

Les conseillers virtuels anthropomorphes et les réactions des internautes : une expérimentation portant sur la voix du conseiller

Recherche et Applications en Marketing
2019, Vol. 34(1) 29–49
© l'Association Française du Marketing, 2018
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0767370118775963
journals.sagepub.com/home/ram



Emna Cherif

IAE Auvergne, Université Clermont Auvergne, Laboratoire CRCGM, France

Jean-François Lemoine

Université Paris I Panthéon Sorbonne, Laboratoire PRISM Sorbonne, ESSCA School of Management, France

Résumé

Les conseillers virtuels se sont fortement démocratisés sur les sites marchands. Compte tenu des bénéfices qu'ils offrent aux entreprises pour améliorer la navigation et l'interaction avec le consommateur, chercheurs et praticiens s'accordent sur l'intérêt de les doter de caractéristiques anthropomorphiques. Cette recherche s'intéresse particulièrement à l'effet de la voix du conseiller virtuel. Bien qu'il existe quelques recherches en interaction homme-machine sur la question, il n'existe pas de travaux traitant le sujet d'un point de vue marketing et comparant l'effet d'une voix humaine vs une voix de synthèse. Les résultats montrent que les consommateurs qui ont interagi avec un conseiller virtuel doté d'une voix humaine éprouvent un sentiment de présence sociale plus fort que ceux interagissant avec un conseiller doté d'une voix de synthèse. La voix humaine permet également de susciter plus de confiance envers le conseiller virtuel et de générer de meilleures intentions comportementales.

Mots-clés

confiance envers le conseiller virtuel, confiance envers le site, conseiller virtuel, médiation multiple en série, présence sociale, voix humaine vs de synthèse

Introduction

Les conseillers virtuels¹ sont de plus en plus présents sur les sites internet. Une étude récente effectuée par le groupe *Grand View Research* (août 2016) montre que le marché mondial des conseillers

virtuels devrait atteindre 12,28 milliards de dollars d'ici 2024. Ils sont des représentations graphiques personnifiées de programmes informatiques qui ont été conçus pour interagir et se comporter comme un

Auteurs correspondants:

Emna Cherif, IAE Auvergne, Université Clermont Auvergne, Laboratoire CRCGM, 11 Boulevard Charles de Gaulle, Clermont-Ferrand, 63000, France.

Email: emna.cherif@uca.fr

Jean-François Lemoine, Université Paris I Panthéon Sorbonne, Laboratoire PRISM Sorbonne, ESSCA School of Management, France.

Email: jean-francois.lemoine@univ-paris1.fr

humain (Holzwarth et al., 2006 ; Viot et Bressolles, 2012). Ils sont régulièrement utilisés afin de rendre les interfaces en ligne plus chaleureuses et de générer des réponses favorables de la part des internautes en rendant l'expérience de magasinage plus interpersonnelle et en augmentant le degré de plaisir et d'immersion (Holzwarth et al., 2006 ; Wang et al., 2007). Il a par ailleurs été démontré que leur présence renforce le sentiment de confiance envers le site (Koh et Sundar, 2010 ; Lemoine et Notebaert, 2011).

Toutefois, malgré ce recours important aux conseillers virtuels, plusieurs conceptions ont été vouées à l'échec et nombre d'entre eux ont disparu puisqu'ils ne répondaient pas aux attentes des consommateurs, manquaient d'autonomie et/ou avaient une apparence inadaptée (Ben Mimoun et al., 2012 ; McGoldrick et al., 2008). Dans ce contexte, une attention particulière a été accordée à l'anthropomorphisme, défini comme la tendance à attribuer des caractéristiques humaines aux objets non animés (Duffy, 2003). En effet, Burgoon et al. (2000) montrent que les utilisateurs ont une meilleure évaluation de l'utilité des agents virtuels et qu'ils éprouvent à leurs égards des sentiments plus favorables lorsqu'ils sont conçus avec des caractéristiques anthropomorphiques. Dans la perspective de faciliter l'interaction avec les conseillers virtuels et de leur permettre de susciter chez les utilisateurs un fort sentiment de présence sociale, définie comme la capacité subjective du média à permettre à un individu de sentir son interlocuteur comme psychologiquement présent (Gefen et Straub, 1997 ; 2004), les recherches suggèrent d'accroître leur ressemblance avec l'humain et de renforcer leur anthropomorphisme (Edlund et al., 2008 ; Lemoine et Chérif, 2012 ; Stern et al., 2006). C'est ainsi que Burgoon et al. (2000) recommandent que les conseillers virtuels soient dotés de caractéristiques humaines telles que la voix, les gestes et les expressions faciales².

Malgré le développement des travaux sur l'anthropomorphisme des conseillers virtuels, les recherches sur la voix sont rares et se sont focalisées en majorité sur les effets de la voix dans le cadre de l'interaction homme-ordinateur. Or, la voix, autant que les expressions faciales et la gestuelle, constitue une caractéristique importante pour interagir

socialement avec autrui (Bartneck, 2001). Nass et Lee (2001) montrent que les utilisateurs peuvent assigner une personnalité à l'ordinateur en fonction de la voix avec laquelle il s'exprime. Short et al. (1976) précisent que les supports audio entraînent un sentiment de présence sociale plus fort que les supports écrits. De même, la voix encourage les utilisateurs à adopter, face à un ordinateur, les mêmes réponses sociales qu'ils ont l'habitude de mettre en place dans leurs relations habituelles entre humains (Reeves et Nass, 1996). Bien que les recherches antérieures montrent que la voix peut avoir un effet sur la présence sociale perçue dans un contexte d'interaction homme-machine, il demeure essentiel d'étudier, d'un point de vue marketing, l'influence de la voix du conseiller virtuel sur la présence sociale perçue sur un site commercial. Un conseiller virtuel s'exprimant oralement permettrait de remédier à l'aspect froid et impersonnel souvent ressenti sur les sites marchands (Holzwarth et al., 2006).

Par ailleurs, très peu d'études se sont interrogées sur la manière de renforcer la confiance des consommateurs envers les conseillers virtuels. Or, elle impacte les intentions de les utiliser (Wang et Benbasat, 2005). Les recherches sur les interactions médiatisées par ordinateur montrent que la confiance est affectée par les modalités de communication utilisées par le média. Ainsi, une communication en audio conférence ou téléphonique entraîne un sentiment de confiance plus fort qu'une communication textuelle (Bos et al., 2002 ; Greenspan et al., 2000). Dans cette optique, cette recherche se propose d'étudier l'effet de la voix du conseiller virtuel sur la confiance éprouvée à l'égard du conseiller virtuel dans le cadre d'une interaction marchande en ligne. L'acquisition de la confiance est en effet un enjeu crucial pour les entreprises (Gefen et Straub, 2003 ; Jarvenpaa et al., 2000). La présence d'un conseiller virtuel s'exprimant oralement pourrait également renforcer la confiance envers le site (Qiu et Benbasat, 2009 ; Wang et al., 2007).

L'objectif de cette étude est d'étudier les effets de la voix du conseiller virtuel sur la présence sociale, la confiance envers le conseiller virtuel, la confiance envers le site et les intentions comportementales. Par l'intermédiaire de notre recherche, nous souhaitons enrichir la littérature en marketing consacrée aux conseillers virtuels, qui, jusqu'à présent, ne s'est pas

intéressée à l'influence de la voix des conseillers virtuels dans un contexte marchand. Bien que l'effet de la voix ait été étudié dans un contexte d'interaction homme-machine, soit pour améliorer le quotidien des personnes ayant une déficience (Loiacono et al., 2013 ; Wongkia et al., 2012) soit pour donner un aspect social aux interfaces médiatisées comme les ordinateurs et les GPS, il n'y a pas eu, à notre connaissance, de travaux en marketing sur le sujet.

Au travers de cette investigation, nous nous proposons d'étudier l'influence de deux types de voix du conseiller virtuel sur les réactions des internautes. En effet, sur les interfaces électroniques, deux types de voix peuvent être utilisées, à savoir une voix de synthèse et une voix humaine préenregistrée. La voix de synthèse est une voix artificielle produite à partir d'un texte numérique. Elle nécessite le recours à des logiciels informatiques sophistiqués appelés synthétiseurs vocaux comme, par exemples, AT et T, Elan Sayso Speech et DecTalk Express, qui sont conçus pour lire le texte disponible et le reproduire oralement. La voix humaine est une voix préenregistrée. Son utilisation passe par l'enregistrement, au préalable, d'un texte lu à voix haute par une personne. Contrairement à la voix humaine, la voix de synthèse permet aux concepteurs d'agents virtuels de bénéficier d'une grande flexibilité et de coûts de conception plus bas dans la mesure où les synthétiseurs ont une capacité infinie à lire des contenus textuels et sont ainsi capables de générer la voix d'une manière continue. Toutefois, malgré les avantages opérationnels et financiers qu'elle peut présenter, la voix de synthèse demeure éloignée de la voix humaine. A l'heure actuelle, les synthétiseurs présentent encore quelques lacunes par rapport à la voix humaine notamment en matière de prosodie et d'expressivité, c'est à dire de capacité à exprimer les émotions, les intentions et les attitudes à travers la modulation de la voix (Beller et Rodet, 2007). Les voix de synthèse présentent des pauses et des discontinuités qui peuvent limiter les signes sociaux que la voix est en mesure de véhiculer (Nass et Scott, 2005).

Cette recherche, réalisée dans le contexte spécifique du marketing, présente les premiers résultats empiriques dédiés aux effets respectifs d'un conseiller virtuel doté d'une voix de synthèse vs d'une voix humaine. Elle se propose, d'un point de

vue théorique, de contribuer au développement des études consacrées à l'anthropomorphisme des conseillers virtuels en marketing et à une meilleure compréhension de l'impact du type de voix du conseiller virtuel sur les réactions des consommateurs en ligne. Sur le plan managérial, cette investigation vise à renseigner les directeurs marketing et les responsables de sites web sur la pertinence d'investir dans l'acquisition de conseillers virtuels aux caractéristiques anthropomorphiques bien précises sur le plan de la voix. Leur installation présente en effet un coût qui peut dissuader les managers de recourir à leurs services. Cette recherche a pour objectif d'apporter aux praticiens des arguments précis quant aux effets de la voix du conseiller virtuel sur des variables potentiellement créatrices de valeur et susceptibles de favoriser la différenciation des sites entre eux telles que la présence sociale perçue, la confiance à la fois envers le site et les conseillers virtuels et les intentions comportementales.

Nous présenterons, dans un premier temps, notre revue de littérature. Puis nous exposerons la méthodologie de recherche mise en œuvre lors de notre étude empirique. Enfin, nous présenterons et discuterons nos résultats avant d'en dégager des contributions théoriques et managériales.

