Virtual human technologies for cognitively-impaired older adults' care : the LOUISE and Virtual Promenade experiments

Pierre Wargnier

THÈSE DE DOCTORAT

de l'Université de recherche Paris Sciences et Lettres
PSL Research University

Préparée à MINES ParisTech

Virtual human technologies for cognitively-impaired older adults' care: the LOUISE and Virtual Promenade experiments

Technologies d'humains virtuels dans le soin aux personnes âgées atteintes de troubles cognitifs : les expériences LOUISE et Virtual Promenade

École doctorale n°521

SCIENCES ET MÉTIERS DE L'INGÉNIEUR

Spécialité MATHÉMATIQUES, INFORMATIQUE TEMPS-RÉEL, ROBOTIQUE

Soutenue par **Pierre WARGNIER**
le 25 novembre 2016

Dirigée par **Pierre JOUVELOT et Anne-Sophie RIGAUD**

COMPOSITION DU JURY :

M Mohamed CHETOUANI
Université Pierre et Marie Curie,
Président

M Stefan KOPP
Universität Bielefeld,
Rapporteur

M Matthias RAUTERBERG
Technische Universiteit Eindhoven,
Rapporteur

Mme Anne-Sophie RIGAUD
Université Paris Descartes,
Membre du jury

M Pierre JOUVELOT
MINES ParisTech, PSL,
Membre du jury

M Samuel BENVENISTE
CEN STIMCO,
Membre du jury

Pierre WARGNIER

Virtual human technologies for cognitively-impaired older adults' care: the LOUISE and Virtual Promenade experiments

Technologies d'humains virtuels dans le soin aux personnes âgées atteintes de troubles cognitifs : les expériences LOUISE et Virtual Promenade

Thèse de doctorat de l'Université de recherche
Paris Sciences et Lettres,
préparée à MINES ParisTech
sous la direction de Pierre JOUVELOT
et d'Anne-Sophie RIGAUD.

Remerciements

Je souhaite en premier lieu remercier mes directeurs de thèse, Pierre Jouvelot et Anne-Sophie Rigaud, ainsi que mon co-encadrant, Samuel Benveniste, pour leur soutien, leurs conseils avisés, leur qualité d'écoute, leur disponibilité et pour les nombreux échanges extrêmement enrichissants que j'ai pu avoir avec chacun d'entre eux au cours de ces trois années.

Je remercie également le président du jury, Mohamed Chetouani, ainsi que mes rapporteurs Stefan Kopp et Matthias Rauterberg, de m'avoir fait l'honneur d'accepter de siéger dans mon jury.

J'adresse des remerciement chaleureux à tous mes collègues du Centre de recherche en informatique de MINES ParisTech, à commencer par son directeur, François Irigoin, pour leur accueil et pour leur aide. Par ordre alphabétique, merci à Corinne Ancourt, Fabien Coelho, Laurent Daverio, Karel De Vogeleer, Emilio Gallego Arias, Pierre Guillou, Florian Gouin, Olivier Hermant, Patryk Kiepas, Catherine Le Caër, Nelson Lossing, Vivien Maisonneuve, Claire Medrala, Bruno Massonni Sguerra, Benoît Pin, Ronan Saillard, Adila Susungi et Claude Tadonki.

Un grand merci aussi aux équipes de LUSAGE et du CEN STIMCO pour leur accueil chaleureux et amical et pour leurs contributions à ces travaux, et en particulier Maribel Pino, qui est à l'origine du projet LOUISE, et Frédéric Bloch et Serge Reingewirtz qui sont à l'origine du projet Promenade virtuelle, ainsi que Giovanni Carletti et Yann Laurent, qui ont contribué directement au projet LOUISE. Merci également à tout le reste de l'équipe du pavillon bleu de l'hôpital Broca, par ordre alphabétique, Victoria Cristancho-Lacroix, Souad Damnée, Manon Demange, Leila Djabelkhir, Grégory Legouverneur, Philippe De Oliveira Lopes, Hélène Kerhervé, Fabienne Marlats, Clotilde Monnet, Chloé Traoré, Caitlin Ware et Ya-Huei Wu.

Je remercie également les stagiaires que j'ai eu le plaisir d'encadrer et/ou qui ont contribué à ces travaux, soit, par ordre d'apparition : Julien Jacquemot, Adrien Malaisé, Michel Bala, Edmond Phuong, Paul-Emile Fauquet et Chenyang Wang.

Je souhaite aussi remercier Matthieu Plichard et Marie-Laure Seux, l'équipe des ré-éducateurs, l'équipe de l'hôpital de jour et l'équipe d'animation de l'hôpital Broca ainsi que les autres membres du personnel de l'hôpital –

médecins, internes, cadres de santé, infirmiers et aides-soignants – pour leur aide.

Mes remerciements vont également à tous les participants de nos études, dont les noms ne peuvent figurer ici, qui ont gracieusement donné de leur temps pour que la recherche puisse progresser. C'est grâce à ces hommes et ces femmes que, je l'espère, les personnes âgées handicapées et leur entourage pourront un jour bénéficier des fruits de ces travaux.

Je remercie le conseil régional et les contribuables de la région Île-de-France qui ont financé ces travaux.

Enfin, je remercie mes parents, Jean-Louis et Véronique Wargnier, sans qui rien de tout cela n'aurait été possible, et mes grand parents, Claudie et Gérard Dufourd, Françoise Wargnier-Khan, Catherine Wargnier, ma troisième grand-mère, et Claude Wargnier, qui nous a quittés trop tôt pour voir l'aboutissement de ces travaux. Je remercie également ma petite sœur, Noémie. Ils ont, chacun à sa manière, contribué à me faire devenir l'homme que je suis. Et, last but not least*, je remercie de tout mon cœur Clémentine Fellah pour tout le bonheur, la joie, la bonne humeur, le réconfort et le soutien qu'elle m'a apportés en partageant ma vie pendant la préparation de cette thèse.*

Contents

[1] Thanks to Anne-Sophie Rigaud for her help in writing this section.

[2]Thanks to Giovanni Carletti and Yann Laurent for their help in writing this section.

List of Figures

List of Tables

Introduction

Depuis le début du XXI^{ème} siècle, les progrès rapides de l'informatique et la large diffusion de l'internet ont provoqué des changements profonds et toujours en cours dans les modes de vie de la plupart des habitants des pays développés. En effet, la proportion d'habitants de ces pays ayant accès à l'internet est passé de 30% en 2000 à 80% en 2015 [94]. Cela n'est pas sans conséquences. En fait, la sur-stimulation des individus, produit d'un véritable bombardement d'information subi par les internautes, généralement à des fin publicitaires, est liée à une réduction de la durée d'attention. D'après une étude menée par Microsoft [133], la durée d'attention moyenne d'un internaute a chuté de 12 secondes en l'an 2000 à 8 secondes en 2015, soit moins que celle d'un poisson rouge, qui est estimée à 9 secondes !

Cette montée en puissance des technologies de l'information et de la communication a significativement transformé nos manière de vivre et de travailler. En particulier, cela a permis la création de mondes virtuels, dans lesquels des millions de personnes s'immergent quotidiennement, principalement pour y tuer zombies, monstres, démons, extra-terrestres, humains ou les autres joueurs à l'aide (respectivement) de battes de baseball, d'arcs et de flèches, d'épées magiques, de pistolets lasers, de fusils à pompe ou de fusils d'assaut. L'industrie du jeu vidéo, l'un des secteurs économiques à plus forte croissance durant la dernière décennie, ainsi que l'industrie cinématographique, ont été les moteurs d'innovations en matière d'imagerie et d'animation tridimensionnelles permettant de créer mondes et personnages virtuels toujours plus réalistes. Pour entrer dans ces mondes virtuels, visibles à travers des écrans de toutes tailles, comme autant de fenêtres sur des lieux imaginaires, les joueurs doivent se mettre dans la peau d'alter-ego virtuels, appelés avatars*, qu'ils contrôlent tels des pantins numériques.*

Il est intéressant de noter que tuer des zombies virtuels en contrôlant un avatar, armé d'un fusil à pompe, et sans se faire écarteler (l'avatar, pas le joueur), requiert d'être alerte et de maintenir son attention pour bien plus longtemps que 8 secondes. Mais est-ce le même type d'attention qui est nécessaire lorsque l'on joue à l'un de ces jeux vidéos immersifs que lorsqu'on lit des pages web ? En fait, jouer à un jeu de tir mobilise un certain nombre de fonctions cognitives de manière intense : cognition spatiale, coordination *sensori-motrice, attention divisée, etc (si vous n'êtes pas convaincu, essayez donc d'avoir une conversation avec un ou une adolescent(e) lorsqu'il ou elle joue à Call of Duty). Plusieurs études menées par des psychologues ont même montré que jouer à des jeux vidéos d'action conduit à des améliorations dans plusieurs domaines perceptuels et cognitifs, tels que la vision périphérique ou la capacité à alterner rapidement entre plusieurs tâches [79].*

À l'autre bout du spectre des consommateurs de média numériques se trouve Marie[3]. Marie a 99 ans et possède une tablette numérique qu'elle utilise pour lire quelques pages web, tâche qu'elle a souvent des difficultés à accomplir. De plus, elle n'a pas connaissance des possibilités presque infinies que comporte l'appareil qu'elle possède. Mais peu de personnes de son âge, ou jusqu'à 15 ans plus jeune qu'elle, soit dans la catégorie du 4ème âge[4], possèdent ou ont accès à un ordinateur (ou sauraient quoi en faire si tel était le cas). On appelle parfois cela "la fracture numérique". En effet, cette classe d'âge comporte le taux le plus faible d'équipement et de compétences informatiques. Cela peut s'expliquer par de multiples raisons : ils ne savent pas ce qu'ils pourraient en faire ; ils ont peur de ne pas savoir s'en servir ; ils n'ont tout simplement pas les moyens financiers. Mais revenons à Marie et aux raisons pour lesquelles elle éprouve des difficultés à utiliser sa tablette : ses doigts sont raides ; la coordination de ses membres n'est plus aussi précise qu'elle ne l'était auparavant ; sa vision décline. Mais surtout, Marie est atteinte de troubles cognitifs légers (TCL) et, si apprendre de nouvelles choses à son âge peut déjà s'avérer difficile, les troubles cognitifs rendent cela d'autant plus ardu. En effet, les TCL affectent plusieurs domaines cognitifs impliqués dans l'apprentissage, tel que la mémoire à court-terme, ou dans l'utilisation d'une tablette, tel que la mémoire de travail ou la cognition spatiale.

Il est fort dommage que les personnes se trouvant dans une situation analogue à celle de Marie n'aient qu'une utilisation très élémentaire des ordinateurs. Imaginez les multiples façons dont cela pourrait les aider : ils pourraient utiliser une application de calendrier pour compenser leurs pertes de mémoire, garder une trace de leur emploi du temps et recevoir des rappels automatiques lorsqu'un rendez-vous est prévu prochainement ; ils pourraient utiliser la visiophonie pour communiquer avec leur famille et rester actifs socialement ; et ils pourraient même bénéficier de programmes stimulants pour la cognition, y compris ceux qui consistent à tuer des zombies virtuels. En fait, le cerveau est comme un muscle et, s'il n'est pas stimulé, à cause de l'isolement social et du manque de tâches stimulantes à réaliser, il s'affaiblit. Cependant, les ordinateurs sous toutes leurs formes

[3]Marie était l'une des participants des études menées au cours de la préparation de cette thèse. Son prénom a été changé pour des raisons de confidentialité.

[4]Généralement 85 ans et plus, mais les catégories d'âge peuvent varier selon les auteurs ou les disciplines. Les personnes rangées dans la catégorie du 4ème âge peuvent n'avoir que 75 ans.

(PCs, tablettes, téléphones) ne sont pas très adaptés pour l'utilisation par des personnes âgées, en particulier celles souffrant de troubles cognitifs. En effet, bien que l'utilisation des ordinateurs soit devenu une seconde nature chez les adolescents tueurs de zombies nés sous l'ère numérique évoqués tout à l'heure, l'interaction avec ce type d'appareils est très complexe et loin d'être naturelle.

C'est pourquoi les travaux de recherche présentés ici visent à étudier comment des logiciels informatique soigneusement conçus peuvent aider les personnes âgées, qu'elles aient des troubles cognitifs ou pas. Deux études de cas ont été menées : la première consiste à proposer un agent conversationnel animé (ACA), un personnage virtuel doté de capacités d'interaction sociale, qui, capable de suivi de l'attention, servirait d'interface facile d'utilisation pour les personnes âgées atteintes de troubles cognitifs ; la seconde consiste à étudier les possibilités offertes par la réalité virtuelle pour traiter le syndrome post-chute (SPC), une combinaison de symptômes psychologiques, de type traumatique, et de symptômes psycho-moteurs résultant d'une chute. Bien que les fondements théoriques de ces deux applications soient assez différents, elles partagent de nombreux points communs : elles s'appuient sur les même technologies, nommément des personnages virtuels tridimensionnels et des moteurs de jeu vidéo ; elles s'adressent au même public, c'est à dire des personnes âgées ayant des troubles cognitifs (même si des personnes souffrant d'un SPC peuvent ne pas avoir de troubles cognitifs, l'application doit être adaptée aux personnes cumulant les deux pathologies) ; elles partagent plusieurs contraintes de conception ; les symptômes psychomoteurs observés dans le SPC sont très liés à la cognition, au niveau du contrôle moteur et de l'équilibre ; et elles doivent toutes deux être peu onéreuses, de sorte que les institution de soin et les personnes âgées puissent se les offrir. De plus, ces deux applications sont même susceptibles d'être combinées, étant donné qu'un ACA d'assistance capable de suivi attentionnel pourrait très bien être inclus dans une application de traitement en réalité virtuelle pour expliquer l'utilisation des commandes, donner les consignes, et effectuer des rappels aux patients lorsqu'ils sont inactifs ou distraits.

Enfin, pour ces deux études de cas, nous adoptons une approche de conception centrée-utilisateur appelée living lab *[22, 4, 160]. Elle est fondée sur les principes de la recherche-action [116] et vise à impliquer toutes les parties prenantes (patients, familles, professionnels de santé, etc.) tout au long du processus de conception. Cependant, même s'il existe des recommandations d'ordre général pour appliquer les principe du living lab, cette approche de conception est relativement récente et chaque équipe a sa propre manière de la mettre en œuvre. Par conséquent, les présents travaux contribuent également à la promotion de cette méthodologie en proposant une manière de l'appliquer et en fournissant un retour d'expérience.*

Dans ce chapitre introductif, les motivations de ces travaux sont exposées. Ensuite, les buts de cette recherche sont présentés et les principales contribu-

tions sont mises en lumière. Enfin, la structure de cette thèse est présentée.

Handicap chez les personnes âgées, perte d'autonomie et exclusion

Grâce aux progrès de la médecine au cours du siècle passé, l'espérance de vie à la naissance s'est considérablement accrue dans les pays développés. Mais cela a un coût : d'après les prédictions de l'ONU, en 2050, les personnes de plus de 60 ans représenteront 20% de la population mondiale et plus de 35% de la population des pays développés. Le coût économique d'une démographie vieillissante est énorme : en 2015, au sein de l'Union européenne, les dépenses publiques liées aux personnes âgées représentaient 25% du PIB, soit 50% du budget des états membres, et il est estimé que ce chiffre connaitra une croissance annuelle de 4% du PIB jusqu'en 2060 [62].

Bien que de nombreuses personnes connaissent un phénomène de vieillissement "normal", d'autres sont moins chanceuses et développent des pathologies sévères liées à l'âge, causant des handicaps. D'après [159] (pp. 14-15), trois modèles de handicap sont fréquemment utilisés dans la littérature spécialisés : un modèle médical, un modèle social et un modèle bio-psychosocial. D'un point de vue médical, un handicap se définit comme "une conséquence directe d'une maladie, d'un trauma ou d'une autre pathologie, qui requiert une prise en charge médicale, sous la forme d'un traitement individuel fourni par des professionnels de santé" *[159]. Le modèle social* "considère le handicap comme un problème d'ordre social et non comme un attribut de l'individu" *[159]. Enfin, le modèle bio-psychosocial fut proposé dans le cadre de la classification internationale Fonctionnement, handicap et santé (*International Classification of Functioning, Disability and Health, ICF*) [225]. Il prend en compte à la fois les facteurs biologiques et sociaux. Selon ce modèle, le terme* handicap *peut aussi bien faire référence à la détérioration d'une fonction ou structure corporelle qu'à une limitation empêchant de réaliser des activités de la vie quotidienne ou à des difficultés à participer aux situations de la vie quotidienne. Étant donné cette dernière définition du* handicap, *le nombre de personnes dans cette situation augmente rapidement avec l'âge, de 6,4% de la population dans la tranche d'âge des 18-50 ans à 29,5% dans celle des plus de 60 ans, dans les pays riches [226].*

Une autre manière de regarder le handicap consiste à considérer à quels types d'activité l'état d'une personne l'empêche de se livrer. Les activités de la vie quotidienne peuvent êtres réparties en deux groupes : les activités instrumentales et les activités élémentaires. Parmi les activités instrumentales se trouvent la gestion des finances personnelles, la préparation des repas ou l'utilisation des transport privés ou en commun, alors que la catégorie des activités élémentaires regroupe manger, boire, s'habiller ou faire sa toilette.

Bien entendu, les personnes n'étant capable d'accomplir aucune de ces activités ne peuvent vivre de manière indépendante, mais les personnes ne pouvant pas accomplir des activités élémentaires ont besoin de plus d'aide et de soins que celles qui n'ont des problèmes que dans l'accomplissement des activités instrumentales [123].

Le soin apporté aux personnes âgées handicapées est évidemment très couteux, mais leurs handicaps causent également l'exclusion de nombres d'entre elles de la société. En effet, des difficultés à se déplacer, par exemple, empêchent les personnes de participer à de nombreuses activités, telles que rendre visite à leurs amis ou leurs familles, emmener leurs petits enfants au zoo, ou simplement aller faire des courses. Les troubles cognitifs, l'une des principales causes de handicap chez les personnes âgées, entrainent un isolement social des personnes, car la communication avec les proches devient plus difficile et cette pathologie peut également causer dépression, apathie ou désinhibition, qui sont généralement mal acceptées dans un contexte social. C'est pourquoi nombre de chercheurs et entreprises à travers le monde travaillent à créer des technologies d'assistance, afin de réduire les coûts de prise en charge et d'améliorer la qualité de vie des patients et de leurs aidants (professionnels ou non).

Les travaux présentés dans cette thèses se concentrent particulièrement sur la création de technologies innovantes fondées sur les humains virtuels qui s'adressent aux personnes atteintes de troubles cognitifs et aux conséquences psychologiques et psycho-motrices des chutes, liées à une perte de mobilité. Mais avant d'en venir aux buts précis de ces recherches, considérons de manière plus détaillée les deux causes de handicap étudiées ici.

Troubles cognitifs et démence chez les personnes âgées

Les troubles cognitifs suffisamment sévères pour affecter la capacité des personnes à accomplir des activités de la vie quotidienne sont pricipalement dues à des pathologies biologiques regroupées sous le terme démence*, ou troubles neuro-cognitifs, depuis la 5ème édition du "*Diagnostic and Statistical Manual of Mental Disorders*" (DSM-5®) [9] de l'Association psychiatrique américaine, l'un des principaux manuels sur les pathologies psychiatriques. Les maladies les plus communes causant la démence chez les personnes âgées sont la maladie d'Alzheimer, la démence vasculaire et la démence à corps de Lewy.*

L'Organisation mondiale de la santé a identifié la démence chez les personnes âgées comme l'une des priorités de santé publique : 35,6 millions de personnes à travers le monde vivaient avec une démence en 2010 et des prévisions suggèrent que ce nombre dépassera les 100 millions en 2050 [227]. D'après une étude récente, environ 25% des personnes de plus de 80 ans sont atteintes de démence et cette proportion augmente rapidement avec l'âge, de 15,7% dans la tranche d'âge des 80-84 ans à 65,9% dans la tranche des 100 *ans et plus [119]. Cette situation entraine souvent l'institutionnalisation des personnes touchées dans les stades avancées de la maladie, jusqu'à leur décès [120].*

À ce jour, il existe peu de traitements pharmacologiques et leurs effets sont, au mieux, de ralentir la progression des symptômes, même lorsque le traitement est commencé tôt [201]. Pour cette raison, les thérapies non-médicamenteuses sont l'objet d'un intérêt de plus en plus grand dans la prise en charge de la démence, car elles sont à la fois plus efficaces et moins couteuses [19]. Cependant, le but des technologies d'assistance n'est pas toujours thérapeutique, mais peut également consister à favoriser le maintien à domicile des personnes et d'apporter une meilleure qualité de vie, à la fois aux patients déments et à leurs aidants, tout en réduisant les coûts de prise en charge, comme détaillé plus loin.

Chutes chez les personnes âgées

Les chutes chez les personnes âgées sont également reconnues comme un problème de santé publique important. En effet, 34 à 40% des personnes de plus de 65 ans chutent et 10% des personnes qui chutent se blessent sévèrement [217]. Les chutes sont également la première cause de mortalité accidentelle chez les personnes de 75 ans, qui représentent environ 85% des personnes qui décèdent des suites d'une chute à leur domicile [148]. Dans les maisons de retraite médicalisées, près de la moitié des résidents chutent au moins une fois par an [215].

Les chutes se soldent fréquemment par des conséquences dramatiques sur la qualité de vie des personnes, principalement à cause de fractures et de guérisons incomplètes, ce qui expose ces personnes âgées à un risque accru de chutes répétées. Par exemple, des études ont rapporté que, chaque année, 4% des résidents de maisons de retraites médicalisées se fracturent la hanche. Parmi eux, dans l'année qui suit, 12% se font une autre fracture et 31% décèdent [215]. Étant donné la mortalité et les coûts de prise en charge occasionnés par les chutes, de nombreux chercheurs se sont penchés sur la prévention des chutes [215] et la rééducation post-chute [125].

Mais les décès et les blessures physiques ne sont pas les seules conséquences des chutes chez les personnes âgées. En fait, de nombreux chuteurs présentent ce qu'on appelle un syndrome post-chute (SPC) [142]. Cela se caractérise par l'association de symptômes psychologiques (anxiété, peur de chuter à nouveau) et de troubles psycho-moteurs, appelé syndrome de désadaptation psycho-motrice *[141]. Cependant, le SPC n'est pas suffisamment étudié dans les recherches actuelles et est mal pris en charge dans les pratiques de soin, et ce alors qu'il rend la rééducation post-chute plus difficile et moins efficace [27].*

Enfin, il est utile de garder en tête, tout au long de la lecture du présent manuscrit, que les personnes âgées sont souvent affectées par plusieurs pa-

thologies, qui causent un certain niveau de handicap. Les pathologies étudiées dans ces travaux ne sont donc pas mutuellement exclusives (et ont en fait été souvent rencontrées chez un même individu), puisque les troubles cognitifs sont un facteur de risque de chute [140, 205]. En plus de la démence et de l'impossibilité de marcher, certaines personnes peuvent également avoir des difficultés pour voir, entendre ou manipuler des objets (à cause de l'arthrose).

Promesses des technologies d'assistance

Dans la plupart des pays développés, la faible natalité et le vieillissement de la population posent des défis sociaux et économiques. En effet, les coûts de prise en charge augmentent à cause du nombre croissant de personnes âgées, non seulement liées aux handicaps, mais aussi à cause de maladies fréquentes et de pathologies chroniques. De plus, une démographie vieillissante est la cause d'un manque d'aidants. C'est pourquoi de nouvelles solutions doivent être trouvées pour prendre en charge des personnes âgées de plus en plus nombreuses avec moins de travailleurs, tout en gardant des standards de qualité de vie élevés, aussi bien pour les aidants (professionnels de santé et aidants familiaux) que pour les bénéficiaires. À cet effet, les technologies d'assistance font l'objet d'un attention croissante.

*Cela pose la question de la relation que les personnes âgées entretiennent avec la technologie. Il existe une croyance répandue selon laquelle les personnes âgées ne savent pas ou sont réticentes à utiliser les nouvelles technologies. Cela s'avère être une idée fausse dans un nombre croissant de cas. Par exemple, des recherches menées dans les années 2000 ont décrit un sous-groupe de personnes âgées, les "Surfeurs d'argent" (*Silver Surfers*), qui ont adopté les technologies de l'information et de la communication tout autant que les jeunes [149, 182]. Pour être plus précis, une étude menée en 2010 par Olson* et al. *sur l'utilisation des technologies par les personnes âgées aux États-Unis a rapporté qu'environ 65% des personnes de plus de 65 ans ont une expérience de l'utilisation des ordinateur, mais que leur utilisation est moins fréquente et moins expertes que chez les personnes plus jeunes [149]. La même étude a également fait apparaitre que la moitié des personnes âgées possédant un ordinateur utilisait l'internet depuis plus de 5 ans. Cela montre clairement que la plupart des seniors ne sont pas réticents à l'utilisation des nouvelles technologies. Cependant, ils ont généralement peu d'expertise dans ce domaine et peuvent avoir besoin d'appareils ou de logiciels simplifiés qui n'incluent que les principales fonctionnalités.*

Ayant démonté les stéréotypes habituels sur les personnes âgées et la technologie, considérons ce que sont les technologies d'assistance, comment elles peuvent venir en aide aux personnes âgées ayant des besoins spécifiques et selon quelles conditions les séniors les adopteront.

Technologies pour le soin et l'assistance

Le terme technologie d'assistance *(TA), dans le contexte des personnes âgées, fait référence à n'importe quel appareil, produit ou équipement qui aide les personnes à réaliser des tâches qu'ils ne pourraient pas accomplir autrement ou qui facilite la vie quotidienne des séniors, que l'objet ait été conçut à cet effet ou que cet équipement soit pris "sur l'étagère" [210]. D'après Sanford [179], ce qui différencie les TA des autres technologies est le fait qu'elles sont "individualisées et suivent généralement la personne" [45]. Cook et Polgar [45] ajoutent que, dans ce cadre, certains appareils qui interviennent dans le soin aux personnes âgées, les technologies de rééducation par exemple, ne peuvent être considérées comme des TA. Étant donné cette définition, les TA peuvent prendre de nombreuses formes, des appareils les plus simples, tels que les lumières à allumage automatique, aux robots personnels les plus sophistiqués.*

La plupart des personnes âgées souhaitent rester autonomes le plus longtemps possible et préfèrent vivre chez elles plutôt que dans des institutions de soin adaptées. Les TA peuvent les aider à poursuivre ce but en leur apportant des moyens de compensation de leurs possibles handicaps. Un certain nombre de produits sont déjà disponibles, l'une des TA les plus répandues étant les appareils auditifs, qui compensent partiellement les pertes auditives. Un autre exemple est celui des fauteuils roulants électriques, qui permettent aux personnes qui ne peuvent plus marcher de sortir et vaquer à leurs occupations. Mais posséder l'un de ces fauteuils roulants ne garantit pas une autonomie complète parce qu'une personne qui ne peut pas du tout se tenir debout a de bonnes chances de ne pas pouvoir se transférer de son lit au fauteuil et, une fois dans le fauteuil, il paraît difficile pour ces personnes de ramasser des objets au sol, de passer l'aspirateur ou de faire la poussière sur leurs étagères. Il apparait donc que les TA disponibles actuellement ne suffisent pas pour permettre aux personnes âgées handicapées de vivre en toute indépendance à leurs domiciles et nous n'avons même pas mentionner la possibilité qu'elles aient en plus des troubles cognitifs.

Les TA pourraient grandement améliorer la situation des personnes âgées, mais celles qui sont disponibles actuellement sont loin d'être suffisantes. Pour cette raison, nombre de chercheurs et entreprises travaillent à la création de TA innovantes. De plus, la taille du marché potentiel pour les produits et services à destination des séniors est évaluée à environ 7000 milliards de dollars par an, ce qui en fait le 3ème secteur économique, d'après un rapport récent de la Commission européenne [62]. Il va donc sans dire qu'il y a des intérêts tant privés que publiques à développer les TA.

Technologies de compensation cognitive

D'après un recensement récent des "prothèses cognitives" *[6] informatisées, quatre catégories de technologies d'assistance pour la cognition peuvent être distinguées, selon les aspects de la vie quotidienne du patient qu'elles soutiennent :*

- *accès à l'information, communication et liens sociaux ;*
- *mobilité, organisation et tâches quotidiennes ;*
- *besoins de stimulation et de divertissement ;*
- *gestion des troubles psycho-comportementaux.*

Cette classification n'est pas la seule disponible. Dans un autre recensement, Gillepsie et al. *proposent de classer les TA selon les fonctions cognitives qu'elles soutiennent [78], en se fondant sur la classification des handicaps de l'ICF mentionnée plus haut.*

Pour chacun des domaines d'activité pour lesquels un soutien cognitif est nécessaire, de nombreux appareils et applications ont été proposés. Par exemple, le dispositif CIRCA, cité dans [6], a pour but de soutenir les échanges sociaux des personnes atteintes de démence [7]. Il s'agit d'un logiciel exécuté sur un ordinateur muni d'un écran tactile qui sert de médiateur de conversation entre la personne démente et son aidant. Grâce à du matériel de réminiscence, sous la forme d'images, de musiques et de vidéos, ils peuvent trouver un sujet de conversation et, grâce aux informations audiovisuelles, la personne démente se rappelle du sujet de la discussion plus facilement, ce qui limite le risque qu'il ou elle perde le fil de l'échange.

L'une des prothèses cognitives les plus réussies est le système COACH [29]. Il s'agit d'un dispositif qui permet de guider une personne démente pas à pas dans l'exécution d'une activité élémentaire de la vie quotidienne. L'activité réalisée par les auteurs pour les besoins de leur étude est celle du lavage des mains. Grâce à une méthode de reconnaissance d'activité fondée sur la vision par ordinateur, des notifications vocales sont faites à la personne pour la guider à chaque étape de la tâche.

*Enfin, on ne pourrait pas écrire sur les prothèses cognitifs sans mentionner l'*intelligence ambiante*. L'idée sous-jacente est la suivante : au lieu de situer l'aide apportée à la personne au niveau d'un appareil, elle est distribuée dans tout le lieu de vie. Ce concept est étroitement lié à celui de maison intelligente. Pour le soutien cognitif, cela pourrait consister à fournir une aide contextualisée dans toutes les pièces de l'habitation du bénéficiaire. Par exemple, des système comme COACH pourraient être installés dans la salle de bain, la chambre, la cuisine et près de la porte d'entrée pour aider les personnes à faire leur toilette, s'habiller, cuisiner et vérifier qu'elles ne sortent pas en n'étant pas suffisamment couvertes, tout en leur rappelant de* *ne pas oublier leurs clés.* Hamada et al. *ont proposé un système similaire à COACH pour aider les personnes démentes à aller aux toilettes [82]. Une maison intelligente pourrait également inclure un système de rappels pour l'organisation, les prises de médicament ou les courses (par exemple, un réfrigérateur intelligent pourrait notifier au système qu'il est vide). Un état de l'art de l'intelligence ambiante pour la santé peut être lu dans [2].*

Dans le contexte de ces travaux de thèse, les maisons intelligentes constituent une application potentielle de notre agent conversationnel. En effet, les notifications en intelligence ambiante pourraient être réalisées par un majordome virtuel, qui pourrait être visualisé sur des écrans placés dans les lieux sus-nommés, plutôt que par une voix désincarnée, comme c'est le cas dans COACH [29].

Technologies pour la rééducation

La rééducation est d'une importance critique pour le bien-être des personnes âgées, puisqu'un rétablissement incomplet à la suite d'une chute ou d'une attaque cérébrale peut causer un handicap, et ainsi engendrer une diminution de la qualité de vie et des coûts de prise en charge élevés. Elle peut être dédiée à des activité cognitives (de l'orthophonie en cas de troubles du langage, par exemple) ou physiques (de la kinésithérapie en cas d'hémiplégie, par exemple). La rééducation n'est pas une nouveauté et est pleinement incluse dans les pratiques de soin standards. En effet, tous les hôpitaux sont dotés d'une équipe de kinésithérapeutes ; des professionnels de la rééducations sont présents dans les maisons de retraites médicalisé, les cabinets privés et certains se déplacent au domicile des patients. La kinésithérapie est fondée sur des activités musculaires qui doivent être effectuées régulièrement (tous les jours ou toutes les semaines) sur une longue période (généralement de plusieurs semaines à plusieurs mois). Parmi les principaux facteurs de succès en termes de rééducation sont la participation active du patient et sa motivation à réaliser les exercices aussi fréquemment et pour aussi longtemps que nécessaire. La rééducation ne nécessite pas toujours l'utilisation d'appareils de haute technologie, mais ces derniers peuvent se révéler utiles de deux manières : pour stimuler la motivation des patients et pour améliorer les conditions de travail des professionnels.

Concernant le premier aspect, de nombreuses recherches se sont intéressées à l'utilisation des jeux sérieux et autres activités ludifiées pour entretenir la motivation de patients sur le long-terme. En effet, le principe même de la ludification de ces activités repose sur la notion de motivation intrinsèque, *définie par Deci et Ryan [53] dans leur théorie de l'auto-détermination, qui peut être employée pour concevoir des jeux motivants [56]. Pour résumer, la motivation intrinsèque dépend d'un équilibre entre le besoin de défi et le besoin de compétence. Cela signifie que les patients devraient se voir proposer des exercices de rééducation de difficulté croissante, afin qu'ils puis-*

sent être récompensés en atteignant des objectifs peu ambitieux dans un premier temps, avant de devoir réaliser des exercices plus difficiles, au fur et à mesure qu'ils recouvrent leurs capacité motrices et d'équilibre, ce qui les rend plus compétents. Pour être plus précis, d'après une revue de littérature de Lohse et al. *[118], six facteurs clés contribuent à la nature engageante des jeux de rééducation : récompenses, difficulté/défi, retours d'information, choix/interactivités, buts et mécanismes clairs et socialisation. La même revue conclut que plusieurs études ont montré que la ludification des activités de rééducation améliore la motivation des patients [118].*

Concernant le second aspect, la kinésithérapie est un travail très physique. En effet, les professionnels passent beaucoup de temps à aider leurs patients à se lever, à tirer ou pousser sur leurs membres pour rendre les exercices plus faciles ou plus difficiles ou même à quasiment porter les patients pour réduire le poids supporté par leurs jambes, afin qu'ils puissent faire quelques pas. Cela est épuisant et peut impacter leur santé. De plus, lorsqu'ils doivent porter un patient, cela mobilise deux voire trois professionnels à la fois. Pour réduire ce fardeau, des chercheurs ont proposé des systèmes mécatroniques, tels que des exosquelettes permettant de faire de la compensation de poids ou, au contraire, de rendre les exercices musculaires plus difficiles. Par exemple, de tels systèmes ont été proposés dans [146] et [14] pour faciliter la rééducation à la marche chez les personnes âgées.

