

Comprendre les profils des voicenautes et leur intention d'achat par les assistants vocaux :

Les apports de la théorie du conflit décisionnel

Amira BERRICHE

Maître de Conférences, IAE de Lille

IAE - 104, Avenue du Peuple Belge, Lille, 59800

Tél : 03. 20. 12. 34. 50

amiberriche@gmail.com

Christophe BENAVENT

Professeur des Universités, Paris Nanterre

200, Avenue de la République, 92000 Nanterre

Tél : 06. 23. 23. 62. 56

christophe.benavent@u-paris10.fr

Comprendre les profils des voicenautes et leur intention d'achat par les assistants vocaux :

Les apports de la théorie du conflit décisionnel

**Résumé :**

Face à l'utilisation fulgurante du consommateur des assistants vocaux et des enceintes connectées chez lui, le commerce vocal intégré à ces dispositifs représente aujourd'hui un nouvel enjeu pour les marques en tant que nouveau canal client. Toutefois, le voicenaute réalise très peu d'achats. Les travaux de recherche utilisent les théories classiques telles que : le modèle de l'acceptation de la technologie et le modèle unifié de l'acceptation et l'utilisation des technologies pour analyser les bénéfices et coûts perçus ce qui ne permet pas de comprendre l'écart entre les bénéfices et coûts en d'autres termes le conflit décisionnel connu sous le nom de la balance décisionnelle (Janis et Mann, 1977). Cette recherche a pour objectif d'identifier différents profils de voicenautes et d'analyser leur conflit décisionnel pour prédire leur intention d'adopter le commerce vocal. L'étude qualitative réalisée montre quatre profils de voicenautes (réfractaires, hédonistes, utilitaristes et non-adeptes). Les résultats obtenus contribuent à l'amélioration des stratégies et actions marketing mises en place par les marques d'assistant vocal ainsi que les marques/ annonceurs auprès de ces profils.

**Mots clés :** v-commerce, voicenautes, conflit émotionnel, conflit décisionnel, intention d'achat par assistant vocal.

Understanding the profiles of voice user and their purchase intention through voice assistants:

The contributions of decision-making conflict theory

**Abstract :**

In face of the rapid consumer use of voice assistants and smart speakers in their homes, voice commerce integrated into these devices represents a new challenge for brands today as a new customer channel. However, the voicenaut makes very little purchases. Previous research use classic theories such as: the technology acceptance model and the unified theory of acceptance and use of technology to analyze the perceived benefits and costs which does not allow to understand the the gap between benefits and costs in other words the decisional conflict known as the decisional balance (Janis and Mann, 1977). This research aims to identify different profiles of voice user and analyze their decision conflict to predict their intention to adopt voice commerce. The qualitative study carried out shows four profiles of voice user (refractory, hedonists, utilitarians and non tech savvy). Results contribute to the improvement of the strategies and marketing actions designed by voice assistant brands as well as the brands / advertisers with these profiles.

**Key words :** v-commerce, voice user, emotional conflict, decision conflict, intention to purchase by voice assistant.

## Introduction et objectifs

Au cours des cinq dernières années, la croissance annuelle du marché des assistants vocaux intégrés à des enceintes connectées a dépassé les 650% soit deux fois plus le taux d'adoption des smartphones (Coskun et Hellriegel, 2020). Plusieurs marques d'enceintes connectées ont réussi à s'introduire à l'intérieur des foyers telles qu'Amazon Echo, Apple HomePod, Google Home, Alibaba Tmall Genie, Xiaomi Xiao AI, and Baidu Xiaodu. D'après Ba, Mendelson et Zhu (2020), le taux de pénétration de ces dispositifs d'Iot (Internet Of Thing)<sup>1</sup> est le plus fort dans les foyers américains : il est de 32% en 2019 et devrait atteindre les 51% d'ici 2022.

Ces dispositifs permettent à l'utilisateur dit voicenaute de réaliser ses achats les mains libres et de leur offrir une nouvelle interface utilisateur « *zéro clic* ». En utilisant uniquement leur voix, les consommateurs peuvent formuler des requêtes de recherche vocale et confirmer des actions d'achat sans avoir besoin d'utiliser des interfaces visuelles ou de saisie manuelle. Par exemple, pour aller au cinéma, aujourd'hui, le consommateur achète ces billets en ligne en quelques clics. Avec la recherche vocale, le consommateur éliminera ces clics et commence le processus d'achat par une simple requête : « *Dis Google, achète une place de cinéma à X pour voir le film X* ».

L'usage de la voix simplifie le processus d'achat et améliore l'expérience client. L'une des principales raisons pour lesquelles les consommateurs préfèrent la voix aux autres formes de communication est qu'elle est plus facile, intuitive et rapide. Une personne moyenne peut généralement parler environ 125-175 mots par minute alors que la plupart des gens ne peuvent taper qu'un peu moins de 40 mots par minute sur un clavier.

Les avancées technologiques en traitement automatique du langage naturel (TALN) et reconnaissance vocale, proposent un nouveau canal client à travers les assistants vocaux. Un assistant vocal fait référence aux agents conversationnels qui effectuent des tâches avec ou pour le consommateur, qu'elles soient de nature fonctionnelle (ex : contrôle de la maison intelligente), transactionnelle (ex : recherche d'informations produit ou passer commande) ou sociale (ex : lire les avis) et sont capables d'auto-améliorer leur compréhension de l'interlocuteur et du contexte (Mari, 2019). Le commerce vocal dit *voice commerce* (v-commerce) ou le shopping commerce est décrit comme la résultante de l'interaction vocale entre un assistant vocal et un consommateur en vue de trouver des produits et de procéder à un achat (Vachaud et Geerts, 2020). Il s'agit d'un processus d'interaction entre le consommateur et la marque. Le commerce vocal génère aujourd'hui 2 milliards de CA et pèsera 40 milliards de dollars dans le monde en 2022 (source : cabinet OC&C Strategy Consultants). Les quatre catégories les plus vendues par le biais de ce nouveau canal sont : les produits d'épicerie (20%), le divertissement (19%), l'électronique (17%) et l'habillement (8%).

En France, très peu de marques ont développé des applications vocales opérationnelles, compatibles avec les assistants vocaux. En novembre 2017, Monoprix est devenue la première enseigne française de la grande distribution à adopter le v-commerce en proposant une expérience permettant de commander ses courses par la voix, via Google Home. Le même service de commande vocale est proposé par Intermarché depuis janvier 2019 et sera bientôt

---

<sup>1</sup> L'IoT est l'acronyme de Internet Of Things (Internet des Objets en français). Il désigne un système où les objets physiques sont connectés à Internet et sont capables de créer et transmettre des données afin de créer de la valeur pour ses utilisateurs à travers divers services (agrégation, analytique...).

accessible sur Alexa. FNAC, Darty, Leclerc, Sephora ou encore Auchan sont les nouveaux acteurs français du v-commerce et proposent des applications permettant principalement d'obtenir des informations telles que l'état d'avancement d'une commande, la proximité d'un magasin ou des précisions concernant un produit. Un autre exemple intéressant est celui d'EDF avec l'application vocale « EDF & Moi » disponible depuis juin 2018 et qui donne des conseils aux clients sur les économies d'énergie et la possibilité aux clients de poser des questions sur leur contrat. Les marques qui offriront une bonne expérience client via les assistants vocaux verront leur chiffre d'affaires augmenter et bénéficieront d'un bouche-à-oreille positif (Capgemini Digital Transformation Institute, 2018).

D'un autre côté, les consommateurs ont une préférence croissante pour les assistants vocaux comme mode d'interaction avec leurs marques. Le recours à la voix simplifie l'expérience client, d'une part, parce que le langage est naturel ; d'autre part, car la situation est choisie par le consommateur. L'achat est déclenché au moment où le besoin est ressenti, qu'il naisse de la lecture d'une recette de cuisine ou corresponde à une disponibilité de la personne. Toutefois, les achats réalisés via les assistants vocaux représentent actuellement seulement 3% du volume total des dépenses de leurs utilisateurs (Capgemini Digital Transformation Institute, 2018). Pour maximiser le potentiel du v-commerce et exploiter la voix en tant que nouveau canal client, les marques doivent offrir des services qui mettent en évidence les avantages et bénéfices des assistants vocaux et des expériences supérieures aux autres interfaces utilisateurs existantes.