Revue de littérature

De la théorie de la réponse sociale à la théorie de la maximisation du réalisme

Dans le domaine des interactions homme – machine, un grand nombre de recherches s'accordent à dire que les gens traitent l'ordinateur comme un acteur social et non seulement comme un intermédiaire (Moon, 2000 ; Nass et Moon, 2000 ; Reeves et Nass, 1996 ; Wang et al., 2007). La théorie de la réponse sociale est à l'origine de ce constat. Elle stipule que les gens appliquent les mêmes règles sociales, telles que la politesse (Nass et al., 1999 ; Tzeng, 2006), les échanges mutuels (Moon, 2000), les stéréotypes liés au genre (Nass et al., 1997), la flatterie (Fogg et Nass, 1997 ; Lee, 2008) et l'attribution de personnalité (Moon et Nass, 1996 ; 1998), généralement appliquées entre les humains, pour interagir avec les machines lorsque celles-ci possèdent des attributs et

des indices humains comme l'apparence visuelle et/ou le comportement (manière de se comporter, de se mouvoir) (Nass et Moon, 2000 ; Reeves et Nass, 1996). Dans cette perspective, Nass et Steuer (1993) présentent quatre caractéristiques pouvant renforcer la présence sociale lors d'une expérience en ligne : le langage, l'interactivité, le rôle et la voix. Ils démontrent que les utilisateurs perçoivent les différentes voix issues du même ordinateur comme provenant de plusieurs interlocuteurs et qu'ils considèrent les différents ordinateurs possédant la même voix comme un seul et même interlocuteur (Nass et Steuer, 1993). Pour leur part, Nass et Scott (2005) montrent qu'une interaction orale permet de dialoguer de manière plus naturelle et plus efficace avec les machines. Les interactions orales avec les machines sont mieux perçues par les utilisateurs que les interactions avec du texte uniquement. Les interactions orales sont capables de renforcer la crédibilité et la compétence perçues et d'accroître la ressemblance avec les interactions humaines (Edlund et al., 2008 ; Stern et al., 2006). La théorie de la réponse sociale précise que des indices minimaux en matière de similitude avec les humains sont suffisants pour amener les utilisateurs à afficher des comportements sociaux généralement réservés aux interactions humaines (Krämer et al., 2013). Ce postulat est toutefois de plus en plus remis en cause suite notamment à l'apparition des travaux portant sur la théorie de la maximisation du réalisme (Groom et al., 2009 ; Kang et Watt, 2013). Cette théorie stipule qu'une ressemblance minimale avec l'être humain n'est pas suffisante pour améliorer l'interaction mais qu'il faut veiller à maximiser le réalisme défini comme le degré de ressemblances visuelles et/ou comportementales du personnage virtuel avec l'être humain (Groom et al., 2009 ; Kang et Watt, 2013). La théorie de la maximisation du réalisme s'est initialement imposée dans les recherches portant sur les jeux vidéo. Les développeurs et les créateurs de jeux vidéo ont donc été parmi les premiers à se demander si le perfectionnement du réalisme des personnages constituait l'une des meilleures stratégies pour améliorer l'expérience de jeu en ligne (Ben, 2005). Pour Ivory et Kalyanaraman (2007), les joueurs qui expérimentent la version très réaliste du jeu, en termes de représentation des personnages et de

reproduction des sons, affichent des évaluations plus positives en termes de présence sociale perçue, d'engagement et d'éveil. Récemment, une attention croissante a été accordée à la maximisation du réalisme des agents virtuels dans le cadre des interactions médiatisées par ordinateur. Ceci est d'autant plus facile que les avancées technologiques permettent aujourd'hui aux concepteurs de créer des agents virtuels de plus en plus anthropomorphes et réalistes (Belk, 2016 ; Burgoon et al., 2000 ; Goudey et Bonnin, 2016 ; Groom et al., 2009). Il en ressort que la conversation avec un agent virtuel réaliste est perçue comme plus naturelle et plus motivante que celle avec un agent virtuel moins ou non réaliste, comme une image statique par exemple (McBreen et Jack, 2001). Un agent réaliste entraîne plus de satisfaction et contribue à une meilleure perception de la présence sociale (Van Vugt et al., 2007 ; Von Der Pütten et al., 2010). L'implémentation d'une voix dans la conception d'un agent virtuel constituerait donc un facteur de maximisation du réalisme des agents qui permettrait de renforcer le sentiment de présence sociale et de rendre l'interaction plus agréable pour l'utilisateur.

La voix dans les interfaces électroniques

Dans les interactions sociales, la voix influence considérablement les réactions physiologiques et affectives des individus (Scherer, 1986 ; 2003). Ceci s'explique par la richesse des informations qu'elle permet de transmettre au-delà du sens littéral du contenu. Il est en effet possible de reconnaître les émotions d'un individu, sa personnalité et ses caractéristiques socio-démographiques à travers les tonalités, la vitesse et l'intonation de sa voix (Scherer, 1986 ; 2003). Dans le cadre des interactions électroniques, les langages verbaux les plus utilisés sont la messagerie instantanée et le « *chatting* » (Galaxhi et Nah, 2007 ; Qiu et Benbasat, 2005). De nombreux travaux montrent que le sentiment de présence sociale est plus fort dans les interactions orales que dans les interactions textuelles (Sallnas, 2005 ; Nass et Gong, 2000). Lester et al. (1997) soulignent que les agents qui communiquent verbalement sont perçus comme plus crédibles. De plus, les utilisateurs interagissant avec une interface

qui utilise la voix lui affectent une personnalité (Evans et Kortum, 2010 ; Lee et Nass 2003).

Les avancées technologiques permettent aujourd'hui aux concepteurs de recourir soit à une voix de synthèse soit à une voix humaine lors de la création des conseillers virtuels. La première est une technologie qui permet une communication plus efficace avec les personnes ayant une difficulté auditive ou vocale (Stern et al., 2006). Fortement étudiée dans les recherches sur les personnes ayant une déficience (Grichkovtsova et al., 2012 ; Loiacono et al., 2013 ; Wongkia et al., 2012), la voix de synthèse attire de plus en plus l'attention des chercheurs en interaction homme – machine comme les GPS, les jeux vidéo et les conseillers virtuels. Comme nous l'avons déjà précisé, elle présente deux avantages. Le premier est lié à sa grande flexibilité à générer du contenu oral à l'infini et le deuxième est lié à ses faibles coûts d'implémentation par rapport à la voix humaine. Toutefois, la voix de synthèse présente encore des limites telles que les discontinuités dans le ton, le tempo et l'intensité qui ne permettent pas de transmettre pleinement les émotions, les intentions et les attitudes. C'est pourquoi elles sont généralement perçues comme non naturelles par les utilisateurs. La voix humaine préenregistrée, quant à elle, permet de pallier ce problème puisque le contenu verbal est enregistré au préalable par une personne pouvant très naturellement transmettre ses émotions, ses attitudes et ses traits de personnalité. La voix humaine naturelle se caractérise par son expressivité encore difficilement reproductible avec une voix de synthèse. Toutefois, l'utilisation d'une voix humaine préenregistrée limite les concepteurs dans la mesure où il n'est pas possible de générer la voix pour un contenu textuel illimité. L'implémentation d'une voix humaine préenregistrée nécessite aussi des investissements plus importants de la part des concepteurs. Malgré ces limites, certaines recherches montrent qu'une voix humaine préenregistrée est souvent jugée comme plus persuasive qu'une voix de synthèse.

Dans le but de comprendre, d'une part, l'effet du type de la voix (voix humaine préenregistrée vs voix de synthèse produite artificiellement) et, d'autre part, l'effet de l'origine de la voix communiquée aux participants (humaine vs ordinateur) sur leur perception, Stern et al. (2006)

conduisent une recherche qui débouche sur les résultats suivants :

- quand l'origine de la voix est humaine, les utilisateurs de l'interface la préfèrent à la voix de synthèse ;
- quand la voix provient de l'ordinateur, la voix humaine et la voix de synthèse sont évaluées de manière similaire.

De même, Nass et Scott (2005) comparent l'effet d'une voix humaine vs une voix de synthèse émanant soit d'un visage humain soit d'un visage synthétique. Les auteurs montrent que les gens divulguent et communiquent plus d'informations personnelles avec un agent présentant un visage clairement synthétique et une voix de synthèse ou avec un agent présentant un visage humain associé à une voix humaine qu'avec un agent présentant un visage et une voix qui ne sont pas cohérents entre eux (Gong et Nass, 2007 ; Nass et Scott, 2005). En cas d'incohérence entre la voix et l'apparence, les participants se sentent moins à l'aise pour interagir avec l'interface et l'évaluent comme grossière, moins fiable et étrange. Ces résultats convergent avec la théorie de la « vallée dérangeante » initiée par Mori (1970), dans le domaine de la robotique, selon laquelle plus un robot est similaire à un être humain, plus les imperfections qu'il peut encore présenter sont susceptibles de conduire à son rejet. En d'autres termes, les individus sont plus à l'aise face à un robot clairement identifié comme non humain que face à un robot tendant à prendre l'apparence humaine tout en présentant encore des imperfections (Chérif et Lemoine, 2014). Mori (1970) souligne qu'au-delà d'un certain niveau de perfection dans la ressemblance humaine, les robots sont mieux acceptés. Il existe donc une zone que les concepteurs doivent franchir (appelée la vallée dérangeante) dans laquelle chaque avancée vers l'imitation humaine est synonyme de rejet. Ainsi, l'ajout d'une voix à un conseiller virtuel ne peut impacter positivement les réactions comportementales de l'internaute que si les concepteurs du conseiller l'ont déjà suffisamment anthropomorphisé et qu'ils ont dépassé cette « vallée dérangeante ».

Qiu et Benbasat (2009), pour leur part, ont comparé l'effet de la voix (humaine vs de synthèse) avec la présence ou non d'un conseiller virtuel. Les auteurs montrent que même s'il existe une prédominance de la voix humaine sur la voix de synthèse, il n'y a aucun effet d'interaction entre les deux variables (voix et présence d'un conseiller virtuel) sur le sentiment de présence sociale. En d'autres termes, la voix, d'un côté, et la présence du conseiller virtuel, de l'autre, agissent d'une manière indépendante. Par ailleurs, les auteurs stipulent que l'addition d'une voix de synthèse (vs texte) n'est pas suffisante pour entraîner une meilleure présence sociale sur le site ; or les travaux de Lee et Nass (2003 ; 2004) montrent que la voix de synthèse est capable d'accroître la présence sociale. Les personnes trouvent la voix plus crédible, plus attractive et plus informative quand la voix synthétisée sur l'ordinateur correspond au contenu textuel auquel ils sont exposés (Lee et Nass, 2003). Les travaux de Qiu et Benbasat (2009) montrent également que la voix de synthèse (vs texte) améliore considérablement le sentiment de confiance bien qu'il n'y ait aucun effet d'interaction également entre la présence (vs absence) du conseiller virtuel et la voix (vs texte). Face à des résultats aussi contradictoires, l'approfondissement des effets du type de voix du conseiller virtuel sur les réactions du consommateur s'avère utile notamment dans un contexte de marketing où il n'existe pas, à notre connaissance, de recherches sur le sujet.