Certains projets vont même jusqu'à jouer sur les deux tableaux : par exemple, Frisoli et al. *ont combiné un exosquelette des membres supérieurs qui réduit la force nécessaire aux mouvements à de la réalité virtuelle pour créer un programme de rééducation post-AVC [71]. En proposant des buts atteignables de difficulté progressive et en réduisant graduellement la contribution de l'exosquelette aux mouvements, le système vise à stimuler la motivation des patients et à être plus efficace que la rééducation classique.*

Des technologies réellement adaptées pour l'utilisabilité et l'acceptabilité

L'un des principaux obstacles au déploiement des technologies d'assistance est leur adoption par les personnes âgées. En effet, de nombreuses personnes qui pourraient bénéficier de ces technologies ne les utilisent pas. On peut trouver plusieurs raisons à cela : manque d'informations sur les produits disponibles ; certains appareils sont trop onéreux ; l'utilisation de TA est considérée comme étant stigmatisante ; certaines personnes ne veulent pas d'un appareil, car elles ne comprennent pas son fonctionnement ou ont peur de ne pas savoir l'utiliser ; etc. Il en résulte que ne nombreuses personnes ne s'équipent pas d'appareils qui pourraient améliorer leur qualité de vie. Une fois qu'une personne a connaissance de l'existence d'un appareil qui pourrait l'aider et qu'elle a les moyens de s'offrir, l'adoption de l'appareil dépend de deux facteurs : son utilisabilité *et son* acceptabilité*.*

Pour qu'une technologie soit utilisable, elle doit être suffisamment simple pour que les utilisateurs cibles puissent s'en servir. Cette condition peut sembler facile à remplir, mais, dès lors que les personnes ciblées sont âgées et/ou ont des handicaps cognitifs, la conception peut s'avérer ardue, comme expliqué dans le chapitre suivant. Par exemple, les ordinateurs sont faciles à utiliser pour les personnes nées après leur invention et n'ayant pas de handicap, mais les personnes malvoyantes, même jeunes, ne peuvent les utiliser que si certaines adaptations sont mises en place. Les personnes âgées sont susceptibles d'avoir des difficultés pour apprendre à utiliser un ordinateur, car ce sont des machines très complexes pour elles, et, même si les modèles à écran tactile constituent une alternative plus accessible, que nombre de personnes âgées ont adoptée, il n'en reste pas moins que celles atteintes de troubles cognitifs auront d'immenses difficultés à connecter l'appareil au Wi-Fi, télécharger des applications, utiliser les applications installées, etc. C'est pourquoi les appareils et logiciels destinés à ces personnes doivent être adaptées à leurs besoins spécifiques, le cas idéal étant un appareil dont l'utilisation ne nécessite aucun apprentissage et dont la forme suggère clairement la fonction. De ce point de vue, les agents conversationnels animés (ACA), tel que celui étudié dans cette thèse, semblent être un mode d'interaction avec les ordinateurs bien mieux adapté pour les personnes atteintes de troubles cognitifs : ils ont une apparence humaine (ou humanoïde) et parlent, ce que les gens ont fait tout au long de leurs vies et qui leur est naturel et facile à comprendre. De plus, les ACA peuvent être actifs et engager la conversation, ce qui est un avantage, car les personnes démentes ont souvent des difficultés à initier les actions.

L'acceptabilité d'une technologies est la volonté d'une personne d'utiliser cette technologie sur le court et le long termes. Un modèle d'acceptation des technologies a été proposé par Davis [52] dans le but d'identifier les facteurs qui influencent l'acceptabilité, dans le contexte de l'introduction de l'informatique sur le lieu de travail, en 1989. Cette première étude a conclu que les facteurs les plus fortement corrélés avec l'utilisation et les intentions d'utilisation futures auto-déclarées sont l'utilité perçue*, c'est à dire "à quel point une personne pense que l'utilisation du système améliorera sa productivité" [209], et* la facilité d'utilisation perçue*, en d'autres termes "à quel point une personne pense que l'utilisation du système se fera sans effort" [209]. Il a été constaté que ce modèle est applicable au cas des personnes âgées, dans une revue de littérature de Chen et Chan [40]. Cependant, il est nécessaire d'adapter la définition de perception d'utilité : pour les personnes âgées, cela correspond à "penser et réaliser que ces technologies peuvent être utilisées pour améliorer leur vie et satisfaire leurs besoins" [40]. Une technologie est dont perçue comme utile si elles en ressentent réellement le besoin. En ce qui concerne la perception de facilité d'utilisation, d'après Chen et Chan, la définition ne change pas. En revanche, cette dernière notion est plus fortement corrélée à l'adoption d'une technologie chez les personnes* *âgées que chez les personnes jeunes. De plus, il a été observé que la perception de facilité d'utilisation pour les personnes âgées est plus fortement liée au succès d'accomplissement de la tâche que chez les personnes jeunes, qui accordent plus d'importance à la rapidité d'exécution. Enfin, il est à noter que, pour les personnes âgées, la perception de facilité d'utilisation influence la perception d'utilité.*

Afin de concevoir des appareils ou des logiciels ayant de bonnes chances d'être adoptés par les utilisateurs ciblés, leurs besoins et capacité d'usage doivent être bien identifiés. Une manière de faire cela (celle que nous pensons être la meilleure) consiste à impliquer les utilisateurs ciblés dans toutes les étapes de la conception d'un produit, de l'idéation au déploiement. Cela s'appelle la conception centrée-utilisateur*. Cependant, les utilisateurs ciblés peuvent ne pas être les seuls décisionnaires de l'achat d'un produit. Par exemple, les personnes démentes peuvent ne plus être capables de prendre ce genre de décisions. Dans ce contexte, l'opinion de l'aidant sur le produit est tout aussi importante que celle du bénéficiaire. Un autre exemple est celui de l'adoption des technologies médicales dans les hôpitaux : les soignants sont les principaux utilisateurs, mais ils ne peuvent pas décider seuls de l'utilisation d'une technologie, puisque cette prérogative incombe à l'administration de l'hôpital.*

C'est pourquoi la méthodologie de conception participative que nous avons retenue dans ces travaux de recherche va au delà de la conception centrée-utilisateur et prend le parti d'impliquer également d'autres parties prenantes. Plus précisément, la méthode que nous avons suivie est celle dite living lab *[22, 4, 160]. Comme la conception centrée-utilisateur, elle repose sur des évaluations d'utilisabilité, des entretiens, des questionnaires et des groupes d'échange avec les utilisateurs ciblés afin d'adapter la conception à leurs capacités et mieux cibler leurs besoins, mais la technologie est aussi montrée et discutée avec d'autres parties prenantes, telles que les aidants, les médecins et professionnels de santé, les compagnies d'assurance, voire même les décideurs publiques. Nos travaux ne se contentent pas de suivre les recommandations sur l'approche de conception participative living lab, mais contribue à son développement, en proposant une implémentation de ces principes et en fournissant un retour d'expérience. Par exemple, appliquer des principes de conception centrée-utilisateur lorsque les utilisateurs ciblés sont des personnes ayant des troubles cognitifs peut se révéler ardu, puisqu'ils ont des difficultés à exprimer leurs souhaits et leurs besoins. C'est pourquoi notre retour d'expérience sur ce sujet peut être utile à d'autres concepteurs. Les détails et principes de la méthode sont présentés dans la section 3.2.*

Objectifs

Le but de ces travaux est d'explorer les possibilités offertes par les technologies du jeu vidéo, les humains virtuels en particulier, pour apporter soin et assistance aux personnes âgées vivant avec une démence ou des troubles psychomoteurs. A cette fin, deux études de cas ont été menées, sous la forme de la conception participative et de l'évaluation de deux prototypes, au sein d'un living lab ayant la particularité d'être situé dans l'enceinte d'un hôpital gériatrique. La première étude de cas est un projet d'agent conversationnel animé (ACA) appelé LOvely User Interface for Servicing Elders*, ou LOUISE, qui est supposé servir en tant qu'interface utilisateur adaptée pour les personnes âgées atteintes de démence. La seconde étude de cas porte sur un projet appelé* Virtual Promenade *(Promenade virtuelle), qui consiste à créer un simulateur de marche combinant un jeu de marche virtuelle, dans lequel les joueurs prennent le contrôle d'un avatar humain pour visiter des environnements virtuels, et d'un fauteuil haptique dont l'assise peut imiter les mouvements du bassin pendant la marche. Cela a pour but de traiter le syndrome post-chute (SPC) chez les personnes âgées, grâce à la thérapie par exposition en réalité virtuelle.*

Ces deux projets font une utilisation différente des humains virtuels : dans LOUISE, l'humain virtuel est utilisé comme non-soi, puisque c'est l'incarnation (virtuelle, certes) d'un agent artificiel, capable de tenir de simili-conversations avec l'utilisateur ; dans Virtual Promenade, les humains virtuels sont utilisés comme soi, sous la forme d'avatars qui représentent le joueur dans le monde virtuel. Bien que nos études de cas répondent à des problématiques assez différentes, elles partagent de nombreuses caractéristiques communes : elles s'appuient toutes deux sur l'imagerie 3D et les moteurs de jeu ; elles sont destinées aux personnes âgées ; elles doivent être adaptées à l'utilisation par des personnes atteintes de troubles cognitifs ; elles doivent être peu coûteuses ; et elles sont toutes les deux conçues, implémentées et testées en suivant une approche participative living lab, fondée sur le prototypage rapide et des cycles d'itération courts.

Des technologies pour compléter l'aide humaine et non la remplacer

Les technologies telles que les robots personnels et les ACA suscitent parfois la peur qu'ils puissent remplacer les humains et que des emplois seront détruits. Bien que cela puisse être une réalité pour les robots dans un contexte manufacturier, notre vision des ACA dans le contexte du soin aux personnes âgées est assez différente. Pour être clairs, nous ne pensons pas qu'ils puissent remplacer l'aide humaine et nous ne les concevons pas non plus dans ce but. En effet, les personnes âgées, comme tous les êtres humains, ont besoin de contacts sociaux, et rien, à notre avis, ne peut remplacer le *contact humain, particulièrement le toucher ou la parole.*

Au lieu de cela, le but poursuivi ici est de donner accès aux personnes âgées atteintes de démence à des services qui ne seraient pas assurés sinon, par manque de temps des aidants, et qui, grâce aux prothèses cognitives, aideraient les personnes à regagner en autonomie et en estime de soi. Par exemple, dans les institutions médicalisés, les pensionnaires peuvent ne pas avoir la possibilité de choisir ce qu'ils mangent parce que le personnel n'a pas le temps de venir présenter à chacun le menu du jour pour leur donner le choix. Un ACA n'a pas de contraintes de temps et pourrait très bien remplir cette tâche. Un autre exemple est celui du questionnement répétitif : les personnes démentes, comme elles sont désorientées, sont souvent anxieuses à propos de certains sujets et cela peut tourner à l'obsession. Par conséquent, ils posent constamment la même question à leurs aidants, puisqu'ils oublient la réponse peu de temps après l'avoir obtenue. Cela peut devenir épuisant pour l'aidant, mais les personnes démentes ont besoin de réponses à leurs questions. Dans ce contexte, un ACA pourrait essayer d'apporter des réponses pour rassurer la personne et également fournir des informations de repérage (heure, jour de la semaine, localisation). Nous pensons que de telles applications des ACA pourraient améliorer significativement la qualité de vie des patients, car priver des personnes de la possibilité de faire des choix et ne pas répondre à des questions qui sont pourtant d'une importance capitale pour leur tranquillité d'esprit s'apparente à de la maltraitance. Cela pourrait également réduire la charge pesant sur les aidants et réduire leur stress. En effet, des soignants qui auraient affaire à des patients qui se sentent mieux, qui sont donc moins agités, plus coopératifs lors des soins et plus autonomes seraient plus disponibles pour mieux s'occuper de chacun. Nous pensons donc que l'introduction d'ACA dans les établissements médicalisés et au domicile des patients pourrait créer un cercle vertueux qui bénéficierait aussi bien aux patients qu'aux soignants.

Une autre vision du type de technologies présentées dans cette thèse, qui s'applique particulièrement à Virtual Promenade, la seconde étude de cas, est qu'elles seraient destinées à l'utilisation par les professionnels de santé dans leurs interventions. Comme énoncé précédemment, les exosquelettes et les jeux de rééducation peuvent être utilisés et prescrits par les kinésithérapeutes dans leurs pratiques. Utilisée de cette manière, la technologie peut améliorer à la fois la qualité du soin et les conditions de travail des professionnels. Cela pourrait même mener à la création de nouvelles professions : il y a bien des zoo-thérapeuthes ou des musico-thérapeutes, alors pourquoi pas des robot-thérapeutes ? En fait, cela pourrait même très prochainement devenir une réalité, puisque l'usage des robots dans le soin aux personnes âgées a déjà produit des résultats très encourageants. Par exemple, un robot bébé phoque japonais, appelé Paro, a été évalué positivement pour le soin aux personnes démentes dans plusieurs études [183].

Ayant clarifié notre vision de l'utilisation des nouvelles technologies dans *le soin aux personnes âgées, considérons de manière plus détaillée les buts spécifiques de cette recherche, qui s'appuie sur deux études de cas : les humains virtuels comme soi (Virtual Promenade) et comme non-soi (LOUISE). Enfin, il est à noter que ces deux cas d'usage ne sont pas mutuellement exclusifs et pourraient être combinés dans une seule application.*

Agents conversationnels animés comme interface Homme-machine pour les personnes âgées ayant des troubles cognitifs

Les agents conversationnels animés (ACA) sont un type particulier d'agents artificiels doté de capacités d'interaction sociale. Les ACA sont des programmes informatiques capable d'interagir grâce à des entrées de l'utilisateur acquises à l'aide d'un jeu de capteurs (caméra, microphone, écran tactile, etc.) et un personnage virtuel, généralement affiché sur un écran, capable de produire des comportements de communication anthropomorphes en parlant et en faisant des gestes, à l'aide d'enregistrement vocaux ou d'une voix de synthèse et d'animations. Ces systèmes dépendent de plusieurs domaines techniques incluant, entre autres, ceux de la reconnaissance vocale, de la synthèse vocale, des images de synthèse, de la vision par ordinateur et de l'intelligence artificielle. De plus, l'imitation réaliste de la communication humaine naturelle par un ordinateur requiert des connaissances dans plusieurs domaines des sciences humaines et sociales comme la psychologie, la linguistique et l'anthropologie.

Bien que d'intenses efforts de recherche sur l'animation des ACA et la perception des "signaux sociaux" [156] ont été produits au cours de la dernière décennie, il reste encore beaucoup de chemin à parcourir avant que des programmes informatiques soient capable d'interagir de manière complètement naturelle. Il reste de nombreuses questions auxquelles il faut apporter des réponses. Quelle apparence doivent avoir les ACA ? Plus précisément, devraient-ils être anthropomorphes ou plutôt ressembler à des animaux ? Devraient-ils être réalistes ou stéréotypés ? Devraient-ils être du genre masculin ou féminin ? Nous pourrions continuer comme cela longtemps et nous n'avons abordé que l'apparence du personnage virtuel.

Comme énoncé précédemment, les ACA présentent plusieurs avantages, détaillés dans la section 2.2.1, pour servir d'interface Homme-machine dans les technologies d'assistance informatisées à destination des personnes âgées atteintes de démence : ils utilisent la communication verbale et non-verbale, ce qui ne requiert pas d'apprentissage ; ils ont une apparence humanoïde, ce qui suggère la manière dont ils s'utilisent ; ils sont faciles à comprendre ; etc. De plus, plusieurs études (voir la section 2.2.1) ont déjà montré des résultats encourageants en termes d'efficacité et de satisfaction dans le contexte des personnes âgées ayant des troubles cognitifs. La technologie des ACA est également bon marché et ubiquitaire, ce qui permettrait de la déployer à *plusieurs endroits du lieu de vie des patients dans des scénarios de maison intelligente.*

*Le principal objectif de cette première étude de cas est de concevoir une interface utilisateur fondée sur un ACA, appelée LOUISE (*LOvely User Interface for Servicing Elders*), qui pourrait être incluse dans diverses technologies d'assistance, en suivant une méthodologie, dite living lab, de conception participative incrémentale. Par exemple, il existe des piluliers électroniques qui sonnent et clignotent lorsqu'il est l'heure de prendre ses médicaments et ne cessent de sonner que lorsque le pilulier est retourné pour prendre les pilules qu'il contient. Avec LOUISE, la sonnerie pourrait être remplacée par des notifications verbales et non-verbales, ce qui serait plus agréable, mais surtout plus explicite, qu'une sonnerie que les personnes démentes risquent de ne pas identifier comme un rappel de prendre leurs médicaments. Ensuite, une fois le pilulier retourné, LOUISE pourrait recevoir un message lui indiquant que les pilules ont été prises dans la boite et se mettre à guider la personne dans la tâche d'ingérer les médicaments. De plus, LOUISE pourrait être dotée de capacités de reconnaissance d'activité pour obtenir un indice de confiance sur le fait que la personne a effectivement ingéré les pilules et pourrait appeler de l'aide en cas de doute. LOUISE est en fait vue comme une boîte à outils qui, grâce à un éditeur de scénarii d'interaction et à la possibilité de communiquer avec des appareils ou logiciels externes, pourrait facilement être utilisée comme interface utilisateur dans plusieurs applications d'assistance, telles que celle présentée ci-dessus. Ce système devrait également être facilement personnalisable en fonction des goûts de l'utilisateur, par exemple en proposant plusieurs modèles d'humains virtuels d'apparences différentes et en rendant facile pour les développeurs d'ajouter de nouveaux modèles. Enfin, le système se doit d'être bon marché, afin que la plupart des patients puissent se l'offrir. Cela peut être réalisé en privilégiant l'utilisation de logiciels libres ou gratuits et de périphériques grand-public dans la construction du système. L'utilisation de logiciels publiquement disponibles et le détournement d'appareils grand-public présentent aussi un intérêt dans notre contexte, puisque cela permet un prototypage rapide et peu coûteux.*

Nos objectifs secondaires sont : d'étudier les interactions entre les personnes âgées démentes et de tels systèmes, afin de proposer une méthode de gestion de l'interaction adaptée et généraliste qui produise une bonne utilisabilité ; et de collecter des informations sur les goûts et les besoins des patients déments qui influencent l'acceptabilité des ACA, afin de mieux cerner les attentes en termes d'aspect visuel et d'identifier les applications les plus utiles des ACA.

Thérapie par exposition en réalité virtuelle pour le traitement du syndrome post-chute

La thérapie par exposition est une méthode de thérapie comportementale dont les bases théoriques sont la théorie du traitement émotionnel [50]. Elle peut être définie comme "le processus consistant à aider un patient à se livrer à un contact répété et prolonger avec un stimulus redouté" [1]. Ce type de thérapie est utilisé dans les cas de phobies, de syndromes de stress post-traumatique (SSPT) et d'anxiété généralisée. Cela consiste à confronter le patient à l'objet de sa peur, en toute sécurité. Dans certaines situations, telles que l'aviophobie (peur de d'avion) ou les SSPT liés au combat, exposer les personnes au stimulus qu'ils craignent est plus difficile que lorsque le stimulus est facile à obtenir, comme pour l'arachnophobie (peur des araignées) par exemple. C'est pourquoi de nombreux chercheurs ont étudié l'utilisation de la réalité virtuelle dans les thérapies par exposition, en remplaçant les stimuli réels, comme les araignées, les serpents ou les clowns, par des stimuli virtuels. C'est ce qu'on appelle la thérapie par exposition en réalité virtuelle (TERV). Il a été montré que la TERV produit de bons résultats dans de nombreux cas [75].

Le syndrome post-chute (SPC) est un trouble anxieux qui ressemble au SSPT, à la différence près qu'il est lié à la peur de chuter ; cette peur peut également être accompagnée de troubles psychomoteurs appelés syndrome de régression psychomotrice, c'est-à-dire une perte des fonctions motrices cognitives (voir section 2.1.3). Le traitement du SPC peut par conséquent s'envisager comme une forme de rééducation cognitive qui prend en compte à la fois les aspects psychologiques et psychomoteurs. De ce point de vue, comme mentionné précédemment, les jeux vidéo sérieux seraient dignes d'intérêt, en termes de renforcement de la motivation. De plus, les conséquences psychomotrices et psychologiques des chutes sont souvent négligées dans les pratiques de soin actuelles.

Pour ces raisons, le but premier de notre seconde étude de cas est d'étudier la conception participative d'un système combinant de la réalité virtuelle, sous la forme d'un jeu sérieux, et un fauteuil haptique, afin de traiter les deux aspects pathologiques du SPC. Ce système a la particularité d'avoir deux utilisateurs finaux : les patients, bien évidemment, puisque ce sont eux qui vont jouer; et aussi les professionnels de santé susceptibles d'inclure ce dispositif dans leurs pratiques de soin. C'est pourquoi notre conception participative devra impliquer des kinésithérapeutes et d'autres professionnels de santé potentiellement impactés par le dispositif. Comme dans le cas de LOUISE, ce dispositif doit rester bon marché et facile à augmenter par de nouveaux contenus.

Nos objectifs secondaires sont : identifier quels aspects de la conception du jeu sont importants aux yeux des personnes âgées, pour qu'elles y adhèrent et soient motivées à continuer de s'exercer ; d'étudier le mode de *commande du jeu et sa pertinence pour les patients ciblés ; et d'évaluer la faisabilité du déploiement d'un tel dispositif dans des environnements de soin fortement contraints et son acceptation par les professionnels et les patients.*

Contributions

Les principales contribution de ces travaux sont les suivantes :

- *Nous avons conçu, implémenté et testé un nouveau système d'ACA, appelé LOUISE, qui pourrait servir d'interface utilisateur dans les technologies d'assistance pour les personnes âgées ayant des troubles cognitifs. En particulier, l'originalité de notre conception réside dans les caractéristiques suivantes :*
  1. *un algorithme simple mais efficace pour le suivi de l'attention de l'utilisateur ;*
  2. *un gestionnaire d'interactions utilitaires qui réagit à l'inattention de l'utilisateur, effectue des rappels pour re-capturer son attention et produit des rappels contextuels pour qu'il puisse reprendre la conversation sans perdre le fil ;*
  3. *un langage XML élémentaire permettant de décrire des scénarii d'interaction d'assistance ;*
  4. *la possibilité d'ajouter aux comportements verbaux et non-verbaux l'affichage d'images présentant des objets et de vidéos montrant des exemples d'actions, afin de créer une redondance et une multimodalité de l'information pour compenser les troubles aphasiques et agnosiques, ainsi que la surdité légère, que les personnes démentes sont susceptibles de présenter ;*
  5. *la possibilité d'ajouter des personnages supplémentaires dans un but de personnalisation.*
- *Après des expérimentations extensives auprès des différentes parties prenantes, nous avons formulé des recommandations pour la création de scénarii d'interaction pour les applications d'assistance aux personnes âgées atteintes de troubles cognitifs incluant LOUISE.*
- *Nous avons conçu, implémenté et testé Virtual Promenade, un dispositif combinant un jeu de marche virtuelle à un fauteuil haptique pour la thérapie en réalité virtuelle du syndrome post-chute chez les personnes âgées. Ce système est bon marché, facilement utilisable et bien accepté par les patients avec ou sans troubles cognitifs et a été co-conçu avec des professionnels de santé (médecins et kinésithérapeutes).*

- *Grâce aux retours sur nos deux systèmes et à l'expérience aquise au travers de la pratique de la conception participative, nous proposons des idées pour la conception de technologies d'assistance aux personnes âgées incluant des humains virtuel.*

- *Nous contribuons au développement de la méthodologie de conception living lab en apportant, en complément des pratiques courantes, des idées venues des mouvements* lean startup *et* do-it-yourself*, ce qui se traduit par du prototypage rapide grâce à la réutilisation ou au détournement de logiciels et matériels existants, et des cycles d'itération courts, notamment en apportant des changements aux prototypes au cours d'une étude d'évaluation.*

- *Nous rapportons des retours d'expérience sur l'application de ces principes de conception innovants au sein d'un living lab situé dans l'enceinte d'un hôpital gériatrique, et incluant des personnes ayant des limitations cognitives et qui éprouvent des difficultés à exprimer leurs besoins et leurs souhaits.*

Structure de la thèse

Ce manuscrit est organisé en 8 chapitres. Dans ce chapitre introductif, nous avons vu qu'à cause de la démocratie vieillissante des pays développés, le nombre de personnes âgées ayant un handicap causé par des troubles cognitifs ou des chutes augmente alors que le nombre de personnes plus jeunes pouvant leur apporter les soins nécessaires diminue. Nous avons également vu que, dans ce contexte, les technologies d'assistance et de rééducation pourraient contribuer à apporter des solutions et améliorer la qualité de vie des patients et de leurs aidants, tout en réduisant les coûts. Cependant, les technologies qui permettraient de prendre pleinement en charge les troubles cognitifs ou la rééducation post-chute ne sont pas disponibles à ce jour, bien que des résultats prometteurs ont été observés en utilisant des technologies informatiques, telles que les ACA ou les jeux sérieux pour la rééducation. De plus, les problématiques sous-jacentes à l'adoption des technologies par les personnes âgées ont été présentées et il a été souligné que les approches de conceptions participatives pourraient améliorer la situation. Le contexte et les motivations de cette thèse ont été introduits, nous proposons d'étudier les ACA comme interface utilisateur dans les aides techniques cognitives à destination des personnes âgées démentes et la thérapie par exposition en réalité virtuelle pour le syndrome post-chute, à travers la conception participative living lab. Enfin, les objectifs de nos travaux ont été détaillés et les principales contributions ont été mises en lumière.

Le chapitre 2 a pour but de présenter et organiser l'information pertinente sur les contraintes de conception de nos projets et leur positionnement *par rapport à l'état de l'art. En premier lieu, ce chapitre présente en détail les aspects du vieillissement, des pathologies auxquelles nous nous intéressons et de l'environnement qui doivent être pris en compte dans la conception pour qu'elle soit adaptée. En second lieu, la littérature sur l'utilisation des ACA et des jeux sérieux dans le contexte du handicap chez les personnes âgées y est résumée et commentée.*

Dans le chapitre 3, les principaux concepts théoriques impliqués dans ces travaux sont définis et discutés. Après quoi la méthodologie et les principes de la conception participative living lab que nous avons adoptés sont détaillés.

Dans le chapitre 4, nous résumons l'information nécessaire pour donner au lecteur une bonne compréhension de la manière dont les humains virtuels sont créés. Pour être plus précis, nous commençons par introduire quelques notions sur la communication humaine et sur le domaine du traitement du signal social. Ensuite, nous présentons l'architecture modulaire caractéristique du logiciel présent dans les ACA et les robots sociaux. Nous fournissons également des détails supplémentaires sur la manière dont les comportements verbaux et non-verbaux peuvent être perçus, analysés, interprétés et générés par un ordinateur.

Les trois chapitres qui suivent sont consacrés aux deux études de cas que nous avons menées, considérant respectivement les humains virtuels sous leurs deux déclinaisons : non-soi et soi. Dans le chapitre 5, nous présentons notre première étude de cas sur LOUISE, ce qui inclut la conception participative de nos prototypes, nos expériences et nos observations. Dans le chapitre 6, nous rapportons nos expériences sur plusieurs cas d'utilisation crédibles de LOUISE. Le chapitre 7, quand à lui, présente la seconde étude de cas. Dans ce chapitre, nous présentons d'abord notre prototype, puis nous rapportons ses évolutions au cours du processus de conception participative et d'une expérience de validation.

Le dernier chapitre (chapitre 8) est consacré aux conclusions, recommandations et perspectives des travaux actuels et futurs. Nous présentons d'abord les conclusions de chacune de nos études de cas et une conclusion générale. Ensuite, des recommandations sur l'application de la méthode de conception participative living lab sont proposées et des idées de travaux futurs sont présentés, avant de terminer ce manuscrit sur des considérations éthiques concernant l'utilisation des humains virtuels dans le soin aux personnes âgées.

Chapter 1

Introduction

SINCE THE BEGINNING OF THE XXI$^{\text{ST}}$ CENTURY, the fast progress of computer technology and the wide spread of the Internet have led to a deep and ongoing change in the lifestyles of most inhabitants of developed countries. Indeed, access to the Internet raised from 30% of the inhabitants in 2000 to 80% in 2015 [94]. This does not go without consequences. In fact, over-stimulation of individuals, through constant information bombing while surfing the Internet, usually for advertising purposes, is linked to a reduced attention span. According to a study conducted by Microsoft [133], the average attention span of an Internet user has dropped from 12 seconds in 2000 to 8 seconds in 2015, or less than the one of a goldfish, which is estimated at 9 seconds!

This rise of information and communication technologies has thus clearly transformed the way we live and work. In particular, it has allowed the creation of virtual worlds in which millions of people immerse on a daily basis, mostly to kill zombies, monsters, demons, aliens, humans or each other with (respectively) baseball bats, bows and arrows, magic swords, laser guns, shotguns or assault riffles. The video game industry, one of the fastest growing economy in the past decade, along with the movie industry, has driven the technical innovations of 3D imaging and animation to create ever more realistic virtual worlds and characters. To enter these virtual worlds, visible through screens of all sizes, as many windows to imaginary places, players have to step in the shoes of virtual alter-egos, called *avatars*, which they control like digital puppets.

Interestingly, killing virtual zombies by controlling an avatar, armed with a shotgun, and without being torn apart (the avatar, not the player) requires being very alert and sustaining one's attention for much longer than 8 seconds. But is that the same kind of attention that is required when playing one of those immersive video games and when reading web pages? In fact, playing a shooting game heavily mobilizes a number of cognitive functions: spatial cognition, sensory-motor coordination, divided attention, etc (if you are not convinced, try having a conversation with a teenager while he or she is playing Call of Duty). Several studies conducted by psychologists have even shown that playing action video games leads to improvements in several sensory and cognitive domains, such as peripheral vision or the capacity to switch rapidly between tasks [79].

On the other end of the spectrum of digital media consumers is Mary[1]. Mary is 99 and owns a tablet computer which she uses to read a few web pages, and often struggles doing so. In addition, she is not aware of the almost infinite possibilities of the device she owns. But not many people her age, or up to 15 years younger than her, in the *older-old* age group[2], own or have access to a computer or the Internet (or would know what to do with it if they did). This is sometimes called "the digital divide". Indeed, this age category has the lowest rate of computer equipment and skills. There are multiple reasons for this: they do not know what they could do with it; they are afraid that they would not be able to use it; they simply cannot afford it. But let us come back to Mary and the reasons why she struggles when using her tablet: her fingers are stiff; her limb coordination is not as precise as it used to; her vision is declining. But most importantly, Mary has mild cognitive impairment (MCI) and while learning new things at her age may already prove difficult, having cognitive impairment makes it even more difficult. Indeed, MCI may affect several cognitive domains involved in learning, such as short-term memory, or in touch-tablet use, such as working memory or spatial cognition.

It is a pity that people in a situation similar to Mary's do not use or only have a very basic use of computers. Think of the many ways in which this could help them: they could use a calendar application to compensate for their memory loss, keep track of their schedule and have prompts sent to them when an appointment is coming soon; they could use video chat to communicate with their family and remain socially active; they could access information easily; and they could even benefit from a number of cognitively stimulating programs, including the ones that involve killing virtual zombies. In fact, the brain is like a muscle and if it is not used, because of social isolation and lack of stimulating tasks to accomplish, it weakens. But standard computer devices of all forms (PCs, tablets, smartphones) are not well adapted for use by the elderly, in particular those with reduced cognitive capabilities. Indeed, although computer use became a second nature for the teenage digital-native zombie slayers mentioned earlier, interaction with this kind of devices is highly complex and far from being natural.

This is why the present research work aims at investigating how carefully designed computer software may help older adults with or without cogni-

[1]Mary was one of the participants of the studies conducted during the preparation of this thesis. Her name was changed for confidentiality.

[2]Usually 85 and older, but this categorization in age groups may vary across authors and study fields. People in the older-old category may be as young as 75.

tive impairment. Two case studies were conducted: the first consisted in proposing an attention-monitoring embodied conversational agent (ECA), a virtual character with social interaction capabilities, as a user-friendly interface for older adults with cognitive impairment; the second case study consisted in investigating the possibility of using virtual reality to treat post-fall syndrome (PFS), a combination of trauma-like psychological symptoms and psycho-motor symptoms resulting from a fall. While the theoretical backgrounds of these two applications are quite different, they share many similarities: they make use of the same technologies, namely 3D animated virtual humans and video game engines; they target the same public, that is to say older adults with cognitive impairment (though some people having PFS do not have cognitive impairment, the application should be adapted for people affected by both conditions); they share several design issues; the psychomotor symptoms observed in PFS are highly related to cognition, on the level of motor and balance control; and they have to be cheap, for them to be easily acquired by care institutions or older adults. In addition, these two applications are even susceptible to be combined, as an assistive attention-monitoring ECA could very well be included in a virtual reality treatment application to explain how to use the controls, give instructions, and send prompts to the patient when he or she is inactive or inattentive.

Lastly, for both of these case studies, we adopt a user-centred design approach called *living lab* [22, 4, 160]. It is founded on the principles of action research [116] and aims at involving all stakeholders (patients, their families, care professionals, etc.) throughout the design process. However, although there are some general guidelines to apply living lab principles, this design approach is quite new and each team that adopted it has its own way of applying it. Therefore, this work also contributes to the development and promotion of this design methodology by proposing a way of applying it and providing feedback on lessons learned.

In this introductory chapter, the motivation for this work is presented in Sections 1.1 and 1.2. Then, the research goals are introduced (Section 1.3) and the key contributions are highlighted (Section 1.4). Lastly, the structure of this thesis is presented in Section 1.5.

1.1 Disabilities in older adults, loss of autonomy and exclusion

Thanks to the progress of health care in the past century, life expectancy at birth has dramatically increased in developed countries. But this comes at a cost: according to predictions of the United Nations, the proportion of people over 60 years old will represent over 20% of the population worldwide and over 35% in developed countries by 2050. The economic cost of aging demographics is huge: in the European Union, public spendings towards elderly people represented 25% of the GDP or about 50% of government expenditures in 2015, and it is expected to grow by more than 4% of the GDP, every year, until 2060 [62].

Though many people experience a healthy "normal" aging phenomenon, others are less lucky and develop severe age-related pathologies, resulting in disability. According to [159] (pp. 14-15), three models of disability are frequently used in the related literature: a medical model, a social model and a bio-psychosocial model. From a medical point of view, a disability can be defined as *"a direct consequence of a disease, trauma or other health condition, which requires medical care in the form of individual treatment provided by health professionals"* [159]. The social model *"considers disability as a socially created problem and not at all an attribute of an individual"* [159]. Lastly, the bio-psychosocial model was proposed in the International Classification of Functioning, Disability and Health (ICF) [225]. It takes both biological and social factors into account. Within this model, the term *disability* may refer to either an impairment in a body function or structure, a limitation to perform daily activities or a restriction to participate in everyday life situations. Given this latest definition of *disability*, the number of people in this situation grows rapidly with age, from 6.4% of the population in the 18-50 years old age group to 29.5% in the 60 and older age group, in high-income countries [226].

Another way to look at disability is to consider what kinds of daily activities the person's conditions prevent him or her to perform. Daily activities may be divided in two groups: instrumental activities of daily living (IADL) and basic activities of daily living (BADL). IADL include management of personal finances, meal preparation or use of public and private transportation, whereas BADL include eating, drinking, dressing or bathing. Of course, people who are not capable of performing either of these groups of activities are not able to live independently, but people who cannot perform BADL require more care and support than people who are only prevented from performing IADL [123].

