

Anthropomorphisation des conseillers virtuels et réactions des internautes : le cas de la voix du conseiller

Emna Chérif

IAE Auvergne, Université Clermont Auvergne (CRCGM)

Jean-François Lemoine

Université Paris 1 Panthéon-Sorbonne (PRISM Sorbonne)

ESSCA School of Management

Introduction

Les conseillers virtuels 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. Les conseillers virtuels¹ sont des représentations graphiques personnifiées de programmes informatiques qui ont été conçus pour interagir et se comporter comme un 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 (Charfi et Volle, 2011 ; Holzwarth et al., 2006 ; Wang et al., 2007). Il a par ailleurs été démontré que leur présence le sentiment de confiance envers le site (Keeling et al., 2010 ; Koh et Sundar, 2010).

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'anthropomorphisme² des conseillers virtuels. En effet, Burgoon et al. (2000) montrent que les utilisateurs affichent une meilleure évaluation de l'utilité et des sentiments favorables envers les agents virtuels quand ils sont conçus avec des caractéristiques anthropomorphiques. Dans une perspective de faciliter l'interaction avec les conseillers virtuels et de susciter les meilleures réactions de la part des consommateurs, les recherches suggèrent d'accroître leur ressemblance avec l'humain et de renforcer leur anthropomorphisme (Edlund et al., 2008 ; Qiu et Benbasat, 2005 ; Stern et al., 2006). Ils suggèrent que les conseillers virtuels devraient être conçus d'une

façon plus anthropomorphe en leur conférant des caractéristiques humaines spécifiques telles que la reconnaissance vocale, une voix de synthèse et une faculté d'animation qui simule les expressions faciales et les gestes humains (Burgoon et al., 2000). Ainsi, plusieurs recherches montrent que les expressions faciales comme le sourire et le regard peuvent agir considérablement 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, 2007 ; Krämer et al., 2013). Okonkwo et Vassileva (2001) ont testé l'effet d'un agent virtuel avec un visage souriant sur les perceptions et les performances des participants. Ils trouvent que l'agent souriant améliore la concentration et la motivation chez les participants et il est également perçu comme plus sympathique que l'agent non souriant. Les travaux antérieurs ont montré, par ailleurs, 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 une relation sociale entre l'agent et l'utilisateur (Imai et al., 2003; Ogawa et Watanabe, 2001). Malgré le développement des travaux sur l'effet des expressions faciales et la gestuelle, les recherches sur la voix sont rares et se sont focalisées en majorité sur les effets de la voix dans une interaction homme-ordinateur. Or, la voix, autant que les expressions faciales, est un puissant élément pour communiquer naturellement dans une interaction sociale (Bartneck , 2001) et donne l'impression aux utilisateurs que les ordinateurs échangent avec eux (Moon, 2000). 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. Dans ce sens, la voix s'avère une caractéristique importante pouvant renforcer le caractère social d'une interaction en ligne (Nass et Steuer, 1993). Short et al. (1976) montrent que les supports audio entraînent un sentiment de présence sociale plus fort que les supports écrits. De même, la voix semble encourager les utilisateurs à adopter, face à un ordinateur, les mêmes réponses sociales qu'ils ont l'habitude d'avoir dans leurs relations habituelles entre humains (Reeves et Nass, 1996 ; Nass et Steuer, 1993). Bien que les recherches antérieures montrent le rôle social que la voix peut jouer 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).

En outre, très peu d'études antérieures se sont interrogées sur la manière de renforcer la confiance des consommateurs dans les conseillers virtuels. Or, la confiance dans le conseiller virtuel impacterait les intentions de l'utiliser (Wang et Benbasat, 2005). Les recherches sur les

interactions médiatisées par ordinateur, montrent que la confiance serait 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 (*chat*) (Bos et al., 2002 ; Greenspan et al., 2000). Dans cette optique, cette recherche s'intéressera à étudier l'effet de la voix du conseiller virtuel sur la confiance dans le conseiller virtuel dans le cadre d'une interaction marchande en ligne. Dans ce contexte, l'acquisition de la confiance est un enjeu crucial pour les entreprises (Gefen et Straub, 2003 ; Jarvenpaa et al., 2000). La présence d'un conseiller virtuel s'exprimant oralement pourrait renforcer la confiance dans le site entièrement (Qiu et Benbasat, 2009 ; Wang et Benbasat, 2005, 2008 ; Wang et al., 2007).

Dans ce sens, 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 étude, 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 GPSs, il n'y a pas eu, à notre connaissance, de travaux en marketing sur le sujet.

Au travers de cette recherche, nous avons essayé 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 (*text-to-speech voice*). Elle nécessite le recours à des logiciels informatiques sophistiqués appelés synthétiseurs vocaux³. Les synthétiseurs 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 des agents virtuels de bénéficier d'une grande flexibilité à des 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 naturelle principalement liées à l'expressivité⁴ de la voix en termes de qualité et de prosodie. Les voix de synthèse présentent des

pauses et des discontinuités qui sonnent "non-humains", et qui pourraient limiter les signaux sociaux que la voix est en mesure de véhiculer (Nass et Scott, 2005).

Cette recherche, réalisée dans le contexte spécifique qui est le marketing, présenterait les premiers résultats empiriques sur les effets d'un conseiller virtuel doté d'une voix de synthèse vs un conseiller virtuel dotée d'une voix humaine. Elle contribuerait à une meilleure compréhension de l'impact de type de la voix du conseiller virtuel sur les réactions des consommateurs en ligne. D'un point de vue managérial, cette recherche se propose de renseigner les créateurs de conseillers virtuels, cherchant à optimiser la conception de ces derniers, sur le choix du type de voix à privilégier puisque les coûts et la qualité varient considérablement entre ces deux types. En identifiant le type de voix le plus efficace, les résultats de cette recherche apportent aux concepteurs des conseillers virtuels et aux webmasters des enseignes souhaitant améliorer l'interaction entre le consommateur et leur site, des arguments robustes sur les effets de la voix du conseiller virtuel.