Cette recherche a pour objectif d'identifier les différents profils de voicenautes, de comprendre et d'analyser les bénéfices et les coûts que chacun des profils souhaite obtenir ou éviter du v-commerce. Les travaux de recherche n'apportent pas une explication suffisante sur les bénéfices et coûts perçus de l'adoption du v-commerce. Certains travaux comparent les bénéfices du v-commerce et du e-commerce (Moriuchi, 2019). D'autres recherches utilisent les théories classiques telles que : le modèle de l'acceptation de la technologie (TAM) (Davis, Bagozzi et Warshaw, 1989) et le modèle unifié de l'acceptation et l'utilisation des technologies (UTAUT) (Venkatesh, Davis et Morris, 2007) pour analyser les bénéfices et les coûts (Rzepka, Berger et Hess, 2020). Elles mettent en évidence plusieurs bénéfices à savoir l'efficacité qui se manifeste par le gain de temps pour passer commande et le multi-tasking c'est-à-dire la capacité à réaliser plusieurs tâches au même temps car l'assistant vocal libère les mains et les yeux ; la commodité (plus besoin de taper sur un clavier) ; l'anthropomorphisme (l'assistant vocal est considéré comme un partenaire conversationnel agréable) et la capacité d'apprendre à travers le feedback. Elles montrent aussi plusieurs coûts perçus tels que : le manque de transparence et de confiance envers les algorithmes de recommandations des produits, les préoccupations relatives au respect de la vie privée et à la confidentialité ainsi que le manque de contrôle perçu en cas de commande accidentel ou piratage.

Ces recherches utilisent des modèles théoriques qui ne tiennent pas compte de la rationalité limitée du consommateur. Ces modèles considèrent que tous les consommateurs suivent le même processus cognitif. En effet, selon la théorie de la diffusion de l'innovation de Roger, différents groupes de clients peuvent correspondre à différentes phases d'adoption d'une innovation et par conséquent, ces groupes n'ont pas le même processus cognitif de l'évaluation des avantages et des inconvénients. Aussi, ces modèles se limitent à analyser les influences des bénéfices et des coûts sur l'intention d'adopter le v-commerce sans étudier l'écart entre les bénéfices et les coûts, en d'autres termes, leur conflit décisionnel connu sous le nom de la balance décisionnelle (Janis et Mann, 1977). Ce modèle est très utilisé en

marketing social et montre qu'une personne est susceptible de changer de comportement problématique et d'adopter un comportement bénéfique pour elle et pour la société lorsque l'évaluation des bénéfices dépasse les coûts perçus (Salerno et al., 2015). En outre, ces modèles intègrent la valeur hédonique sans prendre en compte le conflit émotionnel que nous appelons bien-être subjectif de l'expérience (ratios émotions positives/émotions négatives) et définissons comme *l'évaluation personnelle et subjective d'un état de satisfaction et l'absence d'un état d'insatisfaction au moment de l'expérience avec l'assistant vocal*.

## **Question de recherche**

Les questions de recherche sont par conséquent les suivantes : Quels sont les différents profils de voicenautes et quels critères utilisés pour les identifier ? Comment chaque profil analyse les bénéfices et les coûts et comment leurs conflits décisionnel et émotionnel influencent leur intention d'achat par le biais des assistants vocaux ?

*Au plan théorique*, cette recherche contribue à la compréhension des déterminants de l'adoption du v-commerce. Elle importe du champ du marketing social le modèle du conflit décisionnel pour analyser les bénéfices et coûts perçus ce qui constitue un enrichissement aux modèles théoriques classiques. *Au plan managérial*, les recommandations peuvent servir d'un côté, aux marques d'assistant vocal et d'un autre côté, aux marques / annonceurs. Cette recherche devrait contribuer à l'amélioration des stratégies et actions marketing mises en place par ces marques auprès des voicenautes dans le but de les inciter à adopter le v-commerce et à passer commande par le biais des assistants vocaux.

Le plan de cette recherche commence par la présentation d'une typologie des assistants vocaux développée par nos soins pour une meilleure compréhension des différentes formes qu'ils peuvent prendre. Ensuite, est développée, une synthèse des recherches sur les déterminants d'adoption du v-commerce qui met en évidence les bénéfices et coûts perçus déjà étudiés dans la littérature. La méthodologie de recherche est définie avant de présenter successivement la discussion des résultats, les implications managériales ainsi que les limites et voies de recherche.

## **Revue de littérature**

### **I. V-commerce et typologie des assistants vocaux**

Le commerce vocal dit voice commerce (v-commerce) ou le shopping commerce est décrit comme la résultante de l'interaction vocale entre un consommateur et un assistant vocal en vue de trouver des produits et de procéder à un achat (Vachaud et Geerts, 2020). L'assistant vocal peut prendre différentes formes. Nous proposons une typologie des assistants vocaux en utilisant deux critères de distinction pour une meilleure compréhension de ses formes : (1) le dispositif dans lequel l'assistant vocal est intégré et (2) le ou les sens suscité(s). Les trois formes des assistants vocaux identifiés (screen-first, voice-only (screenless) et voice-first) sont détaillées dans le tableau suivant (Tableau 1).

Les travaux de recherche ne distinguent pas les trois formes et mélangent parfois les bénéfices et les coûts de chaque forme dans une seule étude tels que : l'absence du visuel (spécifique au voice-only), libère les mains et les yeux (spécifique aux voice-only et voice-first), son utilisation par d'autres membres d'un même foyer (spécifique aux voice-only et voice-first), connaît tout sur vous (spécifique au screen-first) etc. C'est pour cela, il nous paraît important

d'établir une typologie des assistants vocaux. La synthèse de la revue de littérature sur l'analyse des bénéfices et des coûts de l'adoption du v-commerce est développée dans la partie suivante.

Tableau 1 : Typologie des assistants vocaux

Critères de distinction	Screen-first	Voice-only (screenless)	Voice-first
Dispositif	Intégré dans des dispositifs (smartphone, tablette, TV)	Intégré à une enceinte connectée	Intégré à une enceinte connectée avec écran tactile
Sens	L'ouïe + le toucher + la vue	L'ouïe	L'ouïe + la vue
Description	L'interaction avec l'assistant vocal est assez limitée car seulement les premières étapes d'une action donnée sont réalisables par la voix. L'interaction continue manuellement à travers une commande écrite (clavier, télécommande, etc.).	L'interaction avec l'assistant vocal se fait uniquement par la voix, car il n'y a pas d'interface graphique. Cela augmente l'effort cognitif de l'utilisateur qui se retrouve à devoir interpréter, comprendre et retenir les informations données par l'assistant. De plus, les informations sont données dans un ordre précis qui ne peut pas être changé, ce qui réduit la liberté de l'utilisateur et la qualité de son expérience (UX). Par exemple, un utilisateur ne peut pas connaître le troisième produit d'une liste sans passer par les deux premiers.	L'interaction avec l'assistant vocal se fait par la voix dans un premier temps. L'interaction continue par un balayage visuel ce qui réduit la charge mentale. Cependant, les interactions tactiles sont limitées. Il n'est pas possible de naviguer sur le site, voir les avis, le contenu de son panier, etc. Il faut demander à l'assistant de vous afficher ces éléments.
Exemple			

## II. Les déterminants de l'adoption du v-commerce : Analyse des bénéfices et des coûts

Pour comprendre les déterminants de l'adoption du v-commerce, les travaux de recherche utilisent le modèle de l'acceptation de la technologie (TAM) (Davis, Bagozzi et Warshaw, 1989) et le modèle unifié de l'acceptation et l'utilisation des technologies (UTAUT) (Venkatesh, Davis et Morris, 2007). Ces modèles montrent que l'utilisation d'une technologie est déterminée par l'intention de son adoption. Cette intention est expliquée par la performance espérée et l'effort attendu qui sont directement tiré du TAM et correspondent successivement à l'utilité perçue et la facilité d'utilisation. Elle est aussi expliquée par l'influence sociale qui correspond dans le modèle du comportement planifié (TCP) (Ajzen, 1991) aux normes subjectives et des conditions facilitatrices spécifiques au modèle UTAUT et définies par la perception du voicenaute quant au degré auquel la marque investit dans les ressources techniques qui facilitent l'utilisation de la technologie.