Taking care of elders with disabilities is obviously very costly but disabilty also causes many older adults to be excluded from society. Indeed, difficulties in mobility, for instance, prevent people from participating in many activities such as visiting family and friends, taking their grandchildren to the zoo, or simply going shopping. Cognitive impairment, which is a very important cause of disability in older adults, also causes people to get isolated, as communication with relatives or friends becomes more difficult and also because this condition may cause depression, apathy or disinhibition, which are not well accepted in a social setting. This is why many researchers and companies around the world work on assistive technologies to reduce care costs and increase the quality of life for patients and people (professional or not) who help them.

The work presented in this thesis specifically focuses on creating innovative virtual human-based assistive technologies to address cognitive impairment and psychological and psycho-motor consequences of falls, related to loss of mobility. But before getting to the goals of this research, let us consider in more details the specifics of these two causes of disability.

1.1.1 Cognitive impairment and dementia in older adults

Cognitive impairment that is severe enough to affect performance in activities of daily living in older adults is mainly caused by organic pathologies grouped under the name *dementia*, or *neurocognitive disorders*, since the 5th edition of the American Psychiatric Association's "Diagnostic and Statistical Manual of Mental Disorders" (DSM-5®) [9], one of the main manuals for psychiatric conditions. The most common diseases causing dementia in older adults are Alzheimer's disease, vascular dementia and Lewy body dementia.

The World Health Organization identified dementia in elderly people as a public health priority: 35.6 million people worldwide were living with dementia in 2010 and predictions suggest that this number will exceed 100 million by 2050 [227]. According to a recent study, about 25% of people over 80 have dementia and this proportion increases rapidly with age, from 15.7% in the 80-84 age group to 65.9% in the 100-plus age group [119]. This situation often ends up with the institutionalization of the dementia sufferers in severe stages of the disease, until their passing [120].

To this day, few pharmacological treatments exist and their effects are, at best, to slow down the progression of the symptoms, even when treatment is started early [201]. For this reason, non-pharmacological therapies are given more and more interest for dementia care because they could be both more efficient and less costly [19]. However, the aim of assistive technologies is not always therapeutic but can also be to support home staying and provide a better quality of life for dementia patients and their helpers, while reducing the cost of care, as detailed in Section 1.2.

1.1.2 Falls in older adults

Falls in older adults are also widely recognized as a very important public health issue. Indeed, 35 to 40% of people over 65 fall and 10% of those who fall suffer serious injuries [217]. Falls are also the main cause of death from injury for older adults and people over 75 account for about 85% of people who die by falling at home every year [148]. In nursing homes about half of the residents fall at least once a year [215].

Falls often result in dramatic consequences on people's quality of life, mostly due to fractures and incomplete recovery, which cause disability and put elders at higher risk of repeated falls. For instance, studies reported that, each year, 4% of nursing home residents suffer a hip fracture. Among these people, 12% incur a new fracture and 31% pass away in the following year [215]. As this generates high health care costs and mortality, many researchers have investigated fall prevention [215] and post-fall rehabilitation [125].

But death and physical injuries are not the only consequences of falls for elderly people. In fact, many fallers experience something called *post-fall syndrome* (PFS) [142]. It is characterized by the association of psychological symptoms (anxiety, worry of repeated falls) and psycho-motor disorders, called *psychomotor disadaptation syndrome* [141]. However, PFS is insufficiently addressed in research and current care practices, whereas it makes post-fall rehabilitation more difficult and less efficient [27].

Lastly, it is worth keeping in mind, throughout the reading of this thesis, that older adults are often affected by more than one condition, causing some level of impairment. The two conditions addressed in this work are not mutually exclusive (and were in fact often met in the same individuals), since cognitive disorders are a risk factor for falls [140, 205]. In addition to dementia or inability to walk, older adults may also have difficulties in hearing and seeing (linked to sensory disorders) or manipulating objects (due to osteoarthrosis).

1.2 Promises of assistive technologies

In most developed countries, low birth rate and aging population cause social and economic challenges. Indeed, care costs are increasing due to the growing number of elderly people, not only because of disability but also because of frequent illness and chronic conditions. In addition, aging demographics cause caregiver shortage. This is why new solutions have to be found to support more and more elders with fewer workers, while keeping up with good quality of life standards, for both caregivers (health professionals and family members) and care receivers. To this aim, assistive technologies are receiving growing interest.

This asks the question of how elders relate to technology. A widespread belief is that older adults are reluctant to use new technologies or that they cannot do so. This turns out to be a misconception for a growing number of older adults. For instance, research conducted in the years 2000 have pictured a subgroup of older adults, the “Silver Surfers”, who embraced information and communication technology just like their younger counterparts [149, 182]. To be more specific, a study conducted by Olson *et al.* on technology use by younger and older adults in the United States in 2010 reported that about 65% of people over 65 had experience with computers, but that their use was less expert and less frequent than the one of younger adults [149]. The same study also reported that about 50% of older computer users had been using the Internet for over 5 years. This clearly shows that most older adults are not necessarily reluctant to technology use. However, they usually have low expertise in using them and may require simplified devices or software that only include key functionalities.

Having gone past the clichés about older adults and technology, let us now consider what are assistive technologies, how they may help older adults with special needs and under which conditions elders will adopt them.

1.2.1 Technologies for care and assistance

The term *assistive technology* (AT), in the context of older adults, refers to any device, product or equipment that helps people to perform a task they would be unable to do otherwise or facilitates seniors' daily lives, whether it was designed for that specific purpose or is an off-the-shelf piece of equipment [210]. According to Sandford [179], what differentiates AT from other technology is that it is "individualized and usually follows the person" [45]. Cook and Polgar [45] also argue that, in that sense, certain devices that intervene in elders' care, rehabilitation technologies for instance, cannot be considered as ATs. Given this definition, assistive technologies may take many forms, from low-tech devices, such as automatically ignited lights, to high-tech personal robots.

Most elderly people want to remain autonomous for as long as possible and live in their homes rather than in adapted care facilities. ATs may help them in pursuing this goal by providing them with compensations for possible disabilities. A number of products are already available, one of the most widespread AT being hearing aids, which partly compensate for hearing impairment. Another example is the electric wheelchair which allows people who cannot walk anymore to go out and do their business. But owning an electric wheelchair does not guarantees that one will stay fully autonomous because people who cannot stand on their feet at all may not be able to go from their bed to their wheelchair and, once on the wheelchair, it seems difficult for them to pickup objects that fell on the floor, to clean their apartment with a vacuum cleaner or to wipe the dust off their shelves. It therefore seems that currently available ATs do not fully allow disabled older adults to live independently in their homes and this example does not even mentions the possibility of cognitive impairment.

ATs appear to be devices that could greatly improve the situation for elders but currently available ones are far from being sufficient. For this reason, many researchers and companies are working on innovative AT solutions. In addition, the market for products and services targeting the elderly, called "Silver Economy", is estimated to be around \$7 trillion per year, which makes it the 3rd largest economy, according to a recent report of the European Commission [62]. Needless to say, there are both private and public interests in developing assistive technologies.

Assistive technologies for cognition

According to a survey on computer-based *"cognitive prosthesis"* [6], conducted in 2011, four categories of assistive technologies for cognition can be defined, depending on which aspects of the patient's daily living are supported: daily activities (this includes washing hands or going to the bathroom for instance), social life, entertainment and creative activities. More broadly, cognitive assistive technologies aim to support:

- access to information, communication and social connexions;
- mobility, planning and daily tasks;
- needs for stimulation and entertainment;
- management of psycho-behavioral symptoms.

This classification is not the only one available. In an other review, Gillespie *et al.* propose to classify assistive technologies depending on which cognitive function is supported [78], according to the ICF classification of disabilities mentioned in Section 1.1.

For each of the domains of activities in which cognitive support is required, many devices and applications have been proposed. For instance, the CIRCA system, cited by [6], aims at supporting social exchanges of people with dementia [7]. It consists in a software application on a touch-screen computer that acts as a conversation mediator between the person with dementia and the caregiver. Thanks to reminiscence material, in the form of pictures, musics and videos, they can find a subject of conversation and, thanks to visual and auditory information, the person with dementia remembers the topic more easily, which limits the risk that he or she loses track of the conversation.

One of the most successful examples of cognitive prosthesis is COACH [29]. It is a system that aims at guiding people with dementia, step by step, through basic activities of daily living. The activity used by the authors in their studies is hand washing. Thanks to computer-vision-based activity recognition, vocal prompts are sent to the person to guide him or her through the task.

Lastly, one could not write about cognitive prosthesis without mentioning *ambient intelligence*. The idea of ambient intelligence is to move away from locating assistance in one device but, instead, distributing it in the whole living space. This is closely related to the idea of assistive smart homes. For cognitive support, it would consist in having contextualized help available in all the rooms of people's homes. For instance, COACH-like systems could be located in the bathroom, bedroom, kitchen and near the front door and help people washing, dressing, cooking or making sure they wear enough clothes to go outside while reminding them not to forget their keys. Hamada *et al.* have even proposed a similar system for assisting people with dementia when they go to the bathroom [82]. A smart home could also include a reminder system for planning, medication or grocery shopping (i.e., an intelligent fridge could notify the system when it is empty). A recent survey of ambient intelligence in healthcare can be found in [2].

In the context of this thesis work, smart homes are a potential target application for our conversational agent. Indeed, prompting in an ambient intelligence scenario could be performed by a virtual butler, which could be displayed on screens located in the aforementioned locations, instead of a disembodied voice, as in COACH [29].

Rehabilitation technologies

Rehabilitation is critical to the wellbeing of older adults as incomplete recovery after a fall or a stroke may cause disability and consequently engender a decrease in quality of life and high care costs. It can be dedicated to cognitive (for instance, speech therapy in case of aphasia) or physical activities (for instance, physiotherapy in case of hemiplegia). Rehabilitation is not something new and it is fully included in standard care practices. Indeed, all hospitals have a physiotherapy staff, and rehabilitation professionals are present in nursing homes, private practices and some do home visitation. Physiotherapy is based on performing muscular activities that need to be repeated regularly (on a daily or weekly basis) and for an extensive time period (usually several weeks or months). One of the key factors to an effective rehabilitation is therefore the patient's active participation and motivation to perform the exercises as frequently and for as long as necessary. Rehabilitation does not always require technology but there are two important ways in which it can be eased by high-tech devices: enhancing patients' motivation and improving work conditions for professionals.

Regarding the first aspect, many researchers have investigated the use of serious games and gamified activities to improve long-term patients' motivation. Indeed, the very principles of gamifying such activities rely on the notion of *intrinsic motivation*, defined by Deci and Ryan [53] in their Self-Determination Theory, which can be harnessed to design motivating games [56]. To sum-up, intrinsic motivation depends on a balance between the need for challenge and the need for competency. This means that patients should be proposed rehabilitation exercises of increasing difficulty, so they can first be rewarded by achieving lowly ambitious goals before moving on to more challenging exercises as they regain motor strength and balance, which makes them more competent. To be more specific, according to a literature review by Lohse *et al.* [118], 6 key factors are responsible for the engaging nature of rehabilitation games: reward, difficulty/challenge, feedback, choice/interactivity, clear goals and mechanics, and socialization. The same review concluded that several studies have shown that gamification of rehabilitation activities improved patients' motivation [118].

Regarding the second aspect, physiotherapy is a very physical work. Indeed, professionals spend a lot of time helping patients to get up, pushing or pulling patients' limbs to make the exercises easier or more difficult or even almost carrying patients to reduce the weight on their legs so they can make a few steps. This is exhausting and may affect their health. In addition, when carrying a patient is required, it mobilizes two or three professionals at the time. To reduce this burden, researchers have proposed mecatronic systems, such as exoskeletons to do gravity compensation or make the muscular exercises harder. For instance, such systems were proposed in [146] and [14] to facilitate walk rehabilitation in older adults.

Some projects even make use of both aspects: for instance, Frisoli *et al.* combined an upper limb exoskeleton that reduces the necessary motor strength to perform movements with virtual reality to create a post-stroke rehabilitation program [71]. By proposing attainable goals of progressive difficulty and progressively reducing the contribution of the exoskeleton to the movement, the system aims at stimulating the patient's motivation and being more effective than ordinary physiotherapy.

1.2.2 Truly adapted technologies to achieve good usability and acceptance

One of the main issues faced by assistive technologies is their adoption by older adults. Indeed, many older adults that could benefit from assistive technologies do not use them. There are a number of reasons why: lack of available information about the existing technologies that could help; some devices are too expensive for many people; using assistive technologies is considered as stigmatizing; people do not want to use a device because they do not understand it and are afraid they will not be able to use it; etc. As a result, many older adults do not get equipped with devices that could improve their quality of life. Once people know about a device that could help them and that they can afford, its adoption depends on two factors: *usability* and *acceptance*.

For a technology to be usable it has to be easy enough for the target users to use it. This condition may seem easy to fill but usability design issues can prove quite tricky as soon as people get old and/or have some form of impairment, as it will be explained in the next chapter. Typically, personal computers are easy to use for people who were born after their invention and with no disability, but visually impaired people, even young ones, cannot use them unless special arrangements are made. Older adults may struggle learning how to use computers, because they are so complicated, and, though touchscreen computers are a more user-friendly alternative that many older adults have adopted, people with cognitive impairment will still struggle learning the procedures to connect the device to the Wi-Fi, download applications, use each application that they have downloaded, etc. This is why devices and software have to be tailored to the special needs of these people, the ideal case being a device that requires no learning and with a shape that clearly suggests what it is for. In that regard, Embodied Conversational Agents (ECA), such as the one studied in this thesis, seem like a good way of interacting with computers for older adults with cognitive impairment: they look like people (or humanoids) and talk, something people have done their whole life, which is natural and easy to understand. In addition, they can be active and start addressing people to engage them, which is an advantage, as people with dementia often have a deficit for initiating actions.

Technology acceptance is the willingness to use a given technology on the short and long terms. A Technology Acceptance Model was proposed by Davis [52] in an attempt to identify the factors that influence technology acceptance, in the context of the introduction of information technology in the workplace, in 1989. This first study concluded that *perceived usefulness*, that is to say "the extent to which a person believes that using the system will enhance his or her job performance" [209] and *perceived ease of use*, which is "the extent to which a person believes that using the system will be free of effort" [209], were the most correlated factors with self-reported technology use and intentions of future use. This model was found to be applicable to the case of older adults in a literature review of empirical evidences by Chen and Chan [40]. However, a shift in the definition of perceived usefulness has been observed: for older adults, it corresponds to "believ[ing] and realiz[ing] that those technologies might be used to improve their lives and satisfy their needs" [40]. Therefore, a technology is perceived useful if they feel like they actually need it. For perceived ease of use, according to Chen and Chan, the definition remains the same as in the original model, but is more strongly correlated with technology adoption in older adults than in their younger counterparts. In addition, it was observed that the perception of ease of use by elders emphasizes task success whereas younger adults consider task completion time more important. Lastly, it can be noted that for older adults perceived ease of use is a predictor of perceived usefulness.

To design devices or software that have good chances of being adopted by the target users, one has to well identify their needs and use capabilities. One way of doing that (which we believe is the best) is to involve target users at all stages of product design, from ideation to deployment. This is called *user-centered design*. However, target users may not always be the only people to make the decision of buying a product or not. For instance, people with dementia might not be able to make such decisions for themselves anymore. In this context, the caregivers' opinion about the product is just as important. Another example is medical technology adoption in hospitals: the care staff are the users but cannot decide on their own of using a technology, for it is the hospital's administration that is responsible for acquiring equipment.

For these reasons, the participatory design methodology adopted in this research work goes beyond user-centered design and involves other stakeholders in the design process than just end users. To be more specific, the method that was used is the *living lab* method [22, 4, 160]. Like user-centered design it relies usability testing, interviews, questionnaires and focus groups with target users to adapt the design to their capabilities and better target their needs, but the technology is also shown to and discussed with other stakeholders, such as caregivers, physicians and health professionals, insurance companies or even public decision makers. This research work does not only follow the guidelines of living lab participatory design but contributes to the development of this approach by proposing an implementation of the living lab principles and providing some feedback on lessons learned in the process. For instance, applying user-centered design principles when dealing with people with cognitive impairment can prove tricky, as they have difficulties in expressing their needs and wishes. This is why our feedback on that matter can be useful to other designers. The details and principles of the method are presented in Section 3.2.

1.3 Goals

The goal of this work is to explore the possibilities offered by video game technologies, especially virtual humans, to provide care or assistance to older adults with dementia or psychomotor disorders. To this aim, two case studies were conducted in the form of the participatory design and evaluation of two prototypes, conducted in a living lab that has the particularity to be located on the premises of a geriatric hospital. The first case study is a project called LOvely User Interface for Servicing Elders, or LOUISE, an embodied conversational agent (ECA) that is meant to be an adapted user interface for older adults with dementia. The second case study is a project called Virtual Promenade, a walk simulator that combines a virtual strolling game, in which players control a human avatar to visit virtual environments, with a haptic chair featuring a moving seat that imitates the movement of the hips in human walk. It is intended for virtual reality exposure therapy to treat post-fall syndrome (PFS) in older adults.

These two projects make a different use of virtual humans: in LOUISE, the virtual human is used as non-self, as it is the embodiment (though it is a virtual one) of an artificial agent, capable of conversation-like interactions with the user; in Virtual Promenade, virtual humans are used as self, in the form of avatars, which represent the player in the virtual world. Although our case studies have quite different purposes, they have a lot in common: they both make use of game engines and 3D imaging; they both address older adults; they both have to be adapted for use by people with cognitive impairment; they both have to be cheap; and they both are designed, implemented and tested, using a living lab user-centered design approach based on fast prototyping and short iteration cycles.

1.3.1 Technology that complements human help but does not replace it

Technologies such as personal robots and ECA sometimes provoke the fear that they will replace humans and that jobs will be lost. While it may be true for robots in the context of product manufacturing, our view of ECA technology in the context of elders' care is quite different. To be clear, we do not think that it can replace human help nor do we design them for that purpose. Indeed, older adults, just like any human being, have a need for social contact, and nothing, in our opinion, can replace human contact, especially touch or speech.

Instead, the goal here is to give to people with dementia access to services that they would not get otherwise because caregivers do not have time to spend on providing them and that, thanks to cognitive prosthesis, help people regain some autonomy and self-esteem. For instance, in care institutions, people may not get to choose what they eat anymore because caregivers do not have time to spend with them to present the day's menu to them so they can have that choice. An ECA has no time constraint and could perfectly perform this task. Another example is repetitive questioning: people with dementia, as they are disoriented, often feel anxious about something and get obsessed with it. Thus, they ask caregivers the same question over and over again, as they forget the answer shortly after they obtained a response. This can become exhausting for the caregiver, but people with dementia need answers to their questions. In that context, the ECA could try to answer that question to reassure the person, and also provide orientation information (time, day of the week, location). We believe that such ECA-based applications could significantly improve patients' quality of life, as denying people the possibility of making choices and not giving answers to questions that are of critical importance for their peace of mind is close to mistreatment. This can also reduce the caregivers' burden and stress. Indeed, caregivers with patients who feel better, who will therefore be less agitated and less reluctant to care, and who are more autonomous will be more available to take better care of them. We therefore believe that the introduction of ECAs in nursing homes or patients' homes could create a virtuous cycle that will benefit both patients and caregivers.

Another view of the kind of technologies presented in this thesis, which applies to Virtual Promenade, the second case study, is that they are also meant to be used by healthcare professionals in their interventions. As stated earlier, exoskeletons and rehabilitation games may be used or prescribed by physiotherapists in their practices. Used in that way, technologies can improve quality of care and work conditions for care professionals. This could even lead to the creation of new professions: there are animal therapists and music therapists, so why not robot therapists? In fact, this could soon become a reality, as the use of robots in elderly care has already shown very encouraging results. For instance, a Japanese baby seal robot, called Paro, was evaluated positively for dementia care in several studies [183].

Having clarified our view of new technologies in elders' care, let us now present in more details the specific goals for this research, which is based on two case studies: virtual humans as self (Virtual Promenade) and virtual humans as non-self (LOUISE). Lastly, it is worth noting that these two uses are not mutually exclusive and could be combined in a single application.

1.3.2 Embodied conversational agents as user interface for older adults with cognitive impairment

Embodied conversational agents (ECAs) are a special kind of software agents with social interaction capabilities. ECAs are computer programs that are able to interact thanks to user input acquired through a set of sensors (cameras, microphones, touch screen, etc.) and a virtual character, usually displayed on a screen, able to produce human-like communication behaviors by speaking and gesturing through recorded-voice playback or speech synthesis and character animation. These systems depend on several technology domains including, but not limited to, speech recognition, speech synthesis, computer graphics, computer vision and artificial intelligence. In addition, realistic imitation by a computer system of natural human communication requires knowledge from several social sciences including psychology, linguistics and anthropology.

Though research efforts on animating ECAs and sensing "social signals" [156] have been intense in the past decade, there is still a long way to go for computer programs to be able to interact in a fully natural way. Many research questions are yet to tackle: What should the ECA look like? More specifically, should it be anthropomorphic or animal-like? Should it be realistic or cartoon-like? Should it look like a male or a female? We could go on, and this is just regarding the embodiment, that is to say, the virtual character's appearance.

As stated earlier, ECAs present several advantages, which will be detailed in Section 2.2.1, to serve as a user interface for computer-based assistive technologies targeting older adults with dementia: they use verbal and nonverbal communication, which does not require learning; they look humanoid, which suggests how they are used; they are easy to understand; etc. In addition, several studies (see Section 2.2.1) have already shown encouraging results in terms of efficacy and appreciation in the context of older adults with cognitive impairment. Moreover, ECAs are a cheap and ubiquitous technology that could be deployed in several rooms of patients' living space in a smart home scenario.

Our main goal in this first case study is to design an ECA-based user interface, called LOUISE (LOvely User Interface for Servicing Elders), that could be included in various assistive technologies, following a living lab incremental and participatory design methodology. For instance, there are electronic pill dispensers that ring and blink LEDs when it is time to take medicine and will not stop ringing until the box is turned upside down to take the pills inside. With LOUISE, the ring would be replaced by verbal and nonverbal prompts, which are more explicit and nicer than a ring which people with dementia may not easily associate with medication time. Then, when the pill dispenser is flipped upside down, LOUISE could receive a message that the pills were retrieved from the box and move on to guide through the task of taking these pills. Furthermore, LOUISE could feature activity detection and attempt to make sure that the person actually took the pills and call help when in doubt. LOUISE is in fact envisioned as a toolbox, which, thanks to an interaction scenario editor and the possibility of communicating with external software applications or devices, can be easily used as user interface in several assistive scenarios, such as the ones we have just mentioned. In addition, it should be easy to personalize to the tastes of the user, for instance by featuring several virtual human models for its embodiment and making it easy for developers to add extra models. Lastly, the system should be cheap, so that most patients can afford it. This can be achieved by privileging the use of open-source or free software and consumer-grade devices to build the system. Using publicly available software and repurposing commercial-grade devices also have advantages in our context, as it allows for fast and cheap prototyping.

Our secondary goals are: to study interactions between older adults with dementia and such systems to propose an adapted, general-purpose, interaction management system which yields good usability; and get information about the tastes and needs of dementia patients that influence ECA acceptance, in order to gain interface and embodiment design insights and identify the most useful applications in which our ECA could be included.

1.3.3 Virtual reality exposure therapy for post-fall syndrome

Exposure therapy is a method of behavioral therapy, theoretically grounded on the emotional processing theory [50]. It can be defined as "the process of helping a patient engage in repeated and prolonged contact with a feared stimulus" [1]. This kind of therapy is used for phobias, post-traumatic stress disorders (PTSD) and generalized anxiety. It consists in confronting the patient, in safe conditions, with the object of their fear. In some situations, such as aviophobia (fear of flying) or combat-related PTSD, exposing people to the stimulus they fear is more difficult than when the stimulus is easy to obtain, as in arachnophobia (fear of spiders) for instance. This is why many researchers have investigated the use of virtual reality to perform exposure therapy, thus replacing real stimuli, such as spiders, snakes or clowns, with virtual ones. This is called virtual reality exposure therapy (VRET). VRET has been proved successful in many cases [75].

Post-fall syndrome (PFS) is an anxiety disorder resembling PTSD, but linked to the fear of falling; it may also be accompanied with psychomotor disorders called psychomotor regression syndrome, that is to say a loss of cognitive motor control functions (see Section 2.1.3). Treating PSF can therefore be seen as a form of cognitive rehabilitation, addressing both anxiety and psychomotor aspects. In that regard, as mentioned in Section 1.2.1, serious video games would be an interesting technology for rehabilitation, in terms of motivation enhancement. In addition, the psychological and psychomotor consequences of falls are overlooked in current care practices.

For these reasons, the primary goal of our second case study is to investigate the participatory design of a system involving virtual reality, in the form of a serious game, and a haptic chair to address both pathological aspects of PFS. This system has the particularity of having two end users: patients, obviously, who will actually play the game; and also care professionals who would include the system in their practices. For that reason, our participatory design will have to involve physiotherapists and other relevant care professionals. As for LOUISE, the system should be cheap and easy to enhance with new content.

The secondary goals are: to identify what aspects of the game design are important for older adults to adhere to it and be motivated to keep training; to investigate the controls of the game and their suitability for the target patients; and to assess the feasibility of deploying the proposed system in highly constrained care environments and its acceptance by professionals and patients.

1.4 Contributions

The main contributions of this work are the following:

- We designed, implemented and tested a new ECA system, called LOUISE, that could be used as user interface in assistive technologies for older adults with cognitive impairment. In particular, the originality of our design lies in the following features:

    1. a simple yet effective algorithm for the attention monitoring of the user;
    2. a conversation manager for task-oriented interactions that reacts to inattention by prompting the user to recapture his or her at-

tention and performs context reminders for him or her to be able to resume the conversation without losing track of the topic;

3. a basic XML-based domain-specific language to specify task-oriented interaction scenarios;
4. the possibility of augmenting verbal and nonverbal behaviors of the avatar with images to present items and videos to show examples of actions, thus creating information redundancy and multimodality to compensate for the aphasic and agnosic disorders as well as partial hearing impairment that people with dementia may have;
5. the possibility of adding extra characters for personalization purposes.

- After extensive experiments with all stakeholders, we formulated guidelines to create task-oriented interaction scenarios that would include LOUISE in assistive applications for older adults with cognitive impairment.
- We designed, implemented and tested Virtual Promenade, a system combing a virtual strolling game and a haptic chair for virtual reality exposure therapy to treat post-fall syndrome in older adults. It is cost-effective, achieves good usability and acceptance for patients with or without cognitive impairment and was co-designed with health professionals (physicians and physiotherapists).
- Using the outputs of our two systems, we provide insights for the design of virtual human-based assistive technologies for older adults, gained through the participatory design practice.
- We contribute to developing the living lab design methodology by complementing existing practices with ideas from the lean startup and do-it-yourself movements, which translate in fast prototyping, thanks to the reuse and repurposing of existing software and hardware elements, and short iteration cycles, by making changes to the prototypes between trial sessions of the same evaluation study.
- We provide feedback on lessons learned in applying these innovative design principles within a living lab, located inside a geriatric hospital, and involving people with cognitive limitations, who have difficulties in speaking out their needs and wishes.

1.5 Organization of this thesis

This manuscript is organized in 8 chapters. In this introductory chapter, we have seen that, due to aging demographics in developed countries, the number of older adults with disabilities caused by cognitive impairment and falls is increasing while the number of young people able to take care of them is decreasing. We have also seen that, in this context, assistive and rehabilitation technologies could contribute to address this issue and improve the quality of life for patients and their caregivers, while reducing costs. However, technologies to fully address cognitive impairment and fall rehabilitation are not yet available, although promising results have been observed when using computer-based technologies, such as ECAs or serious games for rehabilitation. In addition, the issues behind the adoption of technology by older adults were presented and it was emphasized that participatory design approaches could improve the situation. The context and motivations for this research having been introduced, we propose to investigate ECA-based user interface for older adults with dementia and virtual reality exposure therapy for post-fall syndrome through participatory design. Lastly, the specific goals pursued by each of these projects were presented and the main contributions of this work were highlighted.

Chapter 2 aims at presenting and organizing the relevant information to understand the design constraints for these projects and their positioning in the field. Firstly, this chapter presents in details the aspects of aging, of the pathologies we address and of the environments that have to be accounted for in our designs for them to be adapted. Secondly, the literature about the use of ECAs and serious games in the context of older adults with disabilities is reviewed and commented.

In Chapter 3, the main theoretical concepts involved in this work are defined and discussed. After that, the methodology and principles of the living lab participatory design approach we adopted in this work are detailed.

In Chapter 4, we summarize the necessary information to give the reader a good understanding of how virtual humans are created. To be more specific, we start by introducing some notions about human communication and the field of social signal processing. After that, we present the typical modular architecture of the software in ECAs and social robots. We then provide more details about how verbal and nonverbal behaviors can be sensed, analyzed, interpreted and generated by a computer.

The next three chapters are dedicated to the two case studies we conducted. In Chapter 5, we present our first case study about LOUISE, which includes the participatory design of our prototypes and our experiments and findings. In Chapter 6 we report our experiments on several believable use cases for LOUISE. Chapter 7 presents the second case study. In this chapter we first present our prototype, then we report on its evolution through participatory design and on a validation experiment.

The last chapter (Chapter 8) is dedicated to the conclusions, recommendations and perspectives for current and future work. We first present the conclusions for our case studies and a general conclusion. Then, recommendations for the use of the living lab participatory design methodology are proposed and ideas for future work are presented, before ending with some ethical considerations about the use of virtual humans in elders' care.

Chapter 2

Game technology for older adults

Résumé en français. *Les personnes âgées pourraient bénéficier d'un nombre croissant de technologies d'assistance (TA) ; cependant, il existe des obstacles à l'adoption répandue de ces technologies, en termes d'utilisabilité et d'acceptabilité. Comme mentionné dans le précédent chapitre, il y a également de bons exemples de méthodes de conception participative qui pourraient contribuer à une meilleure adoption des TA.*

Lorsque l'on conçoit des produits ou services à destination des personnes âgées, plusieurs déficits physiques, cognitifs et sensoriels, causés par l'âge ou par des pathologies fréquentes, doivent être pris en compte, afin de proposer des dispositifs adaptés. De plus, il y a un important compromis à trouver entre conception universelle/conception pour tous/une-taille-va-à-tous et une approche personnalisée. Dans le premier cas, il s'agit de produire une conception unique, prenant au maximum en compte tout ou partie des déficits que sont susceptibles de présenter les utilisateurs ; dans le second cas, il s'agit d'introduire un certain nombre d'éléments de personnalisation qui permettent d'adapter le dispositif pour un individu donné. Bien que l'approche "conception pour tous" semble plus facile à mettre en œuvre que l'approche personnalisée, elle risque de conduire à un produit ne convenant à personne. Considérons les prothèses auditives, par exemple : il est pratiquement impossible de produire une prothèse à taille unique qui conviendrait à tous. De plus, toutes les limitations possibles dans l'utilisation d'un produit ne peuvent être prévues dans le même produit ; il y aura toujours des personnes qui ne pourront pas s'en servir. L'approche personnalisée n'est pas parfaite non plus. Tout d'abord, il est difficile de produire un dispositif adaptatif. En second lieu, adapter un dispositif à un utilisateur particulier peut nécessiter l'intervention d'un professionnel qualifié. Pour reprendre l'exemple des prothèses auditives, un audio-prothésiste doit systématiquement intervenir pour adapter la forme de la prothèse et régler l'appareil en fonction des troubles

auditifs particuliers de chaque utilisateur. La solution idéale pourrait alors consister à proposer un produit adapté pour la plupart des gens, avec quelques éléments personnalisables, qui s'auto-configureraient. Justement, dans le cas des technologies d'assistance virtuelles, ce but semple atteignable.

Dans ce chapitre, les limitations liées à l'âge impactant l'utilisation de produits sont présentées et des détails sur les pathologies étudiées dans ces travaux sont fournis, ainsi que des informations sur les contraintes liées aux environnements de soin que nous ciblons, avec l'un, au moins, des dispositifs conçus pendant cette thèse (Virtual Promenade), ce afin d'identifier les défis de conception. Ensuite, nous dressons un panorama des travaux sur l'utilisation des agents conversationnels animés (ACA) auprès des personnes âgées et des jeux pour la rééducation post-chute.

Dans la première partie, nous présentons les déclins cognitifs, physiques et sensoriels liés à l'âge : audition, vision, perception kinesthésique, mémoire de travail, cognition spatiale, ralentissement, perte de précision des mouvements et arthrose fréquente. Nous distinguons ensuite deux types de pathologies cognitives liées à l'âge : les troubles cognitifs légers [157] et la démence, dont nous présentons la cause la plus fréquente, la maladie d'Alzheimer. Dans le cas des troubles cognitifs légers, la mémoire à court terme, la mémoire de travail, l'attention, la cognition spatiale et la compréhension du langage peuvent décliner de manière pathologique. Dans le cas de la démence, ces domaines cognitifs peuvent être plus fortement affectés encore, avec en plus des symptômes tels que l'aphasie, l'agnosie, l'apraxie, ainsi que des troubles exécutifs et psycho-comportementaux. Nous détaillons ensuite les symptômes du syndrome post-chute qui se caractérisent principalement par de l'anxiété, une peur de tomber pathologique et un syndrome de désadaptation psycho-motrice [141]. Concernant les environnements de soin, ils se caractérisent par une grande rigidité, liée aux exigences d'hygiène et de sécurité, et par un personnel et un budget restreints. Enfin, nous concluons cette première partie en observant que, bien que les troubles cognitifs de la démence peuvent rendre notre tâche de conception ardue, ils peuvent également, sur le plan technique, faciliter certains aspects, car des applications simples peuvent rendre de grands services.

Dans la seconde partie, nous commençons par lister les principales applications proposées pour l'utilisation des ACA auprès des personnes âgées (assistant, coach, compagnon, majordome) ainsi que les principaux avantages attendus des ACA, par rapport à d'autres modes d'interaction (naturalité, capture de l'attention, facilité de compréhension, crédibilité, motivation et efficacité). De plus, cette revue de l'état de l'art conduit à observer que peu de travaux ont proposé des dispositifs complètement automatisés avec lesquels l'utilisateur interagit de manière verbale et non-verbale. Dans un second temps, nous présentons les travaux sur les jeux de rééducation destinés aux personnes âgées et constatons qu'il existe un très grand nombre de travaux sur le sujet. Les principales informations à retenir sont *que l'efficacité de ces jeux en termes de motivation des patients ressort de plusieurs études, que certaines études s'appuient sur des solutions complètes existantes et que d'autres développent des solutions ad hoc, soit en réutilisant du matériel existant, au niveau des commandes notamment, soit en développant des outils dédiés. Les dispositifs vendus dans le commerce les plus utilisés sont la balance Wii Fit et le tapis de danse Dance Dance Revolution. Enfin, nous avons identifié que quelques équipes de recherche ont adopté un démarche de co-conception [110], voire de co-création [204], impliquant des personnes âgées.*

Older adults could benefit from a growing number of assistive technologies; however, there are obstacles in terms of usability and acceptance to the widespread adoption of these technologies. As stated in the previous chapter, there are also some successful examples of design and participatory design methods that could help achieve better adoption.

