar, S. S., Jung, E. H., Waddell, F. T. and Kim, K. J. (2017) “Cheery companions or serious assistants? Role and demeanour congruity as predictors of robot attraction and use intentions among senior citizens, International Journal of Human Computer Studies, 97, pp.88-97. DOI:10.1016/j.ijhcs.2016.08.006

Velkovska J. et Zouinar M. (2018). « The illusion of natural conversation: interacting with smart assistants in home settings », 2018 CHI Conference on Human Factors in Computing Systems, ACM, DOI:10.1145/3170427.3170619

Vernuccio, M. ; Patrizi, M. ; Pastore, A. (2021). « Delving into brand anthropomorphisation strategies in the experiential context of name-brand voice assistants », Journal of Consumer Behaviour. <https://doi.org/10.1002/cb.1984>

Viot, C. et Bressolles, G. (2012). « Les agents virtuels intelligents : quels atouts pour la relation client ? », Décisions Marketing, (65), pp. 45-56.

Wang W. et Benbasat I. (2007). « Recommendation Agents for Electronic Commerce: Effects of Explanation Facilities on Trusting Beliefs », Journal of Management Information Systems, 23, pp. 217-246.

Zoghaib, A. (2019). « Persuasion vocale : les effets des caractéristiques vocales et du genre de l'orateur sur les réponses du consommateur », Recherche et Applications en Marketing.

Annexes

Variabiles	Caractéristiques	Volume	Pourcentage
Sexe	Homme	8	53%
	Femme	7	47%
Age	Moins de 18 ans	2	13%
	18-25	2	13%
	26-35	3	20%
	36-50	5	34%
	51 ans et plus	3	20%
CSP	CSP+	9	60%
	CSP-	6	40%
Utilisateur aguerri ou nouvel utilisateur	Utilisateur aguerri	11	73%
	Nouvel utilisateur	4	27%
Fréquence d'utilisation	Chaque jour	11	73%
	Chaque semaine	3	20%
	Chaque mois	1	7%

Thème 1 : Délimitation du concept d'assistant vocal et habitudes d'usage

Q1 : Quel type d'assistant vocal possédez-vous (marque, modèle, avec ou sans écran...) ?

Q2 : Racontez-moi vos usages avec votre (ou vos) assistant vocal ?

Thème 2 : Perception de l'assistant vocal

Q3 : Quel est votre opinion sur les assistants vocaux ? Qu'est-ce que cela vous évoque ?

Q4 : Quels sont les avantages que vous trouvez à l'utilisation d'un assistant vocal et pourquoi ?

Q5 : Quels sont les inconvénients que vous rencontrez en utilisant votre assistant vocal ?

Thème 3 : Les effets de la voix de l'assistant vocal sur les réactions utilisateurs

Q6 : Que pensez-vous de la voix de votre assistant vocal / des applications vocales ?

Echangez avec l'application de la marque 1 (puis 2, 3, 4 et 5), puis pendant l'utilisation ;

Q7 : Décrivez ce que vous ressentez en utilisant l'assistant. Pourquoi ?

Q8 : Quels sont les effets de la voix de l'assistant sur votre comportement pendant l'utilisation ?

Q9 : Quels sont les effets de la voix de l'assistant sur vos émotions pendant l'utilisation ?