Les hypothèses de la recherche

Ardelet et Brial (2011) recensent deux catégories de vecteurs susceptibles de générer de la présence sociale. La première fait référence à des indices non anthropomorphiques relatifs à des contenus éditoriaux (le texte, la mise en forme), à des pratiques de marketing direct (envoi d'e-mails de la commande à la livraison, salutation du client par son nom et prénom) et à des composantes du Web (les avis de consommateurs, les forums, les chats, les moteurs de recommandation de produits, le *web call center*). La seconde concerne les indices anthropomorphiques et se compose des agents virtuels, des photos, des vidéos

et de la voix humaine. Dans le cadre de cette recherche, nous nous focalisons exclusivement sur la présence sociale véhiculée par des indices anthropomorphiques et notamment par la voix du conseiller virtuel. Si cette dernière peut renforcer, selon Nass et Steurer (1993), le sentiment de présence sociale au même titre que le langage et l'interactivité, la littérature présente des résultats contradictoires quant aux effets respectifs de la voix humaine et de la voix de synthèse sur la présence sociale (Lee et Nass, 2003 ; 2004 ; Qiu et Benbasat, 2009). De plus, nous savons que les systèmes de synthèse vocale produisent aujourd'hui des voix parfaitement intelligibles mais qui demeurent encore éloignées de la voix humaine naturelle (Beutnagel et al., 1999 ; Vaudale, 2012). Les voix de synthèse présentent en effet des lacunes, telles que des pauses et un accent non naturels, des discontinuités entre les phonèmes et les syllabes (Nass et Lee, 2001), susceptibles de limiter la perception de la présence sociale lors de l'interaction avec les conseillers virtuels. Par ailleurs, la théorie de la maximisation du réalisme (Groom et al., 2009) nous invite à privilégier la conception de conseillers virtuels les plus réalistes possible et, en conséquence, à les doter d'une voix humaine. Afin de clarifier l'impact respectif de ces deux types de voix sur la perception de la présence sociale sur un site web, nous formulons l'hypothèse suivante :

- *H1a : La présence sociale perçue est plus forte sur le site avec un conseiller virtuel s'exprimant avec une voix humaine que sur le site avec un conseiller virtuel s'exprimant avec une voix de synthèse.*

Plusieurs études ont démontré que la présence des conseillers virtuels est en mesure de renforcer le sentiment de confiance envers le site (Lemoine et Notebaert, 2011 ; Qiu et Benbasat, 2009 ; Wang et Benbasat, 2005 ; 2008 ; Wang et al., 2007). En effet, la personification des interfaces électroniques à l'aide d'un agent virtuel se traduit par une augmentation du sentiment de confiance chez les internautes (André et Pelachaud, 2008 ; Cowell et Stanney, 2005 ; Ruttkay et Pelachaud, 2004). Par ailleurs, Stern et al. (2006) précisent que la confiance peut être affectée, dans le cadre d'un contexte médiatisé par ordinateur, par les caractéristiques et

le rythme du discours et par la présence d'erreurs verbales ou d'hésitations dans l'expression du conseiller (Stern et al., 2006). De tels résultats nous amènent à supposer que :

- *H1b : La confiance envers le site est plus forte sur le site avec un conseiller virtuel s'exprimant avec une voix humaine que sur le site avec un conseiller virtuel s'exprimant avec une voix de synthèse.*

Selon les recherches de Holzwarth et al. (2006), Keeling et al. (2010) et Wang et al. (2007), il existe un lien positif entre la présence d'un conseiller virtuel sur un site web et les intentions comportementales des internautes (Holzwarth et al., 2006 ; Keeling et al., 2010 ; Wang et al., 2007). Dans le cadre de cette étude, nous allons au-delà de l'analyse des conséquences de la présence d'un conseiller virtuel en nous concentrant tout particulièrement sur l'impact de son type de voix. Dans la mesure où les travaux de Lee et Nass (2003) démontrent l'effet positif de cette caractéristique anthropomorphique sur les intentions comportementales dans le cadre des interactions homme-machine, nous nous attendons également à ce qu'un conseiller virtuel s'exprimant avec une voix humaine engendre de meilleures intentions comportementales qu'un conseiller s'exprimant avec une voix de synthèse. D'où l'hypothèse H1c :

- *H1c : Les intentions comportementales sont plus fortes sur le site avec un conseiller virtuel s'exprimant avec une voix humaine que sur le site avec un conseiller virtuel s'exprimant avec une voix de synthèse.*

S'il nous semble important de tester l'effet direct de la voix du conseiller virtuel sur les intentions comportementales (H1c), il nous apparaît également utile de nous interroger sur le rôle médiateur de variables susceptibles d'enrichir la compréhension de cette relation, notamment la confiance envers le conseiller puis celle envers le site. Traditionnellement, les effets de ces deux variables sont analysés distinctement. C'est ainsi que la première est souvent présentée comme explicative des intentions comportementales des consommateurs (Komiak et Benbasat, 2006 ; Wang et Benbasat, 2005) et que la

seconde est dépendante de la présence des conseillers virtuels tout en exerçant un impact sur les intentions comportementales (Keeling et al., 2010 ; Lemoine et Notebaert, 2011). Dans le but de compléter la littérature dédiée à la relation voix du conseiller virtuel – intentions comportementales, nous nous proposons d'analyser conjointement l'effet de ces deux variables de confiance et de les présenter dans une relation de médiation multiple en série. D'où notre hypothèse :

- *H1d : L'effet de la voix du conseiller virtuel sur les intentions comportementales est médiatisé par la confiance envers le conseiller puis par la confiance envers le site.*

Pour ce qui est de la confiance envers les conseillers virtuels, elle apparaît la plupart du temps comme agissant positivement sur les intentions comportementales (Komiak et Benbasat, 2006 ; Qiu et Benbasat, 2005 ; Wang et Benbasat, 2005). Il est donc à prévoir qu'il en sera de même dans le cadre de notre recherche dédiée aux effets de la voix du conseiller. C'est pourquoi nous formulons l'hypothèse suivante :

- *H2 : La confiance envers le conseiller virtuel influence positivement les intentions comportementales.*

Comme le soulignent Bressolles et al. (2015), chaque achat en ligne comporte des coûts de transaction liés aux risques et à l'incertitude qui peuvent prendre trois formes : la crainte d'un comportement opportuniste du vendeur qui ne respecte pas ses engagements initiaux, l'asymétrie d'information qui joue en faveur du vendeur et l'exploitation non contrôlée des données personnelles collectées (Lancelot-Miltgen et Lemoine, 2015). Pour tenter de réduire ces facteurs de risque, Gefen et Straub (2004) et Nan et al. (2006) suggèrent de créer un sentiment de présence sociale sur le site web qui va générer par la suite de la confiance envers celui-ci. Afin de vérifier la pertinence de cette recommandation, nous formulons l'hypothèse suivante :

- *H3 : La présence sociale perçue influence positivement la confiance envers le conseiller virtuel.*

Dans le cadre particulier de l'utilisation des conseillers virtuels sur les sites internet, les recherches dédiées à l'influence de la présence sociale sur les intentions comportementales confirment le plus souvent des effets indirects. C'est ainsi que Wang et al. (2007) démontrent que la présence sociale est à même d'influencer favorablement les intentions d'achat par le biais des valeurs hédoniques et utilitaires. Les travaux de Ben Mimoun et Poncin (2015) concluent à des résultats similaires. En analysant l'influence de la conseillère virtuelle présente sur le site de l'enseigne IKEA, les auteurs prouvent que la valeur hédonique médiatise l'influence de la présence sociale perçue sur les intentions comportementales. Qiu et Benbasat (2009), quant à eux, montrent que la présence sociale perçue influence positivement la confiance envers le conseiller virtuel qui, à son tour, agit favorablement sur les intentions comportementales. Dans le cadre de cette recherche, nous souhaitons mettre en évidence, au-delà des relations médiatisées, l'influence directe de la présence sociale sur les intentions comportementales. Pour ce faire, nous formulons l'hypothèse suivante :

- *H4 : La présence sociale perçue influence positivement les intentions comportementales.*

Les effets de la présence sociale sur les différents types de confiance font l'objet de plusieurs recherches. C'est ainsi que Keeling et al. (2010) et Lee et Nass (2004) montrent que la présence sociale influence positivement la confiance envers le site. Selon eux, les caractéristiques verbales et non verbales des interactions homme-machine véhiculent de la sociabilité et de la proximité susceptibles de renforcer le sentiment de confiance. Qiu et Benbasat (2009), pour leur part, mettent en évidence l'impact de la présence sociale sur la confiance envers le conseiller virtuel. Ils précisent qu'un agent virtuel doté de caractéristiques linguistiques analogues à celles des humains génère de la crédibilité et la bienveillance. Afin d'enrichir la littérature en marketing sur la relation susceptible d'unir la présence sociale à la confiance envers le site, nous supposons que celle-ci est médiatisée par la confiance envers le conseiller virtuel. C'est ainsi que nous stipulons que :

- *H5 : La confiance envers le conseiller virtuel médiatise l'influence de la présence sociale perçue sur la confiance envers le site.*

Enfin, notre hypothèse H6 suppose que la confiance envers le site influence positivement les intentions comportementales. Selon Lemoine et Notebaert (2011), plus les internautes éprouvent un sentiment de confiance à l'égard d'un site web, plus leurs intentions d'achat dans le magasin et leur intention de recommander le site et l'enseigne sont élevées. Ces résultats ont été obtenus dans le cadre d'une recherche comparant les effets de la présence (vs de l'absence) d'un conseiller virtuel sur les réactions des internautes. Dans le cadre du présent travail, nous nous demandons si une telle conclusion peut être généralisée au cas d'un conseiller virtuel présentant une voix humaine (vs une voix de synthèse). D'où notre dernière hypothèse :

- *H6 : La confiance envers le site influence positivement les intentions comportementales.*

Toutes nos hypothèses sont représentées dans la figure 1.

La méthodologie de la recherche

Echantillon et procédure expérimentale

Dans le but de tester les hypothèses de la recherche, une expérimentation a été conduite auprès de 640 internautes. Ayant eu l'opportunité de travailler sur le site de la Garantie Mutuelle des Fonctionnaires (GMF), l'échantillon constitué pour l'occasion était représentatif des clients de cette structure (50,4 % de femmes et 49,6 % d'hommes ; âge moyen des répondants : 42 ans (écart type = 9,87) ; salaire moyen des enquêtés : 2 500 euros). Afin de sélectionner les répondants de l'échantillon et de leur administrer un questionnaire en ligne, la société d'études Toluna QuickSurveys, partenaire de l'Association Française du Marketing (AFM), a été sollicitée. Chaque enquêté était affecté, de manière aléatoire, à une des deux versions du site (conseiller virtuel avec voix humaine vs conseiller virtuel avec

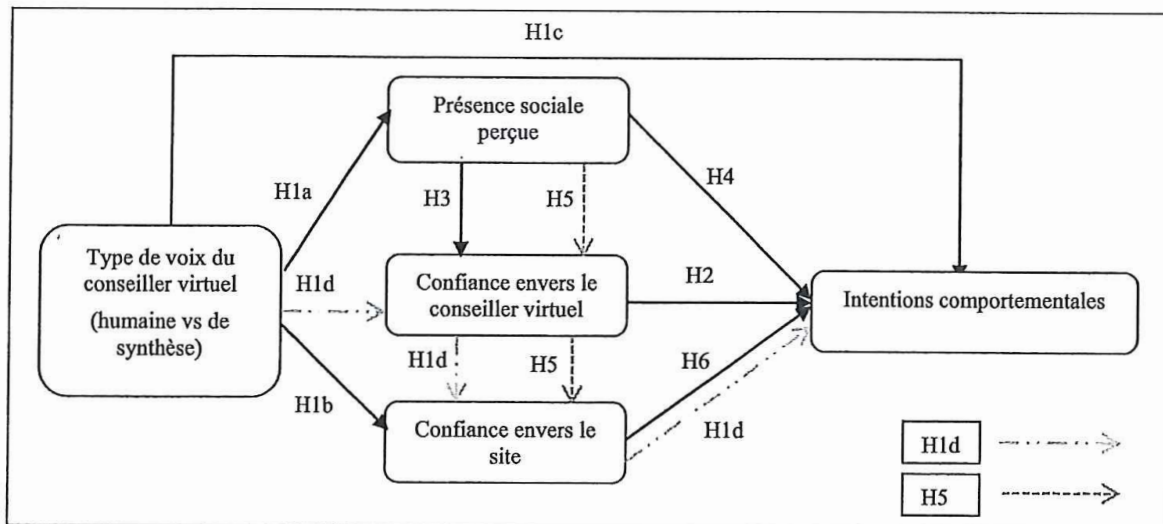


Figure 1. Le modèle de la recherche.

voix de synthèse). A l'aide du test du chi-deux, il a été vérifié que les deux échantillons constitués étaient appariés en termes de genre (chi-deux = 0,094 ; $p = 0,759$), d'âge (chi-deux = 7,655 ; $p = 0,176$) et de revenu (chi-deux = 2,459 ; $p = 0,652$).