When designing for older adults, several physical, sensory and cognitive limitations caused by aging and frequent pathologies have to be considered to produce adapted products. In addition, there is a very important trade-off to make between a universal design/design-for-all/one-size-fits-all approach and a personalized approach. In the first case, one has to produce a design that will account for most or all possible limitations people may have; in the second case, one has to propose a number of personalization features that can be adjusted to adapt a product to a particular user. While the design-for-all approach seems a lot easier to manage than the personalized approach, the risk is to end up with a suitable-for-no-one product. Consider hearing aids, for instance: it is almost impossible to create a one-size-fits-all prosthesis. In addition, not all possible limitations in product use can be addressed in the same device, as there will always be some people who will not be able to use it. The personalized approach is not perfect either. Firstly, it is very challenging to create an adaptable product. Secondly, adapting the product to a particular user may require professional expertise. To take the example of hearing aids again, a professional audiologist has to intervene to adapt the shape of the prosthesis and configure it to fit the hearing loss compensation needs of each patient. The ideal solution might then be to propose a design that is adapted for most people with a few, self-configuring, personalization features. Fortunately, in the case of virtual assistive technologies, this goal seems achievable.

In this chapter, age-related limitations for product use are presented and some details about the pathologies addressed in this work are provided, as well as some information about the hospital care environment we target for at least one project (Virtual Promenade), to identify the design challenges. Then, a review of existing works about embodied conversational agents (ECAs) for older adults and games for fall rehabilitation is presented.

2.1 Challenges in designing for older adults with special needs

It is a common mistake to consider memory loss and disorientation as part of the “normal” aging process when, in fact, these symptoms of cognitive impairment are caused by brain diseases, such as Alzheimer’s disease. The purpose of this section is to identify the physical, sensory and cognitive limitations experienced by older adults in normal aging and in the pathological cases addressed in this work, namely cognitive impairment and post-fall syndrome (PFS).

As already stated in the previous chapter, it is worth keeping in mind that older adults are often affected by more than one pathology. In fact, part of the work of physicians in geriatric care is to decide which pathologies should be treated in priority and to limit as much as possible undesirable side-effects caused by the interactions between all pharmacological treatments received by patients. In our case, PFS and cognitive impairment are not mutually exclusive, and can even be accompanied with other pathologies, such as hearing impairment or osteoarthrosis.

2.1.1 Normal aging

Aging is usually accompanied with subtle declines in cognitive and executive functions. In [67], the authors identified what factors in aging may affect product use and show that ergonomics are more critical when designing for older adults than for younger people. A summary is given here.

Sensory functions

All sensory functions tend to decrease with growing age, though there may be important inter-individual variability. Although all five senses are affected, only vision, hearing and haptics, which are the most relevant in the present work, are detailed here:

Audition By age 65, 50% of men and 30% of women suffer hearing loss, which makes social interaction more difficult. For most people over 70 years old, frequencies above 4kHz become inaudible. Comparatively, the frequencies of the human voice are roughly between 100Hz and 3.5kHz. The standard telephone bandwidth is 4kHz. Hearing impairment is considered severe when an individual’s hearing threshold is higher than 35dBA (acoustic decibels).

Vision Though vision impairments affect many people regardless of age, their prevalence increases with age. 70% of people over 45 need to wear glasses. With age, people tend to develop presbyopia, the inability to see close objects properly. In addition, the eyes’ ability to adapt to darkness and the breadth of visual field decrease. Lastly, visual information processing slows with age.

Haptics The term *haptics* refers to the senses of touch and kinesthetics[1]. Regarding touch, older adults may become less sensitive to temperature and vibrations. A decline of kinesthetic sensitivity is also observed in aging; this partly explains why older adults have a less stable gait than young people and are more likely to fall.

Given these elements, it is necessary to make sure that the output sound is loud enough for most elderly people to hear. In addition, the low frequencies for voices and sounds should be privileged. Lastly, large display sizes should be preferred for people to see properly. That being said, in the context of cognitive impairment, visual acuity might not be the most limiting factor to select display size given the attention disorders affecting dementia patients. Regarding haptics, movement amplitude and vibration intensity should be large enough for older adults to be able to sense them.

Cognitive functions

In addition to sensory declines, cognitive functions also change when people age. Here, the terminology is introduced and some details are given, based on [67].

The main cognitive processing components involved in the interaction with products are the working memory, semantic memory, prospective memory, procedural memory, attention, spatial cognition and language comprehension.

Working memory , as in a computer, is an active memory. It contains information that has just been perceived or retrieved from long-term memory.

Semantic memory , to continue with the computer analogy, is somehow equivalent to the hard drive. It is a long-term memory of acquired knowledge.

Prospective memory corresponds to the ability to plan actions in the future. It allows remembering that an action should be performed at a certain time (meeting Mr. X at 2 p.m.) or what to do in response to an event (turning off the oven when the buzzer goes off, for instance).

Procedural memory is knowledge about how to perform learned task procedures: for example, when riding a bicycle, pushing the pedals harder to go faster and pulling the brake handles to slow down. Some of these actions may even become somehow automatic.

[1] the sensations of movement, balance and muscular activity

Attention can be defined as the process that controls awareness. It is limited and operates selectively on stimuli of the environment. It can be divided between the sources of information and switched between tasks.

Spatial cognition is the ability to manipulate images or patterns mentally, to project into a map for instance.

Language comprehension is the ability to interpret written or spoken verbal information.

Only some of these cognitive functions show age-related declines. For memory, working memory declines the most and prospective memory is affected in a less visible manner. On the contrary, semantic memory shows minimal decline, though the access to the information it contains may be slower and less reliable. Procedural memory is also very well preserved for well-learned procedures, but older adults are slower and less successful at acquiring new procedures than younger adults. In addition, spatial cognition declines with age whereas language understanding is not affected.

Regarding attention, older adults may experience difficulties focusing on one particular stimulus when surrounded by other stimuli. This gets worse as the number of stimuli increases. In addition, they require more time than younger people when switching attention from a stimulus to another. As a result, elderly people may get distracted when too many stimuli are presented at the same time and perform poorly when a situation requires multitasking.

Movement control

Older adults tend to move slower than their younger counterparts (up to two times slower) and their movements are less precise. In addition, many older adults suffer from osteoarthrosis, which results in pain, loss of movement, swelling and stiffness in the joints [152]. This condition is not part of normal aging per see but is frequently associated with aging. These factors have to be considered when designing user interfaces for computer applications and selecting interaction devices for games.

As said earlier, some or all of these declines may occur, at varying degrees, in healthy older adults. In addition to these age-related declines, which are mild and can be addressed by careful design, many older adults are affected by medical conditions that cause physical or cognitive impairment, sometimes both. This thesis work focuses of two cases: neurodegenerative diseases, which is a quite broad domain, and post-fall syndrome, which is a much more specific condition. In the remainder of this section, the symptoms of these conditions and their foreseeable impact on product design are presented.

2.1.2 Cognitive impairment and dementia[2]

Cognitive impairment in elders is usually related to damaged structures in the brain. It is these damages that cause memory losses and other psychological and behavioral symptoms that affect performance in daily activities and social aspects of everyday life. In the following, the most frequent pathologies and their symptoms that affect the ability of an individual to perform activities and use technologies autonomously are detailed. The paragraphs about the symptoms of mild cognitive impairment and Alzheimer's disease are a synthesis of the works of Pino [159], Benveniste [19] and Lapointe *et al.* [111] who have worked on the ergonomics of assistive technologies for cognitively impaired older adults.

Mild Cognitive Impairement (MCI)

Mild Cognitive Impairment (MCI) [157] affects 10 to 20% of people more than 65 years old. Its symptoms consist mostly in memory losses but it may cause other cognitive deficits. It may evolve in Alzheimer's disease after a few years. Two types of MCI can be distinguished: Amnestic MCI (a-MCI) and Non Amnestic MCI (na-MCI). In the first form, episodic memory, that is to say the memory of past events, is affected. This causes deficits in learning and recalling recently acquired information. In na-MCI, episodic memory is not affected but abnormal declines are observed in attention, spatial cognition or language comprehension. In addition, MCI patients can suffer from impairments in one or several of the cognitive functions listed above, regardless of the a-MCI/na-MCI classification. These patterns were established by Winblad *et al.* in [223].

Regarding the ability to perform activities of daily living independently and to use technology products, older adults with MCI have it more difficult than cognitively healthy older adults. In the case of a-MCI, learning how to use a new product may get very difficult. Though MCI patients maintain functional independence, the pathological declines of their cognitive functions cause them to be slow. They may experience difficulties in planning and completing complex tasks. In addition, MCI patients usually have poor judgment. All of this results in a loss of effectiveness in everyday functioning.

Diagnosis of MCI

Currently, MCI is diagnosed by using established neurological and neuropsychological tests, and the following set of clinical criteria [157]:

1. subjective memory complaint – this information should be provided either by the patient or by a knowledgeable informant;

[2]Thanks to Anne-Sophie Rigaud for her help in writing this section.

2. objective memory impairment, characterized by impaired delayed recall performance or difficulty benefiting from semantic cues during learning or recall;

3. normal general cognitive function on measures of general cognition and other nonmemory indexes;

4. intact functional ability, that is to say skills and capacities required to perform activities of daily living (ADL) independently are preserved, though the person can face some difficulties when performing complex instrumental ADL and may require minimal aids or assistance;

5. absence of dementia, i.e. cognitive deficits do not affect significantly functional or social capacities.

However, more recent recommendations consider that the best prediction models should involve a combination of neuroimaging and chemical biomarker measures [5].

Alzheimer's disease (AD)

Alzheimer's Disease (AD) is a neurodegenerative disorder. It is characterized by a progressive pathological decline in cognitive functions. It is the most frequent cause of dementia in older adults as it represents 60 to 80% of dementia patients [159]. Other types of dementia mostly differ by their physiological causes but have more or less the same effects on the person's daily functioning. Here are listed the most common symptoms.

Short-term memory loss AD patients are unable to recall recent events, an ability linked to working memory and episodic memory. This results in repetitive questioning and conversations, recurring loss of personal belongings, forgetting events and appointments, getting lost on a familiar route or not being able to pick up where they left off when interrupted during a task.

Executive dysfunction Demented patients experience great difficulties in planning and executing the sequential steps of a complex task. Even though they are able to execute each of the subtasks involved, they have trouble finding the right order. For instance, they may go into the shower without taking their clothes off. This is linked to a deficit in prospective memory.

Attention disorder Demented patients get distracted very easily and often have trouble focusing on a long task. Visual attention impairment, in particular, makes it very difficult to follow fast-changing images.

Aphasia It is the inability to understand or to produce spoken or written language. This means that AD patients may be temporarily or permanently unable to recognize or produce written or spoken words they normally know. This may result in incoherent speech, misunderstanding, and other difficulties to communicate effectively with others.

Agnosia Older adults with dementia may present a deficit in recognition of visual or auditory stimuli not caused by a sensory deficit. People with AD may not be able to recognize the faces of their relatives, for instance. This can also cause difficulties to get dressed (putting clothes on back-to-front or inside-out), the inability to find objects in direct view or to recognize signals such as warning signs.

Apraxia This terms designates a deficit in voluntary motor control that is not due to a paralysis or motor weakness.

Psycho-behavioral disorders People with dementia usually experience anxiety, agitation, emotional instability, apathy and depression. This results in wandering and socially unacceptable behaviors.

It is important to note that Alzheimer's disease and other dementias have high interpersonal variability. Although the same symptoms are usually observed in most patients, the levels of severity may vary depending on the patient and the progression stage of the illness. The disease's evolution can be very different in speed and intensity from a patient to another. Nevertheless, after several years, most patients become unable to live independently without receiving support from a caregiver on a daily basis. In addition, intrapersonal variability is also observed: a given person with AD usually has good and bad days and may even experience changes within the same day.

However, some functions are well preserved in spite of the illness, even at the advanced stages of the disease. Although demented patients experience great difficulties to consciously access the information stored in their memory, part of the information may still be present. Thus, they are able to achieve learning through implicit and non-conscious memory processes. Procedural memory is one of the best-preserved capabilities and people with AD are still able to perform well-learned routines. Interestingly, it is still possible to condition new behaviors in patients with AD, through repetition and exposure to recurring stimuli, thanks to implicit learning that is better preserved in dementia than explicit learning [105, 30].

Another aspect of cognition that has received much attention in the research on AD is the ability to process and produce nonverbal behaviors. Although verbal communication skills are very affected, nonverbal communication capabilities are well preserved [174]. This includes recognizing and expressing affects through facial expressions, prosody (vocal pitch, rhythm and energy variations), gestures and posture. Details on nonverbal communication modalities are given in Section 4.1. Lastly, although personality traits are altered in AD, people still have their personal tastes and interests as well as life-long habits.

This section should have given the reader some sense of the challenges of designing software for use by this public. Based on the elements given above, a solution could be to adapt the user interface, in a personalized way, based on the remaining capabilities of each patient, as suggested by Lapointe *et al.* in [111] and by Alm *et al.* in [6], and according to their personal preferences.

Diagnosis of Alzheimer's disease

Over the last decades, AD diagnosis has been made based on clinical judgment, by using established international criteria from the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and the Alzheimer's Disease and Related Disorders Association (ADRDA), presented in Table 2.1.

In addition to clinical examination, neuropsychological tests and brain imagery (Magnetic Resonance Imagery or MRI), new criteria for the diagnosis of AD incorporate biomarker measures, such as ß-amyloid and tau levels in the cerebrospinal fluid and blood for diagnosis confirmation at the preclinical stage. At later stages histopathological markers provide a definite confirmation of the clinical diagnosis of AD, whereas clinical criteria only support a probabilistic one [57].

Neuropsychological assessment of cognitive disorders

To ease comprehension of the following chapters, it is worth explaining a few facts about cognitive assessment. Cognitive assessment through psychometric tests is usually conducted by neuropsychologists or physicians. Neuropsychology is a discipline of clinical psychology; it consists in the study of cognitive abilities at the subject's level. In their clinical activities, neuropsychologists work with physicians (psychiatrists and neurologists) to help diagnose cognitive pathologies. They use standardized psychometric tests, which require patients to perform cognitive tasks; this is a way to highlight observable cognitive deficits. The most famous test is called Mini Mental State Examination (MMSE) [68]. This test, thoroughly validated through clinical trials in several languages and in several countries, allows to quickly[3] obtain an overview of a person's cognitive functioning and identify which cognitive domains are affected, if any. It contains 30 questions, each worth one point, which assess spatial and temporal orientation, short-term memory, working memory, attention, language and spatial cognition. The global

[3] five to ten minutes

Table 2.1: Core clinical criteria for Probable AD dementia [130]

Critieria for the clinical diagnosis of Probable AD include:	• dementia established by clinical examination and documented by the Mini-Mental Test, Blessed Dementia Scale, or some similar examination, and confirmed by neuropsychological tests; • deficits in two or more areas of cognition; • progressive worsening of memory and other cognitive functions; • no disturbance of consciousness; • onset between ages 40 and 90, most often after age 65; • absence of systemic disorders or other brain diseases that in and of themselves could account for the progressive deficits in memory and cognition.
Diagnosis of Probable AD is supported by:	• progressive deterioration of specific cognitive functions such as language (aphasia), motor skills (apraxia), and perception (agnosia); • impaired activities of daily living and altered patterns of behavior; • family history of similar disorders, particularly if confirmed neuropathologically; • laboratory results of normal lumbar puncture as evaluated by standard techniques, normal pattern or nonspecific changes in EEG, such as increased slow-wave activity, and evidence of cerebral atrophy on Computerized Tomography with progression documented by serial observation.

score, out of 30, allows to quantitatively estimate the severity of cognitive impairment [199]. For this reason, it is widely used to characterize the participants of a study involving cognitively-impaired older adults. However, it is worth keeping in mind that two people with the same MMSE score may in fact have a quite different cognitive functioning. Indeed, they may not have lost points in the same categories of questions and other factors, such as age, education and cultural background, can influence the MMSE score [199]. There are many other psychometric tests for cognitive assessment. Some tests serve similar goals as the MMSE; for instance the Montreal Cognitive Assessment (MoCA) test [145] is a more sensitive alternative, especially for detecting MCI. Other tests are much more targeted and aim at evaluating functioning in a single cognitive domain.

Research in neuropsychology partly consists in creating and validating such tests. It may also consist in evaluating the therapeutic effects of cognitive interventions on cognition. Neuropsychologists are also well qualified to perform observations in user studies. Lastly, their knowledge of the observable symptoms of neurodegenerative pathologies proves very useful to identify key design constraints when targeting users with cognitive impairment.

2.1.3 Post-fall syndrome

Falls in older adults can be attributed to several risk factors. According to Campbell *et al.*, the most frequently observed fall predictors in men are decreased levels of physical activity; stroke; arthrosis of the knees; impairment of gait; and increased body sway. In women, the dominant risk factors are the total number of drugs; psychotropic drugs and drugs liable to cause postural hypotension; standing systolic blood pressure of less than 110 mmHg; and evidence of muscle weakness [34].

The term “Post-Fall Syndrome” (PFS) was first defined by Murphy and Isaacs [142] as a particular association of motor and psychological symptoms related to falls in older adults. PFS is characterized by the joint observation of three symptoms: anxiety, worry of repeated falls and what is called “Psychomotor Regression Syndrome” [73] or “Psychomotor Disadaptation Syndrome” [141], which consists in acute motor disorders. The symptoms of PFS can resolve spontaneously or develop into fear of falling (FOF) [23, 197, 136], which causes behavioral disorders and limits patients’ ability to perform activities of daily living. Regarding product use, people with PFS may not be able to walk at all, or have balance and gait disorders caused by the psycho-motor consequences of traumas and/or long periods of bed rest. It is also possible that older adults who suffered a fall may not be able to put weight on their feet if they have leg fractures.

2.1.4 Medical applications and hospital environment

Hospital and care environments, where elders in need often stay, are highly constrained, because of the need to provide the high quality of care with limited staff and means and strict hygiene requirements.

Hygiene constraints are easily understandable, as there is a high risk of disease transmissions between patients. As a result, products that are to be deployed in a medical environment have to be easy to clean or even to sterilize. In addition, elders' care incontinence issues have to be properly managed. For instance, the fabric of a seat cover should be waterproof.

Due to staff shortage, care professionals have very little time to spend on training and operating devices. This is why technologies deployed in medical environments have to be very simple to use and to require little or no training. In addition, devices have to be fast and effective, for their operation to take as little time as possible.

Because of limited financial means, care institutions look for cost-effective products, unless their function is critical or saves care staff a lot of time. If a device is not critical, hospitals will only get equipped if products are cheap. Similarly, medical devices should be robust, durable and easy to maintain. Lastly, medical product also have to comply with strict safety requirements, which means they have to be robust and reliable.

2.1.5 Discussion

We have seen in this Section that designing for older adults with cognitive impairment, especially when deployed in highly rigid care environments, can prove quite challenging. Indeed, a user interface targeting this public should be very simple and require very little or no learning. However, in some aspects, creating virtual human applications for older adults with cognitive impairment is easier than for healthy younger people. First of all, as we will see in Chapter 4, creating virtual humans capable of verbal and non-verbal communication is a difficult task, especially regarding conversation management. The fact that our target public has reduced communication capabilities and lack of initiative, due to apathy, can make conversation management simpler in two ways: the virtual human can always have the dialog's initiative, which is easier to manage than mixed-initiative dialog; and they can only be presented with a few possibilities at a time, which reduces the complexity of dialog trees. For instance, Yasuda *et al.* have successfully used an ECA that only asks reminiscence questions and performs active listening with patients with dementia (see Section 2.2.1). Secondly, older adults have little experience of 3D video games, which means they have lower expectations in terms of graphics quality and animation than people who are used to AAA games[4]. Thirdly, people with dementia need support in the basic activities of daily living, which means that even simple applications, such as providing elementary information, can already be useful for them. Lastly, due to their low capabilities in remembering recent events, they can be satisfied by games or applications with reduced content, whereas younger adults would get bored quickly.

2.2 Game technology and serious games for elders' care

In the past few years, new uses for video games and game technologies have spread fast. What was previously seen as pure entertainment has been repurposed to pursue different goals and target serious applications. Serious games are now used in a wide variety of fields, such as military training, education, culture, advertisement and health [165, 220]. As mentioned in the introductory chapter, games have a strong motivational power. This is why their use for health benefits was thoroughly investigated, in particular to foster behavior changes towards healthier lifestyles and do prevention and health education. Indeed, in the past few years, thanks to the availability of cheap sensors, the widespread use of smartphones and the growing market of the Internet of Things, a myriad of connected devices have been made available for the public to pursue better health through self-monitoring [143]. However, according to Munson *et al.* [143], these personal "quantified-self" health tracking systems face challenges for widespread adoption, as many people are not health-concerned enough to start using such devices and people who do use them may be drowned into data and struggle to make sense of it. This is where games come in, to "turn the wash of personal health information into an experience that is meaningful and motivating and to reframe daunting health problems into challenges that are enjoyable to solve"[143]. Older adults, who are known to have deeply-rooted, life-long habits, could thus really benefit from such technologies. In this context, virtual humans could play the role of virtual health coaches to offer older adults user-friendly interfaces for these applications or for other assistive technologies. A good example of this is discussed in Section 2.2.2, Bickmore *et al.* [25] experimented a system including a virtual exercise coach and a pedometer to encourage older adults to walk more.

Encouraging behavior changes for healthier lifestyles is certainly a very interesting application of virtual humans and game technology but it is not the only one, serious health games have also received a lot of attention for rehabilitation purposes, as stated in the previous chapter, and virtual humans

[4] AAA games are the video games produced by major video game production companies. Creating these games requires a team and budget comparable to the ones of Hollywood movies.

can serve as user-friendly interfaces for a much wider range of applications targeting older adults with cognitive impairment.

With our research goals in mind, a survey of the literature about ECAs for the support of older adults and games for fall rehabilitation and prevention are presented. The aim here is to highlight relevant current knowledge about the use of these technologies with older adults, provide an overview of the variety of systems that have been developed and evaluated by researchers around the world to tackle the issues we address, and show the originality of the work presented in this thesis.

2.2.1 Embodied conversational agents as an interface for cognitively impaired older adults

Human beings communicate with each other thanks to a myriad of articulated sounds, voice intonations (prosody), facial expressions and gestures. Learning this very complex mean of interaction starts at birth and quickly becomes very natural, except in some rare cases, such as autism. This is why relying on natural verbal and nonverbal communication for older adults with neurodegenerative diseases to interact with assistive technologies seems like an interesting path to explore. However, computers are not made to support such modalities of interaction and only offer unnatural and complicated user interfaces, which makes their use very difficult for people who are inexperienced and/or have limited cognitive and learning capabilities, due to their illness.

For this reason, researchers around the world investigate the possibility of building machines with social interaction capabilities, in the form of social robots and embodied conversational agents (ECAs). ECAs have many foreseeable applications in elderly care. Compared to assistive robots, who are just beginning to enter the marketplace, the technology is even less mature and thus the applications presented here are only at the research state. This is also due to the fact that robots can be helpful even with limited social interaction capabilities. ECAs, on the contrary, mostly play the role of user interfaces in assistive scenarios.

Globally, current results lead to the conclusion that the use of ECAs could improve the accessibility and acceptance of computer-based assistive technologies, compared to graphical user interfaces and voice interfaces, especially for older adults with cognitive impairment. The main reasons to believe so, invoked in the dedicated literature, are listed below.

Naturalness The way people interact with anthropomorphic ECAs is very similar to the way they interact with other humans, in a fundamentally social manner [151, 137]. This seems to apply even to persons living with cognitive impairment [36]. This implies that researchers can rely on knowledge from social sciences about human interactions to design such Human-Computer interaction.

Attention ECAs attract the user's attention better than other interfaces, including vocal interfaces [42, 151, 137]. This is important because elderly people usually have declining attentional capabilities, compared to their younger counterparts [67]. This is even more critical for cognitively impaired older adults.

Easy understanding It is easier to understand speech when it is co-articulated by a virtual human than when it is disembodied [151] and it also requires less effort to listen to a talking head [137]. This is likely related to the importance of visual information in speech understanding shown by McGurk and MacDonald [129].

Believability ECAs are usually perceived as trustworthy, thanks to their human-like appearance [42, 151].

Motivation The integration of ECAs in a behavior-changing coach or a learning companion has shown to be motivating for long-term use [42, 24, 25].

Effectiveness Several authors have shown that older adults performed better, or at least as effectively, in tasks when guided by a virtual human than with a voice or text interface [151, 137]. This result appears to be more significant with cognitively impaired people. In learning tasks, good effectiveness has also been shown [42].

These elements are as many arguments in favor of using ECAs as user interface for assistive technologies. Even though ECAs may end up not being the ideal solution for all applications targeting older adults with dementia, here is a significant list of use cases for ECAs targeting older adults, with or without cognitive impairment, or their caregivers that have been investigated by researchers around the world:

- virtual coaches that encourage older adults to change life-long habits for a healthier way of life, such as following a diet or doing more exercise [25];
- virtual companions for isolated older adults that aim at providing entertainment and restoring/maintaining social links [208, 92, 172];
- virtual personal assistants that provide help for organizing or taking medication, for instance, or other tasks that get more difficult due to age-related declines or cognitive impairment (in the case of mild cognitive impairment and dementia) [229];
- virtual butlers, that is to say, ubiquitous user interfaces for a smart connected home with ambient intelligence [236, 64];

- training tools to help formal and informal caregivers [121].

Among the studies cited above, some were designed to evaluate if ECAs are a good way of presenting information for older adults with cognitive impairment. In [151], three information presentation modalities were compared: ECA, on-screen text with speech and text only. The results show that the participants, healthy and cognitively impaired older adults, liked the ECA better and performed significantly better in a guided task with the ECA than with the other prompting modalities. In [137], Morandell *et al.* found similar results by comparing a photograph with animated lips with a text and speech prompt in a Wizard of Oz[5] (WoZ) [102] study involving 10 elderly people with cognitive impairment. They later confirmed this result in a larger study involving 12 healthy older adults and 12 seniors with cognitive impairment [138].

Carrasco *et al.* conducted the first work in which people with dementia actually interacted with an ECA, by pressing buttons (green for "yes" and red for "no") on a remote control [36]. More recently, some authors have proposed systems that allow verbal and nonverbal interaction. An effort to design an ECA as a companion for older adults with memory impairment was proposed by Huang *et al.* [92]. The agent is automated, and complemented with a mobile system that can record information from the patient's life to enrich the conversation and help people with dementia remember what they did in the previous days. The ECA proposed in this work mostly performs active listening: the user's prosody is analyzed for power and pitch to make the virtual character produce backchannel behaviors, such as head nods and acknowledgment vocalizations, at appropriate moments. In addition, it performs automatic speech recognition to spot key words in the input speech and find the best response to the user's utterance by searching a database of input-response pairs. This database is personalized for each user. Unfortunately, to the best of our knowledge, no evaluation of this prototype has been published so far. Other authors have proposed an active listening system for older adults with cognitive impairment in [231]. In this study, a cartoon-like animated young boy asked reminiscence questions about their lives to 8 people with dementia. The conversation is simply managed by asking a new question when people are done talking (i.e., a long silence is detected) and producing backchannel behaviors while they talk, finding the appropriate timing through prosody analysis.

As stated earlier, an assistive ECA-based system may also act as a virtual personal assistant. Yaghoubzadeh *et al.* [229] conducted a feasibility study of a daily assistant for older adults and cognitively impaired people. The ECA looked like a male child. This study consisted in two phases, fol-

[5] A protocol consisting in giving the illusion that the system is automatic when it is in fact remotely controlled by an operator. This method is often used in social robotics and ECA research.

lowing a co-conception methodology. The first phase consisted in interviews and a focus group. In the focus group, the participants were shown a prototype of the system, asked about their first impressions and were allowed to perform a small interaction with the ECA in a WoZ setup. The second phase consisted in a usability study involving 11 cognitively impaired adults aged from 24 to 57 and 6 healthy older adults aged from 76 to 85. The participants were asked to interact with the ECA to set appointments and task reminders in a calendar. The system made errors on purpose to see if people were capable of spotting and correcting them. Overall results were very encouraging: ten of the eleven cognitively impaired users successfully entered their appointment and could spot and correct 75% of the introduced errors. All the elderly participants completed the task successfully and could repair about 80% of the errors introduced. Regarding acceptance, feedbacks were mostly positive, though participants with cognitive impairment were more enthusiastic than healthy elderly users, who thought the system would be more useful for others than for themselves. The most interesting aspect of this study is that two error corrections strategies were compared: in the first condition, called "global condition", the calendar entry information was checked in one step, whereas, in the second, "local", condition, each element of the calendar entry (date, time, activity) was checked separately. The local condition was shown to work better, especially for the cognitively impaired subjects.

In a follow-up article [230], the same authors have proposed and evaluated an automated conversation manager to replace the operator in the WoZ protocol. Based on the first study presented above, it was identified that the conversation manager should comply to the following requirements: be flexible enough to accommodate incremental information acquisition and corrections of past entries at a later time; manage uncertainty and be able to ask confirmation questions to resolve uncertain information; structure information in a way that is easy to understand; manage information requests in a hierarchical manner to divide a complex task into subtasks; ask the user for explicit feedback when it is not clear that he or she has properly understood the request. This dialog manager is highly specialized for the calendar management task and does not allow for other applications. This prototype was shown to be effective in a validation study involving 6 healthy older adults.

Lastly, in [222], a WoZ ECA was evaluated with a single user: a brain-injured war veteran. This early study showed good engagement of the participant in the interaction with the ECA.

To the best of our knowledge, except for [230], no other work has proposed a tailored interaction management system for conversations between ECAs and people with cognitive impairment and this first attempt was highly specialized for a calendar management application and did not target people with dementia. As will be presented in Chapter 5, LOUISE

(LOvely User Interface for Servicing Elders) has several innovative features: attention monitoring and management capabilities; context reminders; a general-purpose interaction manager for goal-oriented conversations with older adults with dementia, which allows specifying interaction scenarios in an XML domain-specific language; and illustration images and videos display. In addition, usability testing with older adults with cognitive impairment, in WoZ and with an automated prototype, was conducted so we can provide recommendations about ECA-based user interaction design for this public, which are insufficiently available so far.

2.2.2 Other Applications of ECAs in elderly care

As few precious studies about ECAs and older adults with cognitive impairment were conducted, we found useful to include information about two ambitious ones involving healthy older adults and ECAs that could be extended to elders with dementia. The most extensive study on an assistive technology for older adults featuring an ECA was a randomized controlled trial of a virtual exercise coach for older adults by Bickmore *et al.* [25]. The study was conducted in Boston and included 263 participants, each one being followed for a 12-month period. The aim of the system was to encourage sedentary older adults to walk more. Half of the participants, in the intervention group, were given a touch-screen computer with an ECA and a pedometer. The other half of the participants, in the control group, were only given a pedometer. The intervention group had the ECA installed at home for 2 months and were followed for 10 months after the intervention. Results show that participants in the intervention group walked significantly more on average than participants from the control group during the 2 months of the intervention. This effect persisted over the 10 following months, though differences between the intervention group and the control group progressively diminished. This clearly shows that ECAs have the potential to be actually used by older adults and that they can have a real impact on their habits. However, the results of this study suggest that the system should be installed durably in people's homes for the effects to be maximal.

Another study was conducted by Vardoulakis *et al.*, in which a virtual companion for isolated older adults, intended to be installed permanently in people's homes, was proposed [208]. In this study, a "relational agent" was installed in the homes of twelve older adults for a week. All the participants lived alone but were not particularly lonely. Their age ranged from 56 to 73 years old. The outcome was rather positive, but the results are not very conclusive in terms of technology acceptance. Indeed, as the ECA system installed in people's homes included a video camera and was remotely controlled by an operator in a WoZ-like setup, the participants were informed for obvious ethical reasons. This is a source of bias because it caused some of the participants to be particularly concerned about their privacy. However,

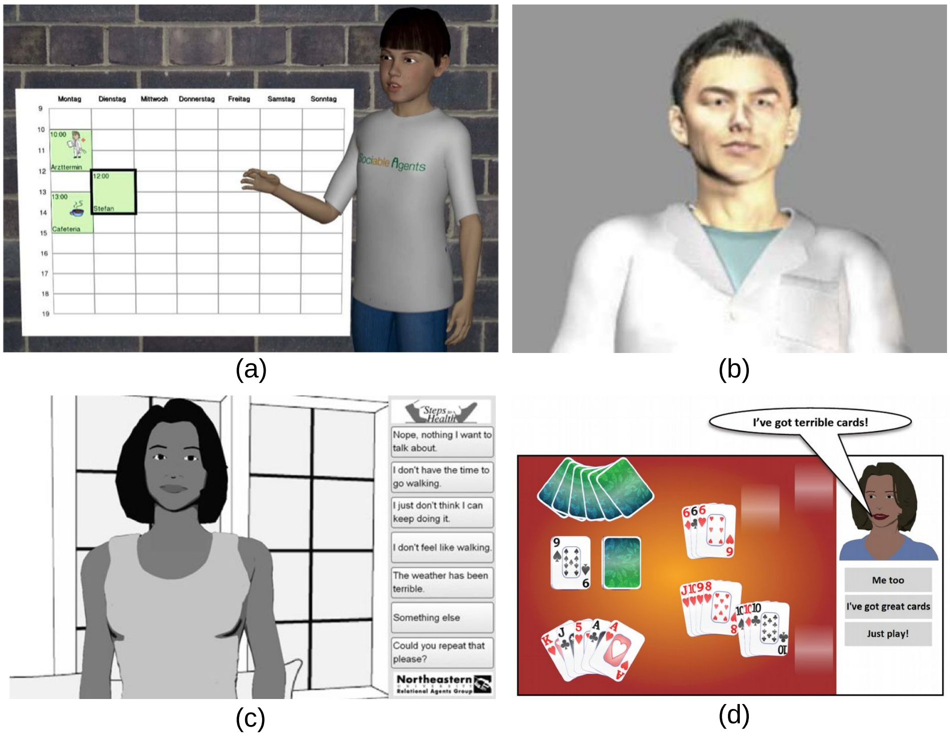


Figure 2.1: Embodiments of some of the ECAs proposed in the literature. (a) Yaghoubzadeh *et al.* [229]. (b) Huang *et al.* [92]. (c) Bickmore *et al.* [25]. (d) Vardoulakis *et al.* [208] (image taken from [184]).

an important point was that the system is not intended to replace human contact. On the contrary, one of its functionalities was to promote social integration by suggesting to the participant to get a walking companion, which was rather effective.

To conclude this section, it is interesting to note that ECAs can be used for dementia care in another way: for care staff training. ECAs and virtual humans are a powerful tool for training, education and e-learning. For instance, in a study of a global care solution for retirement homes relying on innovative technologies, ECAs were used to simulate specific care situations with people with dementia [121]. Care staff members were invited to interact with virtual patients with dementia, in various scenarios, to learn how to handle difficult situations, such as care refusal. This interesting research topic will not be addressed further in this thesis.