Dans un premier temps, nous présenterons 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 sur le fait 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 sociales⁵, 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 (sa manière de se comporter : de s'articuler, ect.) (Nass et Moon, 2000 ; Reeves et Nass, 1996). Dans cette perspective, Nass et Steuer (1993) suggèrent tout particulièrement quatre caractéristiques pouvant renforcer le caractère social d'une expérience en ligne : le langage, l'interactivité, le rôle social et la voix. Il a été démontré que les participants perçoivent les différentes voix provenant du même ordinateur comme provenant de plusieurs interlocuteurs. Par ailleurs, les participants considèrent les différents ordinateurs affichant la même voix comme un seul et même interlocuteur (Nass et Steuer, 1993). En ce sens, Nass et Scott (2005) ont montré qu'une interaction orale permet de dialoguer d'une manière plus

naturelle et efficace avec les machines. Les interactions orales avec les machines sont mieux perçues par les utilisateurs que les interactions avec du texte uniquement. Elles 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 par ailleurs que des indices minimaux en matière de similitude avec les humains sont suffisants pour amener les utilisateurs à afficher des comportements généralement réservés aux interactions humaines (Krämer et al., 2013). Ce postulat est toutefois de plus en plus rediscuté notamment suite 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 des 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 en jeux vidéo. Les développeurs et les créateurs de jeux vidéo ont été parmi les premiers à se demander si le perfectionnement du réalisme des personnages était l'une des meilleures stratégies pour améliorer l'expérience des joueurs (Ben, 2005). Ivory et Kalyanaraman (2007) ont montré que les participants qui ont joué à la version très réaliste du jeu (représentation des personnages et des sons) ont affiché des évaluations plus positives en termes de présence, d'engagement et d'éveil. Récemment, une attention croissante a été aussi accordée par chercheurs et les concepteurs 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 réalistes. Il a été démontré que la conversation avec un agent virtuel réaliste est perçue comme plus naturelle et motivante que celle avec un agent virtuel moins ou non réaliste comme une image statique (McBreen et Jack, 2001). Un agent réaliste entraînerait plus de satisfaction et contribuerait à une meilleure perception de la présence sociale (Van Vugt et al., 2007 ; Von Der Pütten et al., 2010). Plus spécifiquement, il a été démontré qu'un agent discutant oralement engage un sentiment de confiance plus fort qu'un agent s'exprimant uniquement avec du texte (Qiu et Benbasat, 2005). L'implémentation d'une voix dans la conception d'un agent virtuel constituerait un facteur de maximisation du réalisme des agents virtuels 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 possible de reconnaître les émotions d'un individu, sa personnalité et ses caractéristiques sociodémographiques à travers les tonalités, la vitesse et l'intonation (Scherer, 1986, 2003). Dans le cadre des interactions électronique, les langages verbaux les plus utilisés sont la messagerie instantanée et le « chatting » (Galanxhi et Nah, 2007 ; Qiu et Benbasat, 2005). Or, de nombreux travaux ont montré 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) ont montré 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 utilisateurs réagissent aux voix générées par ordinateurs, avec les réponses sociales habituellement réservées aux interactions entre humains (Lee et al., 2000 ; Lee et Nass, 2003 ; Nass et Lee, 2001). A titre d'exemple, ils considèrent différentes identités de voix provenant d'un même ordinateur comme si c'était plusieurs acteurs sociaux distincts et considèrent plusieurs ordinateurs ayant la même identité vocale comme si c'était un seul et même acteur (Nass et Steuer, 1993). 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, de plus en plus réalistes. La voix de synthèse est une technologie qui, parmi ses nombreuses utilisations, permet une communication plus efficace entre 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 GPSs, les jeux vidéo et les conseillers virtuels.

Ce type de voix 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 souffre encore de quelques difficultés liées à l'expressivité. Les discontinuités dans le ton, le tempo et l'intensité ne permettent pas de transmettre pleinement les émotions, les intentions et les attitudes à travers la modulation de la voix. Ainsi, elles sont généralement perçues comme non naturelles par les utilisateurs. La voix humaine préenregistrée permet de pallier à ce problème puisque le contenu verbal est enregistré au préalable par une personne pouvant très naturellement transmettre les émotions, les attitudes et les personnalités. 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 plus 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 ses limites, certaines recherches montrent qu'une voix humaine (préenregistrée) est souvent jugée comme plus persuasive qu'une voix de synthèse. Celles-ci sont généralement moins appréciées car elles sont perçues comme naturelles (Stern et al., 1999). Dans une volonté de comprendre d'une part l'effet du type de la voix (voix humaine préenregistrée vs une voix de synthèse produite artificiellement, et d'une autre part l'effet de l'origine de la voix communiquée aux participants (humaine vs ordinateur) sur leur perception, Stern et al. (2006) arrivent aux 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.

Egalement, Nass et Scott (2005) ont comparé l'effet d'une voix humaine vs une voix de synthèse émanant soit d'un visage humain ou d'un visage synthétique. Les auteurs ont montré que les gens divulguent et communiquent plus d'informations personnelles avec un agent présentant un visage clairement synthétique et une voix de synthèse ou un agent présentant un visage humain jumelé avec une voix humaine qu'avec un agent présentant un visage et une voix qui ne sont pas cohérents (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 grossier, moins fiable et étranger. 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. Mori (1970) souligne qu'au-delà d'un certain niveau de perfection dans la ressemblance humaine, les robots sont mieux acceptés. Il y a donc une zone que les concepteurs doivent franchir (appelée la vallée dérangeante) dans laquelle chaque avancée vers l'imitation humaine est synonyme de rejet (Goudey et Bonnin, 2016). 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 ». Dans la lignée de ces travaux, nous retrouvons l'expérimentation de Qiu et Benbasat (2009) qui 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, linéaire et complémentaire. 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 antérieurs de Qiu et Benbasat (2009) montrent aussi 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 de la voix du conseiller virtuel et de ses types sur les réactions du consommateur semble intéressant. Le développement des recherches en ce sens est d'autant plus nécessaire dans un contexte d'e-marketing.

Les hypothèses de la recherche

La voix compte parmi les quatre principales⁶ caractéristiques pouvant renforcer le caractère social de l'interaction en ligne (Nass et Steuer, 1993) notamment en matière de présence sociale (Lee et Nass, 2003). Gefen et Straub (2003) définissent la présence sociale sur un site comme « la mesure avec laquelle un intermédiaire virtuel permet aux utilisateurs de sentir que les autres sont psychologiquement présents ». Pour de nombreux chercheurs (Ben Mimoun et Poncin, 2011 ; Holzwarth et al. 2006 ; Kohler et al., 2011 ; Qiu et Benbasat, 2009 ; Wang et al., 2007), l'utilisation d'un conseiller virtuel est capable de simuler une présence sociale sur internet, d'humaniser l'interaction entre le client et le site et de rendre l'expérience de magasinage plus interpersonnelle. Cet article s'intéresse donc à l'analyse de l'influence du type de la voix sur la présence sociale perçue. Il permet d'approfondir la littérature relative à la relation « conseiller virtuel – présence sociale » et propose un nouvel axe de recherche en analysant l'effet d'une voix humaine versus une voix de synthèse. Les systèmes de synthèse vocale produisent aujourd'hui des voix de synthèse parfaitement intelligibles mais qui demeurent encore éloignées de la voix humaine naturelle (Beutnagel et al., 1999 ; Vaudale, 2012). Les voix de synthèse présentent encore quelques lacunes⁷ principalement liées à la qualité et à l'expressivité de la voix⁸, (D'Alessandro, 2001, 2006 ; Qiu et Benbasat, 2005). Par conséquent, les voix de synthèse sont encore perfectibles et pourraient limiter le caractère social de l'interaction. Nous supposons ainsi que :

- *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.