Le conseiller virtuel, nommé Prosper, a été conçu et installé sur le site de la GMF par la société Davi spécialisée dans la création d'interfaces digitales (voir annexe 1). L'entreprise utilise le synthétiseur Elan Sayso pour implémenter la voix sur les conseillers virtuels. Il offre une bonne intelligibilité. La voix humaine implémentée sur le conseiller virtuel était préenregistrée par un acteur. Afin de respecter la démarche expérimentale mise en œuvre, les réponses apportées par le conseiller virtuel étaient identiques dans les deux conditions (voix humaine vs voix de synthèse). Elles portaient sur les sept thématiques suivantes : « épargner », « préparer ma retraite », « préparer ma succession/transmettre un capital », « financer un projet », « assurer l'avenir de mes proches en cas de décès », « financer mes obsèques » et « assurer mon projet immobilier ».

Le choix du conseiller virtuel résulte d'une étude qualitative menée par la GMF auprès de ses clients et de ses responsables de la communication. Les résultats obtenus indiquent que le conseiller Prosper est jugé comme le plus cohérent et le plus en phase avec les valeurs de l'entreprise et son positionnement. La voix du conseiller a été choisie de la même manière. Avant de répondre au questionnaire, les répondants étaient invités à simuler une recherche

d'informations sur des produits d'épargne. Il leur était demandé de visiter les sept différentes rubriques du site évoquées précédemment sachant que les possibilités d'interaction avec l'agent qui leur étaient offertes pouvaient avoir lieu à tout moment dès qu'ils le souhaitaient. Un pré-test mené auprès de 50 internautes représentatifs de la clientèle de la GMF a permis d'évaluer le temps moyen nécessaire pour parcourir le site dans sa globalité. Il s'élève à 3 minutes et 50 secondes. Lors de l'administration du questionnaire, un système technique empêchait les répondants de commencer à répondre aux questions tant qu'ils n'avaient pas au moins navigué sur le site pendant ce temps moyen. Ils pouvaient naturellement prendre plus de temps s'ils le souhaitaient. En opérant de la sorte, il était possible de s'assurer que les enquêtés avaient bien pris connaissance du contenu du site et qu'ils avaient eu l'occasion d'interagir suffisamment avec le conseiller virtuel. Grâce à deux questions filtres du questionnaire, seuls les internautes ayant confirmé qu'ils avaient bien entendu la voix du conseiller lors de leurs interactions avec ce dernier ont été retenus dans notre recherche (320 pour la version avec la voix humaine et 320 pour la version avec la voix de synthèse).

Pour vérifier que la manipulation du facteur voix a bien été perçue par les deux groupes de répondants, l'échelle de mesure des caractéristiques de la voix de Stern et al. (2006) a été mobilisée. Elle prend la forme d'un sémantique différentiel et se compose de sept items. Après épuration, et à l'instar de Stern

et al. (2006), les deux dimensions de l'échelle d'origine ont bien été retrouvées : le style de la voix et les caractéristiques physiques de la voix. Pour chacune de ces deux dimensions, l'analyse de variance montre qu'il existe des différences de moyennes statistiquement significatives entre les deux conditions expérimentales (dimension style de la voix : $M_{\text{humaine}} = 4,22$ et $M_{\text{synthèse}} = 3,23$; $F = 77,812$; $p = 0,000$ / dimension caractéristiques physiques de la voix : $M_{\text{humaine}} = 3,88$ et $M_{\text{synthèse}} = 3,66$; $F = 9,079$; $p = 0,003$). Il est donc possible d'en conclure que les deux types de voix affectées au conseiller virtuel sont perçues comme différentes par les deux groupes de répondants.

Echelles de mesure utilisées

Toutes les échelles de mesure mobilisées sont issues de la littérature relative aux conseillers virtuels. La présence sociale perçue sur le site a été mesurée avec l'échelle unidimensionnelle de Qiu et Benbasat (2009) traduite en français par Bressolles et al. (2011). Afin de mesurer la confiance envers le conseiller virtuel, l'échelle de Qiu et Benbasat (2009) a été mobilisée. Il s'agit d'une échelle unidimensionnelle comportant 10 items. Enfin, la confiance envers le site et les intentions comportementales ont été opérationnalisées grâce aux travaux de Keeling et al. (2010) qui proposent respectivement des mesures unidimensionnelles de ces construits en 5 et 3 items. Le choix de l'ensemble de ces échelles s'explique par leur qualité psychométrique d'origine ($\alpha = 0,91$ pour la présence sociale, $\alpha = 0,89$ pour la confiance envers le conseiller virtuel, $\alpha = 0,87$ pour la confiance envers le site, $\alpha = 0,93$ pour les intentions comportementales). Toutes les échelles utilisées dans cette recherche ont fait l'objet d'une rétro-translation à l'exception de celle de Qiu et Benbasat (2009) dont la traduction française était disponible et avait été validée par Bressolles et al. (2011). Les enquêtés faisaient part de leurs réponses à l'aide d'une échelle de Likert en sept points.

A l'issue des analyses factorielles exploratoires réalisées, les résultats relatifs à la structure unidimensionnelle des échelles et à leur fiabilité sont les suivants : présence sociale ($\alpha = 0,936$; 83,8% de la variance), confiance envers le

conseiller virtuel ($\alpha = 0,955$; 71,2% de la variance), confiance envers le site ($\alpha = 0,911$; 74 % de la variance) et intentions comportementales ($\alpha = 0,638$; 59,3% de la variance). Une analyse factorielle confirmatoire a ensuite été conduite afin d'évaluer la fiabilité et les validités convergente et discriminante des échelles de mesure. Les principaux résultats sont consignés dans les annexes 2 et 3. Suite aux analyses, la version épurée de l'échelle d'intentions comportementales a été retenue (l'item « Je me rendrai en agence pour souscrire » présentait en effet une très faible contribution factorielle : 0,254). La fiabilité mesurée par le ρ de Jöreskog est satisfaisante au seuil de 0,90 pour toutes les échelles (sauf pour les intentions comportementales : 0,66). Par ailleurs, toutes les contributions factorielles sont significatives. La validité convergente des échelles est appréciée par le ρ de validité convergente (pvc). Les valeurs étant toutes supérieures à 0,50 pour les échelles, il est donc possible d'attester de leur validité convergente. En ce qui concerne les indices d'ajustement du modèle, ils sont satisfaisants (χ^2 (df) = 3,369 ; CFI = 0,952 ; TLI = 0,945 et RMSEA = 0,067). Enfin, pour ce qui est de la validité discriminante, les carrés des corrélations entre les construits ont été analysés (voir annexe 3). Le carré des corrélations de chaque échelle avec tous les autres construits est inférieur au pvc (Fornell et Larcker, 1981). Par ailleurs, la corrélation de chaque construit avec les autres construits du modèle est comprise entre 0,1 et 0,85 (Kline, 1998). La validité discriminante est donc établie pour toutes les échelles³.

Les résultats des tests d'hypothèses étaient susceptibles d'être affectés par un biais de variance commune lié au fait d'interroger un seul répondant. Dans le but de contrôler l'impact que pourrait avoir ce biais sur nos résultats, nous avons suivi la procédure recommandée par Podsakoff et al. (2003). C'est ainsi que nous avons réestimé le modèle en ajoutant un facteur de méthode. La comparaison des coefficients standardisés lorsqu'un facteur de méthode est ou non ajouté au modèle montre que la significativité et le signe des coefficients restent inchangés. Ces résultats prouvent que les résultats des tests des hypothèses ne sont pas affectés par la présence d'un biais de variance commune.

Tableau 1. Tests de comparaison de moyennes des effets de la voix de synthèse vs de la voix humaine.

	Moyennes		Test de Levene		Test t de comparaison de moyenne		Résultats des hypothèses
	Voix de synthèse	Voix humaine	F	Sig.	t	Sig.	
La présence sociale	3,27	3,77	0,435	0,510	4,715	0,000***	Confirmée
La confiance envers le site	4,30	4,37	0,889	0,346	0,839	0,402	Rejetée
Les intentions comportementales	4,16	4,40	0,077	0,781	2,863	0,004*	Confirmée

*** $p < 0,001$; * $p < 0,05$.

Les résultats de l'étude

Afin de tester les trois premières hypothèses du modèle relatives à l'influence du type de voix du conseiller virtuel sur la présence sociale perçue sur le site, la confiance envers le site et les intentions comportementales (H1a, H1b, H1c), une série d'analyses de variance a été effectuée. Les résultats montrent un effet significatif du type de voix sur la présence sociale (tableau 1). Un conseiller virtuel s'exprimant avec une voix humaine entraîne un sentiment de présence sociale plus fort qu'un conseiller virtuel s'exprimant avec une voix de synthèse ($M_{\text{humaine}} = 3,77$ et $M_{\text{synthèse}} = 3,27$; $F = 0,435$; $p = 0,000$). L'hypothèse H1a est donc confirmée. En revanche, aucun effet statistiquement significatif n'a été relevé en ce qui concerne l'influence du type de voix sur la confiance envers le site ($M_{\text{humaine}} = 4,37$ et $M_{\text{synthèse}} = 4,30$; $F = 0,889$; $p = 0,402$). L'hypothèse H1b est donc rejetée. Enfin, les résultats démontrent un effet significatif du type de voix sur les intentions comportementales. Ainsi, quand le conseiller virtuel s'exprime avec une voix humaine, les intentions comportementales sont plus importantes que lorsqu'il parle avec une voix de synthèse ($M_{\text{humaine}} = 4,40$ et $M_{\text{synthèse}} = 4,16$; $F = 0,077$; $p = 0,004$) L'hypothèse H1c est confirmée.