2.2.3 Games for fall prevention and rehabilitation in older adults

As already mentioned above, games have received much interest for health applications. Aside from healthier lifestyle promotion, they have been used for a variety of health-related applications. Regarding older adults, game-based exercises were proposed and evaluated for stroke rehabilitation, Parkinson's disease and heart failures [187].

As stated in the previous chapter, falls in older adults are a very concerning issue, due to their dramatic consequences in terms of mortality and disabilities. To address this issue, games can intervene in several cases. The best way to deal with falls is probably to prevent them from happening. However, even with the growing availability of cost-effective fall prevention programs, falls will still occur. This is why efforts also have to be put in post-fall rehabilitation, to limit risks of disabilities and repeated falls. In this section, we review the literature to sum up how games can be used in both situations, what are the known benefits of using video games in the context of older adults' falls and which games are helpful.

Skjæret *et al.* recently published a review about game-mediated exercise and rehabilitation for older adults [187]. The studies included in the review had to fulfill the following criteria: involve participants over 65 years old; comply with a pre-intervention evaluation – intervention – post-intervention evaluation methodology; consist in ICT-based weight-bearing activities, use clinical and/or physical assessment indicators for outcome measures. 60 studies were selected. It was found that:

- most studies focused on community-dwelling healthy older adults;
- none reported incidents leading to injury of a participant;
- the majority of the studies made use of commercially available game technology;
- balance is the most frequently used outcome measure;
- a minority of studies reported on adherence and an even smaller number compared adherence to game-based interventions with other types of exercises;
- most studies had methodological weaknesses, such as not performing between group comparisons or not using blinded assessors;
- several works showed clear improvement of balance and/or gait;
- no study found negative effects.

Many researchers have explored ways of preventing falls from happening. Classical approaches necessitate the intervention of physiotherapists and thus are costly and cannot be generalized. Another reason for that is because, as healthcare professionals are involved in these programs, people need to follow a medical path to benefit from such interventions, which means that those at high risk of falling have to be detected first. This is why fall prevention through video game training, which could be performed at home, has been explored as a possible cost-effective alternative. A recent and very comprehensive review of technology-based interventions for fall prevention was done by Hamm *et al.* [83]. They propose to classify interventions according to four broad categories:

- pre-fall prevention, which aims at reducing fall risk before any fall occurs;
- post-fall prevention, which aims at limiting the risk of repeated falls;
- fall injury prevention, which aims at responding to falls after they occur;
- cross-fall prevention, which concerns comprehensive systems that can intervene at all times.

According to this review, the subdivision criteria in each category include application type, user interaction modalities, deployment platform, deployment environment, information sources and collaborative function. To give an idea of the intensity of research in the field, about 150 different interventions, proposed in the previous 5 years, for the most part, are cited in the review. Pre-fall prevention is the category of interventions that uses exercise games, also called exergames, the most, as it focuses mostly on balance, muscle strength and limb coordination. The post-fall prevention systems mostly aim at detecting fall risk factors in patients who have fallen at least once. For this purpose, the use of game technology is least frequent, but a few examples can be found in which game technology was repurposed to assess fall risk, such as Garcia *et al.* [72] who used a Microsoft Kinect sensor, originally destined for games on the Xbox 360 console, to assess physical and cognitive walk-related abilities. The fall injury prevention category has little to do with game technologies, as they mostly consist in fall-detection devices. The cross-fall category is mostly based on smartphones and ubiquitous sensors around the living space.

Regarding post-fall rehabilitation, the exergame systems used in pre-fall prevention may also be used for patients who are past the stage when gravity compensation is necessary, especially for balance control training. However, there are few systems that were evaluated in that context [187]. Other rehabilitation systems that allow use with patients who require force compensation and include games are highly complex exoskeleton-based systems, such as in [71], and do not necessarily address fall rehabilitation. In the remainder of this section, some examples of applications are presented.

They are organized based on the type of gaming platform they make use of, which are depicted in Figure 2.2. No distinction is made between studies using only commercially-available games and studies in which only the game interaction devices are repurposed for use with custom software.

Nintendo Wii

Using commercially-available, consumer-grade software and hardware elements to create fall prevention intervention systems is a good way to achieve short development time and keep products' retail prices low. Several authors have made use of the Nintendo Wii Fit[6] balance board to train balance control. In [233], the Wii balance board was repurposed to create two balance exercise games with increasing levels of difficulty, as the games sold for that platform on the original game console were found too difficult for older adults with limited physical abilities. Both of these games are designed to provide immediate visual feedback of the player's center of pressure (COP), which is used to position an object on the screen. In the first game, the player has to control the medial-lateral position of his or her COP to adjust the position of a virtual basket to collect apples falling from a tree. In the second game, both medial-lateral and anterior-posterior positions of the player are used to move a virtual character to pop bubbles. The games were evaluated with 6 healthy older adults who played for 10 sessions of 20 minutes each. The results showed statistically significant improvements on the anterior-posterior sway variability and a tendency of improvement on the medial-lateral sway variability. In addition, the authors reported that the participants enjoyed the gaming experience and that they observed an 11% mean reduction in the Falls Efficacy Scale (FES), a questionnaire to evaluate fear of falling [198]. Another study has found similar effects using official Nintendo Wii Fit exergames [3]: 7 older adults with impaired balance were asked to play 4 different games in their homes for more than 30 minutes, thrice a week, over a 3-month period. The authors concluded that it is safe and feasible to have patients autonomously perform this activity at home, that participants enjoyed the activity, that their balanced improved significantly, as well as their walking speed. A third study by Bateni [16] compared Wii Fit balance training with physical therapy training. The 17 elderly participants (healthy with a history of falls) were dispatched in 3 groups. The first group received both Wii Fit and physical therapy training, the second group trained only with the Wii Fit and the third group only had physical therapy training. The results show that all trainings improved post-intervention balance evaluation, but that the Wii Fit alone was the least effective and the most effective intervention was the physical therapy, although little differences were observed between physical therapy training alone and physical therapy training combined with Wii Fit training. Lastly, a controlled randomized

[6] http://wiifit.com/

study involving 40 older adults divided in two groups, one that received Wii Fit balance training and one that received no intervention, concluded that, on its own, Wii Fit training yielded improvements in dynamic balance [166]. These results suggest that the costs of fall prevention training could be cut down by reducing the number of sessions with a physiotherapist and replacing the missing sessions with game-based training performed in institution or at home. This would result in equivalent, if not better, effectiveness, as games and varied activities may foster motivation. However, in hospitals, such activities may still need to be monitored by physiotherapists, but, if the activity is safe enough, it could be deployed by less qualified staff.

Other studies have made use of the Wii remotes, the standard game controllers for the Nintendo Wii console, and the Wii Sports[7] game, which consists in 5 sport simulations: tennis, bowling, golf, boxing and baseball. In these games, the Wii remotes, which feature accelerometers and gyroscopes, are used to perform movements resembling the actions of hitting a tennis ball with a racket, throwing a bowing ball, swinging a golf club, punching an opponent boxer and using a baseball bat. For instance, Keogh *et al.* [103] evaluated the Wii Sports game with 34 nursing home residents over 80 years old on average. 19 participants received the intervention whereas 15 were in the control group. Participants in the intervention group played all 5 available games, in groups, for about 30 minutes per week, over 8 weeks. The game play activity was evaluated using three indicators: functional activity (upper limb, balance and walk), physical activity and quality of life. In addition, semi-structured group interviews were conducted to get feedback. The games that were played the most were bowling and tennis, whereas very little play time was spent playing baseball. The authors observed significant improvements in arm mobility and psychological quality of life, as well as a relative increase in the level of physical activity, in the intervention group, compared to the control group. Lastly, the qualitative results of the semi-structured interviews show that participants enjoyed the activity (to their own surprise), had fun and that some participants had a feeling of anticipation between play sessions. Interestingly, by the end of the study, the age care facility in which the intervention group resided purchased a Wii console.

Microsoft Kinect

According to Skjæret *et al.* [187], only two studies made use of the Microsoft Kinect sensor, a device including a color camera, a depth sensor (the combination of which is sometimes called “RGB-D camera”) and microphones, for fall prevention in older adults. This device, which allows full-body tracking in real-time, was originally meant to play games on the Xbox 360 console.

[7] https://www.nintendo.com/games/detail/1OTtO06SP7M52gi5m8pD6CnahbW8CzxE

The first study involving the Kinect [104] consisted in an unsupervised virtual reality-based balance and hip muscle strength training program, over 8 weeks, using the Your Shape: Fitness Evolved game which features a Tai-Chi yoga exercising program. In the game, a virtual human performs movements that the player has to reproduce with his or her avatar, which is controlled thanks to the Kinect's motion tracking capabilities. In addition, the game allows to review the exercise session, compute performance scores, and a virtual coach gives visual and verbal feedback on the player's movements, to ensure they perform the movement sequences correctly, in a motivating way. It involved 32 community-dwelling older adults, split into a control group and an intervention group. Significant improvements in muscle strength and balance were observed in the intervention group. The second study [161] consisted in having older adults with Parkinson's disease play Kinect Adventures, which is a collection of mini-games that involve full-body tracking interaction. 7 participants with Parkinson's disease, aged between 60 and 85, played four different games for 60 minutes, three times a week, over 14 sessions. The authors concluded that this kind of intervention was feasible, as participants could play the game and improved their scores over the sessions, and safe. The also observed improvements in balance, gait and Parkinson's symptoms.

A custom game using the Kinect sensor was proposed by Proffitt and Lange who used user-centered design to produce an exercise game for older adults to prevent falls [163]. To design a new exercise game, they first performed a focus group to inform the design. The implementation of the game was then done iteratively, in a user testing-implementation cycle. Following this method, they produced a design that older adults enjoyed and found easy to use. The game consists in inclining the upper body to move a crash-test-puppet-looking avatar on the screen, who is in a mine wagon, to pick up floating jewels as they get close enough. Lastly, a study involving the Kinect, published after the review by Skjæret *et al.*, evaluated the effects of a custom-built game consisting in playing the goalkeeper in a virtual football stimulation with 43 healthy older adults over 65 years old [38]. After 12 weeks of bi-weekly training, significant improvements in balance were observed.

Dance Dance Revolution

Another type of game that was used in several studies for fall prevention is dance games. In a study by Smith *et al.*, the Dance Dance Revolution[8] (DDR) game's pressure detecting mat was repurposed to serve as input device for a custom-made simple stepping game consisting in stepping on the arrows as instructed on a screen [189]. This system was evaluated with

[8] Originally an arcade dance game, that consists in stepping on the arrows of a pressure-detecting mat or platform in rhythm https://us.konami.com/ddr/

44 participants over 70 years old. The study concluded that the proposed system achieved good usability if the step rate is as low as one step every two seconds and the visual presentation of the target arrows to step on is grouped within a 17° visual angle. To cite another example, a study by Studenski *et al.* made use of a proprietary dancing game called Dancetown by Cobalt Flux, the manufacturer of the DDR dance pad, which is no longer in business. This game is specifically designed for older adults in terms of speed and song selection [193]. The study involved 36 healthy participants aged between 68 and 89, but only 25 of them completed the program. These participants took part in 20 to 24 play sessions over 3 months. People who completed the program showed gains in narrow walk time and reported improved balance confidence. To formally evaluate the therapeutic outcomes of dance games, a randomized controlled trial was conducted by Pichierri, Murer and de Bruin [158]. In this study, 31 older adults over 65 (mean = 86 years old) were randomly assigned to two groups: an intervention group and a control group. Only 22 of them completed the program. Both groups participated in a physical exercise program consisting in balance and strength training, twice a week, over 12 weeks. For the intervention group, physical training was complemented with 10 to 15 minutes of dance game play at each session. The game used in this study was based on a dancing platform with arrows resembling the one in DDR. To evaluate the benefits of the intervention, the following indicators were used: a foot placement accuracy test, a gait analysis, a gaze behavior analysis during the foot placement accuracy test and the Fall Efficacy Scale mentioned earlier. Compared to the control group, the dance group showed more improvements in medial-lateral foot placement, walking speed and walking while performing a concurrent cognitive task.

Custom systems

Some authors think it is preferable to use games that are tailored for use by older adults, and therefore proposed custom systems. For instance, a dance/stepping game for older adults' balance training that uses player video tracking was proposed by Lange *et al.* [110]. In this system, the player's feet are tracked, thanks to LED devices attached to his or her shoes, to control the position of virtual shoes in a 3D environment. The goal is to move the virtual shoes to targets visually represented by green (left foot) and orange (right foot) footprints. The system was evaluated with 3 physical therapists and 4 younger adults for preliminary usability assessment and design recommendations. The freedom of movement offered by this system was appreciated; however, the actual movement amplitude during the game was not judged sufficient by the clinicians, and participants reported that there was not enough variety in stepping patterns.

Other authors used participatory design approaches to co-design fall prevention games with older adults. For instance, Uzor *et al.* [204] have followed a design methodology that involves older adults in the concept stages of the design process. Their initial idea was to start from the instructional booklets and videos for fall-prevention exercises, used in current care practices, and gamify these exercises with the help of older adults, to produce designs seniors would appreciate. The methodology consists in conducting design workshops in small groups (3 or 4 participants at a time). The workshops were organized in 4 phases: discussion of past experiences, scenarios and personas, a game session and user sketches. The scenarios/personas aspect is really interesting, as the authors proposed to have the participants create two fictional potential users of rehabilitation games to reflect about their needs and project how games could help these people. This was also a way to help people come up with ideas, based on personal experiences they would not have felt comfortable to discuss with a group of strangers. The games that participants could try in the workshops were prototypes of muscle strength and balance exercise games, using inertial measurement unit sensors as input devices. Thanks to this method, the key factors that limit adherence to the currently-used, at-home rehabilitation tools could be identified: people were not motivated to perform at home exercises because of the static nature of the instructional material provided and the fact that no charting of their progress was available. In addition, they could help people from the target user group come up with game design ideas for a set of rehabilitation activities. In a first workshop, older adults proposed an idea for a game, based on a balance training exercise that consists in marching on the spot while holding on to a chair. The game, also using inertial sensors, consists in marching in a virtual forest and using the free arm to catch butterflies. Furthermore, one of the games that the authors had created to provide examples to help older adults project and get ideas was improved thanks to the input of the participants. In a second workshop session, both of these games were tested by older adults, who were satisfied with them. More precisely, they thought that seeing improvements in game scores from session to session would foster self-confidence and make exercising more motivating; in addition, participants thought that family members could also play the game, making them social activities, which would also enhance motivation. In addition, participants proposed that the game scores should be based on correct movement repetitions instead of game goals, which would assure people in rehabilitation that they are making progress in terms of balance and strength. In a follow-up work, the same team has used their findings to implement and test a serious game for fall rehabilitation [202]. They found that the resulting design overcomes the major limitations of standard care and is usable by and acceptable for the end users. Finally, in a third paper, Uzor and Baillie studied the long-term use of their exergame in ecological conditions, with people that suffered one or several falls. They show that it yields better adherence to the rehabilitation program than the standard care tools [203]. This work, as well as the work of Proffitt and Lange [163] discussed earlier, are good examples of one of the goals we tried to achieve in this thesis. It stresses the importance of "empowering" [204] older adults by involving them from the ideation phase of the project to maximize usability, acceptance and adherence to the rehabilitation program.

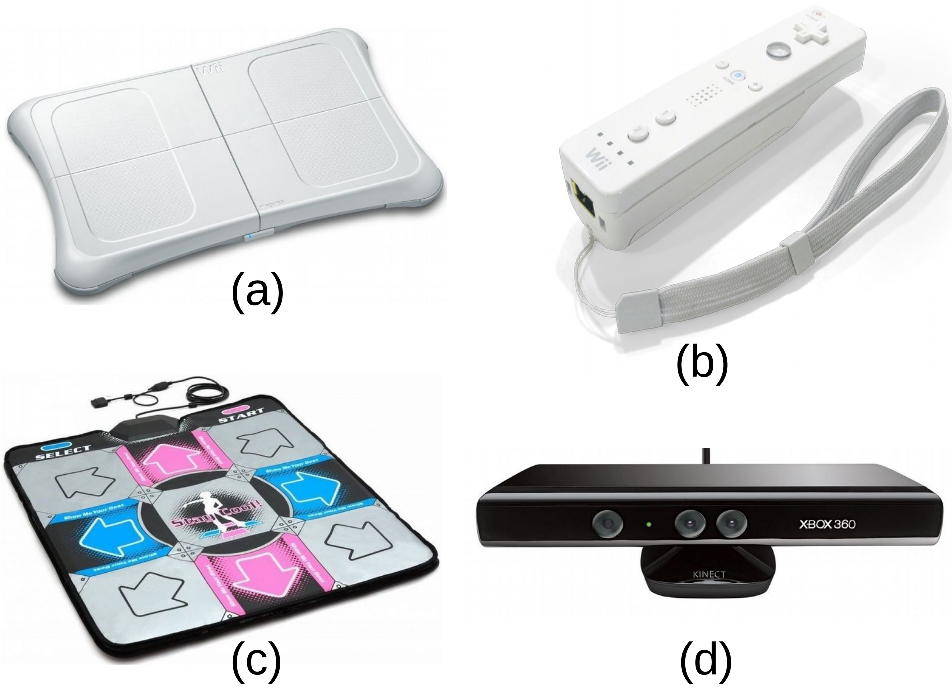


Figure 2.2: Input devices used in studies of games for fall prevention and rehabilitation in older adults. (a) Wii Fit balance board. (b) Wii remote. (c) Dance Dance Revolution pressure mat. (d) Kinect sensor.

As seen in this section, many systems have been proposed to gamify fall prevention and rehabilitation activities and many studies have shown the effectiveness, safety and motivation enhancements of commercial and custom-built serious games. In all of these works, physical activities are involved and, to the best of our knowledge, none have specifically addressed psychological consequences of falls, which means that our attempt to tackle PFS through virtual reality therapy is the first of its kind. However, some works have evaluated the cognitive impact of gamified physical exercise interventions, such as Pichierri *et al.* [158], who have shown improvements in dual-task conditions (walking while performing a concurrent cognitive task). In addition, other works have evaluated the psychological impact, in terms of fear of falling, of a game-based physical training intervention, such as in [233]. These results suggest that playing games may impact cognition and confidence positively. This is another reason to think that a game-based activity could actually help treat PFS.

Chapter 3

Focus points and methodology

Résumé en français. *Le but de ces travaux est de développer des technologies fondées sur les humains virtuels, adaptées à l'utilisation par des personnes âgées atteintes de troubles cognitifs, grâce à la conception participative. Dans le précédent chapitre, nous avons vu que la technologie des jeux vidéos et des humains virtuels, sous la forme de jeux sérieux et d'agents conversationnels animés, peuvent être des solutions bon marché et efficaces pour résoudre les problématiques des chutes et des troubles cognitifs chez les personnes âgées. Il y a, par contre, à cause des limitations sensorielles et cognitives du public que nous ciblons et des contraintes inhérentes aux environnements de soin, des défis de conception à relever. Notre approche pour lever ces difficultés est d'adopter une méthodologie de conception participative living lab qui consiste à impliquer les différentes parties prenantes dans toutes les phases de la conception. C'est, à notre avis, la meilleure manière de produire des technologies qui sont non seulement faciles à utiliser et acceptables pour les utilisateurs mais aussi adaptées à leur environnement de déploiement et validées par des professionnels impliqués dans l'écosystème autour des dispositifs.*

Notre approche pour l'ACA LOUISE est centrée sur les notions d'attention et d'engagement, qui sont des facteurs de succès clés pour l'efficacité de la communication entre un ACA et une personne atteinte de démence. Cela est du au fait que les personnes démentes ont des troubles de l'attention et ont une tendance à l'apathie, qui cause un manque d'engagement dans des activités qui ont du sens. Ces notions sont également d'importance capitale pour la thérapie en réalité virtuelle, qui requiert que les patients soient immergés dans des environnements virtuels. De plus, comme pour la rééducation, ce type d'activité nécessite un engagement soutenu sur une longue période temporelle, ce qui nous pousse à nous intéresser également à la notion de motivation. Enfin, nous avons choisi d'alimenter nos réflections

sur la conception ergonomique à l'aide d'un concept clé dans ce domaine : l'affordance.

Dans ce chapitre, les principaux concepts théoriques impliqués dans ces travaux sont définis et commentés. Puis, les principes du living lab et de la conception participative que nous avons suivis sont introduits.

Le concept d'affordance se définit comme "une relation entre les propriétés d'un objet et les capacités d'un agent qui détermine les possibilités d'utilisation" [147]. Pour qu'un objet soit facile à appréhender, les affordances doivent êtres rendues visibles. Par exemple, la poignée d'une théière permet de la saisir, si l'on a une main, et de la soulever, si l'on a suffisamment de force, et, en voyant la poignée, on comprend facilement que c'est par là qu'il faut la prendre. Dans le cas des humains virtuels, il peut y avoir de fausses affordances, c'est-à-dire des caractéristiques apparentes qui ne correspondent en fait pas à une possibilité d'action. Par exemple, si l'on s'adresse naturellement à un ACA, il se peut qu'il ne soit pas capable de comprendre ce qu'on lui dit, alors que le personnage ressemble à un humain.

Le concept d'attention correspond à la capacité de notre esprit à se focaliser sur un stimulus particulier. Selon la formule de Bundesen et Habekost [32], "nous sommes attentifs lorsque nos esprits sont dirigés vers quelque chose en particulier, comme lorsqu'on cherche un visage dans une foule, qu'on suit une conversation particulière dans une fête bruyante, ou qu'on tente de se concentrer sur l'impression de cette page plutôt que sur l'une des nombreuses autres choses dans la pièce". Les signes observables de la direction de l'attention (direction du regard, hochements de tête, etc.) peuvent être d'une importance capitale lors de la réalisation d'une tâche collaborative ou lors d'une conversation.

L'engagement en général peut se définir comme "l'acte d'être occupé ou impliqué avec un stimulus externe" [44]. De plus, "l'engagement implique une connexion entre l'individu et un stimulus ou une activité qui se caractérise par l'attention et l'intérêt" [171]. L'attention est donc une condition nécessaire, mais pas suffisante, à l'engagement. Il faut y ajouter une participation active, ou une réaction au stimulus, émotionnelle par exemple. Dans le cadre de la conversation, l'engagement se définit comme "le processus par lequel deux (ou plusieurs) participants établissent, maintiennent et mettent fin à leur connexion perçue lors des interactions qu'ils entreprennent conjointement" [185]. Cohen-Mansfield, Dakheel-Ali et Marx [44], le manque d'engagement chez les personnes démentes est problématique, car elles passent la plupart de leur temps sans être engagées dans la moindre activité qui ait du sens.

Enfin, la motivation se définit comme "une construction psychologique qui résulte de la combinaison de deux dimensions : avoir de l'énergie pour effectuer une action, puis déplacer cette énergie dans une direction spécifique" [171]. La direction est liée à ce que nous faisons et l'énergie émane de la raison pour laquelle nous le faisons. La forme de motivation la plus *puissante est la motivation intrinsèque. Elle s'observe lorsque la raison qui pousse une personne à agir émane directement de celle-ci. En dernier lieu, la motivation intrinsèque repose sur un équilibre entre deux besoins : le besoin de défi et le besoin de compétence.*

La méthodologie living lab, quant à elle, découle des principes de conception centrée-utilisateur et élargit le champ des personnes impliquées dans le processus de conception à toutes les parties prenantes autour du dispositif. D'autre part, elle repose sur des principes qui lui sont propres : l'ouverture (la conception est publique et toute personne se sentant impactée peut s'impliquer) ; l'influence équilibrée données aux différentes parties prenantes impliquées ; le réalisme ; la création de valeur (pécunière ou non) ; et la soutenabilité.

Nous proposons également une manière innovante d'appliquer ces principes à notre processus de conception : nous consultons et testons avec des personnes âgées sans troubles dans un premier temps, comme "proxy" de nos utilisateurs finaux ; nous nous autorisons à modifier les prototypes au cours d'une phase d'évaluation, afin de gagner du temps en ne gâchant pas de tests avec des défauts trop saillants qui empêcheraient les participants de voir les autres points d'amélioration ; nous réutilisons ou détournons de leur usage des matériels disponibles dans le commerce et des logiciels libres ou gratuits ; nous profitons d'évènements publiques pour collecter les avis des visiteurs pour garantir le suivi du principe d'ouverture ; nous faisons de l'acculturation en suivant des professionnels de santé dans leur journée de travail, afin de mieux comprendre les besoins et problématiques sur le terrain ; une manière d'impliquer les médecins qui leur prenne un minimum de temps doit être trouvée ; et nous contribuons à la dissémination de la méthodologie living lab en formant des stagiaires et en accueillant des visiteurs au sein du laboratoire.

The goal of this work is to develop virtual human-based technologies adapted for use by older adults with cognitive impairment, through participatory design. In the previous chapter, we have seen that game technology and virtual humans, in the forms of serious games and embodied conversational agents, can be cost-effective and efficient tools to address cognitive impairment and falls on older adults. There are however, due to the sensory and cognitive limitations of our target users, as well as the constrains inherent to rigid care environments, a number of design challenges to tackle. Our approach to solve these issues is to adopt a living lab participatory design methodology that consists in involving several stakeholders in all phases of the design process. This is, in our opinion, the best way to produce technologies that are not only easy to use and acceptable for the users but also adapted to their deployment environment and validated by professionals involved in the ecosystem around the products.

Our approach for the LOUISE ECA is centered around the notions of attention and engagement, which are key success factors for effective communication between an ECA and a person with dementia. This is due to the fact that people with dementia have attentional disorders and apathy, the later causing a lack of engagement in any meaningful activity. These notions are also very important in virtual reality therapy, which requires patients to be immersed in the virtual environments. In addition, like rehabilitation, this kind of activity requires sustained engagement over a long time period, which leads us to also consider the notion of motivation. Lastly, we chose to fuel our reflections on ergonomic design with a key concept of the field: affordance.

In this chapter, the key theoretical concepts involved in this work are defined and discussed. After that, the principles of living lab and participatory design we followed are introduced.

3.1 Key Concepts Involved in this Approach

Through the literature review of how games and virtual humans may help older adults with various issues, presented in the previous chapter, some of the key concepts underlying this work have been met: *attention*, *engagement* and *motivation*. In this Section, these concepts, as well as the concept of *affordance*, are formally defined and discussed.

3.1.1 Affordance

The term *affordance* was first created by J.J. Gibson in 1977, which he defined as follows: "The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill." [77]. An *affordance* is defined by Donald Norman, who introduced this term in the field of ergonomics, in [147] as "a relationship between the properties of an object and the capabilities of the agent that determine just how the object could possibly be used". For instance, a teapot has a handle so it affords grabbing and lifting; it is hollow and impermeable, so it affords filling it with liquid; and it has a spout which allows pouring this liquid into a cup without spilling (too much). However, these possibilities of actions also depend on the user: for instance, a small child who is not strong enough would not be able to lift a full teapot. According to Norman [147], for an object's design to be effective, these affordances have to be visible or perceivable for people to get strong clues to guess how things are meant to be operated. Put simply, the shape of a well-designed object should clearly suggest what it is for and how to use it. In our every day life, many objects comply with this principle. To keep the teapot example, there are teapots of all sizes, shapes and colors – some even are transparent; some are quite sophisticated and include a built-in strainer, which adds an extra affordance – but all have the same basic functionalities and most people would know what to do with them and how to use them (except for young children or older adults with severe dementia).

Virtual humans have many visible affordances; they have a mouth, which suggests they can talk; they have ears, which suggests they can hear; etc. This is an advantage as seeing virtual humans, older adults with cognitive impairment will spontaneously try to talk to them (provided that they are still able to do so), like they would do with a real human. However, the paradox of virtual humans is that their appearance may suggest more capabilities than they actually have. In this case, they have *"false affordances"*, that is to say perceptible features that suggest "nonexistant affordance[s] [...] upon which people may mistakenly try to act" [74]. For instance, they are not yet capable of full natural language understanding, but people may still try to talk to them in a natural way that cannot be understood by the system. This is why, as we will see in the following chapter, interaction management and embodiment of the LOUISE ECA have to be carefully designed to avoid overly-high expectations on the user's side. By the way, this gap between the user's expectations of an artificial social agent's interaction capabilities and the actual possibilities offered by this social agent is sometimes called the "Uncanny Valley" and can lead some people to feel uneasy about virtual humans and social robots [139].

Perception of affordances is also influenced by people's previous experiences with objects. At times, some people may easily identify all affordances of an object and some may only identify parts (or none) of the possibilities offered by it. This is the case of video game controllers: older adults who have never played video games may not know what to do with a game controller that would be handed to them, although they could identify parts of the possibilities, such as how it is supposed to be held, whereas teenagers would perfectly identify it as being a device to play games and hold them correctly without even thinking about it. This means that a serious game targeting older adults, possibly having cognitive impairment, should include a tutorial with detailed gameplay instructions, at least until the current teens get old themselves. Laslty, regarding virtual objects and software applications, it may be a lot more difficult to create visible affordances. In this context, additional *signifiers* can be introduced to highlight affordances. In fact, these two concepts are complementary: "affordances determine what actions are possible" (for instance, touching a touch-screen), and a signifier indicates "where the action should take place" (for instance, on a virtual button) [147].

3.1.2 Attention

Attention is a general word that regroups a variety of phenomena. According to Bundesen and Habekost [32], "we are attentive when our minds are directed at something specific, such as when searching for a face in a crowd, keeping track of a particular conversation at a loud party, or trying to focus on the print on this page rather that one of the many other things in the room". In other words, attention is the cognitive capability allowing us to direct our senses to extract information from a specific stimulus. When considering the concept of attention, the keyword is selectivity. Indeed, human beings have the capability to select a particular stimulus while filtering out all the other stimuli, such as listening to the voice of a person in a crowd to follow a conversation. In psychoacoustics, this phenomenon is sometimes called "the cocktail party effect". This phenomenon was first proved in 1953 by Cherry [41], who conducted experiments in which subjects were asked to repeat what a voice, heard in one ear, said while hearing another voice in the second ear. It turns out that people could perform the task successfully but did not remember any of what the other voice was saying. Visual attention works in a similar manner; it is possible to focus on only one visual stimulus at the time. This is well illustrated by a famous experiment by Simons and Chabris [186] in which subjects were asked to watch a video[1] of 6 people passing a basket ball around and count the number of passes made by the players wearing white shirts. While the players are passing the ball around, a woman in a gorilla suit walks in the middle of the players, faces the camera, taps her chest, then leaves. About 50% of the subjects did not notice the gorilla (sometimes more, depending on the experimental conditions). This phenomenon is called "inattentional blindness".

Visual attention is involved in many activities of daily living: driving, shopping, cooking, etc. In all these situations, it is necessary to visually scan the environment to find objects of interest and perform actions. Doing so, a huge amount of visual information passes through our eyes, but very little of this information is actually processed. In the context of virtual humans and video games, this concept is of great importance, as most of the information (if not all) is conveyed visually, at least initially. According to Bundesen and Habekost [32], in the psychological theory of visual attention, introduced by Bundesen in 1990, attentional mechanisms operate in the following ways:

- elements in the visual field are perceived simultaneously;
- when a particular element is selected, it enters a limited-capacity short-term memory store;
- the element is selected and recognized as part of a particular category at the same time;
- selection is performed by a processing race between the available stimuli;

[1] The video is available at http://www.theinvisiblegorilla.com/gorilla_experiment.html

- the probability of selection for a given stimulus depends on three factors, namely "strength of sensory evidence" (how salient the element in the visual field is), "perceptual decision bias" (predisposition of the subject) and "pertinence" (how important the object is in the current situation).

As mentioned earlier in this chapter, the focus of attention can be switched from a stimulus or a task to another. In addition, when an element of interest is in attentional focus, information about this element gets saved in short-term working memory; as a result, it is possible, by switching rapidly from stimulus to stimulus, to gather information about several elements. Typical visual information search strategies thus consist in quickly looking over several elements, through eye fixations and saccades, to find the ones of interest.

As seen earlier in this chapter, older adults have decreasing working memory; they can manage less information, thus fewer stimuli, at a time than younger people, which results in cognitive overload when too many stimuli are present at the same time. They also are slower at switching from a stimulus to another. When cognitive impairment is present, these deficits in attentional capabilities get pathological and make performing a number of tasks very difficult or impossible. Furthermore, they cannot sustain their attentional focus for as long as cognitively-healthy people. During assistive technology evaluations, it is often observed that people with cognitive impairment get distracted or lose track of what action they are performing. For our purpose, this means two things: information display has to be kept minimal to avoid distractors and cognitive overload; and active attentional monitoring capabilities appear as a key feature to explore, especially in LOUISE as attention also plays an important role in face-to-face communication.

When two or more people participate in a collaborative task, attention direction is necessary for them to coordinate their actions adequately. In particular, when discussing objects that have to be manipulated and what actions will be performed on them, the participants of the interaction all direct their attention on the objects discussed and on each other, by directing their gaze [167, 154]. Attention management in human communication is driven by nonverbal behaviors, in particular body posture, gaze direction and pointing gestures: when a participant of the interaction points and/or gazes towards an object, the other participants will naturally direct their attention towards that object. A study on this type of attention management in ECAs was conducted by Pejsa *et al.* [154]. A similar but possibly adapted to elders attention management strategy has to be designed and implemented in LOUISE to clearly emphasize on what elements of the display users should direct their attention. LOUISE would therefore have two levels of attention management: making sure the user is paying attention to the interaction with it; and, within this interaction, directing his or her attention towards specific elements of the display.

3.1.3 Engagement

Broadly speaking, the term *engagement* can be defined as "the act of being occupied or involved with an external stimulus" [44]. According to Rigby [171], "engagement implies a connection between the individual and a stimuli and activity that is characterized by attention and interest". Attention is therefore a necessary condition to engagement but is not sufficient. For a person to be fully engaged in an activity he or she also has to actively participate or, if passive, react to the stimulus (for instance, by having emotional reactions to a movie). In our context, for our walking simulation through a serious game (Virtual Promenade), engagement means paying attention to the display and playing the game by physically acting on the game controller to control the virtual human. In the case of the LOUISE ECA, we will refer to Sidner *et al.*'s definition of *engagement*, which applies to face-to-face conversation: *engagement* is "the process by which two (or more) participants establish, maintain and end their perceived connection during interactions they jointly undertake" [185]. Even though this definition is more specific than the one presented earlier, it is not incompatible. Indeed, when people have a conversation, they show each other signs of their active attention and interest for the interaction, thanks to a number of communication behaviors, which include gaze direction, body orientation, hand gestures, facial expressions, verbal responses and backchannel behaviors, such as head nods, smiles or nonverbal vocalizations. Lastly, a slightly different definition was given by Benkaouar and Vaufreydaz, who consider *engagement* as "the phase during which one expresses, with [nonverbal] cues, the intention of an interaction" [17]. This last definition is also compatible with the ones given previously with the exception that it only concerns the moment directly preceding interaction.