Les recherches en v-commerce pensent que ces modèles théoriques fournissent un fondement théorique suffisant pour analyser les bénéfices et coûts perçus de l'adoption du v-commerce (Rzepka, Berger et Hess, 2020), car les bénéfices et les coûts influencent les attitudes, intentions et actions. Aussi, ces modèles permettent l'analyse des déterminants des bénéfices et coûts perçus plus spécifiquement les croyances positives des voicenautes renforcent la

perception des bénéfices alors que leurs croyances négatives accentuent leur perception des coûts.

Par conséquent, nous structurons la revue de littérature en deux parties : la première partie met en évidence les croyances positives et bénéfiques de l'adoption du v-commerce et la seconde partie développe les croyances négatives et coûts de son adoption.

## **1. Croyances positives et bénéfiques perçus de l'adoption du v-commerce**

Plusieurs bénéfices et croyances positives sont analysés par les travaux de recherche et sont repris et présentés dans cette partie.

*Efficiencie (utilité perçue).* Renvoie à la performance attendue ou l'utilité perçue à travers l'utilisation du v-commerce. La littérature souligne deux principaux éléments. (1) Le gain de temps dans la réalisation des achats et commande des produits à faibles implications est la principale raison de l'utilisation de l'assistant vocal et de la satisfaction de son utilisation. Selon Moussawi (2018), l'assistant vocal est rapide à fournir les réponses, leur fait gagner du temps surtout lorsqu'ils sont pressés. (2) Le multitasking, c'est-à-dire la capacité de réaliser plusieurs tâches au même temps est rendu possible car l'assistant vocal ne nécessite pas d'avoir un dispositif à la main, ce qui permet de libérer les mains et les yeux (par exemple, cuisiner, commander un produit qui arrive à sa fin et démarrer le minuteur).

*Commodité (utilité perçue).* Est définie comme le degré dans lequel une personne croit que la navigation ou l'engagement dans des transactions par le biais du commerce électronique ou mobile ou assistant vocal est sans effort. Les voicenautes trouvent le commerce vocal plus facile à utiliser que le e-commerce conventionnel (Kraus, Reibenspiess et Eckhardt, 2019). L'assistant vocal permet, en effet de rendre le processus d'achat plus confortable et dénué d'effort. D'après, Luger et Sellen (2016), il est plus facile et pratique d'utiliser la saisie vocale que de taper, car la voix est perçue comme plus rapide. Le processus d'achat est moins compliqué car les consommateurs n'ont pas à naviguer entre différentes pages de produits en ligne et à taper sur le clavier. De plus, l'effort mental est réduit car il confie à l'assistant vocal d'accomplir la tâche à leur place. L'aide à accomplir les tâches quotidiennes difficiles et à simplifier la routine plaît aussi aux voicenautes. En effet, ils n'ont pas besoin d'ouvrir une application et sélectionner un produit parmi une liste fastidieuse de produits. Certains considèrent que l'assistant vocal est davantage plus pratique qu'Amazon Dash Button<sup>2</sup> où il faut avoir un bouton pour chaque produit et chercher le bouton pour enregistrer la commande quand le produit arrive à sa fin.

*Plaisir et connexion émotionnelle (Valeur hédonique).* Les voicenautes éprouvent de façon général du plaisir durant leur interaction avec l'assistant vocal et s'amuse lorsqu'ils parlent avec lui (Moussawi, 2018 ; Rzepka, Berger et Hess, 2020). Certains éprouvent aussi de la gratitude, car ils considèrent que l'assistant vocal les aide dans leur quotidien, leur facilite la vie, les fait rire et toujours disponible pour leur venir en aide. Ils le trouvent sympathique, gentil et bienveillant avec eux et certains le considèrent comme un ami. Selon Jones (2018),

---

<sup>2</sup> L'Amazon Dash Button (bouton) est un objet connecté équipé d'une connexion Wi-Fi pour commander en ligne. Chaque bouton représente un produit de la vie courante : du papier toilette, de la lessive, des rasoirs, de la nourriture pour chien et du café. Il y a plus de 100 boutons. Le bouton ressemble à une sonnette et se fixe n'importe où dans la maison, généralement à côté du produit utilisé. Quand le produit arrive à sa fin, un appui sur le bouton Dash permet de l'ajouter directement à sa liste de courses AmazonFresh, et la commande est enregistrée.

une interaction amicale avec l'assistant vocal a une influence positive sur l'attachement envers la marque.

*Anthropomorphisme (conditions facilitatrices).* Nass et Moon (2000) remarquent que les individus réagissent avec les ordinateurs de manière sociale tout en sachant que les ordinateurs ne sont pas comme des humains. En effet, les humains ont tendance à anthropomorphiser la technologie même en l'absence de caractéristiques humaines telles que la voix. De façon générale, les systèmes de sortie vocale induisent des caractéristiques semblables à l'homme, augmentent l'anthropomorphisme perçu et génèrent une attitude positive envers le système (Luger et Sellen, 2016 ; Jones, 2018 ; Moriuchi, 2019 ; Rzepka, Berger et Hess, 2020). Toutefois, les assistants vocaux sont des systèmes de dialogue dotés de comportements qui ressemblent à ceux d'un être humain et sont développés pour imiter les interactions humaines naturelles. Ils réagissent aux interlocuteurs lorsque leur nom est appelé et utilisent une personnification "je" en référence à eux-mêmes. Tout comme les relations interpersonnelles, ils mémorisent les faits pertinents pendant la conversation pour donner un sentiment de continuité aux interactions suivantes, utilisent un langage informel (par exemple, ils vous appellent par votre prénom), font des blagues, apportent une réponse spontanée et utilisent leur voix pour faire vivre des émotions. Ils sont considérés comme un partenaire de conversation agréable (Mari, Mandelli et Algesheimer, 2020).

*Personnalisation, apprentissage à travers l'interaction et feedback (utilité perçue).* Certains voicenautes déclarent que l'assistant vocal leur fournit des réponses selon leurs préférences et habitudes de consommation. Quelques-uns constatent son évolution et apprentissage au fil du temps (Kraus, Reibenspiess et Eckhardt, 2019). En effet, l'assistant vocal discute avec eux pour connaître leurs préférences, interprète, élargit ses connaissances et traite les demandes de manière contextuelle, apprend des erreurs et intègre les feedbacks pour améliorer ses futures requêtes. L'assistant vocal collecte, transforme et stocke ce qu'il connaît et entend pour prédire leurs futurs comportements. Il peut identifier des interactions insatisfaisantes ou des échecs de compréhension et corriger automatiquement ces erreurs. L'utilisation des techniques d'auto-apprentissage lui permet de corriger ses erreurs, de réinitialiser les requêtes non abouties jusqu'à réussir à fournir la bonne réponse pour réduire les frictions pendant l'interaction et met à jour automatiquement les corrections pour un grand nombre de requêtes (Mari, 2019). Son raisonnement automatisé donne au système la capacité à utiliser l'information stockée pour répondre aux prochaines requêtes. L'apprentissage automatique le rend capable de s'adapter aux nouvelles circonstances auxquelles il peut être confronté. L'intelligence artificielle est perçue comme l'art de créer des systèmes fonctionnels qui utilisent des modèles informatiques pour reproduire une intelligence humaine, des facultés mentales et un comportement intelligent. Elle imite le raisonnement humain et dispose de facultés rationnelles (Moussawi, 2018).

## **2. Croyances négatives et coûts perçus de l'adoption du V-commerce**

Dans cette partie, la synthèse des coûts et croyances négatives perçus de l'adoption du V-commerce est développée.

*Manque de transparence et de confiance (croyances négatives).* La transparence perçue fait référence à la compréhension par l'utilisateur du fonctionnement interne d'un système, de ses motivations sous-jacentes et caractéristiques qui déterminent son comportement. Les voicenautes craignent que lorsqu'ils commandent un produit, l'assistant vocal choisit un produit par défaut sans comparer ou croiser les avis des consommateurs. L'étude qualitative

réalisée par Rzepka, Berger et Hess (2020) montrent que certains répondants souhaitent comprendre comment l'assistant vocal prend sa décision d'achat : « *Je trouve aléatoire comment il prend une décision* » ; « *Je ne sais pas sur quelle base il choisit un produit. Choisit-il le meilleur ? Comment décide-t-il ?* ». Ils déclarent avoir besoin des représentations visuelles des produits et de consulter les avis clients. En effet, ils se méfient des choix produits proposés par l'assistant vocal et préfèrent se référer aux évaluations clients (Rzepka, Berger et Hess, 2020). Ils ne considèrent pas la technologie comme suffisamment fiable pour lui confier leurs achats. Aussi, ils ne font pas confiance à l'assistant vocal pour leurs achats de produits. Ils remettent en question sa compétence et bienveillance à leur proposer le meilleur choix et s'interrogent sur l'éthique de l'assistant vocal et sa tendance à favoriser les vendeurs, marques et produits les plus rentables ou encore les produits de sa propre marque comme Amazon.