De multiples études se sont intéressées au concept de la confiance en ligne notamment suite au développement du commerce électronique au cours des dernières années. Ainsi, plusieurs chercheurs en interaction homme – machine ont essayé de comprendre le sentiment de confiance au niveau des interfaces homme-machine et tout particulièrement face aux agents virtuels (Komiak et Benbasat, 2004 ; Muir et Moray, 1996). Il ressort de ces investigations qu'il n'y a pas de différence significative entre les composantes de la confiance envers un individu et celles envers un agent virtuel (Cassell et Bickmore, 2000 ; Jian et al., 2000). Par ailleurs, Bos et al. (2002) ont comparé l'effet de quatre situations de communication sur le sentiment de confiance. Les auteurs ont montré que la communication en face à face est largement plus efficace que la communication textuelle. Il a été aussi démontré que la communication moyennant des supports médiatisés socialement riches (audio ou vidéo) affecte autant la confiance que la communication en face à face. Ainsi, en comparant l'effet des messages textes et un agent virtuel discutant avec une voix de synthèse, Qiu et Benbasat (2005) montrent que la voix de synthèse augmente la confiance envers l'agent virtuel. Par ailleurs, la confiance peut être affectée dans un contexte médiatisé par un ordinateur par les caractéristiques du discours, le rythme du discours et la présence d'erreurs verbales ou des hésitations qui se dégagent habituellement de la voix humaine (Stern et al., 2006). Nous pensons donc que :

- *H1b : La confiance envers le conseiller virtuel 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.*

Par ailleurs, plusieurs études ont démontré que la présence des conseillers virtuels est en mesure de renforcer le sentiment de confiance envers le site (Qiu et Benbasat, 2009 ; Wang et Benbasat, 2005, 2008 ; Wang et al., 2007). En effet, la personnification 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). Il a été démontré que des patients interagissant avec un agent virtuel relationnel, se sentent en confiance pour discuter avec lui de leurs problèmes de santé (Bickmore et Cassell, 2001 ; Cassell et Bickmore, 2000) et pour entretenir une relation d'échange avec lui sur la durée. Ces résultats nous laissent supposer que :

- *H1c : 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.*

Les hypothèses (H1b) et (H1c) ont donc pour objectifs de corroborer la littérature portant sur les conseillers virtuels et la confiance puis de mettre en évidence l'effet du type de voix du conseiller virtuel sur la confiance envers le conseiller virtuel et la confiance envers le site.

De nombreuses recherches font ressortir le lien positif entre la présence des conseillers virtuels et les intentions comportementales (Holzwarth et al., 2006 ; Keeling et al., 2010 ; Wang et al., 2007). Cette recherche prend en considération non seulement la présence d'un conseiller virtuel mais aussi l'effet du type de voix de celui-ci. A l'instar des travaux de Lee et Nass (2003) qui démontrent l'effet positif de la voix sur ordinateur sur les intentions comportementales, nous nous attendons ainsi à ce qu'un conseiller virtuel s'exprimant avec une voix humaine engendre des meilleures intentions comportementales par rapport à un conseiller s'exprimant avec une voix de synthèse d'où l'hypothèse H1d.

- *H1d : 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.*

Cette recherche tend également à compléter les travaux sur le rôle joué par le conseiller virtuel dans le développement du sentiment de confiance à la fois envers le conseiller virtuel lui-même ainsi qu'envers le site. Il s'agit ainsi d'étudier l'effet médiateur simple et multiple en série de la confiance envers l'agent virtuel respectivement dans les relations « voix du conseiller virtuel - intentions comportementales » et « voix du conseiller virtuel - confiance envers le site - intentions comportementales » d'où notre hypothèse H1e pour la médiation multiple en série:

- *H1e : L'effet de la voix du conseiller virtuel sur les intentions comportementales est médiatisé par la confiance envers le conseiller puis la confiance envers le site.*

Bien que beaucoup de recherches aient étudié l'impact de la présence du conseiller virtuel sur la confiance envers le site (Keeling et al., 2010), très peu se sont intéressées à la confiance envers les conseillers virtuels (à l'exception de Qiu et Benbasat, 2005 et de Wang et Benbasat, 2005, 2008). Or, si la confiance vis-à-vis du site internet peut dépendre de la présence du conseiller virtuel, la confiance envers du conseiller peut découler en amont. Ainsi, nous supposons qu'elle peut influencer la confiance envers le site ainsi que les intentions comportementales. Les hypothèses H2 et H3 relatives respectivement aux effets de la confiance envers le conseiller virtuel sur la confiance envers le site et les intentions comportementales sont ainsi formulées afin d'analyser leurs rôles dans un contexte électronique français.

- *H2 : La confiance envers le conseiller virtuel influence positivement la confiance envers le site.*

- *H3 : La confiance envers le conseiller virtuel influence positivement les intentions comportementales.*

Les hypothèses (H4) et (H5) s'intéressent à l'analyse du rôle de la présence sociale sur la confiance envers le conseiller virtuel et les intentions comportementales. Keeling et al. (2010) se sont intéressés à l'étude du rôle de l'orientation sociale du conseiller virtuel sur la confiance envers le site. Par ailleurs, Qiu et Banbasat (2009) mettent en évidence le rôle de la présence sociale dans le développement de la confiance envers le conseiller virtuel. Nous postulons, ainsi, que la présence sociale perçue peut impacter positivement la confiance envers le conseiller virtuel.

- *H4 : La présence sociale perçue influence positivement la confiance envers le conseiller virtuel.*

Parmi les recherches portant sur les intentions comportementales, les recherches sont mitigées. Ben Mimoun et Poncin (2011) n'ont pas réussi à confirmer l'effet significatif de la présence sociale sur les intentions comportementales alors que Wang et al. (2007) confirment le rôle significatif de la sociabilité perçue du site sur les intentions d'achat. Nos hypothèses sont ainsi posées afin de vérifier ces relations :

- *H5 : La présence sociale perçue influence positivement les intentions comportementales.*

Concernant l'hypothèse H6, cette recherche propose d'analyser également l'effet de médiation de la confiance envers le conseiller virtuel entre la relation « présence sociale – confiance envers le site ». La littérature existante montre que la présence sociale influence positivement la confiance envers le site (Lee et Nass, 2004). Par ailleurs, Qiu et Banbasat (2009) mettent en évidence l'effet de la présence sociale sur la confiance envers le conseiller virtuel. Nous supposons ainsi qu'il existe un effet médiateur de la confiance envers le conseiller entre la présence sociale et la confiance envers le site.

- *H6 : La confiance envers le conseiller virtuel médiatise l'influence de la présence sociale perçue sur la confiance envers le site.*

Enfin, l'hypothèse H7 pose l'influence de la confiance envers le site sur les intentions comportementales. La littérature existante met majoritairement en avant l'effet positif de cette relation. Lemoine et Notebaert (2011) montrent que la confiance envers le site avec un conseiller virtuel influence plus les intentions comportementales qu'un site sans conseiller virtuel. Egalement, Keeling et al. (2010) mettent en évidence le rôle positif de la confiance sur les intentions comportementales. A l'instar de ces travaux, nous souhaitons vérifier que la confiance envers le site engendre un effet positif sur les intentions comportementales, d'où notre hypothèse :

- H7 : La confiance envers le site influence positivement les intentions comportementales.