Suite à l'étude des effets directs du type de voix (H1a à H1c), nous nous sommes demandés s'il n'était pas possible d'envisager un effet médiateur simple et multiple en série de la confiance envers le conseiller virtuel et de la confiance envers le site entre le type de voix et les intentions comportementales (H1d). Pour tester cette hypothèse

de médiation simple et de médiation multiple en série, la macro PROCESS utilisée sur SPSS a été employée (modèle 6 de médiation multiple en série)⁴. Les résultats qui suivent sont interprétés suivant les préconisations de Zhao et al. (2011).

L'examen du tableau 2 montre que l'effet total et l'effet direct sont significatifs. Toutefois, l'observation du premier effet indirect ($a1 \times b1 \times c$) est positif confirmant l'existence d'une médiation complémentaire de la confiance envers l'agent virtuel pour l'effet de la voix sur les intentions. Ainsi, les répondants exposés à une voix de synthèse évaluent plus négativement la confiance dans l'agent virtuel que ceux exposés à une voix humaine, et ceci aura un effet négatif sur les intentions comportementales. La médiation complémentaire, telle que préconisée par Zhao et al. (2011), suggère que le modèle théorique est incomplet et qu'il est possible d'intégrer d'autres variables médiatrices omises.

La non significativité du coefficient « $a2 \times b2$ » ne nous permet pas de confirmer l'effet médiateur simple de la confiance envers le site dans le cas de l'influence de la voix sur les intentions comportementales. Dans la mesure où il existe un effet direct « c » de la voix sur les intentions comportementales, nous pouvons donc intégrer d'autres variables médiatrices omises aux modèles théoriques. D'où l'intérêt d'étudier la médiation multiple en série de la confiance envers le conseiller virtuel et de la confiance envers le site. Le coefficient « $a1 \times a3 \times b2$ », correspondant à l'effet de la voix sur les intentions comportementales en passant par la confiance envers l'agent virtuel puis la confiance envers le site, est significatif. Un fort sentiment de

Tableau 2. Résultats des effets médiateurs de la confiance envers l'agent virtuel et de la confiance envers le site.

Effet total			Effet direct			Effet indirect			Résultat
c'	t	p	c	t	p	Coefficients	Valeurs	p	IC
-0,246	-2,863	0,0043*	-0,157	-2,441	0,0149*	a1	-0,2003	0,0179*	
						a2	0,0812	0,1281	
						a3	0,7515	0,000***	
						b1	0,2953	0,000***	
						b2	0,4429	0,000***	
						a1 × b1	-0,0592		[-0,121 ; Médiation simple -0,013] complémentaire
						a1 × a3 × b2	-0,0722		[-0,129 ; Médiation multiple en -0,015] série complémentaire
						a2 × b2	0,0324		[-0,008 ; Médiation simple 0,084] seulement indirecte (non-médiation)

*** $p < 0,001$; * $p < 0,05$; Codage de la variable *muette* : voix de synthèse (1) et voix humaine (0).

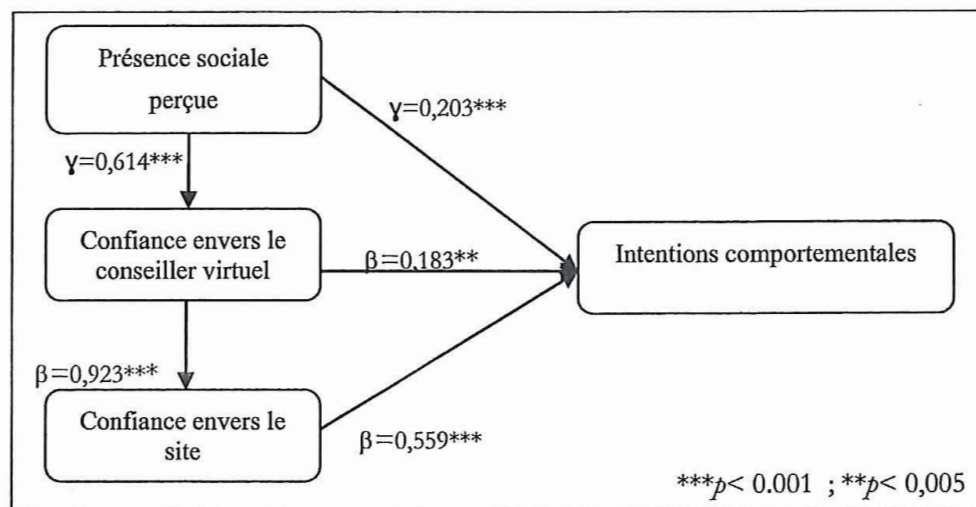
« a1 » : l'effet de la voix sur la confiance envers l'agent virtuel.

« a2 » : l'effet de la voix sur la confiance envers le site.

« a3 » : l'effet de la confiance envers l'agent virtuel sur la confiance envers le site.

« b1 » : l'effet de la confiance envers l'agent virtuel sur les intentions comportementales.

« b2 » : l'effet de la confiance envers le site sur les intentions comportementales.

**Figure 2.** Les relations du modèle structurel.

confiance dans le conseiller virtuel résultant d'une voix humaine se traduit par un fort sentiment de confiance dans le site qui, à son tour, influencera positivement les intentions comportementales. Il s'agit ici aussi d'une médiation complémentaire positive ($a1 \times a3 \times b2 \times c$) suggérant l'existence d'autres médiateurs tels que la présence sociale. L'hypothèse H1d est ainsi confirmée.

La suite des hypothèses (H2, H3, H4, H6) a été testée à l'aide des équations structurelles sous le

logiciel AMOS. Comme l'indique la figure 2, toutes les relations entre les variables du modèle structurel sont significatives. La présence sociale influence positivement la confiance envers le conseiller virtuel ($\gamma = 0,614$; $t = 14,779$; $p = 0,000$) et les intentions comportementales ($\gamma = 0,203$; $t = 5,223$; $p = 0,000$). La confiance envers le conseiller virtuel, pour sa part, influence positivement la confiance envers le site ($\beta = 0,923$; $t = 19,089$; $p = 0,000$) et les intentions comportementales ($\beta = 0,183$; $t = 2,829$; $p = 0,005$).

Tableau 3. Résultat de l'effet médiateur de la confiance envers le conseiller virtuel.

Effet total			Effet direct			Effet indirect				Résultat
c'	t	p	c	t	p	A	b	a x b	IC	Médiation complémentaire
0,387	14,930	0,000***	0,068	2,890	0,0004***	0,4574	0,696	0,318	[0,260 ; 0,382]	

*** $p < 0,001$.

Tableau 4. Synthèse des résultats de l'étude empirique.

Hypothèses de la recherche : C = confirmée, R = rejetée		
H1a	Voix humaine/voix de synthèse – Présence sociale perçue	C
H1b	Voix humaine/voix de synthèse – Confiance envers le site	R
H1c	Voix humaine/voix de synthèse – Intentions comportementales	C
H1d	Voix humaine/voix de synthèse – Confiance envers le conseiller virtuel – Confiance envers le site – Intentions comportementales	C
H2	Confiance envers le conseiller virtuel – Intentions comportementales	C
H3	Présence sociale perçue – Confiance envers le conseiller virtuel	C
H4	Présence sociale perçue – Intentions comportementales	C
H5	Présence sociale perçue – Confiance envers le conseiller virtuel – Confiance envers le site	C
H6	Confiance envers le site – Intentions comportementales	C

Enfin, la confiance envers le site influence positivement les intentions comportementales ($\beta = 0,559$; $t = 9,187$; $p = 0,000$). Les hypothèses H2, H3, H4 et H6 peuvent donc être confirmées.

Par ailleurs, en appliquant la démarche de Zhao et al. (2011), il est possible de prouver l'existence de l'effet médiateur de la confiance envers le conseiller virtuel entre la présence sociale perçue et la confiance envers le site. Le produit « $a \times b \times c$ » étant positif (tableau 3), il est possible de conclure à une médiation complémentaire suggérant l'existence d'autres médiateurs dans le modèle théorique. De ce fait, l'hypothèse H5 est confirmée.

Le tableau 4 présente la synthèse relative à la confirmation ou au rejet des hypothèses de la recherche.

Apports, implications, limites et voies de recherche

Sur le plan théorique, le premier intérêt de cette recherche est de s'intéresser aux phénomènes d'interaction entre un internaute et un agent virtuel aux caractéristiques anthropomorphiques bien spécifiques (voix humaine vs voix de synthèse) dans

un contexte marchand. S'il existe déjà en robotique, en informatique et en systèmes d'information des recherches dédiées au rôle de la voix dans les interactions homme-machine, force est de constater qu'il n'en est pas de même en marketing. Ce travail vient donc enrichir la littérature en marketing et plus particulièrement celle traitant de l'impact des facteurs sociaux sur le comportement des internautes.

Le deuxième apport théorique de ce papier est de démontrer, d'une part, l'effet de la voix d'un agent virtuel sur les réactions de l'internaute, d'autre part, la supériorité de la voix humaine sur la voix de synthèse. Une première série d'hypothèses stipule l'influence des caractéristiques de la voix sur la présence sociale, la confiance envers le site et les intentions comportementales. Les résultats montrent que les deux types de voix sont en mesure d'impacter positivement la présence sociale et les intentions comportementales corroborant ainsi les travaux de Lee et Nass (2003 ; 2004) selon lesquels la voix de synthèse sur un ordinateur est capable d'accroître la présence sociale. Néanmoins, l'analyse des données collectées fait ressortir que les effets sont nettement meilleurs dans le cas d'une voix humaine que dans le cas d'une voix synthétique. Plus précisément, la

présence sociale perçue est plus forte sur un site avec un conseiller virtuel s'exprimant avec une voix humaine que sur un site avec un conseiller virtuel s'exprimant avec une voix de synthèse (H1a). Ces résultats viennent conforter la supériorité de la voix humaine préenregistrée sur la voix de synthèse et confirment les travaux antérieurs sur la capacité des conseillers virtuels à simuler une présence sociale sur internet, à humaniser l'interaction entre le client et le site et à rendre l'expérience de magasinage plus interpersonnelle (Ben Mimoun et Poncin, 2015 ; Holzwarth et al., 2006 ; Kohler et al., 2011 ; Qiu et Benbasat, 2009 ; Wang et al., 2007). Ils apportent également un nouvel éclairage quant aux effets de la voix du conseiller virtuel sur la présence sociale. Contrairement aux travaux de Qiu et Benbasat (2009), les résultats de cette recherche confirment qu'un conseiller virtuel doté d'une voix humaine génère davantage de présence sociale qu'un agent doté d'une voix de synthèse. Si ces conclusions confortent les résultats obtenus par Nass et Scott (2005) et par Stern et al. (1999) en interaction homme-machine, ils sont mis en évidence pour la première fois dans un contexte marketing. Par ailleurs, les résultats mettent en avant l'effet significatif de la voix sur les intentions comportementales (H1c). Un conseiller virtuel doté d'une voix humaine suscite davantage d'intentions comportementales qu'un conseiller virtuel avec une voix de synthèse. Si plusieurs auteurs ont constaté l'effet positif de la présence d'un conseiller virtuel sur les intentions comportementales (Holzwarth et al., 2006 ; Keeling et al., 2010 ; Lemoine et Notebaert, 2011 ; Wang et al., 2007), très peu de recherches ont étudié l'influence spécifique de la voix du conseiller sur ces dernières. Le présent travail vient combler ce manque tout en confirmant les conclusions des rares investigations, notamment celles de Lee et Nass (2003) et de Qiu et Benbasat (2009), dédiées à la relation voix du conseiller-intentions comportementales. Concernant l'effet des caractéristiques de la voix sur la confiance envers le site, les résultats obtenus ne permettent pas en revanche d'en déduire une supériorité de la voix humaine sur la voix de synthèse (H1b). D'une manière générale, la confirmation des hypothèses H1a, H1b et H1c permet de conclure à la primauté de la voix humaine sur la voix de synthèse, non pas dans

le cadre d'une interaction homme-ordinateur, comme cela a déjà été démontré en robotique, informatique et systèmes d'information, mais dans le contexte précis d'une interaction marchande entre un internaute et un agent virtuel. Ce résultat confirme, par ailleurs, l'intérêt de la maximisation du réalisme des conseillers virtuels dans le but d'offrir aux internautes une meilleure expérience en ligne.