*Capacité de contrôle limitée (contrôle perçu)*. Le contrôle perçu fait référence à la perception des personnes que les résultats obtenus sont les conséquences de leur propre comportement. Les voicenautes signalent une perte de contrôle perçue sur l'assistant vocal. Ils craignent que d'autres personnes (invités, amis) utilisent leur enceinte connectée et déclenchent un achat accidentel ou encore des craintes de piratage comme commander des produits sans leur consentement et à leur insu. Certains préfèrent une modalité de saisie manuelle qui leur permet de parcourir différentes options, de choisir consciemment un produit et de confirmer la commande. De plus, ils craignent le risque d'incompréhension pouvant conduire à un "mauvais achat" de produits c'est-à-dire si le consommateur n'a pas pu atteindre ses objectifs d'achats (par exemple, mauvais produit ou mauvaise commande) (Rzepka, Berger et Hess, 2020).

*Faible maturité technique (facilité d'utilisation)*. Selon Rzepka, Berger et Hess (2020), les voicenautes trouvent la maturité technique de l'assistant vocal encore faible. Sa capacité d'interagir de manière réactive est limitée. Il a du mal à comprendre le contexte d'une conversation et à savoir de quelles informations ils ont besoin. Pour certains, il n'est pas encore très intelligent et n'a pas encore atteint son plein potentiel (Moussawi, 2018). Quelques-uns reportent l'émergence de tension au cours de leur interaction avec lui. Certains d'entre eux se sentent ennuyés lorsqu'ils doivent répéter la commande plusieurs fois. D'autres déclarent se sentir frustrés lorsqu'il ne répond pas à leurs attentes. D'autres encore se sentent irrités lorsqu'il semble incohérent. Ils ont exprimé leur frustration à la suite de l'obtention de réponses ou de réactions différentes à des demandes similaires. Pour se faire comprendre par l'assistant vocal, ils doivent s'adapter à celui-ci et non l'inverse. Ils doivent réduire le nombre de mots utilisés "*faire des économies de mots*", utiliser des mots-clés et éviter les mots complexes ou le langage naturel (mots familiers), utiliser des termes spécifiques, changer l'énoncer, parler plus lentement / clairement et parfois changer d'accent (Luger et Sellen, 2016).

*Anthropomorphisme (conditions facilitatrices)*. L'étude réalisée par Rzepka, Berger et Hess (2020) montre que la ressemblance entre l'assistant vocal et l'être humain peut mettre mal à l'aise certains voicenautes qui déclarent : « *Je préfère parler à un ordinateur qui ne ressemble pas à un humain, [...] C'est moins effrayant !* ». Ils préfèrent une expérience de type humain en termes de réactivité et compétences, mais ils s'abstiennent des imitations humaines, par exemple une interaction avec une voix humaine. Les auteurs expliquent ces réactions par la théorie de la vallée dérangeante ou vallée de l'étrange (uncanny valley) de Masahiro (1970), selon laquelle plus un robot androïde est similaire à un être humain, plus ses imperfections

sont remarquées et ils incitent les futures recherches à étudier le niveau d'anthropomorphisme optimal pour une meilleure acceptation de l'assistant vocal.

*Préoccupations relatives au respect de la vie privée et à la confidentialité (éthique).* Elles renvoient à la réticence des voicenautes à partager des informations personnelles avec l'assistant vocal par souci de manque de sécurité et de confidentialité. Certains s'inquiètent de communiquer à leur assistant vocal des informations bancaires pour faire un achat ou des informations personnelles et que ces informations soient subtilisées (hacking) ou utilisées par le vendeur à d'autres fins sans leur consentement (Mari, 2019 ; Rzepka, Berger et Hess, 2020 ; Vachaud et Geerts, 2020).

L'analyse de la revue de littérature met en évidence plusieurs bénéfices et coûts de l'adoption du v-commerce sans apporter de précision sur les différents profils de voicenautes et leur perception de ces bénéfices et coûts. Nos questions de recherche sont les suivantes : Quels sont les différents profils de voicenautes et quels critères utilisés pour les identifier ? Comment chaque profil analyse les bénéfices et les coûts et comment ce conflit décisionnel influence leur intention d'achat par les assistants vocaux ?

## Méthodologie

Pour apporter des réponses aux questions de recherche, une étude qualitative a été réalisée par entretiens semi-directifs en face-à-face. *La suffisance théorique* (Dey, 1999) c'est-à-dire le recueil de données et analyse s'arrêtent lorsque les thèmes obtenus offrent un cadre explicatif suffisant compte tenu des données collectées- a été atteinte à onze entretiens (Tableau 2). La durée moyenne d'entretien a été de 30 minutes. *L'échantillon* comprend 7 femmes et 4 hommes en activité âgés de 21 à 51 ans. 8 répondants possèdent une enceinte connectée (Amazon Echo ou Google Home) et ont déjà interagit avec l'assistant vocal et 3 répondants l'ont désactivé sur leur smartphone.

*Le guide d'entretien* comportait des questions ouvertes sur la fréquence et la nature d'utilisation de l'assistant vocal ainsi que l'utilité perçue et la satisfaction de l'expérience utilisateur avec cette technologie. Les questions se poursuivaient sur leur expérience d'achat avec les marques vocales : les motivations et freins dans le processus d'achat par l'assistant vocal. Tous les entretiens ont été enregistrés et retranscrits en intégralité.

*L'analyse de contenu* a été réalisée en respectant les phases de l'analyse verticale et l'analyse horizontale. Une lecture flottante de l'ensemble des entretiens a donc été faite et les unités d'analyses et les rubriques de la grille de lecture ont été déterminées. Les fréquences des différents thèmes ont été calculées.

Tableau 2 : Description de l'échantillon

Répondants	Age	Genre	Nombre de personnes dans le foyer	Marque de l'assistant vocal possédé
Répondant 1	23	H	2	Assistant téléphone (Siri) désactivé
Répondant 2	21	F	2	Google Home
Répondant 3	21	F	2	Google Home
Répondant 4	21	F	2	Amazon Echo
Répondant 5	24	F	4	Google Home
Répondant 6	21	H	2	Assistant téléphone (Siri) désactivé
Répondant 7	26	F	2	Google Home

Répondant 8	21	H	2	Assistant téléphone (Siri) désactivé
Répondant 9	22	F	2	Google Home
Répondant 10	25	H	3	Amazon Echo
Répondant 11	51	F	4	Google Home

## Résultats

L'analyse de contenu de l'étude qualitative, nous a permis d'identifier quatre profils (Figure 1) en tenant compte simultanément de :

(1) La perception et croyances de la facilité d'utilisation de l'assistant vocal définie comme le degré dans lequel le voicenaute pense que l'utilisation de l'assistant vocal est facile et dénuée d'efforts

(2) Le niveau de connaissances de cette technologie renvoie au savoir, à la compréhension et à la conscience du mode d'utilisation de l'assistant vocal (dire « Ok, Google ! » ou « Alexa » pour commencer une interaction avec l'AV ou dire « Parler à X (le nom de la marque) » pour accéder au site de la marque, ainsi qu'aux différentes fonctionnalités, commandes et usages possibles. Les quatre profils sont les suivants :

Figure 1 : La matrice des profils de voicenautes identifiés à l'aide de l'étude qualitative

Niveau des connaissances	Facilité d'utilisation	
	Faible (-)	Elevée (+)
Faible (-)	<p><b>Profil 1 : Les réfractaires</b></p> <p><i>« Trouvent l'AV nul »</i></p>	<p><b>Profil 2 : Les hédonistes</b></p> <p><i>« Considère l'AV comme un simple gadget »</i></p>
Elevé (+)	<p><b>Profil 4 : Les non-adeptes</b></p> <p><i>« Pensent que l'AV les prive du plaisir de faire les choses d'eux-mêmes »</i></p>	<p><b>Profil 3 : Les utilitaristes</b></p> <p><i>« Pensent que l'AV leur rend la vie plus simple, car il fait les choses à leur place »</i></p>

- Profil 1 : Les réfractaires ont à la fois un faible niveau de connaissances du mode d'utilisation de l'assistant vocal, ses fonctionnalités, commandes et usages possibles et de faibles croyances envers sa facilité d'utilisation. La majorité des personnes appartenant à ce segment ne possède pas d'assistant vocal ou le désactive sur leur smartphone et pourtant juge l'interaction avec celui-ci complexe. En effet, ces personnes pensent que l'assistant vocal est incapable de comprendre leur requête, peu intuitif et difficile à utiliser. Par exemple, *« Je ne sais pas comment ça fonctionne, de toute façon je préfère vérifier de moi-même sur internet. Je trouve ça nul ! »*

(Répondant 1) Ces personnes se créent ainsi des barrières psychologiques injustifiées et font un transfert de leurs faibles connaissances et compétences sur l'assistant vocal.