Toutes nos hypothèses sont schématisées dans la figure 1.

La méthodologie de la recherche

Echantillon et procédure expérimentale

Afin de tester la validité de nos hypothèses, nous avons mené une expérimentation 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)⁹.

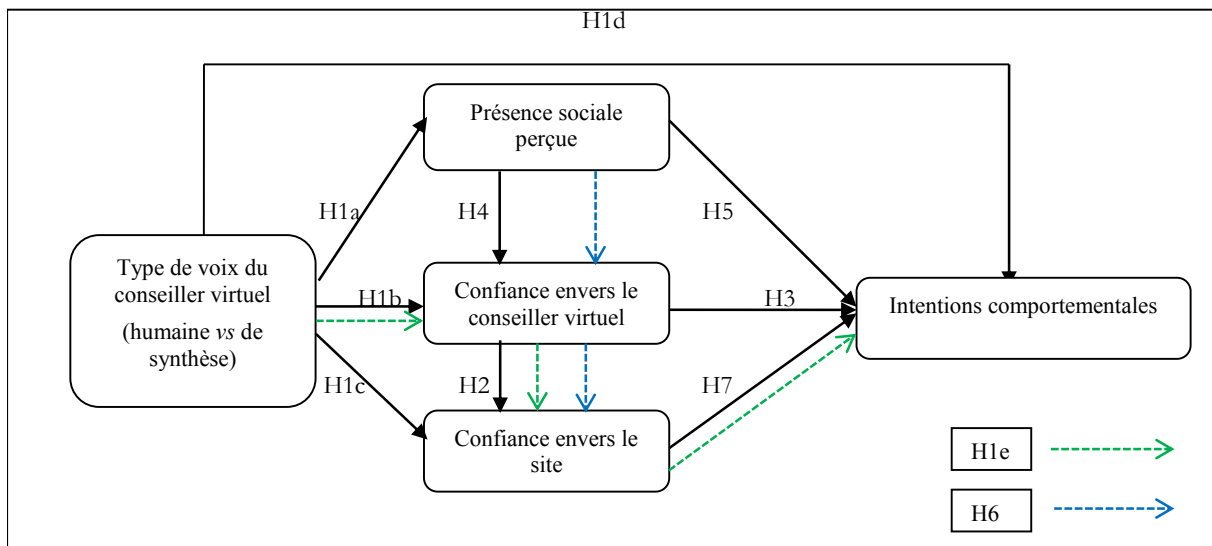


Figure 1. Le modèle de la recherche.

Pour composer notre échantillon et administrer notre questionnaire en ligne, nous avons eu recours au service de la société d'études Toluna QuickSurveys, partenaire de l'AFM. Chaque enquêté était affecté, de manière aléatoire, à une des deux versions du site (conseiller virtuel avec voix humaine et conseiller virtuel avec voix de synthèse). 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 capture d'écran en annexe 1). L'entreprise utilise le synthétiseur Elan Sayso pour implémenter la voix sur les conseillers virtuels. C'est un synthétiseur qui offre une bonne intelligibilité. La voix humaine implémentée sur le conseiller virtuel était préenregistrée par un acteur. Le choix de ce conseiller virtuel résulte d'une étude qualitative menée par la GMF auprès de ses clients et des responsables de la communication de cette structure. Ils ont trouvé

que ce conseiller est jugé comme le plus cohérent et congruent avec l'entreprise. Le choix de la voix a été également choisi de la même manière. Avant de répondre à notre questionnaire, les répondants étaient invités à simuler une recherche d'informations sur des produits d'assurance. Grâce à deux questions filtres, seuls les internautes nous ayant confirmé qu'ils avaient bien entendu le conseiller ont été retenus dans notre recherche (320 pour la version avec la voix humaine et 320 pour la version avec la voix de synthèse).

Les vérifications expérimentales ont été effectuées moyennant l'échelle de la perception de la voix de Stern et al. (2006). Nous avons confirmé une différence de perception entre les deux types de voix (Annexe 2).

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). Il est à noter que la version anglaise de l'échelle comporte cinq items. A l'instar de Bressolles et al. (2011), nous avons utilisé la version épurée de cette échelle qui se compose de quatre items¹⁰. Afin de mesurer la confiance envers le conseiller virtuel, nous avons mobilisé aussi l'échelle de Qiu et Benbasat (2009). C'est 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 de ces construits à 5 et à 3 items. Toutes les échelles sont présentées en annexe 3. Elles ont toutes pris la forme d'une échelle de Likert à 7 points.

A l'issue de notre analyse factorielle exploratoire, les résultats relatifs à la structure unidimensionnelle de nos échelles et à leur fiabilité sont les suivants : présence sociale ($\alpha=0,936$; 83,865% de la variance), confiance envers le conseiller virtuel ($\alpha=0,955$; 71,246% de la variance), confiance envers le site ($\alpha=0,911$; 74,025% de la variance) et intentions comportementales ($\alpha=0,638$; 59,377% de la variance).

Les résultats de l'étude

Afin de tester nos quatre premières hypothèses (H1a, H1b, H1c, H1d) relatives à l'influence du type de voix du conseiller virtuel sur la présence sociale perçue sur le site, la confiance envers le conseiller virtuel, la confiance envers le site et les intentions comportementales, 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). La présence d'un conseiller virtuel s'exprimant avec une voix humaine entraîne un sentiment de

présence sociale plus fort ($m = 3,77$) qu'un conseiller virtuel s'exprimant avec une voix de synthèse ($m = 3,27$). Nous confirmons donc l'hypothèse H1a. Par ailleurs, le type de voix a un effet significatif sur la confiance envers le conseiller virtuel. Dans le cas d'une voix humaine, la confiance envers le conseiller virtuel est plus forte ($m = 4,48$) que dans le cas d'une voix de synthèse ($m = 4,28$). H1b est aussi 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 ($p = 0,402 > 0,1$). Nous rejetons donc l'hypothèse H1c. Enfin, nos 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 ($m = 4,40$) que lorsqu'il parle avec une voix de synthèse ($m = 4,16$). L'hypothèse H1d est validée.