La troisième contribution théorique de cette recherche est la mise en évidence d'effets de médiation simple et de médiation multiple en série entre les variables de notre modèle. Ainsi, la confiance envers le conseiller virtuel apparaît comme un médiateur de la relation « présence sociale perçue – confiance envers le site » (H5). A notre connaissance, ce résultat n'a pas été mis en évidence jusqu'à présent dans la littérature en marketing. Par ailleurs, les résultats de la recherche permettent de conclure à l'existence d'un effet médiateur multiple en série de la confiance envers le conseiller virtuel puis de la confiance envers le site dans la relation « type de voix – intentions comportementales » (H1d). Ils sont l'occasion de confirmer que la confiance envers le site est impactée par la confiance envers le conseiller virtuel. Ainsi, un fort sentiment de confiance dans le conseiller virtuel résultant d'une voix humaine se traduit par un fort sentiment de confiance dans le site qui, à son tour, influence positivement les intentions comportementales. La mise au jour de cette médiation multiple en série enrichit la littérature dédiée à l'analyse du type de voix du conseiller virtuel sur les intentions comportementales.

La dernière contribution théorique de ce papier repose sur la confirmation, dans un contexte marketing, des hypothèses de notre modèle structurel. Les résultats montrent que toutes les relations envisagées sont statistiquement significatives. Ils mettent en évidence le lien significatif de la confiance envers le conseiller virtuel sur la confiance envers le site ainsi que sur les intentions comportementales (H1d et H2). Tel que vérifié précédemment au niveau de la médiation, plus l'internaute éprouve un sentiment de confiance envers l'agent virtuel, plus il fera confiance au site et plus ses intentions comportementales seront fortes. Ces résultats viennent enrichir la littérature sur les effets de la confiance envers les conseillers virtuels qui ont été

très peu étudiés sauf dans le cadre des travaux de Qiu et Benbasat (2005) et de Wang et Benbasat (2005 ; 2008). Par ailleurs, l'influence positive de la présence sociale perçue sur la confiance envers le conseiller virtuel (H3) et sur les intentions comportementales (H4) a également été démontrée. Ces résultats corroborent les investigations de Qiu et Benbasat (2009) et de Wang et al. (2007) qui mettent en évidence le rôle de la présence sociale dans le développement de la confiance envers le conseiller virtuel et dans l'explication des intentions d'achat. Enfin, la présente recherche confirme un effet significatif positif de la confiance envers le site sur les intentions comportementales (H6) confortant ainsi les conclusions de Keeling et al. (2010) et de Lemoine et Notebaert (2011).

Sur le plan managérial, des recommandations sont susceptibles d'être formulées suite aux résultats obtenus. Ces derniers peuvent être utilisés afin de convaincre les directeurs marketing et les responsables de sites web de l'intérêt de recourir à des conseillers virtuels. Le coût de leur installation représente souvent un frein pour les managers qui s'interrogent sur la pertinence d'investir dans de telles interfaces. Grâce aux résultats de cette recherche, il est possible de les renseigner précisément sur les avantages marketing que présentent les agents virtuels et donc de réduire leurs réticences à les utiliser. Par ailleurs, les conclusions de cet article permettent d'aider les professionnels dans leur choix du type de conseiller virtuel à installer. Plus précisément, il ressort de notre étude qu'il est préférable de se doter d'un agent virtuel à la voix humaine plutôt qu'à la voix de synthèse. En opérant de la sorte, les managers peuvent accroître la présence sociale perçue sur leurs sites et améliorer la confiance que les internautes éprouvent vis à vis des conseillers virtuels. Les professionnels sont également en mesure d'influencer davantage les intentions de recommandation du site lorsqu'ils installent un agent virtuel à voix humaine plutôt qu'à voix de synthèse. Utiliser des conseillers virtuels à voix humaine est donc synonyme, pour les responsables de sites web, d'acquisition d'avantages concurrentiels en matière de présence sociale perçue, de confiance ressentie à l'égard des agents virtuels et de recommandation du site par les internautes. En

d'autres termes, les effets des agents virtuels à voix humaine sur les réactions des consommateurs constituent des leviers de différenciation utiles pour les managers dans un contexte de forte concurrence entre les sites web. De telles implications managériales ne vont pas sans soulever des questions d'ordre éthique dont il faut avoir conscience. En dotant les conseillers virtuels d'une empathie artificielle à l'aide de la voix, il existe un risque pour certains consommateurs de se trouver en situation de dépendance affective vis-à-vis de ce type d'interface et de se laisser manipuler (Tisseron, 2015). Ces mêmes individus peuvent aussi finir par trop s'attacher à ces outils interactifs, les préférer à la compagnie humaine et attendre des autres qu'ils présentent les mêmes qualités, à savoir la serviabilité, la gentillesse et l'absence d'ego (Devilleers, 2017 ; Tisseron, 2015).

Notre recherche comporte des limites qui peuvent constituer des pistes de recherche pour l'avenir. Tout d'abord, cette étude présente une validité externe réduite compte tenu de la spécificité du site sur lequel nous avons travaillé (la GMF). Pour pallier ce problème, il conviendrait de répliquer notre expérimentation sur des sites web appartenant à divers secteurs d'activité (alimentation, culture, transport, etc.). Par ailleurs, notre recherche s'est limitée à l'analyse des effets du type de voix du conseiller virtuel. Nous n'avons pas tenu compte de ses autres caractéristiques anthropomorphiques telles que, par exemple, ses expressions faciales, ses gestes, sa posture corporelle. Nous pourrions envisager de tester, dans de prochains travaux, les effets combinés de ces caractéristiques sur les réactions des internautes. Nous pourrions également nous intéresser aux autres composantes atmosphériques des sites web et voir comment les caractéristiques anthropomorphiques des conseillers virtuels interagissent avec les dimensions sensorielles du site et son design (couleur, musique, navigabilité, accessibilité des produits, etc.). Une telle investigation permettrait d'identifier les composantes atmosphériques des sites web qui ont le plus grand impact sur les réactions des internautes et d'en déduire des facteurs de différenciation pertinents pour ces derniers (Lemoine, 2008). Enfin, l'intégration de variables modératrices dans notre modèle de recherche, telles que la familiarité avec internet et avec les conseillers virtuels, pourrait aussi être

envisagée afin de le rendre plus réaliste et nous permettre d'affiner notre compréhension de l'impact des caractéristiques anthropomorphiques des conseillers virtuels sur les réactions des internautes. Plus précisément, nous pourrions tenter de vérifier si les effets de la voix humaine, par rapport à la voix de synthèse, sur la présence sociale perçue, la confiance envers le conseiller virtuel, la confiance envers le site et les intentions comportementales sont d'autant plus forts que les consommateurs présentent une faible familiarité avec internet et les conseillers virtuels. On peut en effet supposer que les internautes familiers avec internet et avec les agents virtuels prêtent moins d'attention à leurs caractéristiques anthropomorphiques car ils sont habitués à les rencontrer lors de leur navigation. En revanche, un individu peu coutumier des achats en ligne et des interfaces homme-machine pourrait y accorder une attention plus forte en raison, d'une part, de son besoin d'être aidé lors de ses acquisitions en ligne et, d'autre part, de son souhait de ne pas se sentir seul durant sa navigation.

Remerciements

Les auteurs remercient la rédactrice en chef et les relecteurs anonymes pour leurs précieux commentaires et suggestions. Ils remercient également Pascal Arbault, Président Directeur Général et Fondateur de la société Davi pour leur avoir permis d'utiliser le conseiller virtuel Prosper dans le cadre de l'étude empirique de ce papier.

Notes

1. Dans la littérature en marketing et en systèmes d'information, ils sont également nommés agents conversationnels incarnés (Cassell et Bickmore, 2000), agents virtuels, agents synthétiques, personnages interactifs, agents pédagogiques animés, agents de shopping et agents animés (McGoldrick et al., 2008).
2. Pour ce qui est du sourire et du regard, ils agissent fortement sur la crédibilité et la confiance éprouvées vis-à-vis du conseiller virtuel ainsi que sur les attitudes et les comportements des utilisateurs (Beale et Creed, 2009 ; Cowell et Stanney, 2005 ; Gong et Nass, 2007 ; Krämer et al., 2013). Okonkwo et Vassileva (2001), pour leur part, montrent qu'un agent virtuel avec un visage souriant améliore la concentration et la motivation chez les participants et qu'il est perçu comme plus sympathique qu'un

agent non souriant. De même, il a été montré que la gestuelle adoptée par l'agent virtuel (mouvements des bras, hochements de la tête) permet non seulement d'attirer l'attention des utilisateurs, d'améliorer et de synchroniser la communication avec ces derniers mais aussi d'établir avec eux une relation sociale (Imai et al., 2003 ; Ogawa et Watanabe, 2001).

3. Afin d'établir clairement la validité discriminante de la confiance envers le conseiller virtuel et celle de la confiance envers le site, nous avons mis en oeuvre l'approche de Bagozzi et al. (1991) qui recommandent de comparer un modèle libre où la corrélation entre les deux construits est laissée libre à un modèle contraint où cette corrélation est fixée à 1. Il ressort des analyses que le modèle libre présente un meilleur ajustement que le modèle contraint et que la différence entre les deux est statistiquement significative selon le test du Chi-deux ($\Delta\chi^2 = 6,232$; $p = 0,01$). Les validités discriminantes de la confiance envers le conseiller virtuel et de la confiance envers le site sont donc bien confirmées.
4. Par convention, le nombre de *bootstraps* est fixé à 5000. Pour plus de détails sur la méthode, voir Cadario et al. (2017), Hayes (2012) et Preacher et Hayes (2008).