- Profil 2 : Les hédonistes trouvent l'assistant vocal facile d'utilisation, ludique, divertissant et amusant. Ils l'utilisent comme un simple gadget pour par exemple, écouter la musique, activer le minuteur, raconter une histoire aux enfants, demander la météo etc. Ils l'utilisent aussi pour se divertir « *Parfois je lui demande de faire des blagues, des quizzes, des énigmes, ça passe le temps pendant ce confinement* » (Répondant 3) Dans leur jugement de la persuasion de l'assistant vocal, ils se concentrent sur les éléments périphériques plutôt que sur des éléments centraux tels que son design, sa valeur et son apparence (objet de décoration), la qualité des enceintes pour écouter la musique et la rapidité de ses réponses. Ils ont de faibles connaissances sur les fonctionnalités et usages possibles de l'assistant vocal telles que accéder au site de la marque, commander directement son produit, ajouter des produits à une liste de course, commander et se faire livrer au drive ou chez soi. Ils sont prêts à découvrir les possibilités d'usage de l'assistant vocal. Par exemple, « *Actuellement, il n'est pour moi qu'un simple gadget divertissant et rien d'autre, parce que je n'ai jamais vraiment ressenti le besoin de l'utiliser, à part euh en tant qu'enceinte pour la musique ou des recherches, en plus je ne l'ai pas forcément depuis longtemps donc euh je n'ai jamais pensé à regarder toutes ses fonctionnalités [...] Ah oui j'aimerais bien découvrir tout ce qu'il peut faire, oui.* » (Répondant 3)
- Profil 3 : Les utilitaristes disposent d'un niveau de connaissances élevé sur les fonctionnalités et usages de l'assistant vocal et trouvent son utilisation très intuitive. Ils sont favorables à répéter leur requête jusqu'à ce qu'elle soit comprise par l'assistant vocal. Ils comprennent les limites de cette technologie et tolèrent les erreurs à partir du moment où l'assistant vocal les décharge de certaines tâches, leur fait gagner du temps et surtout pour le plaisir de faire les choses à leur place. Par exemple, « *Je l'utilise principalement pour effectuer des tâches à ma place [...] Ça facilite plutôt la vie encore plus si on va plus loin et que l'on achète les prises connectées pour pouvoir euh allumer la lumière à distance, allumer la cafetière ou ce genre de chose* » (Répondant 7) ; « *c'est super pratique et ça peut faire gagner du temps pendant que je suis occupé à faire autre chose par exemple* » (Répondant 8). Les personnes appartenant à ce groupe ont déjà acheté au moins une fois par l'assistant vocal et ont un grand espoir envers le développement du commerce vocal dans un futur proche. Toutefois, ils déclarent acheter seulement des produits d'achat courant (par exemple, livre, les courses, etc.) et qu'ils ne sont pas encore prêts à réaliser des achats de produits réfléchis. Pour ce type de produits, l'assistant vocal devient un canal de découverte du produit ou de la marque qui mène vers son site e-commerce. Par exemple, « *Je n'aurais pas acheté un truc nouveau via l'assistant : j'aurai juste écouté les informations concernant mon produit et ensuite je serai allé sur mon ordinateur pour me faire une meilleure idée [...] mais je vais bien garder à l'esprit le premier résultat proposé par l'assistant vocal* » (Répondant 4) Par conséquent, la marque a tout intérêt à optimiser son site pour la recherche par la voix, autrement dit sa stratégie SEO vocal.
- Profil 4 : Les non-adeptes ont de très bonnes connaissances des fonctionnalités et usages possibles offerts par l'assistant vocal. Cependant, ils préfèrent accomplir les tâches d'eux-mêmes et pensent que l'assistant vocal les prive du plaisir de faire les choses d'eux-mêmes. Par exemple, « *mais après ça voudrait dire qu'on a plus rien à*

*faire de notre journée dans la maison en fait ... Je veux dire ... On se lèverait le matin et voilà il suffirait d'appuyer sur un bouton et hop tout est fait. Je trouve ça un peu dommage et niveau activité physique c'est pas ce qu'il y'a de mieux. J'ai besoin d'un minimum d'activité physique donc je sais qu'à un moment j'en aurais marre. »* (Répondant 2) La difficulté d'utilisation de l'assistant vocal est constatée lorsqu'on leur demande de réaliser des tâches transactionnelles telles que renseignement produit, passer commande, acheter un produit etc. Le manque d'envie d'utiliser l'assistant vocal réduit la facilité d'utilisation perçue.

## Discussion

Cette recherche met en évidence quatre profils de voicenautes. L'objectif de cette recherche est d'une part, de proposer un modèle conceptuel pour comprendre les déterminants de l'intention d'achat par l'assistant vocal et d'autre part, de réaliser une analyse comparative de ces déterminants entre les quatre profils de voicenautes.

Trois principaux déterminants de l'intention d'achat par l'assistant vocal sont identifiés : (1) les émotions suscitées (positives et/ou négatives) durant l'expérience de l'assistant vocal, (2) les bénéfices perçus de l'expérience et (3) les coûts perçus.

Les résultats montrent que l'interaction avec l'assistant vocal peut générer simultanément des émotions positives (*gadget, divertissant, amusant*) et des émotions négatives (*Mince ! Il se moque de moi ou quoi ?*) et le voicenaute peut aussi être conscient à la fois de ces bénéfices ou utilité définis ici comme la mesure dans laquelle une personne croit que l'aide de l'assistant vocal renforcerait sa performance (*C'est le côté pratique, gain de temps. Tu peux faire autre chose au même temps quand tu as les mains prises, ou tu cuisines, ou les mains sales...*) mais aussi de ces coûts définis ici comme l'ensemble des évaluations subjectives des conséquences de l'utilisation de l'assistant vocal (*Je sais que Google Home risque de récupérer tout ce qu'on dit à notre insu*). En effet, un article publié par Le Figaro dénonce des employés d'Amazon qui écoutent les conversations des utilisateurs avec Alexa dans le but d'améliorer les performances de la reconnaissance vocale (Diallo, 2019)

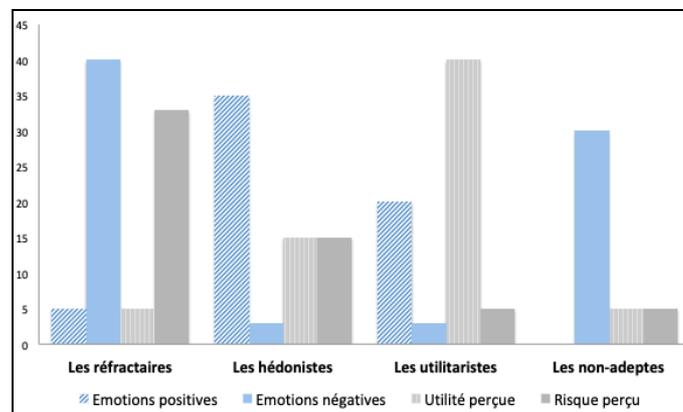
L'intention d'achat par l'assistant vocal est constatée lorsque les émotions positives suscitées sont supérieures aux émotions négatives. De la même façon, cette volonté d'acheter par l'assistant vocal est observée lorsque l'évaluation des bénéfices est supérieure aux coûts. Cela implique que l'intention d'achat par l'assistant vocal se produit à travers un calcul volontaire délibéré et rationnel qui correspond au modèle de conflit décisionnel connu sous le nom de la balance décisionnelle (Janis et Mann, 1977). Ces auteurs ont été les premiers à décrire un modèle de prise de décision en utilisant l'image d'une balance. L'idée maîtresse est qu'une personne ne modifie son comportement que lorsque les avantages qu'elle associe au changement surpassent les inconvénients qui en résultent. Ainsi, toute personne qui envisage de modifier intentionnellement son comportement se livre essentiellement à une analyse des effets probables de ce changement.