According to Cohen-Mansfield, Dakheel-Ali and Marx [44], lack of engagement is problematic in people with dementia, as they spend most of their time without engaging in any meaningful activity. The same authors also argue that "prolonged lack of stimulation can be particularly detrimental to persons in nursing homes who suffer from dementia, as it magnifies the apathy, boredom, depression, and loneliness that often accompany the progression of dementia" [44]. This is why the engaging characteristics of games and ECAs are key to this research work.

Several works in the literature have focused on characterizing and quantifying engagement, either in interaction with objects or in face-to-face conversation with robots or ECAs. For instance, Cohen-Mansfield, Dakheel-Ali and Marx [44] proposed a concept of engagement based on five observable dimensions: rate of stimulus refusal; duration of involvement or occupation with the stimulus; level of attention towards the stimulus; attitude towards the stimulus; and actions on the stimulus (touching it, talking to it, talking about it). This model applies to patients with dementia in interaction with physical objects. Another engagement scale was proposed by Judge, Camp and Osulic-Jeras [98]. It consists in four levels to quantify engagement in dementia patients: constructive engagement, which consists in motor or verbal reactions exhibited in response to a stimulus or activity; passive engagement, which is defined as an exhibition of looking and/or listening behaviors in response to the stimulus; non-engagement, which is characterized by starring off into space, looking away from the stimulus or sleeping; and self-engagement, which is characterized by exhibiting motor, verbal, looking and/or listening behaviors exhibited during transition phases or when the patient was participating in any activity. This is called the Menorah Park Engagement Scale, named after the institution in which the authors worked. Finally, there are computational models of engagement. They do not concern people with dementia in particular and are used in human-robot or human-ECA face-to-face interaction, such as in [168]. These models all consist in audio/video-based nonverbal behavior analysis of the human in interaction with the machine and mostly make use of gaze patterns, speech, hand gestures and facial expressions. More details about these automatic engagement measures are provided in Section 4.3.

3.1.4 Motivation

The goal of using serious games in healthcare is to provide therapeutic tools that foster sustained engagement, for patients to perform activities regularly, over an extended time period. To achieve this goal, the theoretical background about motivation can be of great help. According to Rigby [171], building on the work by Deci and Ryan [54], *motivation* is "a psychological construct that is a combination of two dimensions: having *energy* to take action and then moving that energy in a specific *direction*". Put simply, the direction dimension relates to "what" we do and the energy dimension relates to "why" we do it. Rigby also argues that engagement and motivation are closely related concepts, often used interchangeably in the gamification literature, although they are not synonymous. According to him, engagement can be seen as the "behavioral expression or manifestation of a motivated state" [171]. In other words, motivation gives one the energy to engage in a particular activity.

The most commonly used theoretical background of motivation used in gamification is Deci and Ryan's Self-determination Theory (SDT) [55]. This theory aims at explaining the underlying psychological processes that drive people's choices and actions. SDT is founded on the idea that human beings act in the pursuit of satisfying "innate psychological needs", which is "essential for ongoing psychological growth, integrity, and well-being":

competence, *autonomy* and *relatedness* [55].

Competence is our need to feel effective and successful when performing activities. This fundamental psychological need also pushes us to seek out new challenges that will allow us to improve our skills and become more competent [171]. However, overwhelming challenges, leading to failure, tend to have a negative effect on motivation, as the need for competence is not satisfied in that situation. This is why, when gamifying an activity, the level of difficulty has to be adjusted to keep challenging the player while making him or her feel that the goals are within his or her reach. Giving the players a feeling of competence also means providing them with feedback on their actions and rewards to acknowledge their successes in game tasks.

The concept of autonomy relates to our need for choosing freely our path. On the contrary, feeling constrained or controlled by other people or circumstances have a negative impact on motivation [171]. Interestingly, this principle is also present in the literature about manipulation, in which the feeling of freedom is a key factor to success and constraints or strong justifications have negative effects on the effectiveness of strategies to influence other's behaviors [97]. Regarding games, excessive rewards may also undermine the feeling of autonomy [171].

Relatedness is the need to feel supported by others and that one matters to them. According to Rigby [171], computer-generated, non-player characters in game are capable of satisfying this need, if they are scripted to display supportive behaviors towards the player. This could also be harnessed in assistive ECA scenarios, by having the virtual human display empathy and support for the user.

Lastly, in SDT, as summarized by Denis and Jouvelot [56], motivation qualities are classified in three categories: amotivation, which is the absence of motivation; extrinsic motivation, which is a form of motivation driven by factors that are external to the activity or the individual, such as reward or threat; and intrinsic motivation, which is the strongest form of motivation as it is self-induced and it pushes us to act freely. This distinction between extrinsic and intrinsic motivations lies in the "locus of causality"[55]. Being intrinsically motivated to do something means that we think doing this activity is interesting. In addition, there is a mechanism, called *internalization*, that allows individuals to assimilate external regulations, for them to become self-regulations. For instance, social conventions can be transformed into personally endorsed values. This results in doing things because we think it is important. Finally, there is a continuum in the perceived locus of causality for extrinsic motivation which ranges from external to internal [55]. The more internal this perceived locus of causality is, the stronger the motivation gets.

To sum up, the key concepts of affordance, attention, engagement and motivation were defined and their relations to the design of acceptable and usable virtual human-based assistive applications were highlighted. There are also interesting relations between these concepts. In fact, they intervene in technology use in a hierarchical way: motivation to use a product leads to adherence; this adherence is manifested by the individual through sustained engagement in the activity or with the device; being engaged implies focusing one's attention on the stimuli that are relevant to the activity or device; eventually, this attention needs be driven towards objects of interest within the display or environment through the use of visible affordances, which can be highlighted by signifiers, including attention management strategies through nonverbal behaviors, in the context of face-to-face conversation with a virtual human.

Let us now consider, in the following section, the design approach we chose to tackle the challenge of creating accessible and acceptable virtual human-based applications for the support of older adults, given the limitations in physical and cognitive capabilities of our target audience.

3.2 Living lab participatory design

As seen earlier in Sections 2.2.1 and 2.2.3, some authors have reported on the user-centered, participatory design of an ECA-based application [229, 230] and of serious games for fall prevention and rehabilitation [204, 163], which led to good designs that the target users enjoyed and were able to use easily. In these works, people from the target user groups were involved at all steps of the design process, from ideation (particularly in [204]) to the completion of a prototype, or even until the deployment of the system for a pilot study. These are good examples of what we would like to achieve in our two case studies. The main difference lies in the fact that, in our living lab design approach, aside from end users, other stakeholders, such as physicians, physiotherapists, psychologists or patients' informal caregivers, have to be consulted in the process. The closest example of what we would like to achieve in terms of design practices is the work of Benveniste [19], who relied on the principles of action research and user-centered design to develop serious games for music therapy: one for children with autism spectrum disorders and one for older adults with cognitive impairment.

3.2.1 Principles

Our design approach is rooted in *action research*, a methodology used in social sciences for a long time. It has first been introduced by Lewin [116] in 1946. As summarized by Benveniste [19], it consists in moving away from purely observational strategies, which were traditionally used in sociology and anthropology, to more active research practices. The main idea behind action research is to voluntarily introduce transformations in the groups or environment of interest and carefully study the consequences of these transformations. Benveniste argues that this may raise ethical questions, which can be addressed by making sure that taking part in the research is as beneficial as possible for the participants [19].

A more recent definition of action research was proposed by Winter and Munn-Giddings: "action research is the study of a social situation carried out by those involved in that situation in order to improve both their practice and the quality of their understanding"[224]. Action research is therefore a reflective practice. According to them, it consists in "alternating between inquiry and action [...] in a spiral of practical decision-making and evaluative reflection". This is in line with Lewin's original theory, as the key principle is to progress in an iterative way, introducing a small transformation, evaluating its impact and repeating these two steps until the goal is reached. This allows to progress carefully, step by step, and making sure that things are moving in the right direction and that the changes will not be detrimental to the participants.

This incremental construction approach is particularly suitable for software development and is in fact used in the trendy *agile* methodology [127], which consists in identifying functionalities, organizing them by order of importance and implementing the most important ones first, then testing and iterating to improve these functionalities, before adding others. This kind of practice is also common in the video game industry, in which game designers quickly build elementary prototypes of their games and have *playtesting* sessions [178]. These approaches are also present in the *Lean Startup* methodology [170], with the concept of *minimum viable product*, which consists in making a product with only the core functionalities and start selling it, before improving it and adding extra functionalities. This allows to reduce the time-to-market and quickly evaluate the product's market impact and iterate to improve it or change strategy ("pivot") if it turns out that the product is not bought by consumers or the feedbacks from early customers are negative.

According to Bergvall-Kåreborn *et al.* [21], the other parent of the living lab is participatory design. Participatory design, according to the same authors, is "based on the foundational principle of democracy"[21]. It consists in giving users the possibility to directly influence, or even drive, the design. They propose to organize participatory processes in three categories.

Design for users The design process is conducted by developers (engineers, technical experts). Users are consulted in the process and get to express their opinion by providing feedback on requirements and prototypes. In this approach, developers drive the design and users' opinions influence it. In this case, users are only seen as designated consumers.

Design with users The design is undertaken jointly by developers and users. In this framework, users are given more influence than in *design for users* and intervene in the design throughout the process, which is a "continual iteration between the developers and the users with a focus on knowledge sharing" [21]. Developers still lead the design process, especially regarding the technical aspects, and users give the direction, by providing a detailed understanding of their needs and of the context of use. In this case, we speak of co-construction.

Design by users In this approach, users drive the design and developers provide assistance and support throughout the process. This means that all ideas and design choices are made by the users and technical experts only play the role of performers, who only implement what users have decided.

In the examples mentioned earlier, Uzor *et al.* [204] adopted a *design by users* approach to create serious games for fall prevention, by involving older adults in the ideation phase, and had them propose ideas of games and even make drawings. Then, the best ideas were selected, implemented and tested. In [229] and [163], the participative design approach is closer to *design for users.*

Design by users is obviously the approach that empowers users the most in constructing a product or service for themselves. It is therefore the approach that has the best chances of success in terms of acceptance. In our case, as much as we would like to use this methodology, it is quite difficult for three reasons: most of the target users have cognitive impairment, which makes it difficult for them to express their needs and opinions or to project and propose ideas; the technology is very complex, particularly in LOUISE, and it is difficult for older adults to get a deep enough understanding of the possibilities for them to drive the design completely; and ideation of these projects came from care professionals and domain experts. For these reasons, design for users and design with users are more adapted to our situation. However, although we cannot fully empower users in the design process, our design approach goes beyond user-centered design and involves other stakeholders in the design process, to account for the whole environment around the products and not only the end user. In the LOUISE case study, we used a design for stakeholders process and, in the Virtual Promenade case study, we used a design with stakeholders process.

3.2.2 The living lab: a place and a methodology

According to Bergvall-Kåreborn and Ståhlbröst, [22], the term *living lab* is an umbrella concept that can refer to a place, a methodology or even a system. In the environment perspective, it consists in constructing home-like environments to perform in-laboratory evaluations of new technologies in simulated ecological conditions, which allows more control than true ecological conditions. In its methodological form, "living lab" consists in the processes of involving users and other stakeholders in the design process.

According to the French association of health and autonomy living labs[2] [69], this methodology is based on five principles:

Openness the conception process is public and anyone considering himself or herself impacted can get involved;

Influence all participants should have a balanced influence on the design – once people are involved, their opinion has to impact the design based on their importance regarding the product's deployment;

Reality the conception and evaluation of the solution should be performed as ecologically as possible;

Value creation the design process should create value for everyone involved, not only from an economic perspective but also at the social level;

Sustainability once developed, the solution should keep evolving and being re-evaluated thanks to new data from the field.

The living lab methodology, as presented earlier, is also organized around the idea of user/stakeholder-centered participatory design. It consists in involving as many stakeholders as possible in all phases of the product design. Design is conducted in a cyclical process, as depicted on Figure 3.1 [69]. It is composed of four stages:

1. *stakeholder mapping and needs analysis*, which consists in identifying who the stakeholders are, who the users are, what their respective needs are, and what the relationships between them are;

2. *conception*, which is the ideation and specification phase and can also include partial iterative prototyping and testing in laboratory or ecological conditions;

3. *development*, which is the iterative implementation and testing phase of a fully functional prototype and ends with a small-scale final evaluation to validate the solution;

4. *deployment*, which consists in elaborating a deployment strategy, and applying this strategy on a small-scale, at first, before moving on to a larger scale and performing ecological evaluation of the solution.

Each of these stages is further decomposed in four steps: ideation, exploration, experimentation and evaluation.

In this research work, we do not go through all of the stages, as the stakeholder mapping and needs analysis were already performed beforehand, and

[2] http://www.forumllsa.org/

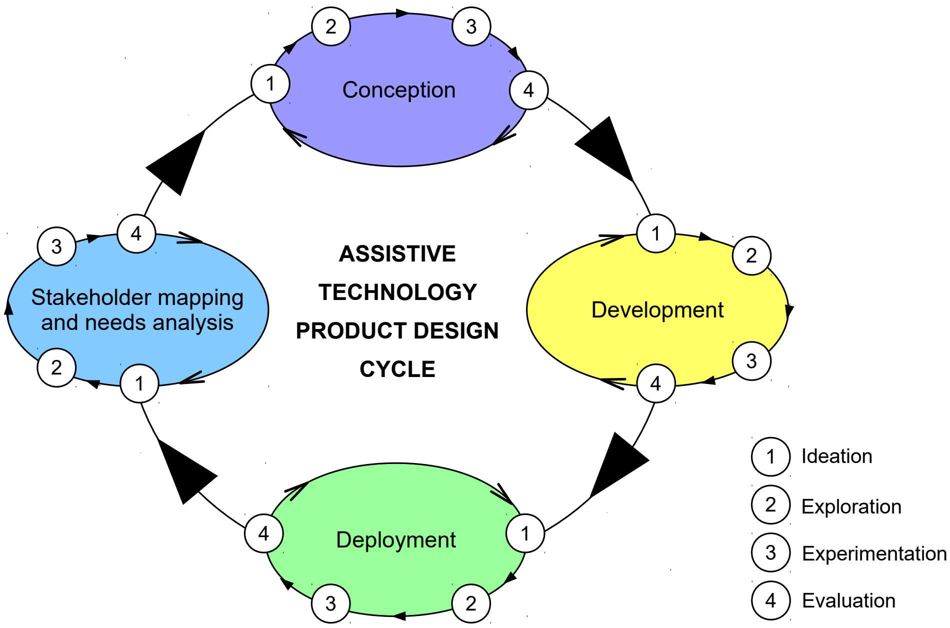


Figure 3.1: The living lab product design cycle.

we are not concerned with the deployment phase, as we are positioned on upstream explorations. Although the principles of living lab were followed to conduct this work, we also contribute to the development of this methodology in several ways, by proposing the following adjustments, which we have experimented in the context of this thesis:

- to deal with the cognitive limitations of older adults with dementia, and to avoid bothering ill patients with very early prototypes, healthy older adults should be involved in the early implementation (exploration) phases, as proxy to the target users, and assistive technology experts should be consulted;

- as the goal of the user testing sessions is to obtain the best possible design, changes should be made to the prototypes as soon as clear shortcomings appear, even in the middle of an evaluation phase, to speed up the process and avoid wasting time on testing a prototype with flaws so obvious that it will prevent participants to detect other issues;

- inspiration should be taken from the do-it-yourself and fab lab trends by reusing or repurposing cheap, commercially available hardware and free or open-source software, as detailed in Section 3.2.4;

- the design process should be made as open as possible by doing demonstrations to the public in events, such as the annual open house day of the hospital, to collect information about people's needs, expectations and preferences (see Section 5.3.1) and conduct informal playtesting (see Section 7.3);

- acculturation[3] should be performed through shadowing of various professionals in their day-to-day practices, such as physicians' day hospital consultations, neuropsychologists' cognitive assessments and rehabilitation with physiotherapists – this allows to establish close collaborations between the research team and the care staff;

- a way to involve physicians in the research has to be found, despite the limited amount of time they can afford to spend to help, by intervening in staff meetings and "going the extra mile", that is to say doing extra work to make things easier and quicker for them, so they can contribute with minimum involvement;

- lastly, contributions should be made to the dissemination of the living lab methodology by training interns and welcoming visitors in our lab.

3.2.3 The Broca Hospital's living lab

This thesis work was conducted at the LUSAGE (*Laboratoire des usages en gerontologie*) living lab [160], located in the facilities of the Broca Hospital in Paris, France. The Broca Hospital is a small geriatric hospital, part of *Assistance Publique – Hôpitaux de Paris*, the biggest group of hospitals in France. It builds on both the environmental view of the living lab, by including a room simulating an apartment's living room and having access to the hospital's premises, for projects in which the target environment is a care institution and not the home, and the methodological view, by working in close collaboration with the hospital's medical staff and being open to patients, caregivers, health insurance companies, non-governmental organizations and public deciders. It hosts a multidisciplinary team, composed of physicians (specialized in geriatrics and psychiatry), neuropsychologists, a social psychologist, anthropologists and engineers. Together, they work on various aspects of the design and evaluation of assistive technologies for cognition, mainly directed towards older adults.

The LUSAGE team conducts research and works on national and European collaborative research programs with technological partners, or on public health and social innovation. For instance, projects hosted in the living lab in the past few years include Project *PRAMAD*, a national collaborative research program about an assistive companion robot for the elderly, in which a small contribution of this thesis work was made (see

[3] By "acculturation" we mean acquiring the culture of the professionals in the field, getting know them and their practices to better understand their points of interests and concerns.

Figure 3.2: Picture of the LUSAGE living lab's building in Paris (France)

Section 8.3.2); the *Café Mutimédia* project [228], which focuses on social innovation for e-inclusion and technology awareness of elderly people; and the *Diapason* project [51], which consisted in creating an e-learning platform for therapeutic education about Alzheimer's disease targeting people who help a family member having this disease. It is the collaboration with these professionals of diverse expertise that allowed to conduct the work presented in this thesis.

The *Café Multimédia* project is particularly interesting and contributes a lot to the animation of the living lab because, though it is originally meant to train older adults to use current technologies, tactile tablets in particular, it is a great opportunity to get feedback from "expert" older adults regarding technologies developed in the living lab. During the course of this thesis work, a focus group on assistive embodied conversational agents (see Section 5.3.2), a user testing of a virtual strolling game (see Section 7.3) and parts of the evaluation of a moving head with facial emotions for the *PRAMAD* project's robot (see Section 8.3.2) were organized within this activity.

Lastly, due to our location within the hospital, one of the ways we see value creation for the participants of our research is that patients benefit from getting involved in this type of research. Many patients at the hospital do not get visits very often and some do not get any. Participating in the research therefore benefits them in at least three ways: it distracts them from their boring hospital stay; they can have a bit of conversation (we spent much more time than required for our purposes with very lonely patients who told us stories about their lives); and they get to discover new things (for instance, a 99-years old lady played a video game for the first time in her life). In addition, in the Virtual Promenade project or similar pathology-oriented games, if there are any therapeutic effects, participants to the study will benefit from it. Lastly, potential negative effects have to be avoided as much as possible, thanks to preliminary usability studies with healthy older adults and the consultation of physicians, experts and/or an ethics committee. For care professionals, getting involved in the research means having a change from their usually repetitive work, discovering new things and having a say in the development of new tools that may impact the way they will work in the future.

3.2.4 Off-the-shelf hardware, software and graphical elements

As mentioned in Chapter 1, we do not only follow the living lab methodology, we also contribute to it by relying on prototyping methods inspired by the do-it-yourself and fab lab trends. The implementation of our prototypes is thus based on reuse or repurposing of easily available, consumer-grade hardware elements, as well as open-source of freely available software elements. More specifically, LOUISE makes use of the Microsoft Kinect sensor, for user sensing, with the Microsoft software development kit, which is free to use for developers and can be redistributed without charge to end users. It is also built on open-source software, namely SmartBody[4] for character animation, Panda3D[5] (open-source game engine) for rendering and display and the Boost[6] library for the conversation manager. Virtual Promenade, on the other hand, makes use of an existing haptic chair, that was originally developed by the Backwell company[7] to treat or prevent back pains in workers who sit at their desk all day. In addition, we experimented with several game controllers, including a Sony Playstation controller and a Nintendo 64 controller. Lastly, the software part uses the cheap (or free), but proprietary, game engine Unity[8], which is associated with a content sharing platform, called Unity Asset Store[9], on which many ready-for-use graphical elements can be downloaded for free or purchased for a few dollars.

Doing so allows to gain a lot of time in prototyping and to focus on design aspects that will create the most value for the stakeholders. Another advantage of this approach is to create cheap products that most people or care institutions can afford. In addition, regarding deployment, using commercially-available hardware, instead of custom-built elements, potentially allows consumers to buy the parts at their local store and simply

[4] http://smartbody.ict.usc.edu/
[5] http://www.panda3d.org/
[6] http://www.boost.org/
[7] http://www.back-well.com/
[8] http://unity3d.com/
[9] http://www.assetstore.unity3d.com/

download the software (provided that it is distributed), whereas custom-built elements would require some kind of manufacturing and shipping logistic [19]. Lastly, our software allows for easy personalization and people with some technical skills could modify it to tailor it to the needs of a particular patient.

Chapter 4

Creating virtual humans

Résumé en français. *Vers 328 avant J.C., Aristote écrivit : "L'Homme est par nature un animal social" ; depuis, des scientifiques de différents domaines ont identifié des structures biologiques qui tendent à confirmer cette affirmation [213]. En fait, un certain nombre de structures dans le corps humain semblent particulièrement adaptées pour la communication interpersonnelle. Nos oreilles sont sensibles aux fréquences de la voix humaine ; nos conduits vocaux permettent la production de sons articulés pour former des mots ; et nos cerveaux sont équipés de neurones "miroirs", qui ne semblent avoir d'autre fonction que d'améliorer notre conscience de nos congénères [213]. C'est pourquoi se reposer sur la communication verbale et non-verbale semble être un choix pertinent pour concevoir des interfaces Homme-machine (IHM) évoluées. Les machines, au contraire, bien que conçues par des humains et pour des humains, et aussi "intelligentes" qu'elles soient, sont pratiquement dénuées des capacités de communication les plus élémentaires. Il en résulte que les humains doivent s'adapter à leur fonctionnement et à leurs modalités d'interaction pour les utiliser, via des interfaces compliquées.*

Au tout début, les ordinateurs se programmaient à l'aide de cartes perforées ou d'autres moyens rudimentaires. Au fil de l'évolution des technologies de l'information, les modalités d'IHM ont changé. De nos jours, l'IHM la plus commune est l'interface utilisateur graphique (IUG). Elle requiert un écran et peut se contrôler grâce à un clavier et une souris ou un écran tactile, ces derniers constituant aujourd'hui l'interface privilégiée pour les appareils mobiles [100]. Comme ils sont considérés comme plus faciles à utiliser que les IUG classiques (avec un clavier et une souris), leur utilisation par les personnes âgées a été étudiée en profondeur et il a été montré que ces dernières arrivent à s'en servir avec autant d'efficacité que les personnes jeunes, y compris avec des petites tailles d'affichage [191].

Afin d'évoluer vers des interactions plus naturelles, reposant sur la communication verbale et non-verbale, une partie de la solution pourrait venir

d'un domaine de l'ingénierie informatique inattendu : celui des jeux vidéo. Par exemple, le capteur Microsoft Kinect, initialement créé pour jouer sur la console Xbox 360, permet très facilement de suivre le corps et le visage des utilisateurs. Ces capacités de perception permettent d'effectuer de la reconnaissance de gestes ou d'expressions faciales. C'est pourquoi le Kinect a été utilisé par de nombreux chercheurs en vision par ordinateur pour effectuer de la reconnaissance de gestes de la main ou d'activités, entre autres choses [84]. Comme les gestes, les postures et les expressions faciales jouent un rôle essentiel dans la communication interpersonnelle naturelle, ce type de capteurs facilite grandement l'analyse automatique des comportements non-verbaux. Le Kinect est particulièrement intéressant, car il contient également des microphones, ce qui le rend suffisant pour l'implémentation d'interactions multimodales. La disponibilité et le faible coût de cet appareil et d'une poignée d'autres, telles que les cartes Arduino[1], rendent faisable et économiquement viable la diffusion de solutions capables d'interactions sociales qui, précédemment, restaient confinées à des laboratoires de pointe. Nous nous sentons donc pertinents en étudiant le potentiel de tels appareils pour l'utilisation quotidienne en tant que solutions d'assistance.

Dans ce chapitre, nous résumons les fondements théoriques et techniques qui rendent possible la création de programmes informatiques interactifs sociaux. Nous commençons par présenter des généralités sur la communication humaine dans la section 4.1. Dans la section 4.2, nous donnons un aperçu de la conception commune des ACA. Puis, nous détaillons comment les ordinateurs peuvent avoir des interactions simili-sociales avec les humains, grâce à l'analyse des comportements de l'utilisateur (section 4.3), la génération des comportements sociaux artificiels (section 4.4) et la gestion de dialogues (section 4.5). Nous terminons ce chapitre par une présentation de l'usage des langages spécifiques à un domaine dans les ACA (section 4.6) et une discussion sur les éléments présentés ici que nous avons choisi d'inclure dans notre système d'ACA (section 4.7).

La communication humaine se compose de deux parties : verbale et non-verbale. La partie verbale correspond simplement aux mots échangés entre interlocuteurs et la partie non-verbale regroupe tout ce qui entoure le discours, c'est-à-dire l'apparence, la distance interpersonnelle, la posture, les gestes, la direction du regard, les expressions faciales et les comportements vocaux (prosodie et vocalisations non-verbales). La communication non-verbale représente entre 70 et 90% du contenu informationnel des échanges lors d'une interaction face-à-face entre deux ou plusieurs personnes. La discipline consistant à analyser automatiquement et produire artificiellement de tels comportements s'appelle "traitement du signal social" [156].

Pour réaliser les fonctions de perception et de production de ces comportements, les ACA se composent généralement de 5 éléments : des cap-

[1] http://arduino.cc

teurs, qui permettent de percevoir l'utilisateur ; un étage d'analyse et d'interprétation des comportements de ce dernier ; un gestionnaire d'interaction, qui détermine comment réagir aux comportements de l'utilisateur ; un étage de production de comportements ; et des périphériques de sortie, généralement des hauts-parleurs et un écran, dans le cas des ACA, ou un système mécatronique, dans le cas des robots sociaux.

L'analyse et l'interprétation des comportements sociaux se réalise à l'aide de programmes de reconnaissance vocale, d'estimation d'engagement/d'attention, de caractérisation des émotions, etc. Dans le cadre des ACA, la génération de comportements consiste principalement à animer un personnage. Pour cela, des chercheurs ont créé des logiciels spécialisés, appelés "réalisateurs de comportements", qui, couplés à un synthétiseur vocal, permettent de générer les animations et sons nécessaires en temps réel. Dans nos travaux, nous avons utilisé un de ceux-ci, appelé SmartBody [196].

Concernant la gestion d'interaction, il existe plusieurs méthodes. Les plus simples consistent à créer un graphe ou un arbre conversationnel prédéfini, généralement implémenté sous la forme d'une machine à états finie. Les plus complexes consistent à faire apprendre à la machine à converser, à partir d'exemples de conversations, selon des méthodes de Machine Learning.

Enfin, plusieurs langages XML spécifiques à un domaine ont été proposés dans la littérature pour permettre de représenter les données au sein d'un système d'ACA : le Behavior Markup Language (BML), qui permet de représenter les comportements devant être produits par un ACA, décomposés en actions élémentaires (gestes, posture, direction du regard, paroles, etc.) qui peuvent être combinées et synchronisées ; le Function Markup Language (FML), qui permet de décrire des intentions de communication ; le Perception Markup Language (PML), qui permet de représenter les comportements de l'utilisateur au niveau perceptif (position spatiale, pose de la tête, énergie du signal vocal, etc.) et au niveau analytique (niveau attentionnel, émotion exprimée, etc.) ; et le Speech Synthesis Markup Language, qui permet de donner des instructions à un synthétiseur vocal (intonation, phonétique, etc.).

Dans nos travaux, nous nous sommes appuyés sur l'estimation d'attention, la reconnaissance vocale, la description de comportements en BML et leur réalisation par SmartBody [196], couplé à un moteur de jeu, chargé du rendu, et à un synthétiseur vocal. La gestion d'interaction est faite à l'aide de machines à états finies.

Around 328 B.C., Aristotle wrote: "Man is by nature a social animal"; since then, scientists in various domains have identified biological structures that tend to confirm this statement [213]. In fact, a number of structures in the human body seem to be particularly adapted to interpersonal communication. Our ears are sensitive to the frequencies of the human voice;

our vocal tract allows us to produce articulated sounds to form words; and our brains are equipped with “mirror” neurons whose main function is to improve our awareness of other people [213]. This is why relying on people’s natural verbal and nonverbal interpersonal communication capabilities seems to be a good choice for conceiving advanced human-computer interactions (HCI). Machines, on the contrary, though designed by humans and for humans, and as “intelligent” as they might seem to be, are practically devoid of the most elementary communicative capabilities. As a result, people need to adapt to their functioning and interaction modalities by learning how to use them through complicated interfaces.

In the very beginning, computers were programmed using punch cards or other rudimentary means. As computing technologies evolved, HCI modalities changed as well. Today, the most common human-computer interface is the Graphical User Interface (GUI). It requires a display and can be controlled using a keyboard and a mouse or a touchscreen, which is today the preferred interface for mobile devices [100]. As it is considered user-friendlier than classical GUIs (with a keyboard and a mouse), its use by older adults has been deeply investigated and has proven to be as effective as it is for younger adults, even when using a small display size [191].

To go towards more natural interaction, through verbal and nonverbal communication, part of the solution might come from an unexpected field of computer engineering: video games. For instance, Microsoft’s Kinect sensor, which was initially introduced to play with the Xbox 360 gaming console, allows user body and face tracking almost out of the box. These sensing capabilities allow to perform gesture and facial action recognition. This is why the Kinect has been used by many researchers in computer vision for human activity recognition and hand gesture recognition, among other things [84]. As gestures, postures and facial expressions are essential in natural interpersonal communication, this kind of sensor eases greatly the work of automatic nonverbal behavior analysis. The Kinect is particularly interesting as it also features microphones and is thus in itself sufficient to implement multimodal interaction. The availability and low price of this device and a few others, such as the Arduino[2] board, makes it feasible and economically sound to disseminate socially interactive solutions that previously remained confined to high-end laboratories. Thus, we feel justified in studying the potential of such devices for everyday use as assistive solutions.

In this chapter, we review the theoretical and technological background that makes the creation of socially interactive computer programs (or social robots) possible. We start by presenting some generalities about human communication in Section 4.1. In Section 4.2, we give an overview of the common system design for ECAs. Then, we detail how computers can have social-like interactions with people, thanks to user behavior analysis

[2] http://arduino.cc

(Section 4.3), artificial social behavior generation (Section 4.4) and dialog management (Section 4.5). We end this chapter with a focus on the use of domain-specific languages in ECAs (Section 4.6) and a discussion about the elements presented here that we chose to include in our ECA system (Section 4.7).

4.1 Human communication and social signal processing

The last decade has seen the appearance of a new discipline of computer science and engineering called “Social Signal Processing” (SSP) [156]. This recent field aims at sensing, measuring, analyzing and modeling, in a computational and automated manner, human social behaviors for the machines to reach some kind of understanding of these “social signals” and give them the ability to interact in a social-like way with humans. A quite comprehensive review of the field can be found in [213].

Nonverbal social signals have been defined as follows: “Social signals and social behaviors are the expression of one’s attitude towards social situation and interplay, and they are manifested through a multiplicity of nonverbal behavioral cues including facial expressions, body postures and gestures, and vocal outbursts like laughter” [212]. We present here a short taxonomy based on Salah, Pantic and Vinciarelli [177] and Vinciarelli *et al.* [212, 213].

The first distinction to make when talking about social signals is between verbal and nonverbal cues. Verbal social signals are just spoken words. Nonverbal social signals consist in all other aspects of human social interaction conveying information. This includes physical appearance (e.g., height, somatotype), gestures, body postures, space (e.g., interpersonal distance), facial expressions, eye gaze and vocal behavior (e.g., prosody, linguistic vocalizations, non-linguistic vocalizations, silences and turn-taking). These are illustrated on Figure 4.1. It is worth noting that “verbal” and “semantic” are not synonyms as nonverbal cues may contain semantic information. In fact, nonverbal messages are just as important as verbal messages, if not more.

Nonverbal messages have been classified by Ekman and Friesen [58] in the 1960’s as follows:

Affective/attitudinal/cognitive state: information about how one feels (e.g., joy, stress, etc.), what attitude that a person adopts in a given social context (e.g., disagreement, politeness, etc.) or how available to converse the person is (e.g., inattention, fatigue, etc.);

Emblems: culture-specific actions that convey semantic information, such as a wink or a raised thumb;

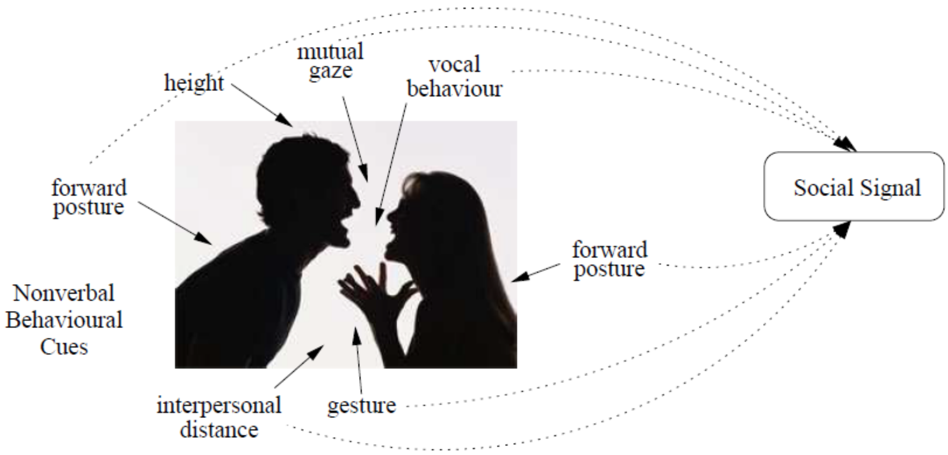


Figure 4.1: Behavioral cues in human communication: two people having an argument. Image taken from [212]

Manipulators: actions used to act on objects of the environment or self-manipulation such has head-scratching or lip-biting;

Illustrators: actions accompanying speech, such as pointing or gazing at an object to designate it or raising eyebrows;

Regulators: conversational mediators, such as head nods or smiles, that coordinate the timing of other signals during conversation and are mostly used to give feedback in a conversation (this feedback is usually called backchannel).

This classification shows how rich and complex nonverbal communication is in social interaction between people, as it the mean by which a lot of information that is complementary to a verbal message can be conveyed. For instance, illustrators, such as pointing gestures, are very useful to lift any ambiguities about the object under discussion. Using embodied social agents, such as robots or ECAs therefore allows to convey more information to the user than with a disembodied voice. Similarly, sensing, analyzing and interpreting the nonverbal communication of the user allows artificial agents to behave more socially, by responding to these additional cues.