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

	Moyennes		Test de Levene		Test t de comparaison de moyenne	
	Voix de synthèse	Voix humaine	F	Sig.	t	Sig.
La présence sociale	3,27	3,77	0,435	0,510	4,715	,000*
La confiance envers le conseiller virtuel	4,28	4,48	0,090	0,764	2,373	,018**
La confiance envers le site	4,30	4,37	0,889	0,346	0,839	,402
Les intentions comportementales	4,16	4,40	0,077	0,781	2,863	,004**

* $p < 0,000$; ** $p < 0,05$

Suite à l'étude des effets directs du type de voix (H1a à H1d), nous nous sommes demandé s'il n'était pas possible d'envisager un effet médiateur simple et multiple en série de la confiance envers le conseiller virtuel et la confiance envers le site entre le type de voix et les intentions comportementales. L'hypothèse de médiation simple et médiation multiple en série est donc analysée par le test de bootstrap de Preacher et Hayes (2008)¹¹. Les résultats ci-dessous sont interprétés suivant les préconisations de Zhao et al. (2011). Tel que renseigné sur le tableau 2, le coefficient « a1 » correspond à l'effet de la voix sur la confiance envers l'agent virtuel. Le coefficient « a2 » est relatif à la relation entre la voix et la confiance envers le site. Enfin, le coefficient « a3 » est lié à l'effet de la confiance envers l'agent virtuel sur la confiance envers le site. Les coefficients « b1 » et « b2 » sont respectivement liés à l'impact de la confiance envers l'agent virtuel sur les intentions comportementales et la confiance envers le site sur les intentions comportementales. L'examen du tableau 2 montre que l'effet total et l'effet direct sont significatifs. Toutefois, c'est l'observation des effets indirects « a1 x b1 », « a1 x a3 x b2 » et « a2 x

b2 » qui a plus d'intérêt pour nous. La significativité du coefficient « a1 x b1 » nous renseigne sur l'existence d'une médiation complémentaire « a1 x b1 x c>0 » 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 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.

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

Effet total			Effet direct			Effet indirect				Résultat	
c'	t	p	c	t	p	Coefficients	Valeurs	P	IC		
-,2469	-2,8634	,0043**	-,1570	-2,4414	,0149**	a1	-,2003	,0179**			
						a2	,0812	,1281			
						a3	,7515	,000*			
						b1	,2953	,000*			
						b2	,4429	,000*			
						a1 x b1	-,0592		[-,121 ; -,013]	Médiation	simple
										complémentaire	
						a1 x a3 x b2	-,0722		[-,129 ; -,015]	Médiation	multiple en
										série complémentaire	
						a2 x b2	,0324		[-,008 ; ,084]	Médiation	simple
										seulement indirecte (non-médiation)	

*p< 0,000 ; **p< 0,05 ; Codage de la variable *muette* : voix de synthèse (1) et voix humaine (0)

La non significativité du coefficient « a2 x b2 » ne nous permet pas de valider l'effet médiateur simple de la confiance envers le site pour l'influence de la voix sur les intentions comportementales. Nous concluons ainsi à une médiation seulement indirecte (non-médiation) puisque l'effet direct est significatif. Ainsi, encore une fois, il est possible d'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 la confiance envers le site. Le coefficient « a1 x a3 x 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 confiance dans le conseiller virtuel résultant d'une voix humaine est 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 (a1 x a3 x b2 x c>0) suggérant l'existence d'autres médiateurs tels que la présence sociale. Nous confirmons ainsi notre hypothèse H1e. La suite des hypothèses (H2, H3, H4, H5, H7) a été testée à l'aide des équations structurelles sous le logiciel AMOS. Nous avons, en premier lieu, procédé à une analyse

factorielle confirmatoire afin d'évaluer la fiabilité et les validités convergente et discriminante de nos échelles de mesure. Les principaux résultats sont consignés dans l'annexe 3 et 4.

La fiabilité mesurée par le ρ de Jöreskog est satisfaisante au seuil de 0,90 pour toutes les échelles (sauf pour les intentions comportementales = 0,684). Par ailleurs, toutes les contributions factorielles sont significatives. La validité convergente des échelles est appréciée par le ρ de validité convergente (ρ_{vc}). Les valeurs étant supérieures à 0,50 pour nos échelles (très proche de 0,50 pour les intentions comportementales), nous pouvons donc confirmer leur validité convergente. En ce qui concerne les indices d'ajustement du modèle, ils sont satisfaisants (χ^2 (df) = 3,369 ; GFI = 0,893 ; CFI = 0,952 ; TLI = 0,945 et RMSEA = 0,067). Enfin, pour ce qui de la validité discriminante, nous avons examiné la matrice des corrélations entre les construits présentée en annexe 4. Nous remarquons que le carré des corrélations de chaque échelle avec tous les autres construits est inférieur au ρ_{vc} (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 nos échelles¹². En ce qui concerne nos résultats, comme nous pouvons le voir avec la figure 2, toutes les relations entre les variables de notre modèle sont significatives au seuil de 1%.

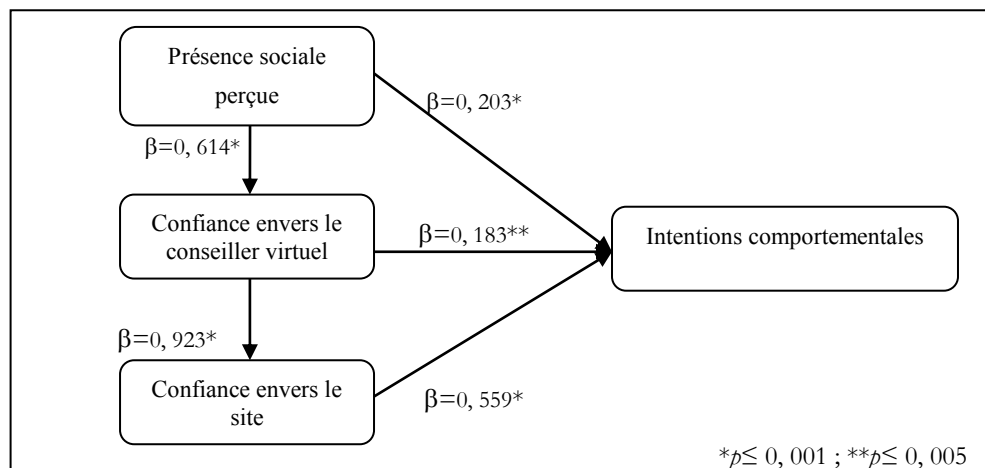


Figure 2. Les relations du modèle structurel

La présence sociale influence positivement la confiance envers le conseiller virtuel ($\beta = 0,614$; $t = 14,779$; $p = 0,000$) et les intentions comportementales ($\beta = 0,203$; $t = 5,223$; $p = 0,000$).

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 ($\beta = 0,183$; $t = 2,829$; $p = 0,005$). Enfin, la confiance envers le site influence positivement les intentions

comportementales ($\beta = 0,559$; $t = 9,187$; $p = 0,000$). Nous pouvons donc valider nos hypothèses H2, H3, H4, H5 et H7.

Par ailleurs, en appliquant la démarche de Zhao et al. (2011), nous sommes en mesure 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), nous pouvons donc conclure à une médiation complémentaire suggérant l'existence d'autres médiateurs dans le modèle théorique. En ce sens, nous validons notre hypothèse H6.

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	$a \times b$	IC	Médiation
,387	14,930	,000	,068	2,890	,0004	,4574	,696	,318	[0,260 ; 0,382]	complémentaire

Afin de résumer la présentation des résultats, le tableau 4 présente la synthèse relative à la validation ou au rejet de nos hypothèses de recherche.