Ardelet et al. (2017) se sont intéressés à analyser l'ambivalence des consommateurs envers les objets connectés. Ils expliquent que le consommateur ambivalent n'a pas une attitude stable, positive ou négative à l'égard de l'objet connecté. Il n'est pas indifférent non plus (état dans lequel il n'éprouverait pas d'émotions). Il oscille entre un jugement et son contraire et que la présence d'une pensée ambivalente ne permet pas de distinguer les potentiels acheteurs des

non acheteurs. Cette recherche permet d'aller plus loin en adoptant les apports du modèle de conflit décisionnel qui explique que lorsque les émotions et les jugements positifs envers l'assistant vocal dépassent les émotions et les jugements négatifs, le voicenaute a l'intention de faire des achats à travers l'assistant vocal.

Le processus du conflit décisionnel a ensuite été analysé et comparé auprès des quatre profils de voicenautes (Figure 2). Les résultats montrent que les réfractaires ressentent davantage d'émotions négatives comparé aux émotions positives et qu'ils se focalisent surtout sur les coûts perçus au lieu des bénéfiques. Etant donné que les émotions et les jugements négatifs sont supérieurs, ce profil n'a pas l'intention de passer des achats via l'assistant vocal. Les hédonistes révèlent des émotions positives supérieures aux émotions négatives mais il n'y a pas de différence entre la perception des bénéfiques et des coûts. Ce profil n'est pas prêt dans l'immédiat à passer un achat via l'assistant vocal. Son intention d'achat peut être déterminée par une évaluation des bénéfiques devenant supérieure aux coûts. Toutefois, les utilitaristes pensent que l'assistant vocal présente plus de bénéfiques que de coûts et éprouvent des émotions positives élevées quant à son utilisation. Ce profil confirme son intention de faire des achats de produits d'achat courant par l'assistant vocal. Enfin, les non-adeptes déclarent seulement ressentir des émotions négatives et ne perçoivent aucune différence entre bénéfiques et coûts de l'utilisation de l'assistant vocal. Pour que ce profil pense acheter via l'assistant vocal, il est important de réduire ses émotions négatives envers celui-ci et d'augmenter les bénéfiques perçus de son utilisation.

Figure 2 : Analyse comparative des profils psychologiques des voicenautes



L'étude qualitative a aussi permis d'identifier trois déterminants du conflit émotionnel (émotions positives et négatives suscitées durant l'interaction avec l'assistant vocal) et conflit décisionnel (bénéfices et coûts perçus de l'utilisation de l'assistant vocal).

Le premier antécédent renvoie à la fonction symbolique que le voicenaute donne à l'enceinte connectée. En effet, sa valeur, son design et apparence représentent pour certains répondants des ressources et des moyens symboliques pour la construction de soi et l'inscription dans des significations sociales. Par exemple, certains répondants déclarent l'utiliser comme objet de décoration dans leur pièce de vie pour renvoyer auprès de leurs invités une image de personne ouverte à l'innovation et à l'affût des dernières tendances. L'analyse de contenu montre une relation positive entre la fonction symbolique et les émotions suscitées durant l'interaction avec l'assistant vocal.

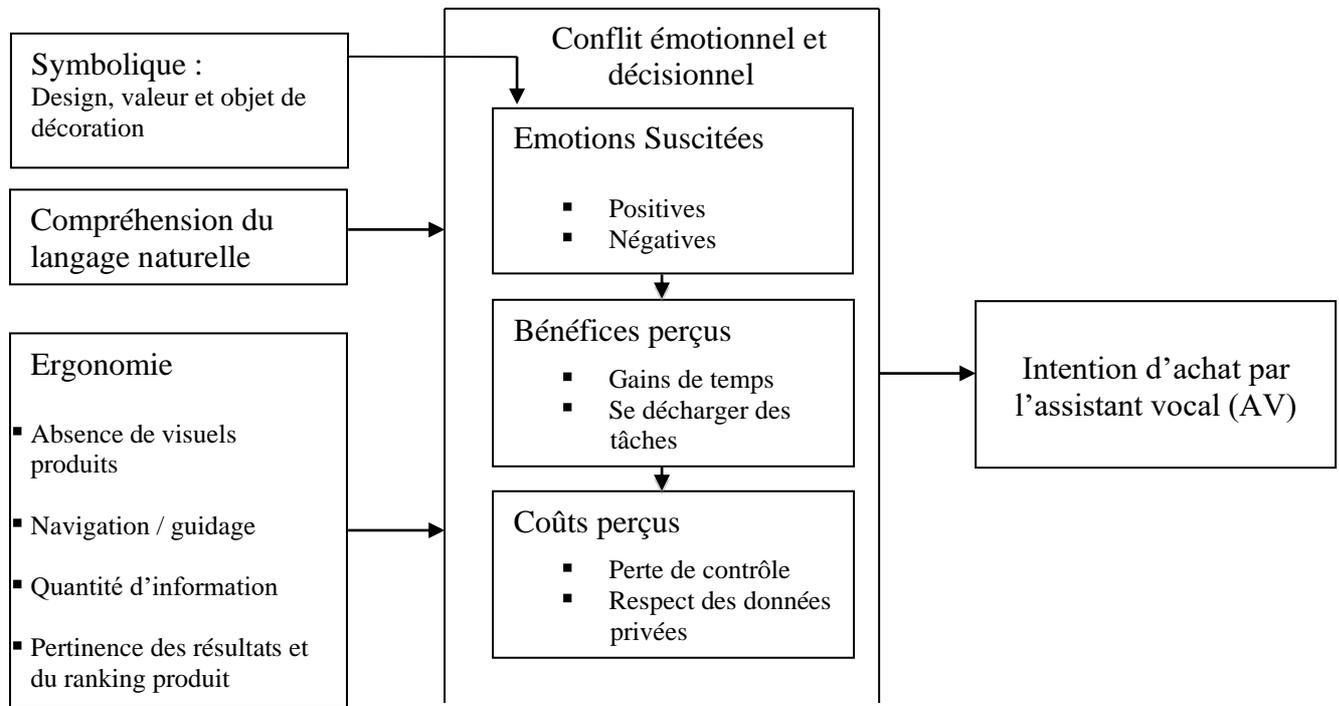
Le deuxième antécédent tient compte de la capacité de l'assistant vocal à comprendre le langage naturel pour répondre à la demande du voicenaute sans le faire répéter. L'assistant vocal peut ne pas comprendre la demande « Google Home : *Excusez-moi je ne comprends pas* » ou ne pas entendre la demande « Google Home : *Je n'ai pas entendu ce que vous avez dit ?* ». Cela est dû à cause de problèmes d'articulation, de hauteur de la voix ou un bug. L'analyse thématique met en évidence une relation négative entre l'incompréhension du langage naturel et les émotions positives et les bénéfiques et une relation positive entre les émotions négatives et les coûts. « *J'ai dû répéter plusieurs fois ma demande sans que ça n'aboutisse à quelque chose. C'est à la fois agaçant, ça me fait perdre mon temps [...] Je n'ai pas l'impression d'avoir le contrôle !* » (Répondant 11)