Social interaction, as the prefix *inter* suggests, requires both listening actively and responding. Hence, sensing the social signals and interpreting them is not sufficient. This is why, in the following sections, social behavior analysis and artificial social behavior generation are both discussed.

4.2 Embodied conversational agents design

The software in ECAs and social robots that allows them to interact with humans using natural communication can be built from models from the aforementioned disciplines. Figure 4.2 shows the typical structure of an artificial socially interactive system. The main difference between robots and ECAs when performing social interaction resides in the fact that the outputs, that is to say gestures and other expressions, are not implemented in the same way. ECAs display communicative behaviors by animating a virtual character displayed on a screen, thanks to computer graphics, whereas robots physically execute actions using a mechanical system involving actuators and electronic control systems. Unlike gestures and facial expressions, speech is produced the same way in both systems, thanks to speech synthesis and speakers. As a result, robots may be limited in terms of expressiveness by their mechanical capabilities, whereas ECAs may display very realistic behaviors, albeit in a virtual way.

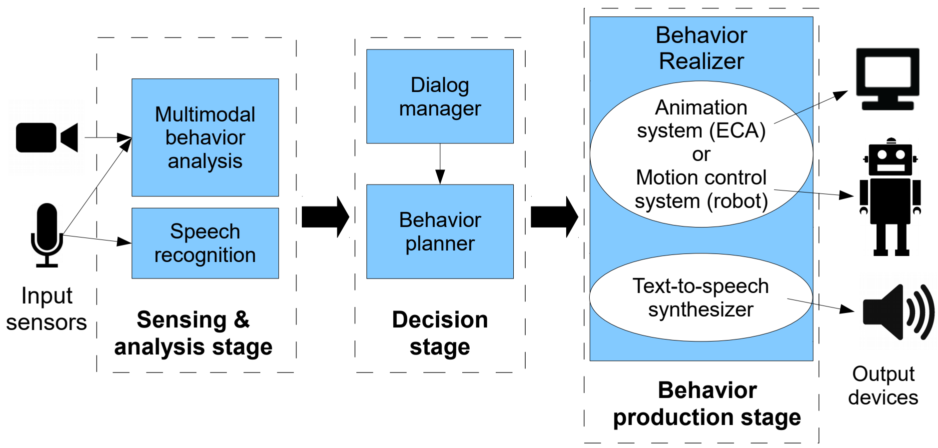


Figure 4.2: Common architecture of social interaction software in robots and embodied conversational agents

To design autonomous socially interactive systems, five elements are required.

1. Input sensors. The first step consists in acquiring data from the surrounding world. For social interaction, the most common sensors are video cameras and microphones. Sometimes, more sophisticated sensors, such as motion sensors, can be used.

2. Behavior analysis module. The input data is analyzed to extract information about the user's behavior. This stage corresponds to the perceptive capabilities of the system and processes user input.

3. Decision module. This stage takes the semantic information extracted in the sensing stage and decides what to do; that is to say, what social behavior the system should produce in response to the user's input. The decision stage corresponds to the cognitive capabilities of the artificial socially interactive system.

4. Behavior-production stage. This stage produces the desired gestures, expressions and sounds of the artificial social system, which compose the output sent to the user. This stage allows the behavioral capabilities of the ECA or robot.

5. Output devices. The calculated behavior is then transmitted to the user through speakers (for the voice and sounds), and a display in the case of ECAs or the robot's body.

These five elements are put together as illustrated in Figure 4.2. Note that the process is sequential. However, there may be feedbacks from a module to the previous one. This figure synthesizes the most common interactive system design frameworks found in the literature. The interested reader can refer to Kopp *et al.* [107] and Scherer *et al.* [181] about ECAs and to Huang and Mutlu [91] about robots. This type of software architecture may rely on domain-specific XML-based behavior representation languages, as proposed in the SAIBA (Situation, Agent, Intention, Behavior, Animation) framework [107, 211]. This aspect is covered in details in Section 4.6. The most popular language in the SAIBA framework is Behavior Markup Language (BML). It allows describing social behaviors by assembling elements of nonverbal communication (gestures, facial expressions, head movements and gaze) and synchronizing them with each other and with speech. This can be interpreted by a behavior realizer, as explained in Section 4.4.

4.3 User behavior analysis

Computerized analysis of social signals is essentially composed of three steps: (a) sensing the environment (mainly through video and audio recording); (b) detecting people in the incoming signals; and (c) interpreting their behaviors. Interpreting social behaviors through audio and video data, sometimes combined with other information sources, is a challenging task. It often requires knowledge of the context, which is not usually available to a computer, unless it is known a priori, when performing a specific task for instance. In addition, models to characterize human social behaviors may be difficult to formalize in a software application. This is why social signal processing researchers often rely on machine learning, which consists in analyzing large amounts of data to infer classification rules, to automatically recognize social behaviors.

Wagner *et al.* [216] have proposed a generic framework for SSP, called Social Signal Interpretation[3]. It is a C++ library that allows to acquire and synchronize input data from various sensors, then extract features, train machine-learning classifiers for real-time recognition and perform multimodal data fusion. It includes a number of signal processing libraries, as well as some implemented machine-learning algorithms that can be trained to build custom models, if the ones already available in the library do not fit your needs.

There are several aspects of human communication that were studied by social signals researchers. Some works have focused on having computers understand highly abstract notions of context and social behaviors. According to Vinciarelli *et al.* [213], these include social relations and role recognition, that is to say determine what role or status each participant of an interaction plays, social attitudes (dominance and personality) and social emotions, such as empathy, envy or admiration. However, most works in the field have focused on measuring and modeling basic aspects of non-verbal behaviors. These include appearance (beauty in particular), facial expression analysis (for emotion recognition), vocal analysis (for emotion recognition, turn-taking, laughter and hesitation detection) and gaze directions [212, 177].

In the context of this thesis work, we are mainly interested by two aspects: the verbal messages and the attention payed to the ECA or conversational engagement of the user. In the following section, we provide some details about automatic speech recognition, which allows to transcribe the verbal content of speech into text, to address the first point. Then, in Section 4.3.2, we address the second point by reviewing the literature on detecting and measuring attention and engagement in human-machine interaction. Lastly, in Section 4.3.3, we provide some details about emotion recognition methods, which could be an interesting source of information in our context, as people with dementia may have difficulties in expressing their feelings.

4.3.1 Automatic speech recognition

Automatic Speech Recognition (ASR) consists in extracting the verbal content of a speech signal and translating it into text. This is why it is sometimes called "Speech-to-Text". Today, though research is still conducted in that area, ASR has moved from research labs to large industry companies such as Google or Microsoft. State-of-the-art research platforms include CMU Sphinx[4] [114] and Julius[5] [113]. However, though these systems are

[3]This software is distributed under the open-source GNU General Public License (GPL) at http://www.openssi.net/

[4]http://cmusphinx.sourceforge.net/

[5]http://julius.osdn.jp/en_index.php

based on phoneme or diphone recognition, and therefore are agnostic to the language to recognize, the main challenge of performing reliable ASR (especially when supporting several languages) is to rapidly analyze a speech segment against a very large database. This is why big companies or specialized companies, such as Nuance[6], are taking over in this field. Very recently, Google and Microsoft opened pay-as-you-go commercial ASR cloud services (accessible for free when creating applications for their respective operating systems), which allow developers to easily integrate high-end ASR in any software, provided that an Internet access is available. Furthermore, the first few hundreds or thousands API calls are free of charge, which allows to benchmark these solutions for free. Lastly, ASR, which usually only relies on audio processing, can be combined with video-based lip reading [109] but machine lip-reading techniques are not very reliable at the moment.

4.3.2 Attention and Engagement

To detect and measure attention and engagement, several indicators have been used in the literature. The most common are eye gaze, head pose, body posture and the presence of backchannel behaviors or emotional reactions, which can be combined, through multimodal fusion, to achieve better robustness. According to [11], the most important features are eye gaze and head pose. However, eye gaze is sometimes difficult to observe, especially when using low-definition cameras, or from a distant location. In addition, systems for precise eye gaze estimation often require calibration. This is why, in [190] and [192], head pose is used as a proxy for mean eye gaze direction, to estimate one's visual focus of attention. Several studies show how to determine the face's orientation to estimate the focus of attention in various contexts: in front of a computer screen [12], during interaction with an ECA [95, 39] or a mobile robot [192]. Some methods are only dedicated to estimating the user's level of attention [31, 11, 117, 12]. Other methods try to estimate the more abstract notion of conversational engagement, when a person is interacting with a robot or an ECA [167, 192, 95, 39], or defined as the intention of interacting with a robot [17]. As seen earlier in Section 3.1.3, attention is a necessary condition for engagement. Therefore, most methods to estimate the user's engagement in human-machine interaction use behavioral cues of attention display. This is why we include works on both notions in this quick review. Let us now go into the details of some of the methods found in the literature, which inspired the attention estimation method we propose (see Section 5.2.2).

In [117], Li, Xu and Tan have proposed to use attention estimation, for a robot to select its addressee actively, based on who is the most attentive, in a multi-party interaction setting. The user's attention is estimated

[6]Service provider for Apple's SIRI and Aldebaran Robotics. Official website: http://www.nuance.com/index.htm

using several features: proximity from the robot, upper-body orientation, face detection, gaze detection, mouth motion detection and facial expression recognition. Each of these features is extracted separately, thanks to a specialized algorithm. A weighted sum of the features (after normalization) is then computed to determine the attention level for each person in sight of the robot's stereo camera. Interestingly, this method does not require machine learning. In [192], the direction and level of attention are estimated by measuring the user's head pose on the horizontal and vertical axes. Direction of attention is estimated by computing the average orientation over several video frames, whereas the level of attention is estimated by computing the variance over the same time interval. High variance is considered to be related to low attention. To go from attention to engagement, the authors add nod detection and emotion valence (positive or negative) classification, obtained through facial expression analysis. Lastly, in [18], engagement, that the authors define as the intention to interact, is estimated with the help of a Microsoft Kinect sensor. They show, through a statistical feature reduction, that only 7 features are sufficient, without loss of precision: the relative orientation of the body and shoulders, distance from the sensor's vertical plane, the speed on the axis parallel to the sensor's vertical plane, the size of the head seen by the sensor, the x position of the head in the color image produced by the sensor, the activated sound beams and the sound source localization angle.

4.3.3 Emotion recognition

In the previous section, we have seen that the presence of emotional response is an indicator of engagement in face-to-face conversation. In fact, the expression of emotions plays an important role in human communication, as it conveys information about a person's beliefs, desires and intentions [126]. It also informs about how one feels, which may trigger empathy in other people. As it is a valuable source of information in human-computer interaction, many researchers have investigated the automatic recognition of emotion, to enrich conversation with social robots and virtual humans. This means building computational models of emotions and developing methods to recognize emotional and affective states from perception data.

According to Gunes *et al.* [81], there are three main types of psychological models of emotions that have been used to build computational models:

- categorical models, which consist in classifying emotions according to discrete basic emotion categories, such as Ekman and Friesen's six universal basic emotions [59], which are anger, happiness, fear, sadness, disgust and surprise, to which we can add a neutral state;
- dimensional models, which consider emotions to be a continuous, rather than discrete, phenomenon and characterize them according to two

axes, valence (pleasure–displeasure) and arousal (arousal–sleepiness), as in Russell's circumplex model of affect [175], or up to five axes, adding dominance, expectation and intensity;

- appraisal-based approaches, which are founded on "attempts to detail the mental processes underlying the elicitation of emotions" [126] and work under the assumption that all affective states are linked to an emotion-inducing situation, they therefore consists in evaluating how one feels about a situation given his or her goals and concerns (is this situation desirable?, what/who caused it?, etc.) – the interested reader can find more detailed information in [126].

Regarding the modalities of emotion and affect recognition, some authors have worked on facial expressions [135], speech [106] and body movements [99]. The references given here are surveys of the methods proposed for each of these modalities the interested reader can refer to. Lastly, an overview of computational models of emotion and their applications is presented in [126].

4.4 Artificial social behavior generation

To produce artificial social behaviors, one needs to create virtual characters and animate them. This can be done in a cartoon-like fashion, with two-dimensional drawings, or in a more realistic way, using three-dimensional models, as it was done for LOUISE. Animated 3D characters are usually composed of three elements, as illustrated on Figure 4.3: a 3D mesh, which defines the spatial geometry of the model; a texture or "skin", which defines the visual aspect of the mesh's surface; and an animation skeleton or "rig". The animation skeleton is in fact a hierarchy of 3D points called "joints", bound to one another by "bones". The joints' positions and organization in an oriented-graph-like representation allows defining where the mesh deforms when the skeleton is moved. For instance, when modeling a human arm, joints will be positioned at the shoulder, elbow, wrist and at the base of each phalanx of the fingers. In this configuration, the shoulder's joint is the root of the graph, the elbow is a child of the shoulder, the wrist is a child of the elbow and each finger is a child of the wrist. The character constructed in such manner is then animated by recording a sequence of rotations and translations of the skeleton's joints. Animations are usually created by placing key positions and rotations of the joints, called "key frames", on a timeline and interpolating between these key frames. The modeling and animation process is done thanks to dedicated software such as Autodesk Maya[7] and Cinema4D[8], which are professional industry-standard software products or

[7] http://www.autodesk.com/products/maya/overview

[8] http://www.maxon.net/products/cinema-4d-studio/who-should-use-it.html

Blender[9], a credible open-source alternative.

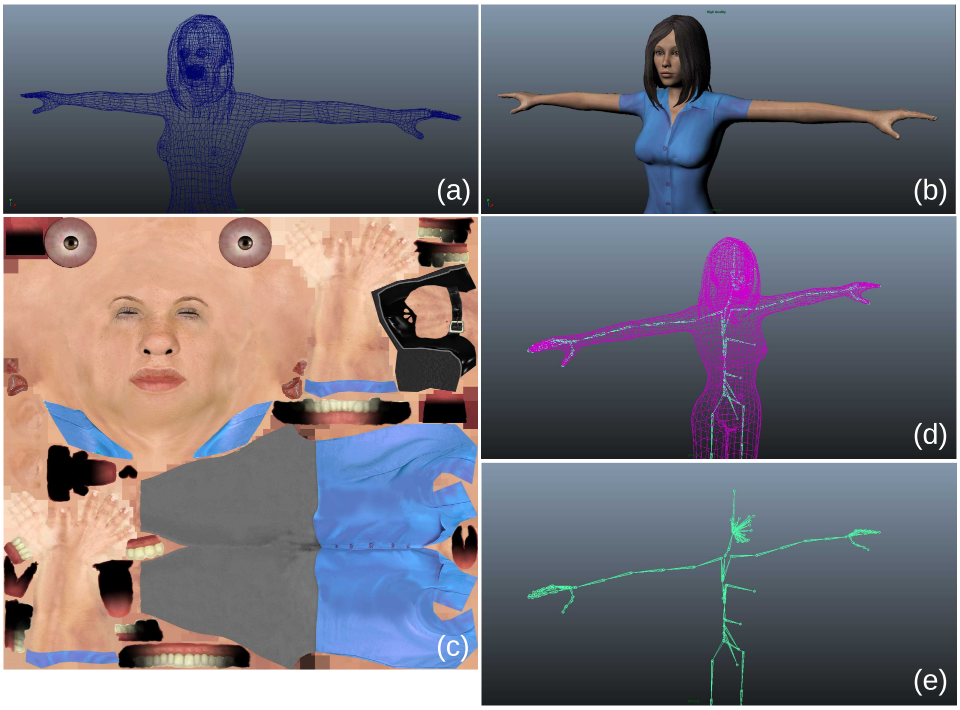


Figure 4.3: Elements of a 3D character model. (a) 3D mesh. (b) 3D mesh with texture. (c) Texture map. (d) Skeleton in mesh. (e) Skeleton.

4.4.1 Behavior realizers

Animating virtual characters is a very time-consuming process that requires professional expertise. In addition, once animations are created, it may prove quite tricky to combine them to create artificial social behaviors. To solve this issue, several research teams have created software elements that allow describing character animation in terms of gestures, facial expressions, head movements, gaze directions, etc. These are called BML realizers, as they use Behavior Markup Language for behavior representation. The key feature of these tools is to perform the synchronization of the lips' animations and other movements with speech playback. Several BML realizers proposed in the literature are listed in Table 4.1. Some of them are publicly available for research.

Other BML realizers were proposed in the literature, such as Elkerlyck [219], which is a former version of AsapRealizer, or RealActor [37]. Among the systems made available for the research community, SmartBody is the most comprehensive tool: it allows to add character models and behaviors;

[9] http://www.blender.org/

Table 4.1: State-of-the-art behavior realizers for ECAs

Name	Reference	Comments	Publicly available
Greta	[155]	Greta is the first SAIBA compliant ECA. It can interpret both FML and BML. It is implemented in Java.	Yes
MARC	[49]	MARC is a research platform written in Java. It features a BML editor with a GUI and facial and body animation editors.	Yes
SmartBody	[196]	SmartBody is a very comprehensive ECA creation tool. It comes with a large database of behaviors, which can be applied to custom character models, thanks to its animation retargeting feature. It is written in C++.	Open source
AsapRealizer	[206]	AsapRealizer is a BML 1.0 realizer. It is made for incremental behavior construction and interactional coordination.	No

it is natively compatible to work with several speech synthesis engines; it can run on Android and iOS; it comes with a large database of behavior animation files, which can be used with custom characters, thanks to its motion retargeting capabilities; and there are a some behaviors (head movements, gaze direction and eye saccades) that are computed procedurally (instead of depending on keyframe animation files). It is also the most flexible, as it is open-source; it can be integrated with several game engines and its application programming interface (API) is available in C++ and Python. In addition, it is well documented, which is rare enough in research platforms to be mentioned. For our purpose of explorations in interactions between ECAs and older adults, it is the most practical tool. This is why it was used to build the second prototype of LOUISE. SmartBody is also distributed as part of the Virtual Human Toolkit [85], which includes other tools to build Human-ECA interactions. Lastly, it is able to communicate with other software modules through the Apache ActiveMQ[10] middleware, which allows to build easy-to-extend ECA systems, in a modular architecture. Figure 4.4 shows the results of characters' rendering in some of the BML realizers mentioned in Table 4.1.

Though this is out of the scope of the present research work, it is worth mentioning that a peculiar kind of behavior realizer was proposed by Heloir and Gibet [87]. Indeed, this virtual human animation system is meant to perform communication in sign language with deaf or hard-of-hearing people.

4.4.2 Speech synthesis

As for ASR, many speech synthesis engines are built by large companies or smaller, specialized companies. The best engines currently available are Microsoft Speech API, Cereproc's CereVoice SDK[11], which is specialized in regional accents and emotional voices, and Acapela[12]. Microsoft Speech API is free to use on Windows systems and Cereproc offers a free academic license for CereVoice SDK.

4.4.3 Evolving human-agent relationships and personality models

Personal robots and ECAs may be used as artificial companions. This use case requires installing the system in people's homes for a long period of time. Some researchers suggested that in this context the relationship between the system and the user should evolve, following human-like social constructs [46]. In human societies, the topics of conversation and the kind

[10] http://activemq.apache.org/
[11] http://www.cereproc.com
[12] http://www.acapela-group.com/

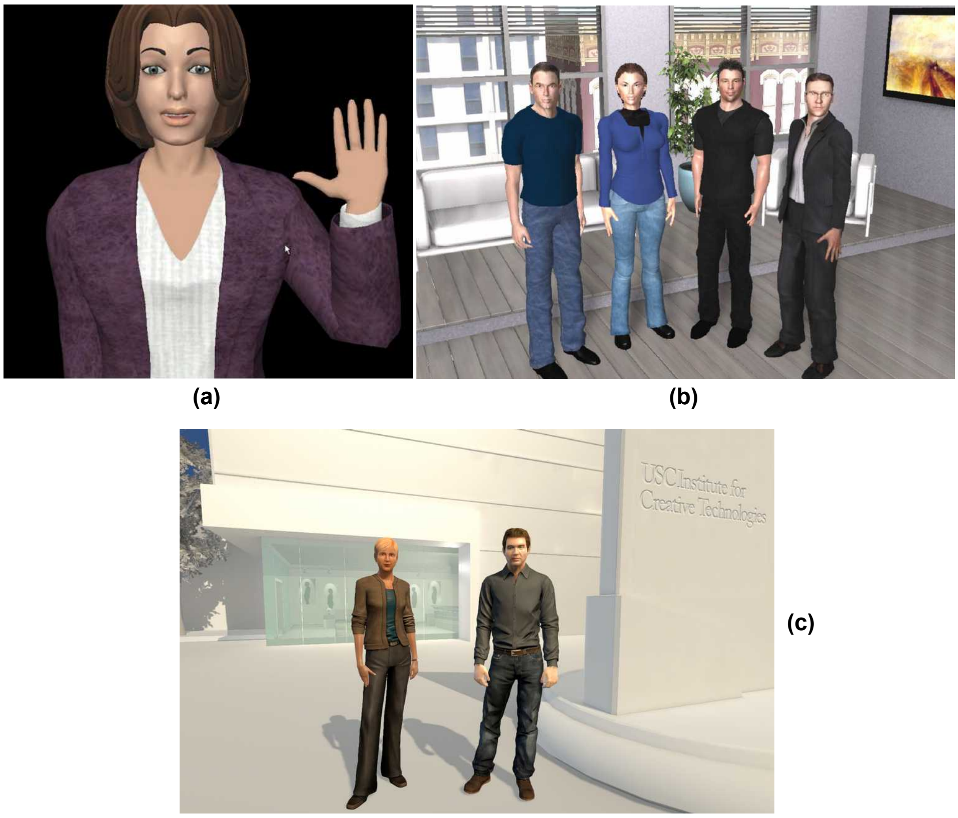


Figure 4.4: Virtual humans from BML realizers. (a) Greta (image taken from [155]). (b) Characters in MARC (image taken from [48]). (c) Screenshot of the demonstration scene of the Virtual Human Toolkit.

of activity a person would share with someone else depend on the relationship existing between the interaction partners. For instance, strangers would not discuss highly personal issues with each other. Instead, they would keep the conversation to "social chit chat" such as discussing the weather or talking about sports. They could also share casual activities such as playing cards.

This would ideally go the same for human-robot or human-virtual agent relationships, and after some time the interaction possibilities could get more extensive. In the model proposed by Coon, Rich and Sidner [46], the relationship can evolve through three stages: (a) stranger, (b) acquaintance, and (c) companion. Each transition from a stage to the next unlocks new interaction possibilities. Another model was proposed by Pecune [153], based on research in social sciences. This model has four levels of intimacy: (a) superficial, (b) intimate, (c) personal, and (d) care.

This kind of relationship evolution principles have already been applied in the social simulation game The Sims[13]. In this game, the player controls a family of virtual humans and can act on every aspect of their lives, from birth to death. As two "Sims" (the name given to the virtual characters in the game) develop a closer relationship with each other, more actions become available and the risk of rejection decreases. For instance, as they become friends, they will be able to tickle the other "Sim" or tell it salacious jokes and the interaction partner gets more likely to react positively. In the aforementioned game, the characters also have virtual personalities that influence their behaviors. Giving a personality to a robot or an ECA allows making it more sociable and more believable to the user. Some authors have explored the possibility of integrating a personality model in social machines. Giving them a personality, "a set of stable and individual behaviors" is believed to endow the interactive system with "consistency and coherence in behaviors" [66].

Faur *et al.* [66] have proposed a model of personality for virtual agents based on socio-cognitive theories of personality: two self-regulatory processes (self-discrepancy and regulatory focus) are formalized and implemented in a computational model. A more common approach is to implement a traits-based model of personality. The "Five Factors Model" by Costa and McCrae [47], also called "The Big Five", or a subset from it, is often used for this purpose. This model uses five dimensions to represent an individual's personality: neuroticism, extraversion, openness, agreeableness and conscientiousness. For instance McRorie *et al.* [131] based their model of personality on Eysenck's model [63] which features only three dimensions: extraversion and neuroticism, which are similar notions and named the same as in the Five Factors Model, and psychoticism, which somehow corresponds to agreeableness and conscientiousness combined.

[13] http://www.thesims.com

4.5 Dialog management

The first and most important element in the decision stage is the dialog manager. Its task is to follow the course of the interaction along a dialog tree, either established a priori or constructed dynamically. Then, given the current state of conversation and the user's input, the system may decide what behavior to produce next, that is to say what utterances to speak out and what gestures and expressions to display.

As summarized by Leonhart [115], dialog, seen as a "connected sequence of information which provides coherence over the utterances", presents four key characteristics:

- turn-taking behavior, a set of rules that allow determining who should talk and when;
- utterances, that are the basic elements to construct a dialog and can be of different types (questions, answers, greetings, etc.);
- grounding, the common beliefs shared by the speaker and the listener, which can be established collaboratively during the interaction;
- and conversational implicature, a set of rules that help producing and interpreting utterances.

Leonhardt [115], citing Rajman[164], adds that managing interaction in ECAs requires handling several dialog control tasks: recognizing the type of dialog act, lifting ambiguities, confirming with the user in case of difficulties, handling errors, inferring missing information, switching context when managing more than one task at a time, grounding information and responding to the user.

According to McTear, Callejas and Griol [132], the simplest strategy to manage dialog is to represent all possible entries of the agent's utterances and the user's responses in a graph, usually implemented in the form of a finite state machine. The interaction is managed by navigating the graph, thanks to transitions between states, based on handcrafted rules. The authors argue that this method, used in many voice interfaces in the industry, as it is simple and cheap, has three major limitations: the interaction needs to be directed by the system, which does not allow the user to take the initiative of the dialog; it lacks versatility, as only a small number of choices can be available on each node; and it has poor domain portability. To go beyond the manual authoring of these conversational structures, Rich and Sidner [169] proposed a model to partially automate dialog tree authoring. This method combines hierarchical task networks with traditional conversation trees and is based on collaborative discourse theory. It consists in establishing the high-level hierarchical goal structure for the dialog, then formalizing its structure and control flow in a hierarchical task network, before adding application-specific subdialogs at the fringe. The generated tree is then iteratively refined. Lastly, the automatically generated utterances can be modified to add "color" to them.

Grios *et al.* [80], as well as McTear, Callejas and Grios [132], stated that another classical approach is the frame-based dialog management. It consists in representing the goal of the conversation as a frame, containing slots, each corresponding to a piece of the information that has to be gathered for the goal to be fulfilled. The typical use of this method is for form-filling scenarios, in which the system asks the user questions to obtain information. This allows for mixed-initiative dialog, as the user can provide several pieces of information at a time, which will be stored in the corresponding slots. In addition, the order in which the information is acquired does not matter. The same authors cited a more advanced version of frame-based dialog management called "information state" dialog management [200], which consists in storing all previous information provided by the user during the dialog and using it to determine what the next conversational action of the system should be.

Lastly, in their overview of the field, McTear, Callejas and Grios [132] mention a number of machine learning-based approaches to dialog management. The idea is to learn conversational strategies through statistical methods, by feeding the system corpora of dialogs, or through automated trials-and-errors methods; this is called reinforcement learning. The interested reader can find quite comprehensive information on the subject in [132], Chapter 10.

As seen in Section 2.2.1, to the best of our knowledge, only one work [92] focused on managing interactions between people with dementia and ECAs. In the system proposed in the article, conversation is performed by looking for keywords in the input speech and searching a database of conversational pairs to find an appropriate response. In addition the virtual agent performs pitch variation analysis and silence detection to perform active listening, that is to say, produce backchannel behaviors at appropriate times. An interesting work on dialog management in ECAs, tailored for older adults and cognitively-impaired people, was proposed in [230]. They created a dialog manager called *flexidam* that consists in a set of information collection tasks, called issues, to create calendar entries. Each task is composed of subtasks to collect elements of information (date, time and activity), in a similar manner as in the frame-based approach. They also add extra flexibility by allowing to leave information frames incomplete and fill them in later. Lastly, other works have focused on speech interfaces for older adults, with or without cognitive impairment, but with no embodiment. For instance, the COACH system [29], a task guidance system for older adults with cognitive impairment (they implemented hand-washing as an example), uses a flexible task model and probabilistic activity recognition to send prompts at appropriate times. The system allows for a degree flexibility regarding the order in which some of the actions are executed and can handle uncertainty in the activity recognition. Another example is the work of Vipperla, Wolters and Renals [214], who studied the specificities of vocal interaction for older adults. They compared various strategies for menu navigation (the old issue of depth versus width) and confirmation (implicit, explicit or no conversation).

4.6 Domain-specific languages in embodied conversational agents

As seen earlier in this section, the SAIBA framework consists in using domain-specific languages (DSLs) to describe verbal and nonverbal behaviors in ECA systems, namely BML and FML. There also are other XML-based DSLs that could intervene in ECA design: Perception Markup Language (PML) [180], which allows describing the user's nonverbal behaviors; and Speech Synthesis Markup Language (SSML) [195], which is mostly intended to control prosody and language parameters of speech in speech synthesis engines.

4.6.1 XML

XML (eXtensible Markup Language) is a general formalism to structure data, used in many computer applications. For instance, HTML is used to describe the contents of web pages. XML languages are quite verbose and rely on only one type of entries: markups. The markups then have a hierarchy. An XML markup is easy to identify; it always starts with a “<” and finishes with a “>”. There are two types of markups: single tag and double tags.

Single tag markups use the following syntax:

```
<tag_name/>
```

Note that the “/” character is used to tell it is the end of the markup. Double tag markups have a start tag and an end tag. They use the following syntax:

```
<tag_name> content </tag_name>
```

The start tag is characterized by beginning with “<” and ending with “>” and the end tag begins with “< /” and ends with “>”.

Most XML markups allow for parameters. The parameters are put in the first tag for double tag markups. In XML, the order of the parameters does not matter; parameters are used as follows: name_of_the_parameter = “value”. Also note that the name of the parameter has to be followed by an “=” sign and the value is written between single or double quotation marks ('value' and "value" are equivalent). If a markup contains more than one parameter, they are simply separated by white spaces, for instance

```
<tag_name parameter1="value1" parameter2="value2"/>
```

for single tag markups and

```
<tag_name parameter1="value1" parameter2="value2">
    content
</tag_name>
```

for double tag markups. For a given markup, some parameters may be optional. If they are not used, they can be simply omitted.

Lastly, all XML files must start with a markup that indicates the version of XML used and the file encoding scheme. For instance, in

```
<?xml version="1.0" encoding="UTF-8"?> ,
```

we are using version 1.0 of XML specification and the document is encoded using the Unicode UTF-8 (variable bit length) character set.

4.6.2 BML

As stated earlier, Behavior Markup Language (BML) [107, 211] is dedicated to describing communication behaviors by putting together elementary actions: speech, gaze, head movements, facial expressions, posture and gestures. There are currently two versions of the BML specification: BML 1.0 [134] and SmartBody BML [196] which are not fully compatible [207]. This is due to the fact that the team developing SmartBody made extensions to the early specification of the language, before the BML 1.0 specification was issued.

In the context of this work, as LOUISE, in its latest version, is based on SmartBody for behavior realization, we present the SmartBody version of BML. There are a number of behaviors that can be produced by our ECA. It can perform gestures, through animations, speech with lips synchronization, head movements, gaze direction, face expressions, eye saccades and selecting an "idle" animation used when the character is at rest (for posture). This section is only meant to give a quick reference of the typical use of these commands in our software. There is also an event system that allows for extensions. Its use in our program is detailed in Section 5.4. SmartBody allows other behaviors, such as reaching, grabbing and locomotion, but they are not implemented in our system. Also note that most of these behaviors are performed once, when the command is called. However, three of them are persistent behaviors (the effects of the command will continue until another command of the same type is used): Gaze, Idle and Eye saccades.

Simple commands

Here is a list of the BML commands used in our system with their syntax and an example of use.

- To set a body posture, or idle pose the BML markup is

  ```
  <body posture="animation_name"> .
  ```

  For example:

  ```
  <body posture="ChrHarmony@IdleHandOnHip01"/> .
  ```

- For eye saccades, SmartBody allows for 3 modes, which are performed procedurally, based on a computational model [196]:

  ```
  <saccade mode="chosen_mode"> .
  ```

  Valid modes are “listen”, “talk” or “think”. For example:

  ```
  <saccade mode="talk"/> .
  ```

- For gaze, the markup is the following:

  ```
  <gaze target="target" sbm:joint-range="list of joints"/> .
  ```

  This command will make the character gaze at the specified object. It must be an object known in the SmartBody scene. In addition, the joint range parameter is a custom feature of SmartBody, compared to standard BML. For example:

  ```
  <gaze target="front" sbm:joint-range="EYES NECK"/> .
  ```

  Valid joints that can be listed in the joint range argument are: EYES, NECK, CHEST, BACK. It is possible to list all of the joints you want to be affected by the command or the range. “EYES BACK”, for instance, will include EYES, NECK, CHEST and BACK. This is very useful to finely control gaze behaviors to provide attention direction cues during interaction.

- Animations are used to perform gestures, and invoked as follows.

  ```
  <animation name="animation_name"> .
  ```

  For instance:

  ```
  <animation
      name="ChrBrad@Idle01_InclusivityPosBt01"/> ,
  ```

where "ChrBrad@Idle01_InclusivityPosBt01" is an animation file, represented in the format imposed by SmartBody. This allows playing key frame animations. It is also possible to define gesture types by associating key frame animations to the type of communication intent the gestures they allow to perform correspond to. Once this association is specified in a gesture lexicon, it is possible to invoke gesture behaviors as follows:

```
<gesture type="type_of_gesture"/> .
```

This is the preferred way of calling gesture behaviors in the BML 1.0 specification, which advises to create a gesture lexicon.

- Head movements are performed procedurally, unlike other gestures, which rely on keyframe animations. Here is the markup to invoke these behaviors:

  ```
  <head type="type_of_mvt" repeats="N"
      velocity="v" amount="a"> .
  ```

  The only mandatory parameter is type. Valid types are NOD, SHAKE, TOSS (side-side), WIGGLE (nods of decreasing intensity) and WAGGLE (shakes of decreasing intensity). Here are examples of head movement markups:

  ```
  <head type="NOD" amount=".5"/> ,
  <head type="WIGGLE" repeats="3"/> .
  ```

- Face expressions are invoked as follows:

  ```
  <face type="chosen_type" au="au_number"
      side="side" amount="number">
  ```

  Face expressions in SmartBody BML are described using Ekman and Friesen's Facial Action Coding System (FACS) [60]. The type parameter will therefore always be "FACS". Animation Units (AUs) are numbered and correspond to the contraction of a particular muscle in the face. The side parameter has to be set to "right", "left", or "both", depending on how the corresponding keyframe animation for that AU was created (two separate keyframe animations, one for each side of the face, or a single animation, which acts on both sizes). Here is an example of the corresponding code:

  ```
  <face id="sm" start="0" end="2" type="facs"
      au="12_left" amount="1.1"/>
  <face start="0" end="2" type="facs" au="12_right"
      amount="1.1"/>
  <face start="0" end="2" type="facs" au="6"
      amount="1"/> .
  ```

This example will produce a frank smile. This system is extensible in SmartBody and it mostly depends on what keyframe animations the designer provides. In our case, for instance, the 6 basic emotional face expressions can be used by calling only one AU "side" parameter set to "both". They are numbered with figures that are out of the range of the FACS specification, to avoid any confusion. The start and end parameters can be set to values (in seconds) to synchronize the elements of the face expression.