Tableau 4. Synthèse des résultats de l'étude empirique

Hypothèses de la recherche : V = validée, R = rejetée		
H1a	Voix humaine/voix de synthèse – Présence sociale perçue	V
H1b	Voix humaine/voix de synthèse – Confiance envers le conseiller virtuel	V
H1c	Voix humaine/voix de synthèse – Confiance envers le site	R
H1d	Voix humaine/voix de synthèse – Intentions comportementales	V
H1e	Voix humaine/voix de synthèse – Confiance envers le conseiller virtuel – Confiance envers le site – Intentions comportementales	V
H2	Confiance envers le conseiller virtuel – Confiance envers le site	V
H3	Confiance envers le conseiller virtuel – Intentions comportementales	V
H4	Présence sociale perçue – Confiance envers le conseiller virtuel	V
H5	Présence sociale perçue – Intentions comportementales	V
H6	Présence sociale perçue – Confiance envers le conseiller virtuel – Confiance envers le site	V
H7	Confiance envers le site – Intentions comportementales	V

Apports, implications, limites et voies de recherche

L'importante croissance accordée à l'anthropomorphisme des conseillers virtuels nous a amené à étudier l'effet du type de voix du conseiller virtuel sur les réactions des consommateurs. En effet, la première série des hypothèses de la recherche met en exergue l'influence du type de voix (humaine vs de synthèse) sur la présence sociale, la confiance envers le conseiller, la confiance envers le site et les intentions comportementales. Les résultats montrent tout d'abord que les deux types de voix sont en mesure d'influencer positivement la présence sociale, la confiance envers le conseiller virtuel et les intentions comportementales. Ces résultats convergent avec les travaux de Lee et Nass (2003, 2004) selon lesquels la voix de synthèse sur un ordinateur est capable d'accroître la présence sociale et de renforcer la crédibilité. De plus, au-delà de

l'interaction avec l'ordinateur, cette recherche a confirmé l'effet positif de la voix dans une interaction avec un conseiller virtuel dans un cadre marketing. Ces conclusions sont donc conformes à la théorie de la réponse sociale qui stipule qu'un faible anthropomorphisme, en l'occurrence une faible ressemblance avec l'humain est capable de renforcer le caractère social de l'interaction et de générer des réponses positives (Krämer et al., 2013 ; Moon, 2000 ; Reeves et Nass, 1996).

Néanmoins, les résultats font ressortir que les effets sont nettement meilleurs dans le cas d'une voix humaine que dans le cas d'une voix synthétique. Nous pouvons ainsi déduire que la théorie de la maximisation du réalisme n'est que le prolongement logique de la théorie de la réponse sociale (Groom et al., 2009 ; McBreen et Jack, 2001). En effet, la présence sociale perçue ainsi que la confiance envers le conseiller virtuel 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 (H1a et H1b). Ces résultats apportent donc l'assurance aux concepteurs des conseillers virtuels sur la primauté de la voix humaine enregistrée sur la voix de synthèse.

Nous confirmons par ailleurs les précédents travaux sur la capacité des conseillers virtuels à simuler une présence sociale sur internet, d'humaniser l'interaction entre le client et le site et de rendre l'expérience de magasinage plus interpersonnelle (Ben Mimoun et Poncin, 2011 ; Holzwarth et al., 2006 ; Kohler et al., 2011 ; Qiu et Benbasat, 2009 ; Wang et al., 2007).

Par ailleurs, les résultats mettent en avant l'effet significatif de la voix sur les intentions comportementales (H1d). Le conseiller virtuel doté d'une voix humaine suscite davantage les intentions comportementales que le conseiller virtuel avec une voix de synthèse. 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 ; Wang et al., 2007) mais très peu de recherches ont étudié l'effet spécifique de la voix du conseiller sur les intentions comportementales. Ces résultats sont donc en accord avec les conclusions de Lee et Nass (2003) et Qiu et Benbasat (2009) qui démontrent l'effet positif de l'utilisation de la voix sur les intentions comportementales. De plus, cette recherche permet de confirmer l'influence positive de la voix dans un contexte d'interaction avec un conseiller virtuel. Elle permet également de constater que la voix humaine a plus d'effets sur les intentions comportementales que la voix de synthèse.

Concernant l'effet de la voix sur la confiance envers le site, les résultats ne montrent pas une influence significative (H1c). Ainsi, le type de voix du conseiller virtuel n'a aucun effet direct sur la confiance éprouvée envers le site. Ces résultats contredisent les travaux montrant l'influence positive de la présence des conseillers virtuels sur le sentiment de confiance envers le site (Qiu et Benbasat, 2009 ; Wang et Benbasat, 2005, 2008 ; Wang et al., 2007). Toutefois, lors de la

formulation des hypothèses de la recherche, il était question de savoir s'il existe un effet médiateur multiple en série de la confiance envers le conseiller virtuel puis la confiance envers le site dans la relation « type de voix – intentions comportementales » (H1e). Les résultats du tableau 2 permettent de constater qu'il existe un effet sur la confiance envers le site par le biais de la confiance envers le conseiller virtuel (coefficient a_3). Plus la confiance envers le conseiller virtuel est forte, plus l'internaute aura confiance dans le site. Les résultats confirment aussi l'hypothèse de médiation multiple en série de la confiance envers le conseiller virtuel puis la confiance envers le site dans la relation « type de voix – intentions comportementales ». Ainsi, un fort sentiment de confiance dans le conseiller virtuel résultant d'une voix humaine est traduit par un fort sentiment de confiance dans le site qui à son tour influencera positivement les intentions comportementales. La mise à jour de cet effet médiateur constitue l'une des contributions théoriques de ce papier. Par ailleurs, la médiation complémentaire ($a_1 \times a_3 \times b_2 \times c > 0$) suggère l'existence d'autres médiateurs comme la présence sociale. En effet, Lee et Nass (2004) ont montré que la présence sociale peut agir comme une variable médiatrice dans la relation « voix – confiance envers le site ». Les résultats avancent aussi l'existence d'une relation médiatisée par la confiance envers le conseiller virtuel entre la présence sociale et la confiance envers le site (H6). Les conclusions mettent également en évidence qu'il s'agit d'une médiation complémentaire permettant de prendre en compte d'autres médiateurs.

Au sujet des hypothèses relatives au 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 (H2) ainsi que sur les intentions comportementales (H3). 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 dans le site et ses intentions comportementales seront fortes. Très peu d'études se sont focalisées sur l'étude de la confiance envers les conseillers virtuels (Qiu et Benbasat, 2005 ; Wang et Benbasat, 2005, 2008). Ce travail de recherche fait en effet ressortir le rôle de la confiance envers le conseiller virtuel dans la formation de la confiance envers le site.