Enfin, le troisième antécédent repose sur l'ergonomie et la qualité d'expérience avec l'assistant vocal. Certains répondants déclarent que l'absence de visuels produits augmente leur frustration et le risque d'erreur et de ne pas faire le bon choix de produits, réduit les bénéfiques perçus et les dissuade d'acheter via l'assistant vocal : « *Oui ça peut aider d'avoir un visuel parce que si on veut acheter quelque chose, par exemple, un tableau ou un cadre photo, si on n'a pas de visuel, on ne sait pas si ça va aller dans notre salon, on ne sait pas les dimensions, donc il n'y a pas le descriptif, on a juste les grosses lignes de l'objet.* » (Répondant 3). Selon Mari, Mandelli et Algesheimer (2020) l'absence du visuel pourrait devenir obsolète dans l'avenir. En effet, le e-commerce a ouvert la voie au v-commerce, amenant les consommateurs à surmonter l'asymétrie d'information comparé au store et la réticence d'acheter sans directement voir, toucher ou sentir un produit. Les technologies vocales limitent les sens et incitent à prendre des décisions d'achat sans parcourir les photos, les vidéos ou tout autre contenu visuel. Les répondants expriment aussi les limites de navigation et de guidage de l'assistant vocal, par exemple : « *J'ai envie d'écouter de la musique sans passer par Google ou en allant sur YouTube, l'assistant vocal pourrait la mettre directement.* » (Répondant 3)

Quelques répondants soulignent que la quantité d'information donnée verbalement par l'assistant vocal réduit les bénéfiques perçus et l'intention d'achat « *Je vais demander un produit, il me donne plusieurs résultats après je vais oublier le prix du premier résultat, car si ça me les donne d'affiler* » (Répondant 9) Ils proposent que les résultats doivent être analysés et comparés de façon plus poussée en amont pour réduire la liste des choix de produits. Ils souhaitent que l'assistant vocal devienne connu par son expertise et compétence, tout comme un bon vendeur en magasin capable de vous proposer une liste réduite de produits performants. L'assistant vocal devient alors un outil de recommandation produits et d'aide à la décision d'achat au lieu d'un simple outil d'information. Selon la théorie de la richesse des médias de Daft et Lengel (1986), plus le média est riche, moins les utilisateurs sont disposés à faire des recherches. L'effort cognitif est, en effet, plus élevé dans un environnement à faible richesse multimédia tel que le v-commerce en comparaison au e-commerce. Kraus, Reibenspiess et Eckhardt (2019) proposent de réduire intentionnellement la complexité des recommandations dans le v-commerce en termes de quantité de produits proposés et la longueur du descriptif d'un même produit pour éviter une surcharge cognitive du voicenaute. Par ailleurs, d'autres répondants souhaitent que l'assistant vocal soit capable de donner des résultats avec un ranking produit pertinent en tenant compte de leur historique d'achat et de navigation comme c'est déjà le cas sur desktop, tablette et smartphone.

Les résultats de l'étude qualitative ont permis le développement d'un modèle conceptuel pour comprendre les déterminants de l'intention d'achat par l'assistant vocal (Figure 3).

Figure 3 : Proposition de modèle conceptuel pour comprendre les déterminants de l'intention d'achat par l'assistant vocal



### Implications théoriques et managériales

Les travaux de recherche sur l'intention d'adopter le v-commerce utilisent les théories classiques de l'acceptation de la technologie telles que le TAM et l'UTAUT en insistant sur le rôle des bénéfices et des coûts perçus. Cette recherche apporte un nouvel éclairage par l'utilisation du modèle du conflit décisionnel (ratios bénéfices/coûts) connu sous le nom de la balance décisionnelle (Janis et Mann, 1977). Ce modèle est très utilisé en marketing social et montre qu'une personne est susceptible de changer de comportement problématique et d'adopter un comportement bénéfique pour elle et pour la société lorsque l'évaluation des bénéfices dépasse les coûts perçus (Salerno et al., 2015). L'intérêt est de tenir compte de l'écart entre les bénéfices et les coûts perçus pour prédire l'intention d'achat par l'assistant vocal et de ne pas se limiter à mesurer les bénéfices et les coûts séparément. Le modèle du conflit décisionnel pourrait aussi apporter un enrichissement aux recherches sur l'ambivalence du consommateur envers les objets connectés (Ardelet et al., 2017) par le calcul de la différence de l'intensité entre deux réactions ambivalentes pour pouvoir distinguer les potentiels acheteurs des non acheteurs.

Cette recherche met aussi en évidence l'intérêt de tenir compte du conflit émotionnel que nous appelons bien-être subjectif de l'expérience (ratios émotions positives/émotions négatives) et définissons comme *l'évaluation personnelle et subjective d'un état de satisfaction et l'absence d'un état d'insatisfaction au moment de l'expérience avec l'assistant vocal* et de ne pas se limiter à la valeur hédonique apportée par celui-ci. Selon Page et Vella-

Brodrick (2009, p.443), le bien-être subjectif s'obtient par la recherche d'émotions positives et du plaisir et l'évitement des émotions négatives et de la souffrance (2009, p.443). Elles soutiennent la question de bipolarité de la structure de l'émotion en stipulant qu'un individu peut se trouver dans un état plus ou moins positif et dans un état plus ou moins négatif à un moment donné (Clark et Watson, 1991). Cette recherche montre l'intérêt de tenir compte de l'écart entre les émotions positives et négatives de la même façon que les travaux sur le bien-être général.

Enfin, cette recherche propose d'autres déterminants des bénéfices et coûts perçus de l'utilisation des assistants vocaux en complément des croyances positives et négatives tels que la valeur symbolique, la compréhension du langage naturel et l'ergonomie (absence du visuel, quantité d'information, pertinence de la recommandation et du ranking produit).

*Les apports pour l'action* peuvent servir d'un côté, aux marques d'assistant vocal et d'un autre côté, aux marques / annonceurs. Les résultats montrent l'intérêt que portent les voicenautes à ce nouveau canal. Pour satisfaire les voicenautes, les marques d'assistant vocal et les marques / annonceurs peuvent mettre en place des partenariats stratégiques pour développer des services accessibles par les assistants vocaux et définir les règles d'une stratégie SEO vocal efficace. Par exemple, le partenariat "Panier Vocal E.Leclerc" entre Google et Leclercq où la personne peut dicter ses produits préférés et les retrouve ensuite directement dans son Panier en se connectant à son compte E.Leclerc DRIVE ou E.Leclerc Chez Moi.

Cette recherche met en évidence quatre profils de voicenautes (personas) que les marques peuvent prendre en compte pour proposer des stratégies et actions adaptées. Par exemple, les utilitaristes ressentent un bien-être durant l'expérience avec l'assistant vocal, sont conscients des bénéfices qu'il leur apporte et ont déjà effectué un achat de produits à faible implication. Les marques d'assistant vocal peuvent utiliser des algorithmes de recommandation produits qui prennent en compte leur historique d'achat de type "automated match", leur poser la question s'il souhaite découvrir une autre marque avant de passer commande pour éviter le sentiment de manquer de choix. Elles peuvent aussi utiliser les techniques de vente additionnelle « cross-selling » et demander s'il souhaite ajouter un autre produit complémentaire. Par ailleurs, les marques / annonceurs peuvent proposer des coupons et offres promotionnelles pour inciter leurs clients à utiliser ce nouveau canal voire oser acheter des produits réfléchis (à forte implication).

Quant aux hédonistes, ils ressentent certes un bien-être durant l'expérience avec l'assistant vocal, toutefois, ils ne sont pas encore conscients des bénéfices qu'il peut leur apporter. Les marques d'assistant vocal peuvent leur envoyer un cadeau après avoir participé à un jeu sur l'assistant vocal, leur demander de suivre sa livraison sur l'assistant vocal et leur proposer une offre promotionnelle qu'ils découvrent à la réception du cadeau pour les inciter à faire un premier achat. Ces marques peuvent utiliser des algorithmes de recommandation produits de type "exact match" lorsqu'ils citent le nom de la marque et de type "broad match" s'ils ne citent pas le nom de la marque. Dans ce cas, ces marques se trouvent confronter à un phénomène du « coldstart » c'est-à-dire le démarrage à froid où le système ne dispose pas de données suffisantes. Elles doivent alors éviter de proposer leur propre marque (exemple, produits Amazon) et mettre plutôt en avant des produits offrant un bon rapport qualité-prix tout en tenant compte des avis clients. Aussi, les marques / annonceurs peuvent proposer un cadeau ou une réduction pour inciter leurs clients hédonistes à faire un premier achat par le biais de l'assistant vocal.

Les réfractaires et les non-adeptes ressentent un mal-être durant l'expérience avec l'assistant vocal, trouvent l'interaction avec lui complexe et ignorent complètement les bénéfices qu'il peut leur apporter. Les marques d'assistant vocal peuvent dans un premier temps leur proposer différents divertissements (énigme, jeux, raconter des histoires, etc.) pour les aider à développer des émotions positives à l'égard de l'assistant vocal et à se familiariser avec lui pour ensuite les convertir au shopping vocal.