Références bibliographiques

- André E et Pelachaud C (2008) Interacting with embodied conversational agents. In: Chen F et Jokinen K (éds) *New Trends in Speech Based Interactive Systems*. New York: Springer.
- Ardelet C et Brial B (2011) Influence des recommandations d'internautes : le rôle de la présence sociale et de l'expertise. *Recherche et Applications en Marketing* 26(3): 45–69.
- Bagozzi RP, Yi Y et Phillips LW (1991) Assessing construct validity in organizational research. *Administrative Science Quarterly* 36(3): 421–458.
- Bartneck C (2001) How convincing is Mr. Data's smile: affective expressions of machines. *User Modeling and User-Adapted Interaction* 11(4): 279–295.
- Beale R et Creed C (2009) Affective interaction: how emotional agents affect users. *International Journal of Human Computer Studies* 67(9): 755–776.
- Beller G et Rodet X (2007) Content-based transformation of the expressivity in speech. In: *16^{ème} Congrès ICPHs*, Saarbruecken, Allemagne, 8–10 août: 2157–2160.
- Belk R (2016) Comprendre le robot : commentaires sur Goudey et Bonnin (2016). *Recherche et Applications en Marketing* 31(4): 89–97.

- Ben S (2005) Escaping the world: high and low resolution in gaming. *IEEE MultiMedia* 12(4): 4–8.
- Ben Mimoun MS et Poncin I (2015) A valued agent: how ECAs affect website customers' satisfaction and behaviors. *Journal of Retailing and Consumer Services* 26: 70–82.
- Ben Mimoun MS, Poncin I et Garnier M (2012) Case study: embodied virtual agents: an analysis on reasons for failure. *Journal of Retailing and Consumer Services* 19(6): 605–612.
- Beutnagel M, Conkie A, Schroeter J, Stylianou Y et Syrdal A (1999) The AT&T next-gen TTS system. In: *Joint Meeting of ASA, EAA, and DAGA*, Berlin, Allemagne, mars.
- Bos N, Olson JS, Gergle D, Olson GM et Wright Z (2002) Effects of four computer-mediated communications channels on trust development. In: *Congrès du SIGCHI Conference on Human Factors in Computing Systems CHI'02*, Minneapolis MN, 20–25 avril 2002: 135–140.
- Bressolles G, Durrieu F et Viot C (2011) Les effets de la présence sociale perçue sur la personnalité, la qualité du site web et la satisfaction. In: *27^{ème} Congrès de l'Association Française du Marketing*, Bruxelles, 18–20 mai.
- Bressolles G, Vignolles A et Viot C (2015) La présence sociale : quels effets sur la personnalité, la réputation et la confiance envers le site web ? In: *31^{ème} Congrès de l'Association Française du Marketing*, Marrakech, 19–22 mai.
- Burgoon J, Bonito JA, Bengtsson B, Cederberg C, Lundeberg M et Allspach L (2000) Interactivity in human-computer interaction, a study of credibility, understanding and influence. *Computers in Human Behavior* 16(6): 553–574.
- Cadario R, Butori R et Parguel B (2017) *Méthode expérimentale : analyses de modération et médiation*. Louvain la Neuve : De Boeck Supérieur.
- Cassell J et Bickmore T (2000) External manifestations of trustworthiness in the interface. *Communications of the ACM* 43(12): 50–56.
- Chérif E et Lemoine JF (2014) La cohérence du langage verbal et non verbal des conseillers virtuels et les réactions des internautes : une application de la théorie de la cohérence. *17^{ème} Colloque Etienne Thil*, Paris, 15–17 octobre.
- Cowell AJ et Stanney KM (2005) Manipulation of non-verbal interaction style and demographic embodiment to increase anthropomorphic computer character credibility. *International Journal of Human-Computer Studies* 62(2): 281–306.
- Devillers L (2017) *Des robots et des hommes. Mythes, fantasmes et réalité*. Paris: Plon.
- Duffy BR (2003) Anthropomorphism and the social robot. *Robotics and Autonomous Systems* 42(3–4): 177–190.
- Edlund J, Gustafson J, Heldner M et Hjalmarsson A (2008) Towards human-like spoken dialogue systems. *Speech Communication* 50(8–9): 630–645.
- Evans RE et Kortum P (2010) The impact of voice characteristics on user response in an interactive voice response system. *Interacting with Computers* 22(6): 606–614.
- Fornell C et Larcker DF (1981) Evaluating structural equations models with unobservable variables and measurement error. *Journal of Marketing Research* 18(1): 39–50.
- Fogg BJ et Nass C (1997) Silicon sycophants: the effects of computers that flatter. *International Journal of Human-Computer Studies* 46(5): 551–561.
- Galanxhi JH et Nah H (2007) Deception in cyberspace. A comparison of text-only vs. avatar-supported medium. *International Journal of Human-Computer Studies* 65(9): 770–783.
- Gefen D et Straub DW (1997) Gender differences in the perception and use of email: an extension of the TAM. *MIS Quarterly* 21(4): 389–400.
- Gefen D et Straub DW (2003) Managing user trust in B2C e-services. *e-Service Journal* 2(2): 7–24.
- Gefen D et Straub DW (2004) Consumer trust in B2C e-Commerce and the importance of social presence: experiments in e-Products and e-Services. *The International Journal of Management Science* 32(6): 407–424.
- Gong L et Nass C (2007) When a talking-face computer agent is half-human and half-humanoid: human identity and consistency preference. *Human Communication Research* 33(2): 163–193.
- Goudey A et Bonnin G (2016) Un objet intelligent doit-il avoir l'air humain ? Etude de l'impact de l'anthropomorphisme d'un robot compagnon sur son acceptation. *Recherche et Applications en Marketing* 31(2): 3–22.
- Greenspan R, Goldberg D, Weimer D et Basso A (2000) Interpersonal trust and common ground in electronically mediated communication. In: *The 2000 ACM Conference on Computer Supported Cooperative Work (CSCW'00)*. Disponible en ligne: <https://dl.acm.org/citation.cfm?id=358916>
- Grichkovtsova I, Morel M et Lacheret A (2012) The role of voice quality and prosodic contour in affective speech perception. *Speech Communication* 54(3): 414–429.
- Groom V, Nass C, Chen T, Nielsen A, Scarborough JK et Robles E (2009) Evaluating the effects of behavioral

- realism in embodied agents. *International Journal of Human-Computer Studies* 67(10): 842–849.
- Hayes AF (2012) *Process: A Versatile Computational Tool for Observed Variable Mediation, Moderation, and Conditional Process Modeling*. White paper. Disponible en ligne: <http://www.afhayes.com/public/process2012.pdf>
- Holzwarth M, Janiszewski C et Neumann MM (2006) The influence of avatars on online consumer shopping behavior. *Journal of Marketing* 70(4): 19–36.
- Imai M, Ono T et Ishiguro H (2003) Physical relation and expression: joint attention for human-robot interaction. *IEEE Transactions on Industrial Electronics* 50(4): 636–643.
- Ivory JD et Kalyanaraman S (2007) The effects of technological advancement and violent content in video games on players' feelings of presence, involvement, physiological arousal, and aggression. *Journal of Communication* 57(3): 532–555.
- Jarvenpaa SL, Tractinsky N et Vitale M. (2000) Consumer trust in an Internet store. *Information Technology and Management* 1(1–2): 45–71.
- Kang SH et Watt JH (2013) The impact of avatar realism and anonymity on effective communication via mobile devices. *Computers in Human Behavior* 29(3): 1169–1181.
- Keeling K, McGoldrick P et Beatty S (2010) Avatars as salespeople: communication style, trust, and intentions. *Journal of Business Research* 63(8): 793–800.
- Kline RB (1998) *Principles and Practice of Structural Equation Modeling*. New York: The Guilford Press.
- Koh YJ et Sundar SS (2010) Effects of specialization in computers, web sites, and web agents on e-commerce trust. *International Journal of Human-Computer Studies* 68(12): 899–912.
- Kohler T, Fueller J, Stieger D et Matzler K (2011) Avatar-based innovation: consequences of the virtual co-creation experience. *Computers in Human Behavior* 27(1): 160–168.
- Komiak SYX et Benbasat I (2006) The effects of personalization and familiarity on trust and adoption of recommendation agents. *MIS Quarterly* 30(4): 941–960.
- Krämer N, Kopp S, Becker-Asano C et Sommer N (2013) Smile and the world will smile with you – the effects of a virtual agent's smile on users' evaluation and behavior. *International Journal of Human-Computer Studies* 71(3): 335–349.
- Lancelot-Miltgen C et Lemoine JF (2015) Mieux collecter les données personnelles sur Internet. Une étude qualitative auprès d'internautes français. *Décisions Marketing* 79: 35–52.
- Lee EJ (2008) Flattery may get computers somewhere, sometimes: the moderating role of output modality, computer gender, and user gender. *International Journal of Human-Computer Studies* 66(11): 789–800.
- Lee KM et Nass C (2003) Designing social presence of social actors in human computer interaction. In: *Congrès du SIGCHI Conference on Human Factors in Computing Systems CHI'03*, Ft. Lauderdale FL, 5–10 avril: 289–296.
- Lee KM et Nass C (2004) The multiple source effect and synthesized speech: doubly disembodied language as a conceptual framework. *Human Communication Research* 30(2): 182–207.
- Lemoine JF (2008) Atmosphère des sites web marchands et réactions des internautes. *Revue Française du Marketing* 217: 45–61.
- Lemoine JF et Chérif E (2012) Comment générer de la confiance envers un agent virtuel à l'aide de ses caractéristiques ? Une étude exploratoire. *Management et Avenir* 58: 169–188.
- Lemoine JF et Notebaert JF (2011) Agent virtuel et confiance des internautes vis-à-vis d'un site web. *Décisions Marketing* 61: 47–53.
- Lester JC, Converse SA, Kahler SE, Barlow ST, Stone BA et Bhogal RS (1997) The persona effect: affective impact of animated pedagogical agents. In: *The 1997 SIGCHI Conference on Human Factors in Computing Systems*. New York: ACM Press, 359–366.
- Loiacono ET, Djasasbi S et Kiryazov T (2013) Factors that affect visually impaired users' acceptance of audio and music websites. *International Journal of Human-Computer Studies* 71(3): 321–334.
- McBreen HM et Jack MA (2001) Evaluating humanoid synthetic agents in e-retail applications. *IEEE Transactions on Systems, Man and Cybernetics: Part A Systems and Humans. Special Issue: The Human in the Loop* 31(5): 394–405.
- McGoldrick PJ, Keeling KA et Beatty SF (2008) A typology of roles for avatars in online retailing. *Journal of Marketing Management* 24(3–4): 433–461.
- Moon Y (2000) Intimate exchanges: using computers to elicit self-disclosure from consumers. *Journal of Consumer Research* 26(4): 323–339.
- Moon Y et Nass C (1996) How “real” are computer personalities? Psychological responses to personality types in human-computer interaction. *Communication Research* 23(6): 651–674.
- Moon Y et Nass C (1998) Are computers scapegoats? Attributions of responsibility in human-computer interaction. *International Journal of Human-Computer Studies* 49(1): 79–94.