Par ailleurs, dans le cadre de cette recherche, nous confirmons aussi l'influence positive de la présence sociale perçue sur le site sur la confiance envers le conseiller virtuel et les intentions comportementales (H4 et H5). Ces résultats sont dans la lignée des travaux de Qiu et Banbasat (2009) qui mettent en évidence le rôle de la présence sociale dans le développement de la confiance envers le conseiller virtuel et les intentions d'achat (Wang et al., 2007)

Enfin, les résultats montrent un effet significatif de la confiance envers le site sur les intentions comportementales. Ces résultats sont en accord avec les conclusions de Keeling et al. (2010)

montrant que la confiance envers le site avec un conseiller virtuel influence plus les intentions comportementales qu'un site sans conseiller virtuel.

D'une manière générale, nos résultats corroborent les recherches antérieures menées en robotique, en informatique et en systèmes d'information qui démontrent l'importance des facteurs sociaux sur internet et, en particulier, l'influence de la présence d'un conseiller virtuel sur la présence sociale, la confiance et les intentions comportementales (Holzwarth et al., 2006 ; Keeling et al., 2010 ; Qiu et Benbasat, 2009 ; Wang et al., 2007). Dans cette recherche, il était question d'étudier l'effet de la présence d'un conseiller virtuel avec une voix humaine vs un conseiller virtuel avec une voix de synthèse sur les réactions des consommateurs. Au-delà des résultats déjà validés dans les travaux antérieurs, les résultats de cette expérimentation, ont permis de confirmer, d'une part, l'effet positif de la voix du conseiller virtuel dans le cadre d'une interaction avec ce dernier et non seulement dans une interaction avec un ordinateur. D'autre part, cette recherche a permis de montrer que l'interaction avec un conseiller virtuel interagissant avec une voix humaine entraîne de meilleures réactions comportementales par rapport à une interaction avec un conseiller virtuel interagissant avec une voix de synthèse. Ce résultat confirme l'intérêt de la maximisation du réalisme des conseillers virtuels dans le but d'offrir aux internautes une meilleure expérience en ligne. De plus, nous avons vérifié le rôle de la confiance envers le conseiller virtuel puis la confiance envers le site à médier l'effet de la voix du conseiller virtuel sur les intentions comportementales. La médiation multiple en série est complémentaire suggérant l'existence d'autres médiateurs dans cette relation.

D'un point de vue managérial, des préconisations concernant la conception des conseillers virtuels sont à prendre en considération. Les résultats de l'expérimentation mise en œuvre confirment l'importance du type de voix à produire lors de l'élaboration des conseillers virtuels. Les concepteurs de sites web qui cherchent à renforcer la perception de la présence sociale, la confiance et les intentions comportementales ont intérêt à réfléchir à la meilleure façon de concevoir leurs conseillers virtuels. En effet, les consommateurs s'attendent aujourd'hui à ce qu'ils soient performants et capables d'offrir une interaction semblable à l'interaction entre humains (Viot et Bressolles, 2012). Doter ces conseillers d'une voix humaine, par opposition à une voix de synthèse, apparaît donc comme un moyen pertinent pour maximiser l'humanisation perçue de l'interface et pour impacter positivement les intentions comportementales des internautes.

Notre recherche comporte des limites qui peuvent constituer des pistes de recherche pour l'avenir. Tout d'abord, cette étude présente une faible validité externe 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 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 d'un site web qui ont le plus grand impact sur les réactions des internautes. Enfin, l'intégration de variables modératrices dans notre modèle de recherche (familiarité avec internet et avec les conseillers virtuels) pourrait 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.

Notes

¹ Plusieurs termes considérés comme synonymes sont employés dans la littérature en marketing et en systèmes d'information pour désigner les conseillers virtuels. Nous retrouvons les termes : avatars (Holzwarth et al., 2006), agent conversationnel incarné (Cassell et Bickmore, 2000), agent virtuel, agent synthétique, personnage interactif, agent pédagogique animé, agent de shopping, agent animé (McGoldrick et al., 2008).

² Duffy (2003) définit l'anthropomorphisme comme « la tendance à attribuer des caractéristiques humaines aux objets non animés dans le but de rationaliser leur actions ».

³ Exemples de synthétiseurs : ATet T, Elan Sayso Speech, DecTalk Express

⁴ L'expressivité de la voix est définie comme « la possibilité d'exprimer et de comprendre les émotions, les intentions et les attitudes à travers la modulation de la voix » (Beller et Rodet, 2007).

⁵ La politesse (Tzeng, 2006 ; Nass et al., 1999), les échanges mutuels (Moon, 2000), les stéréotypes liés au genre (Nass et al., 1997), la flatterie (Lee, 2008 ; Fogg et Nass, 1997) et l'attribution de personnalité (Moon et Nass, 1996).

⁶ Avec le langage, l'interactivité et le rôle social.

⁷ Des pauses et un accent non naturels, des discontinuités entre les phonèmes et les syllabes (Nass et Lee, 2001).

⁸ L'expressivité de la voix est « la possibilité d'exprimer et de comprendre les émotions, les

intentions et les attitudes à travers la modulation de la voix » (Beller et Rodet, 2007).

⁹ Cette composition socio-démographique de notre échantillon se retrouve au niveau des deux sous-échantillons constitués pour les besoins de l'expérimentation.

¹⁰ Suite à notre analyse factorielle exploratoire, nous avons retrouvé les 4 items que Bressolles et al. (2011) ont retenus dans leur recherche.

¹¹ Un bootstrap à 5 000 et 95% pour l'intervalle de confiance.

¹² Il est à noter que la validité discriminante entre la confiance envers le site et les intentions comportementales est validée suivant la condition de Kline (1998).

Références bibliographiques

André E et Pelachaud C (2008) Interacting with embodied conversational agents. Dans: Chen F et Jokinen K (coord) *New Trends in Speech Based Interactive Systems*. Springer.

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.

Ben Mimoun MS et Poncin I (2011) Anna peut-elle m'aider ? Apport des ACA à la productivité des e-consommateurs: les effets modérateurs des caractéristiques individuelles. Dans: *27^{ème} Congrès de l'Association Française du Marketing*, Bruxelles, Belgique, 18-20 mai 2011.

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.

Beller G et Rodet X (2007) Content-based transformation of the expressivity in speech. Dans: *16^{ème} Congrès ICPbS*, Saarbruecken, Allemagne, 8-10 août 2007, pp. 2157–2160.

Beutnagel M, Conkie A, Schroeter J, Stylianou Y et Syrdal A (1999) The AT&T next-gen TTS system. Dans: *Joint Meeting of ASA, EAA, and DAGA*, Berlin, Allemagne, mars 1999.

Bickmore T et Cassell J (2001) Relational Agents: A Model and Implementation of Building User Trust. Dans: *Congrès du SIGCHI Conference on Human Factors in Computing Systems CHI'01*, Seattle, WA, USA, 31 mars - 5 avril 2001, pp. 396-403.