### **Conclusion : Limites et futures voies de recherche**

Cette recherche a proposé, à partir d'une étude qualitative un modèle conceptuel pour comprendre les déterminants de l'intention d'achat par l'assistant vocal. Une modélisation par équations structurelles serait naturellement intéressante à réaliser. Aussi, une analyse comparative auprès des quatre profils de voicenautes à l'aide d'une analyse multi-groupes pour comprendre les différences entre les influences des bénéfices et coûts sur l'intention d'adoption du v-commerce pourrait aussi être mise en place. Ces analyses sont à compléter par une comparaison des moyennes et un calcul des écarts respectifs entre les émotions positives et négatives ainsi qu'entre les bénéfices et coûts perçus pour chacun des profils.

Cette recherche a identifié 3 formes d'assistants vocaux à savoir : screen-first, voice-only (screenless) et voice-first. Une analyse comparative entre au moins 2 formes des assistants vocaux en termes de conflit émotionnel (ratios émotions positives/émotions négatives) et de conflit décisionnel (ratios bénéfices/coûts) reste aussi à étudier. Cette recherche a réalisé une étude pour comprendre les réactions des voicenautes (croyances, attitudes et intentions). Une recherche holistique auprès des marques / annonceurs ou encore des marques d'assistant vocal mériterait d'être examinée. En effet, ces marques s'introduisent dans l'espace privé de leur client et la possibilité d'interagir avec sa marque depuis son propre domicile contribue au développement d'une relation intime plutôt que transactionnelle avec la marque (Mari, 2019). Il serait intéressant de comprendre les conséquences sur l'expérience de marque.

Il serait également intéressant de réaliser une étude interculturelle pour comprendre les différences entre les profils de voicenautes issus de différentes cultures où le taux de pénétration du v-commerce est plus élevé comparé à celui en France comme aux Etats-Unis ou encore dans les pays asiatiques. Par exemple, selon la société d'études eMarketer, aux Etats-Unis, le nombre d'utilisateurs d'enceintes connectées dépassera 83 millions d'ici fin 2020, en progression de plus de 13% par rapport à 2019. L'étude de Humphry et Chesher (2020) montre que les assistants vocaux tels que Google Home et Amazon Echo sont préparés pour le consommateur à travers un processus de "pré-domestication", c'est-à-dire l'utilisation de connotations sociales pré-existantes pour créer des associations positives envers ces dispositifs et les accepter à l'intérieur de son foyer. En effet, leur analyse a permis de montrer que les assistants vocaux sont dotés de voix inspirée d'un imaginaire culturel à partir de robots extraterrestres ou robots emblématiques qui ont marqué l'histoire du cinéma et des films hollywoodiens.

Enfin, des recherches dans une perspective éthique constituent aussi un important domaine de recherche. Il serait pertinent de s'intéresser aux préoccupations d'ordre éthique qui peuvent se poser telles que l'usage des data collectées et les enjeux de ces dispositifs d'Iot (Internet of Things, IoT) ou le manque de transparence et confiance envers les systèmes de recommandation produit, par exemple, la crainte de proposer des systèmes de recommandation produit qui favorisent les vendeurs, marques et produits qui sont plus

rentables ou encore la capacité à proposer des systèmes qui prennent en compte les avis clients. Ces préoccupations d'ordre éthique méritent d'être considérées dans une approche multi-perspectives (téléologique, déontologique, éthique de la technologie persuasive).

## **Bibliographie**

Ardelet C., Veg-Sala N., Goudey A. et Haikel-Elsabeh M. (2017), Entre crainte et désir pour les objets connectés : Comprendre l'ambivalence des consommateurs, *Décisions Marketing*, 86(2), 31-46.

Azjen I. (1991), The Theory of Planned Behavior, *Organizational Behavior and Human Decision Process*, 50(2), 179-211.

Ba W., Mendelson H. et Zhu M. (2020), Sales Policies for a Virtual Assistant, *Computer Science and Game Theory*, 1(1): 1-37

Capgemini Digital Transformation Institute (2018), Conversational Commerce : Why consumers are embracing voice assistants in their lives, *Cabinet de conseil Capgemini, US, UK, France, and Germany*.

Clark L.A. et Watson D. (1991), Tripartite model of anxiety and depression: Psychometric evidence and taxonomic implications, *Journal of Abnormal Psychology*, 100(3): 316-336.

Coskun A.N. et Hellriegel O.T. (2020), The Potential Of Voice Commerce: Attitudes And Behaviors In Ordering Via Voice Assistant, *AMA Summer Academic Conference, San Francisco, CA*.

Daft R.L. et Lengel R.H. (1986), Organizational Information Requirements, Media Richness and Structural Design, *Management Science*, 32(5), 554-571.

Davis F.D., Bagozzi R.P. et Warshaw P.R. (1989), User Acceptance of Computer Technology: A Comparison of Two Theoretical Models, *Management Science*, 35(8), 982-1003.

Dey I. (1999), *Grounding Grounded Theory: Guidelines for Qualitative Inquiry*, San Diego: Academic Press.

Diallo K. (2019), Des employés d'Amazon écoutent les conversations des utilisateurs avec Alexa, *Le Figaro, Tech & web*.

Humphry J. et Chesher C. (2020), Preparing for smart voice assistants: Cultural histories and media innovations. *New Media & Society*, 1(1), 1–18.

Janis I.L. et Mann L. (1977), *A psychological analysis of conflict, choice, and commitment*, New York: Free Press.<sup>[1]</sup><sub>[SEP]</sub>

Jones V.K. (2018), Voice-Activated Change: Marketing in the Age of Artificial Intelligence and Virtual Assistants, *Journal of Brand Strategy*, 7(3), 239–251.

- Kraus D., Reibenspiess V. et Eckhardt A. (2019), How Voice Can Change Customer Satisfaction: A Comparative Analysis between E-Commerce and Voice Commerce, *14th International Conference on Wirtschaftsinformatik, Siegen, Germany*.
- Luger, E. et Sellen, A. (2016), Like Having a Really Bad PA: the Gulf Between User Expectation and Experience of Conversational Agents, *Proceedings of the CHI Conference on Human Factors in Computing Systems, San Jose Convention Center : San Jose, CA, USA*.
- Mari A. (2019) Understanding shopping-related voice assistants and their effect on brands, *IMMAA Annual Conference, Doha, Qatar*.
- Mari A., Mandelli A. et Algesheimer R. (2020), The Evolution of Marketing in the Context of Voice Commerce : A Managerial Perspective, *22nd International Conference on Human-Computer Interaction, Copenhagen (Danemark)*.
- Mori M. (1970), The Uncanny Valley, *Energy*, 7(4), 33–35.
- Moriuchi E. (2019), Okay, Google!: An empirical study on voice assistants on consumer engagement and loyalty, *Psychology and Marketing*, 36(5), 489-501.
- Moussawi S. (2018), User Experiences with Personal Intelligent Agents: A Sensory, Physical, Functional and Cognitive Affordances View, *Proceedings of the 2018 ACM SIGMIS Conference on Computers and People Research, Buffalo-Niagara Falls, NY, USA*.
- Nass C. et Moon Y. (2000), Machines and mindlessness: Social responses to computers, *Journal of Social Issues*, 56(1), 81–103.
- Page K.M. et Vella-Brodrick D.A. (2009), The ‘what’, ‘why’ and ‘how’ of employee well-being: A new model, *Social Indicators Research* 90(3): 441-458.
- 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, Hawaii*.
- Salerno F., Berriche A., Crié D. et Martin A. (2015), Aider les consommateurs à changer de comportement financier : les apports d’un modèle de changement multiphases, *Recherche et Applications en Marketing*, 30(3), 63-87.
- Vachaudéz A. et Geerts A. (2020), Le voice commerce pour les marques de luxe : Revue de littérature et proposition d’un modèle conceptuel, *19ème édition du colloque sur le Marketing Digital, Panthéon-Sorbonne, Paris, France*.
- Venkatesh V., Davis F. et Morris M.G. (2007), Dead Or Alive? The Development, Trajectory And Future Of Technology Adoption Research, *Journal of the Association for Information Systems*, 8(4), 267-286.