- Mori M (1970) The uncanny valley. *Energy* 7(4): 33–35.
- Nan N, Anghelcev G, Myers J, Sar S et Faber R (2006) What if a web site can talk ? Exploring the persuasive effects of web-based anthropomorphic agents. *Journalism and Mass Communication Quarterly* 83(3): 615–631.
- Nass C et Gong L (2000) Speech interfaces from an evolutionary perspective. *Communications of the ACM* 43(9): 36–43.
- Nass C et Lee KM (2001) Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *Journal of Experimental Psychology Applied* 7(3): 171–181.
- Nass C et Moon Y (2000) Machines and mindlessness: social responses to computers. *Journal of Social Issues* 56(1): 81–103.
- Nass C, Moon Y et Carney P (1999) Are people polite to computers? Responses to computer-based interviewing systems. *Journal of Applied Social Psychology* 29(5): 1093–1110.
- Nass C, Moon Y et Green N (1997) Are machines gender-neutral? Gender stereotypic responses to computers with voices. *Journal of Applied Social Psychology* 27(10): 864–876.
- Nass C et Scott S (2005) *Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship*. Cambridge MA: MIT Press.
- Nass C et Steuer J (1993) Voices, boxes, and sources of messages: computers and social actors. *Human Communication Research* 19(4): 504–527.
- Ogawa H et Watanabe T (2001) InterRobot: speech-driven embodied interaction robot. *Advanced Robotics* 15(3): 371–377.
- Okonkwo C et Vassileva J (2001) Affective pedagogical agents and user persuasion. In: *4th International Conference on Universal Access in Human-Computer Interaction* (ed Stephanidis C), Beijing. New York: Springer, 5–10.
- Podsakoff PM, MacKenzie SB, Lee JY et Podsakoff NP (2003) Common method biases in behavioral research: a critical review of the literature and recommended remedies. *The Journal of Applied Psychology* 88(5): 879–903.
- Preacher KJ et Hayes AF (2008) Contemporary approaches to assessing mediation in communication research. In: Hayes AF, Slater MD et Snyder LB (éd.) *The Sage Sourcebook of Advanced Data Analysis Methods for Communication Research*. Thousand Oaks CA: SAGE, 13–54.
- Qiu L et Benbasat I (2005) Online consumer trust and live help interfaces: the effects of text-to-speech voice and three-dimensional avatars. *International Journal of Human-Computer Interaction* 19(1): 75–94.
- Qiu L et Benbasat I (2009) Evaluating anthropomorphic product recommendation agents: a social relationship perspective to designing information systems. *Journal of Management Information Systems* 25(4): 145–182.
- Reeves B et Nass CI (1996) *The Media Equation*. Stanford CA: CSLI Publications.
- Ruttkay Z et Pelachaud C (2004) *From Brows to Trust: Evaluating Embodied Conversational Agents*. Dordrecht: Kluwer Academic.
- Sallnas E (2005) Effects of communication mode on social presence, virtual presence, and performance in collaborative virtual environments. *Presence* 14(4): 434–449.
- Scherer KR (1986) Vocal affect expression: a review and a model for future research. *Psychological Bulletin* 99(2): 143–165.
- Scherer KR (2003) Vocal communication of emotions: a review of research paradigms. *Speech Communication* 40(1–2): 227–256.
- Stern SE, Mullennix JW, Dyson C et Wilson SJ (1999) The persuasiveness of synthetic speech versus human speech. *Human Factors* 41(4): 588–595.
- Stern SE, Mullennix JW 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): 43–52.
- Short J, Williams E et Christie B (1976) *The social psychology of telecommunications*. Londres: John Wiley & Sons.
- Tisseron S (2015) *Le jour où mon robot m'aimera. Vers l'empathie artificielle*. Paris : Albin Michel.
- Tzeng JY (2006) Matching users' diverse social scripts with resonating humanized features to create a polite interface. *International Journal of Human-Computer Studies* 64(12): 1230–1242.
- Vaudale C (2012) *Analyse et reconnaissance des émotions lors de conversations de centres d'appels*. Thèse de Doctorat en informatique, Laboratoire LIMSI, Université Paris 11.
- Van Vugt HC, Konijn E, Hoorn JF, Keur I et Eliëns A (2007) Realism is not all! User engagement with task-related interface characters. *Interacting with Computers* 19(2): 267–280.
- Viot C et Bressolles G (2012) Les agents virtuels intelligents : quels atouts pour la relation client ? *Décisions Marketing* 65: 45–56.
- Von Der Pütten AM, Krämer NC, Gratch J et Kang SH (2010) It doesn't matter what you are! Explaining

social effects of agents and avatars. *Computers in Human Behavior* 26(6): 1641–1650.

Wang LC, Baker J, Wagner JA et Wakefield K (2007) Can a retail web site be social ? *Journal of Marketing* 71(3): 143–157.

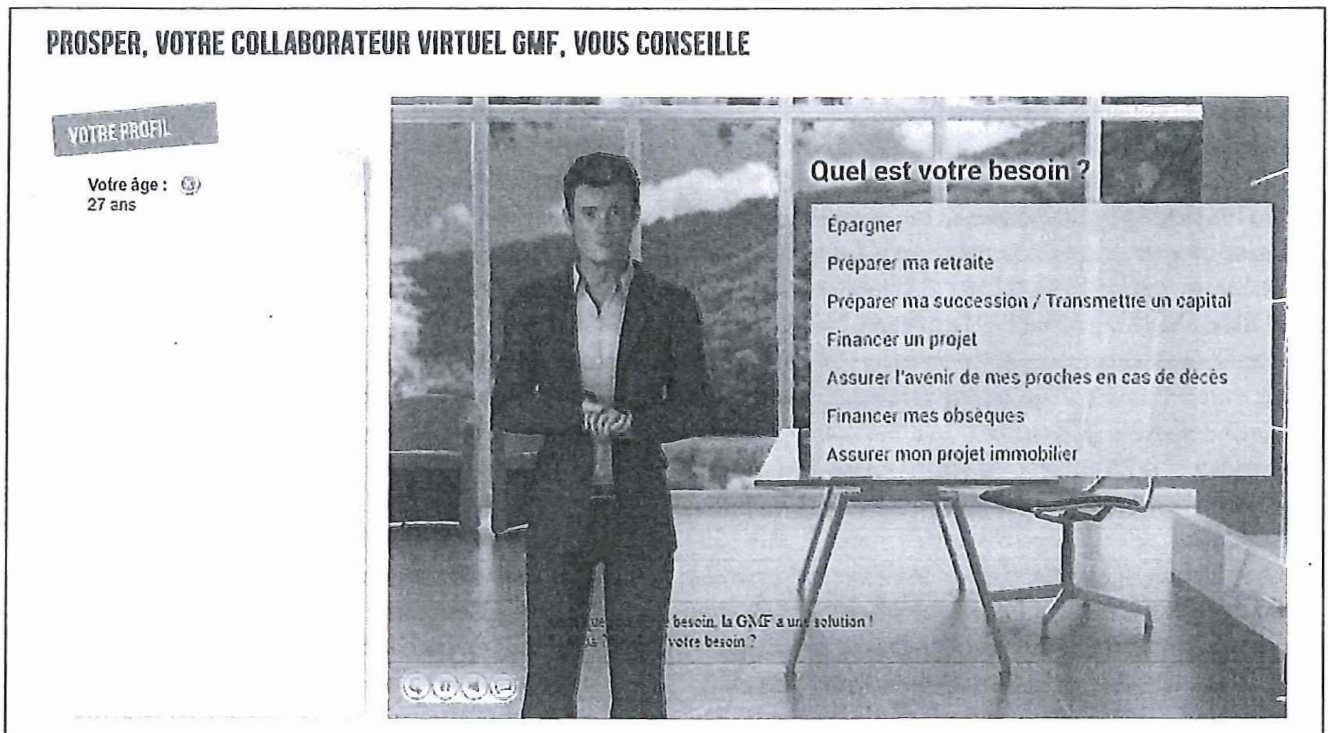
Wang W et Benbasat I (2005) Trust in and adoption of online recommendation. *Journal of Association for Information Systems* 6(3): 72–101.

Wang W et Benbasat I (2008) Attributions of trust in decision support technologies: a study of recommendation

agents for e-commerce. *Journal of Management Information Systems* 24(4): 249–273.

Wongkia W, Naruedomkul K et Cercone N (2012) i-Math: automatic math reader for Thai blind and visually impaired students. *Computers & Mathematics with Applications* 64(6): 2128–2140.

Zhao X, Lynch JG et Chen Q (2011) Reconsidérer Baron et Kenny: mythes et vérités à propos de l'analyse de médiation. *Recherche et Applications en Marketing* 26(1): 81–95.



Annexe 1. Exemple de capture d'écran du site de la GMF.

Annexe 2. Echelles de mesure et résultats de l'analyse factorielle confirmatoire.

Echelles	Items	λ_i	t- value	SMC	Fiabilité	pvc
Présence sociale (Qiu et Benbasat, 2009)	J'ai ressenti de la chaleur humaine sur ce site	0,907		0,822	0,939	0,793
	J'ai ressenti un contact humain sur ce site	0,888	34,504	0,789		
	J'ai ressenti de la sociabilité sur ce site	0,837	30,145	0,700		
	J'ai ressenti une sensibilité sur ce site	0,928	38,386	0,861		
Confiance envers le conseiller virtuel (Qiu et Benbasat,2009)	Le conseiller virtuel a une bonne connaissance des assurances	0,741		0,550	0,954	0,677
	Le conseiller virtuel a fait passer mes intérêts en premier	0,800	21,017	0,640		
	Le conseiller virtuel est compétent dans ses recommandations et explications	0,859	22,786	0,738		

Annexe 2. (Cont.)

Echelles	Items	λ_i	t- value	SMC	Fiabilité	ρ_{vc}
Confiance envers le site (Keeling et al.,2010)	Le conseiller virtuel a accompli son rôle de conseiller efficacement	0,861	22,847	0,742	0,913	0,679
	Le conseiller virtuel est capable de fournir les conseils appropriés	0,851	22,527	0,723		
	Le conseiller virtuel est sincère	0,835	22,054	0,697		
	Le conseiller virtuel aurait fait de son mieux pour m'aider si j'en avais eu besoin	0,819	21,579	0,671		
	Les recommandations effectuées par le conseiller virtuel étaient dans mon intérêt en priorité et non dans le sien (ou celui de l'entreprise)	0,838	22,155	0,703		
	Le conseiller virtuel est honnête	0,806	21,190	0,649		
	Le conseiller virtuel a fourni des recommandations objectives et non biaisées sur les assurances	0,812	21,383	0,660		
	Je crois aux informations données sur ce site	0,868		0,753		
	Je ferai confiance au processus de paiement proposé par ce site	0,758	23,438	0,575		
	Je suis sûr que ma commande sera bien traitée	0,818	26,611	0,669		
Intentions comportementales (Keeling et al.,2010)	Je suis confiant dans les recommandations de ce site	0,913	32,644	0,834		
	Ce site semble être plus fiable que d'autres sites que j'ai visités	0,750	23,031	0,569		
	Je recommanderais ce site à mes proches	0,895		0,801	0,795	0,663
	Je retournerai sur ce site pour chercher plus d'informations	0,724	18,947	0,525		

Annexe 3. Matrice des carrés des corrélations entre les construits.

	AVE	Présence sociale	Confiance envers le conseiller virtuel	Confiance envers le site
Présence sociale	0,793			
Confiance envers le conseiller virtuel	0,677	0,389		
Confiance envers le site	0,679	0,292	0,675	
Intentions comportementales	0,663	0,381	0,592	0,652