Bos N, Olson JS, Gergle D, Olson GM et Wright Z (2002) Effects of four computer-mediated communications channels on trust development. Dans: *Congrès du SIGCHI Conference on Human Factors in Computing Systems CHI'02*, Minneapolis, MN, USA, 20-25 avril 2002, pp. 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. Dans: *27^{ème} Congrès de l'Association Française du Marketing*, Bruxelles, Belgique, 18-20 mai 2011.

- 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.
- Cassell J et Bickmore T (2000) External manifestations of trustworthiness in the interface. *Communications of the ACM* 43(12): 50-56.
- Charfi A et Volle P (2011) Valeur perçue et comportements en ligne en état d'immersion : le rôle modérateur de l'implication et de l'expertise. Dans: *27ème Congrès de l'Association Française du Marketing*, Bruxelles, Belgique, 18-20 mai 2011.
- 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.
- D'Alessandro C (2001) 33 ans de synthèse de la parole à partir du texte: une promenade sonore (1968-2001). *Traitement Automatique des Langues* 42(1): 1-29.
- D'Alessandro C (2006) Qu'est-ce que la synthèse vocale, *Technolanguage: le portail des technologies de la langue*. Disponible sur http://www.technolanguage.net/article.php?id_article=275 (consulté le 24 janvier 2017).
- 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
- Galaxhi 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 D (2003) Managing user trust in B2C e-services. *e-Service Journal* 2(2): 7-24.
- Greenspan R, Goldberg D, Weimer D et Basso A (2000) Interpersonal Trust and Common Ground in Electronically Mediated Communication. Dans: *The 2000 ACM Conference on Computer Supported Cooperative Work (CSCW'00)*
- 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.
- 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.
- 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.
- Jian J, Bisantz AM et Drury CG (2000) Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics* 4(1): 53-71.
- 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 SX et Benbasat I (2004) Understanding customer trust in agent-mediated electronic commerce, web-mediated electronic commerce, and traditional commerce. *Information Technology and Management* 5(1-2): 181-207.
- 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.

- 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. Dans: *Congrès du SIGCHI Conference on Human Factors in Computing Systems CHI'03*, Ft. Lauderdale, Floride, USA, 5-10 avril 2003, pp. 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.
- Lee E, Nass C et Scott S (2000) Can computer-generated speech have gender? An Experimental Test of Gender Stereotypes. Dans: *Congrès du SIGCHI Conference on Human Factors in Computing Systems CHI'00*, The Hague, Netherlands, 1-6 avril 2000.
- 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. Dans: *The 1997 SIGCHI Conference on Human Factors in Computing Systems*, New York, USA, pp. 359–366. ACM Press.
- 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.
- Mori M (1970) The uncanny valley. *Energy* 7(4): 33-35.
- Muir B M et Moray N (1996) Trust in automation: Part II. Experimental studies of trust and human intervention in a process control simulation. *Ergonomics* 39(3): 429-460.
- Nass C, Fogg BJ et Moon Y (1996) Can computers be teammates?. *International Journal of Human-Computer Studies* 45(6): 669-678.
- 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 and 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: 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. Dans: *4th International Conference on Universal Access in Human–Computer Interaction* (ed C Stephanidis), Beijing, China, pp. 5–10. Springer.

Preacher K J et Hayes A F (2008) Contemporary Approaches to Assessing Mediation in Communication Research. Dans: Hayes AF, Slater MD et Snyder LB (coord.) *The Sage Sourcebook of Advanced Data Analysis Methods for Communication Research*, Thousand Oaks, CA: Sage Publications, pp: 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 Publisher.

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. London: John Wiley & Sons
- 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*, Université Paris 11, Laboratoire LIMSI, France.
- 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 L C, 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 Cerccone 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 Application en Marketing* 26(1): 81-95.

Annexes

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



Annexe 2. Vérifications expérimentales

Nous avons utilisé l'échelle de la perception de la voix de Stern et al. (2006) pour vérifier la validité de nos manipulations expérimentales. Cette échelle comprend sept items auxquels nous avons ajouté l'item : voix artificielle ↔ voix réaliste. Après épuration, nous avons gardé à l'instar des auteurs 2 dimensions : le style de la voix et les caractéristiques de la voix (tableau 1).

Tableau 1. Résultats de l'analyse factorielle exploratoire de l'échelle de la perception de la voix

Items	Le style de la voix	Les caractéristiques de la voix
Une voix artificielle ↔ Voix réaliste	0, 817	-0, 303
Voix monotone ↔ Voix pleine de vie	0, 798	-0, 362
Voix profonde ↔ Voix aigue	0, 409	0, 706
Voix forte ↔ Voix douce	0, 484	0, 512
α de Cronbach	0, 570	

L'analyse de variance montre qu'il existe des différences de moyenne significatives (tableau 2) entre les deux conditions expérimentales, et ce pour les deux dimensions de l'échelle de la perception de la voix.

Tableau 1. Résultats de l'analyse de variance entre les deux groupes

	Le style de la voix			Les caractéristiques physiques de la voix		
	Moyenne	F	P	Moyenne	F	P
Voix humaine	4, 22	77, 812	0, 000	3, 88	9, 079	0, 003
Voix de synthèse	3, 23			3, 66		

Annexe 3. Echelles de mesure et résultats de l'analyse factorielle confirmatoire

Echelles	Items	λ_i	t- value	SMC	Fiabilité	gvc
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 un sentiment de sensibilité humaine dans le 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 mis mes intérêts en premier	0,800	21,017	0,640		
	Le conseiller virtuel est compétent dans ses recommandations et explication	0,859	22,786	0,738		
	Le conseiller virtuel a accompli son rôle de conseiller efficacement	0,861	22,847	0,742		
	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'avais besoin	0,819	21,579	0,671		
	Les recommandations effectuées par le conseiller virtuel étaient dans mon intérêt en priorités et non pas 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		
Confiance envers le site (Keeling et al., 2010)	Je crois aux informations données sur ce site	0,868		0,753	0,913	0,679
	Je ferais confiance au processus de paiement donné sur ce site	0,758	23,438	0,575		
	Je suis sûr que ma commande sera bien traitée	0,818	26,611	0,669		
	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		
Intentions comportementales (Keeling et al., 2010)	Je recommanderais le site à mes proches	0,878		0,771	0,684	0,454
	Je retournerai sur ce site pour chercher plus d'information	0,726	19,365	0,528		
	Je me rendrais en agence pour souscrire	0,254	6,095	0,065		

Note : Selon (Garson, 2007), l'analyse factorielle confirmatoire est acceptable pour une échelle comportant 3 items tel qu'il est le cas pour l'échelle des intentions comportementales. Ainsi, pour le premier item « IC3 », la valeur SMC est faible. Par ailleurs, la fiabilité et la validité convergente sont proches de 0,7 et 0,5.

Annexe 4. Matrice 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	0,611		
Confiance envers le site	0,679	0,535	0,822	
Intentions comportementales	0,454	0,610	0,765	0,815

Note : Les valeurs présentées correspondent aux corrélations de chaque échelle avec les autres construits et non le carré des corrélations.