- Speech is the only behavior that uses double-tag markup:

```
<speech type="chosen_type">
    content of the speech
</speech> .
```

Valid types are "plain/text" and "application/ssml+xml". The former only allows to write text, whereas the later allows to give additional instructions to the speech synthesizer in SSML (see dedicated paragraph in Section 4.6.5). If the type parameter is unspecified, "plain/text" is chosen by default. Here is a working example of speech command:

```
<speech type="application/ssml+xml">
    My name is John
    <break time="0.5s"/>
    I am here to help you
</speech> .
```

This will make the character say "My name is John", then pause for half a second before saying "I am here to help you".

- Lastly, BML allows for easy extensions, thanks to the use of events:

```
<event message="message_to_send"> .
```

In our prototype, we have made use of events to display images, for instance. This is detailed in Section 5.4.

Composite commands

To describe complex behaviors, it is possible to use several commands at the same time by simply putting them one after the other. For instance,

```
<speech type="plain/text"> Hello! </speech><head type="NOD"/>
```

will make the character say "hello" while nodding.

All BML commands can be identified thanks to an "id" parameter. This allows naming a component to synchronize another behavior with it. For instance, in

```
<speech id="sp" type="plain/text"> Hello! </speech> ,
```

the speech behavior will be identified by the name “sp”.

All commands allow for a “start” parameter. The value for the “start” parameter can be a number, which stands for a time delay after the command was received by the BML realizer, expressed in seconds. For instance:

```
<speech id="sp" type="plain/text">
    Hello!
</speech>
<head id="hd" start="1" type="NOD" repeats="2"/> .
```

This command will make the character say “Hello!” as soon as the command is received and nod two times, starting one second after. The start parameter can also be used relatively to another behavior, using synchronization points. For instance,

```
<speech id="sp" type="plain/text">
    Hello!
</speech>
<head id="hd" start="sp:end" type="NOD" repeats="2"/>
```

will have the character nod two times after the speech is over. For all behaviors, start and end synchronization points are available; they can be referred to by writing “given_id:start” or “given_id:end”, “given_id” being the name given to the behavior.

Lastly, a command can be synchronized to an other command with an additional delay by writing start=”given_id:start+N”, given_id being the name of the behavior and N being the time delay expressed in seconds. For instance,

```
<speech id="sp" type="plain/text">
    Hello!
</speech>
<head id="hd" start="sp:start+0.5" type="NOD" repeats="2"/>
```

will cause the character to nod two times, half a second after it started saying “Hello!”.

Instead of imposing the start time of a behavior, it is possible to set its end time, or both start and end. A head movement or an animation can be accelerated or slowed down by setting the start and end parameters. For a face expression, it is used to control its duration. For instance the command

```
<face id="f1" type="facs" au="6" side="both" start="0"
end="1"/>
```

makes the character perform Action Unit 6 on both sides of its face for 1 second.

Some behaviors have other synchronization points than just start and end. Indeed, “animation” and “head” behaviors also have “ready”, “stroke” and “relax” parameters. Between “start” and “ready”, the character prepares to execute the movement; “stroke” is the most intense moment of the animation. For instance, when performing a pointing animation, the stroke time will be the time where the arm is fully directed towards the target, before it goes back to its original position. The “relax” is when the movement is almost over but the character is not yet back in its idling pose. This synchronization points can be used either to control the behavior or to synchronize other behaviors with it. This is illustrated on Figure 4.5. For instance,

```
<speech id="sp" type="application/ssml+xml">
    Go to the right.
</speech>
<animation id="a1" stroke="sp:end"
    name="ChrBrad@Idle01_IndicateRightRt01"/>
<gaze start="a1:stroke" target="right"
    sbm:joint-range="EYES NECK"/>
```

will make the character point to the right-hand side and gaze towards the right, at the same time, immediately after saying the word “right”. The movement of the pointing animation will have started before the speech is over but the gaze behavior will only start at the moment when the character’s hand is fully pointing.

Some synchronization marks can be added to the speech command, using the <mark id=“name”> markup. Lastly, to work properly, BML instructions will always have to start with the <speech>content</speech> markup, if there is speech in the desired behavior.

4.6.3 FML

FML [88] stands for Function Markup Language. Like BML, it is part of the SAIBA framework, but is not used as much. In fact, for now, the only behavior realizer to use it is Greta [155]. The reason for this is that it is experimental and, until recently, was not fully specified[14] [33]. When BML allows to describe communication behaviors in terms of actions (gestures, gaze, speech, etc.), FML describes them in terms of intents, that is to say what the agent wants to achieve. This DSL is in fact meant to specify the communication intent, to be translated into actions, described in BML.

FML allows encoding the following communicative functions:

- the degree of certainty the agent intends to express (doubt, uncertain, certain, etc.);

[14] The FML 1.0 specification is available at http://secom.ru.is/fml/

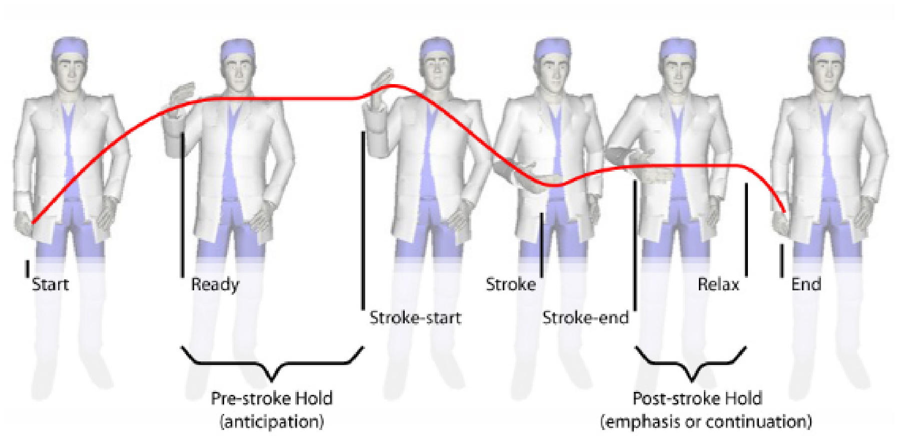


Figure 4.5: Synchronization points for BML behavior synchronization. Image taken from [107].

- the performative value of the intent (suggest, approve, disagree);
- the goal of stating the relationship between two elements of the discourse;
- turn-taking (giving or taking a conversational turn);
- emotional state;
- emphasis;
- backchannel attitude;
- referring to objects.

Given our use of SmartBody and the experimental nature of FML, we decided not to use it for our projects.

4.6.4 PML

Perception Markup Language (PML), proposed by Scherer *et al.* [180], allows to represent the nonverbal behaviors of a person in interaction with a conversational agent or robot. It is meant to be a standard way of representing behavior perception data, obtained through various sensing and analysis algorithms. PML takes inspiration from two World Wide Web Consortium (W3C) representation languages: VoiceXML, which allows speech and language representation, and EmotionML, which is dedicated to user emotional state coding. PML has three particularities:

- it allows to deal with uncertainty, inherent to behavior sensing systems;

- it allows to represent data according to three interpretation layers, namely sensing, behaviors and functions;
- it is easily extensible, to account for future new possibilities in automatic behavior analysis.

The possibility of representing data according to three layers of interpretation, of increasing abstraction levels, enables the use of PML in several stages of behavior interpretation and allows to use it to represent the input of a dialog manager.

A PML message contains two main sections: <header> and <person>. The former allows to pass meta-data, such as the source of information or the time stamp, and the later is meant to represent the behavior of a particular user. The person section can be subdivided in three sections, which correspond to the three interpretation layers, marked respectively with the <sensingLayer>, <behaviorLayer> and <functionLayer> tags. The sensing layer contains low-level information about the current audiovisual state of the user. It includes elements such as <gaze>, <headPose> or <posture>. These elements allow for fields that contain numerical values, such as the head's vertical and horizontal orientation angles for <gaze>. The behavior layer contains information about the interpretation of the perception data. This information is represented in terms of <type> of behavior and strength of this behavior, either numerically, in the <value> markup, or qualitatively, in the <level> markup. The function layer is not specified yet. Lastly, all PML elements of each behavior description layer allow for a <confidence> markup, which take numerical values, between 0 and 1.

Here is an example of behavior representation in PML, given by Scherer *et al.* in [180]:

```
<person id="interlocutorA">
    <sensingLayer>
        <headPose>
            <position z="223" y="345" x="193"/>
            <rotation rotZ="15" rotY="35" rotX="10"/>
            <confidence>0.34</confidence>
        </headPose>
        ...
    </sensingLayer>

    <behaviorLayer>
        <behavior>
            <type>attention</type>
            <level>high</level>
            <value>0.6</value>
            <confidence>0.46</confidence>
```

```
    </behavior>
    ...
  </behaviorLayer>
</person>
```

Regarding the actual use of PML, a framework for multimodal behavior analysis making use of this language was included in the Virtual Human Toolkit [85].

4.6.5 SSML

Speech Synthesis Markup Language, originally proposed in [195], is now a W3C standard. It is used to specify speech and pronunciation parameters instructions for speech synthesis engines. The parameters that can be specified include, but are not limited to, language, phonemes (for phonetic specification of the content to synthesize), prosody and pauses. The full specification is available on the corresponding W3C's website[15]. We chose not to get into the details of this DSL because this norm is not really implemented in most speech synthesis systems. In this research work, we have only used the *break* tag, which allows to create pauses in the synthesized speech signal.

4.7 Discussion

The aim of this chapter is to give the reader an overview of the classic design of ECAs and the current possibilities offered by the state-of-the-art technologies for the generation and analysis of verbal and nonverbal behaviors. For this research work, we have made use of some elements only, which are:

- attention estimation, which is used to make sure the user is always paying attention during the interaction with our ECA, as people with dementia may get distracted easily;
- automatic speech recognition, to process verbal inputs;
- behavior generation through character animation;
- the SmartBody BML realizer, to easily synchronize and execute behaviors;
- text-to-speech synthesis, to produce the ECA's voice;
- the BML and SSML domain-specific languages, to specify the behaviors to be produced by the ECA;

[15] http://www.w3.org/TR/speech-synthesis/

- and simple graph-based or finite state machine-based conversation management.

More detail about how each of these elements fits in the design are given in Chapter 5. We have also noted that some other elements, such as emotion recognition, could be useful in our context, but are left out for future work.

Chapter 5

LOUISE, a LOvely User Interface for Servicing Elders

Résumé en français : *LOUISE (*LOvely User Interface for Servicing Elders*) est un agent conversationnel animé (ACA), un personnage humain (ou humanoïde) capable d'interagir avec une personne, par voies de communication verbale et non-verbale. Il a pour but de servir comme interface utilisateur accessible à intégrer dans les technologies d'assistance pour les personnes âgées ayant des troubles cognitifs.*

Comme vu dans le chapitre 2, les ACA possèdent plusieurs avantages, en comparaison avec d'autres modes d'interaction Homme-machine, pour produire une bonne utilisabilité pour ce public. En effet, les humains commencent à développer des capacités conversationnelles dès le plus jeune âge et les utilisent tout au long de leur vie. Ce mode d'interaction est donc naturel et ne requiert pas d'apprentissage, ce dont les personnes âgées démentes ne sont pratiquement plus capables. Les capacités cognitives liées au langage peuvent être affectées dans les stades avancés de la démence, mais les capacités de communication non-verbales sont mieux préservées [174]. C'est pourquoi les ACA pourraient être mieux adaptés pour l'interaction avec ce public que des voix désincarnées, par exemple. Les résultats des expérimentations sur l'utilisation des ACA comme interface utilisateur auprès des personnes âgées atteintes de troubles cognitifs rapportées dans la littérature confirment majoritairement cette hypothèse [151, 137, 138].

Les ordinateurs, bien qu'il soient conçus par des humains pour être utilisés par d'autres humains, ont été initialement créés pour traiter rapidement de grandes quantités d'information numérique et réaliser des tâches répétitives. C'est pour cette raison qu'ils ne sont pas nativement capables d'interaction sociale. Pour combler cet écart, comme vu dans le précédent chapitre, au cours de la dernière décennie, de nombreux efforts de recherche se sont employés à créer des programmes informatiques capables d'imiter la communication humaine. A cet effet, de nombreux ACA et robots sociaux

expérimentaux ont été créés. Grâce à l'augmentation rapide de la puissance de calcul et de la mémoire disponible, l'objectif de donner aux machines des compétences sociales semble aujourd'hui plus atteignable. Cependant, les ACA sont des logiciels complexes, à cheval sur plusieurs domaines, de l'informatique, de l'intelligence artificielle à l'imagerie de synthèse, et leur conception requiert une grande expertise. De plus, les prototypes les plus évolués à ce jour ne prennent pas en compte les besoins spécifiques des personnes ayant des troubles cognitifs. En dernier lieu, peu de connaissances sur la manière dont les personnes démentes peuvent réagir à cette technologie sont disponibles. C'est pourquoi nous avons choisi d'étudier cet aspect.

Ce chapitre est dédié à décrire la première étude de cas que nous avons menée sur l'utilisation des humains virtuels dans le soin aux personnes âgées. Pour commencer, dans la section 5.1, nous présentons les objectifs de cette recherche, l'approche de conception de ce prototype d'ACA et les spécifications envisagées pour LOUISE. Puis, dans la section 5.2, nous décrivons la première étude que nous avons menée, consistant à développer et tester un prototype préliminaire de LOUISE. Ensuite, dans la section 5.3, nous montrons de quelle manière nous avons impliqué des personnes âgées, le public et des professionnels des technologies d'assistance dans la conception, à travers des démonstrations dans des évènements publiques, des tests et des groupes d'échange. Enfin, dans la section 5.4, nous présentons le système automatisé que nous proposons.

Le principal objectif est de concevoir de manière participative un ACA capable d'interagir avec les personnes âgées démentes. Les objectifs secondaires consistent à identifier les cas d'usage les plus utiles pour ce public et les principaux facteurs d'acceptabilité.

Comme point de départ, nous nous sommes intéressés à la gestion des déficits attentionnels des personnes âgées démentes : nous avons créé un algorithme d'estimation d'attention que nous avons intégré avec un ACA semi-automatique qui réagit à cette information de manière autonome mais dont la gestion de dialogue est faite en "magicien d'Oz", c'est-à-dire qu'un opérateur doit sélectionner les répliques produites par l'ACA, à l'insu de la personne en interaction avec ce dernier. Nous avons évalué les performances de notre estimateur d'attention à travers des expériences d'interaction, au cours desquelles des distractions étaient volontairement introduites, avec un groupe de professionnels des technologies d'assistance et un groupe de personnes âgées dont la plupart avaient des troubles cognitifs. Les interactions entre l'ACA et les personnes âgées ont ensuite été analysées par des anthropologues pour tirer le maximum d'indices pouvant informer la conception dans les itérations suivantes. Cette première expérimentation a permis de mesurer que notre estimateur d'attention simple, rapide et peu coûteux produit un taux de bonnes reconnaissances de près de 85%, que les participants aux tests ont été majoritairement satisfaits, sauf concernant l'animation du personnage, jugée trop peu expressive, et que les personnes démentes ten-

dent à interagir de manière "sociale" avec l'ACA, ce qui oriente vers des solutions de reconnaissance vocale de type détection de mots clés.

*Lors des tests, des questionnaires ont été remplis par les professionnels et par les personnes âgées. Le même questionnaire a été rempli par des visiteurs au cours de deux évènements : la journée portes-ouvertes de l'hôpital et une journée d'information sur les technologies d'assistance organisée par Old'up, une association pour le vieillissement actif. Les informations ainsi recueillies ont permis de préciser les spécifications de LOUISE. Pour compléter, nous avons animé une séance d'échange (*focus group*) sur les ACA avec un petit groupe de personnes âgées. Nous avons ainsi identifié que le personnage devrait avoir l'apparence d'une jeune femme ou d'un robot humaoïde et que l'utilisation la plus plébiscitée serait comme assistant virtuel (effectuer des rappels, donner des informations), qui pourrait centraliser plusieurs applications d'assistance (accès à la communication, etc.).*

Grâce à cela nous avons conçut un prototype entièrement automatique de LOUISE dont les fonctionnalités principales sont les suivantes : gestion de l'attention, gestion des tours de paroles, rappels contextuels, et gestion de scénarios d'assistance décrits dans un langage XML ad hoc que nous avons créé.

LOUISE (LOvely User Interface for Servicing Elders) is an Embodied Conversational Agent (ECA), a virtual human (or humanoid) character capable of interacting with people, through verbal and nonverbal communication. It is meant to provide an accessible user interface to be integrated in assistive technologies for older adults with cognitive impairment.

As seen in Chapter 2, ECAs present several advantages, compared to other human-machine interaction modalities, to achieve good usability for that public. Indeed, people start developing conversation skills from a very young age and use them all their life. This type of interaction is therefore natural and does not require learning, which older adults with dementia are hardly capable of. Cognitive capabilities linked to language may be affected in advanced stages of dementia but nonverbal communication is preserved longer [174]. This is why ECAs could be more adapted for interaction with this public than disembodied voices, for instance. The outcome of experiments about the use of ECAs as user interfaces for older adults with cognitive impairment, reported in the scientific literature, mostly supports this assumption [151, 137, 138].

Computers, although they are designed by humans, for use by other humans, were initially created to rapidly process large amounts of numerical data and perform repetitive tasks. For this reason, they are not natively capable of social interaction with people. To bridge this gap, as seen in the previous chapter, in the past decade, a lot of research effort have been put in creating computer programs capable of imitating human communication. To this aim, many experimental social robots and ECAs were created. Thanks to the fast increase of available computing power and memory, the goal of giving machines social skills now seems within our reach. However, ECAs are complex softwares, which involve several domains of computer science, from artificial intelligence to computer graphics, and their design requires much expertise. In addition, the most advanced existing prototypes do not account for the specific needs of people with cognitive impairment. Lastly, little is known about how patients with dementia may react to this new technology. This is why we chose to investigate this aspect.

This chapter is dedicated to reporting on the first case study that was conducted to investigate the use of virtual humans in elders' care. To start off, in Section 5.1, we present the goals of this research, the design approach for this ECA prototype and the intended specifications for LOUISE. Then, in Section 5.2, we describe the first study we conducted, consisting in developing and testing an early prototype of LOUISE. After that, in Section 5.3, we show how we involved older adults, the public and assistive technology professionals in the design, through demos in open events, tests and a focus group. Lastly, in Section 5.4, we present the fully automatic system we propose.

5.1 Goals and design approach

Our main goal in this project is to propose an ECA system that is able to interact with people with dementia, in a way that is easy to use for them and that they will like. As a starting point, it was observed in several assistive technology trials, conducted by our team at the Broca hospital, that people with dementia tend lose track of what they are doing or get easily distracted. This observation is in line with the current knowledge of dementia, which causes losses in working memory, episodic memory, and attentional capabilities. This is why the first feature we investigate is the capability for LOUISE to recapture the user's attention whenever he or she gets distracted or stops interacting because he or she lost track of the ongoing activity. From there, a first prototype can be built and the participatory design may start.

LOUISE is seen as a user interface for patients with cognitive impairment to access computer-based assistive services, and we set ourselves the goal of making a multi-purpose system that can be used as a plugin element in various assistive scenarios. It therefore has to be flexible enough for developers to create their own assistive applications and have it communicate with external devices. Therefore, as one of our secondary goals, we want to identify for which applications this type of user interaction is suitable and which applications would be the most useful for our target users.

The design process of LOUISE includes usability testing with patients. To conduct these tests, some of the most the useful applications identified have to be implemented as test-cases. Through our experiments, we hope to gain information about how people with dementia interact with LOUISE. This knowledge could then be used to implement an interaction management program able to handle simple goal-oriented conversations with people with dementia.

These user testing sessions are also a good occasion to meet people individually to ask them about their preferences regarding LOUISE (embodiment, voice, gender, the way it addresses them, etc.), through semi-directed interviews. This is probably the best way to get such feedback from people with cognitive impairment, as collective focus groups are likely to be too difficult for them to follow.

In this case study, we concentrate on the conception and development phases of our living lab participatory design. The first phase consists in exploring solutions serving our goals by developing small parts of the application, experimenting with them and evaluating the results. In fact, given the complexity of ECA software, these parts are not that small and already require a great amount of work. In our case, this consisted in conducting a Wizard of Oz [102] study, that is to say, experimenting the interaction between the ECA and our target users without having to develop a full-fledged system by having an operator control the ECA to simulate the dialog manager. To this aim, we developed a basic application with simplistic animation capabilities and, as the only automated part, our attention management feature. The second phase consists in building the full system and experimenting with it. Note that the experimentations of the fully automated system, based on believable application scenarios, are presented in Chapter 6.

5.2 Phase 1: Wizard of Oz study

In this design phase, the goal is to experiment with our idea of making LOUISE capable of attention monitoring. This is also the occasion to collect feedback from two stakeholders, older adults and assistive technology professionals, regarding this feature and the whole prototype, and get insights about what a good design would be. We started by putting together a first prototype and we proposed a simple attention estimation method. We then conducted two series of tests, one with our technology professionals and one with patients and healthy older adults, in order to

- evaluate the performances of our attention estimation method,
- get information about the usefulness of this feature,
- witness first-hand some interactions between patients and our ECA,

Figure 5.1: The LOUISE virtual character

- and obtain design requirements through semi-directed interviews and questionnaires.

Lastly, the video recordings of the experiment with the group of older adults were analyzed by linguistic anthropologists, to get insights about the requirements for the system's automation. These professionals are able to finely analyze the interaction between people and the ECA on the linguistic level, as well as the experiment session as a whole, thanks to an ethnographic approach.

5.2.1 Prototype description

The first version of LOUISE is embodied by a cartoon-like female character. It is animated and displayed with a neutral background and includes speech synthesis. The layout is presented in Figure 5.1. A Microsoft Kinect sensor is used to monitor the user during the interaction to perform attention estimation, thanks to the algorithm described in Section 5.2.2. The interaction is managed thanks to a written scenario, consisting in a list of utterances. The utterances are spoken in a predefined order. When the user answers a question LOUISE asks, a hidden human operator has to press a key to move on to the next utterance. For transitions, LOUISE performs an acknowledgment utterance, randomly selected in a dedicated list. When a loss of attention is detected, LOUISE automatically sends a prompt, also randomly chosen from a list. Lastly, when the user pays attention again, a transition phrase is spoken, before asking the last unanswered question again. This basic dialog management is illustrated in Figure 5.2.

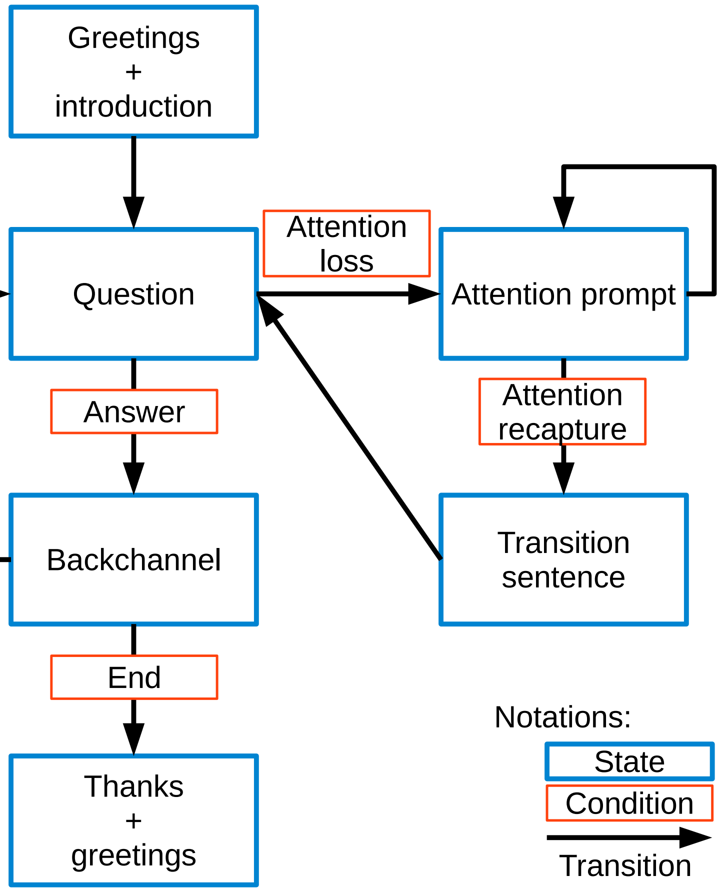


Figure 5.2: Interaction management in the first prototype of LOUISE

The prototype is built from three software modules, as depicted on Figure 5.3, where the arrows stand for data exchanges.

Attention estimator The Kinect's body and face tracking data is obtained using the Microsoft Kinect for Windows Software Development Kit (SDK) v1.8[1] and a Kinect sensor (Xbox 360 version). This data is then processed to estimate the user's attention in real time as detailed in Section 5.2.2.

Interaction manager A scenario file in XML, keyboard presses by a hidden operator and the user's attentional state are used as inputs to

[1] https://www.microsoft.com/en-us/download/details.aspx?id=40278

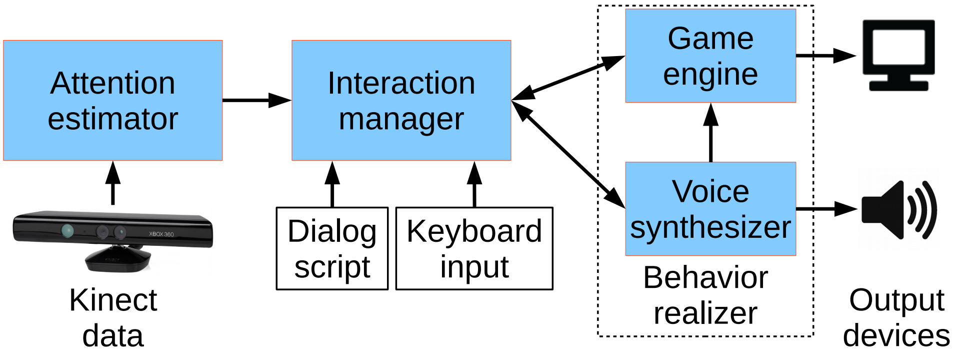


Figure 5.3: Functional diagram of the first LOUISE prototype

manage the conversation. The dialog starts when the user is detected and tracked by the Kinect. The course of the dialog is automatically interrupted to perform attention prompting when inattention of the user is detected.

Behavior realizer This part of the program animates and displays the character and performs voice synthesis and lip synchronization. The character animation was implemented using the Unreal Engine 4 game engine and the voice synthesis is done thanks to the Cereproc Cerevoice speech synthesis engine with the "Suzanne" French voice.

5.2.2 Attention estimator

We implemented an attention estimator to monitor the user's gaze direction throughout the interaction. Our method relies solely on determining, in real time, if the user is gazing towards the screen or away from it. It combines the measures of orientation of the user's shoulders, as done in [18] and [117], and head pose, as in [192], seen here as proxies for his or her intensity of attention towards LOUISE.

This method relies on the Kinect's skeleton and face tracking data. Only the 3D positions of the shoulders and the yaw and pitch rotations of the head are used (see Figure 5.4 for notation). It assumes that the sensor is placed on top and in the middle of the screen displaying the ECA.

The azimuth of the user is defined as $\theta = arctan(N_x/N_z)$ and the angle of the upper-body α as

$$\alpha = \frac{\pi}{2} - arccos(\frac{N_z - L_z}{\sqrt{(L_x - N_x)^2 + (L_z - N_z)^2}}). \tag{5.1}$$

The posture feature f_1 is defined, at each time t as:

$$f_1 = \varphi = \alpha - \theta. \tag{5.2}$$

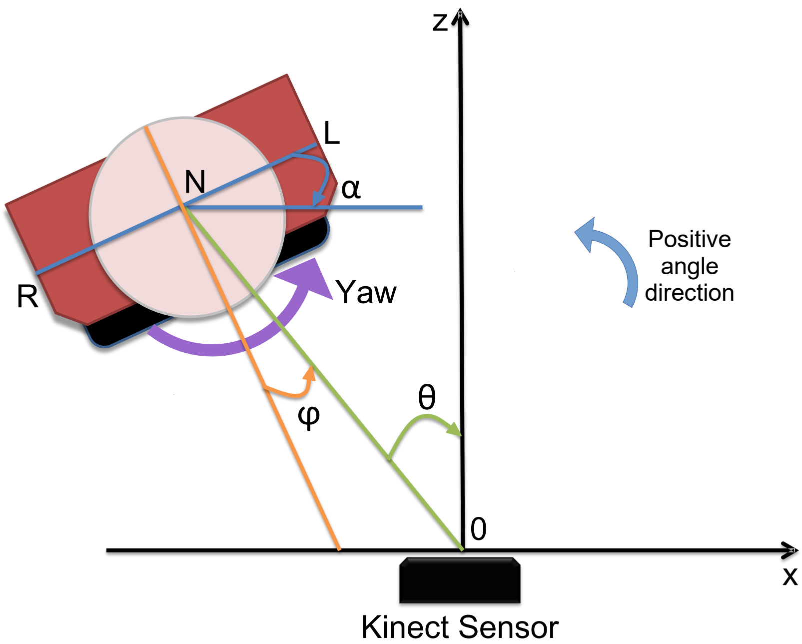


Figure 5.4: Estimation of the angle φ between the shoulder line and the Kinect sensor. θ is the angular position of the user; R, the position of the right shoulder; L, the position of the left shoulder; and N, the position of the neck (computed as the center of the segment $[R,\ L]$).

The Kinect's face tracker outputs, for every sampling time t, the three rotation angles of the tracked head: pitch $\gamma_{pitch}(t)$, yaw $\gamma_{yaw}(t)$ and roll. To make the estimation more stable and more robust to noise and to the failure of the face tracker over a few frames, these angle values are averaged over a period $T = 30$ frames, which roughly corresponds to one second, given the sensor's sample frequency (see Equations 5.3 and 5.4). This calculation produces Values f_2 and f_3 for the monitoring algorithm, at each time t:

$$f_2 = yawMean = \frac{1}{T} \sum_{k=t-T+1}^{t} \gamma_{yaw}(k), \tag{5.3}$$

$$f_3 = pitchMean = \frac{1}{T} \sum_{k=t-T+1}^{t} \gamma_{pitch}(k). \tag{5.4}$$

The three features f_j $(j = 1, 2, 3)$ are normalized as $\overline{f_j}$ in the same way:

$$\overline{f_j} = \frac{cos(f_j) - cos(Max_j)}{1 - cos(Max_j)}, \tag{5.5}$$

where Max_j represents the maximum value for each feature f_j; these values correspond to the Kinect's tracking limitations (30° for yaw, 20° for pitch and 60° for upper-body pose).

A sum of the $n = 3$ normalized features $\overline{f_j}$, weighted by coefficients ω_j (see Equation 5.6), is computed to assess the *AttentionLevel* for each sampling instant. The face's horizontal rotation has the heaviest weight to account for the importance of the face's orientation in the attention estimation. This corresponds to using this information as a proxy to the user's gaze direction. For normalization purposes, the sum of the weights is equal to 10:

$$AttentionLevel = \sum_{j=1}^{3} \omega_j \overline{f_j}. \tag{5.6}$$

The obtained attention level values range from 0 to 10, 10 being the maximum level, when the user's body and face are directly oriented towards the sensor. These values are then used to decide the user's attentional state, i.e., whether the user is engaged or not, using a hysteresis threshold rule: the user is considered engaged if the attention value is more than 8. Transition from engaged to disengaged is triggered when the attention value decreases below 6. Two more states are used: "user detected" (at the beginning, when the user has been detected and is not engaged yet) and "no user".

5.2.3 Experiment protocol

The validation of the attention estimation method and the usefulness of this feature were evaluated in two phases: first with assistive technology professionals, then with older adults, with and without cognitive impairment. The first phase was conducted to make sure that the system was reliable enough, before testing with the actual targeted public, to refine the contents of the scenarios, and to get professionals' opinion about the use of ECAs for older adults with dementia.

Participants were seated in front of a large display, between 22 and 46 inches, depending on the test room used, on which the ECA was displayed. This unfortunate difference in test conditions was due to room availability and practical constraints. Participants were told that the character on the screen was going to talk to them and to ask them questions, to which they were instructed to answer.

A typical interaction consisted of 3 utterances for the introduction, 7 questions, 3 acknowledgment utterances, 3 prompting utterances in the first phase, with assistive technology professionals. The interaction script of the second phase, with older adults, included 2 utterances for the introduction, 9 questions, 5 acknowledgment utterances and 5 prompting utterances. In addition, both scripts included a transition utterance and 2 utterances for

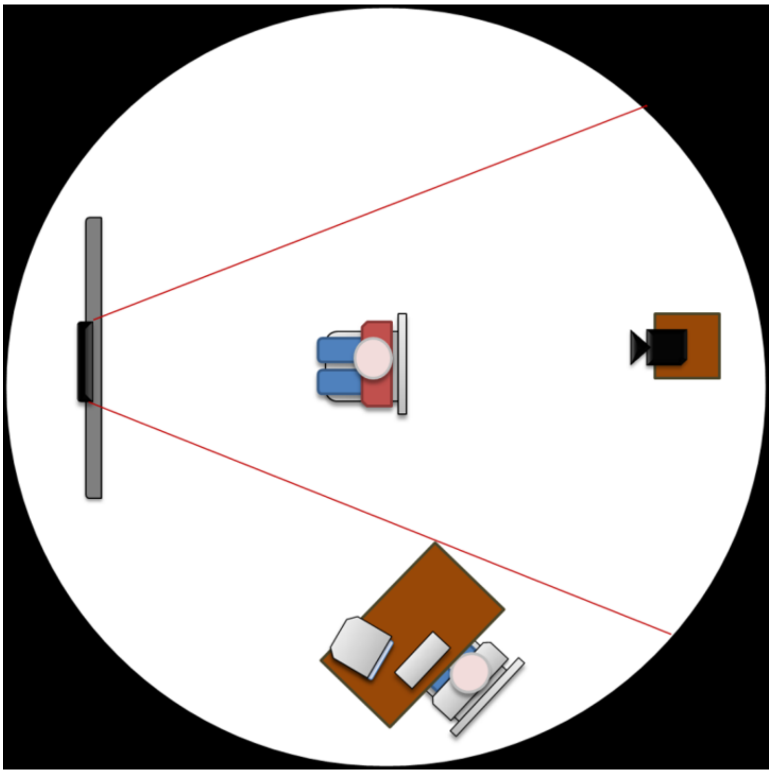


Figure 5.5: Room setup for the attention estimator validation experiments. The participant was placed in front of the screen; the Kinect sensor was placed on top of the screen; and the video camera was placed behind the participant.

the conclusion. When necessary, the acknowledgment and prompting utterances were randomly selected from the corresponding list in the scenario; the full lists of utterances (one for the first phase with assistive technology professionals, one for the second phase with older adults) are given in Appendix B.1. When a user got distracted, LOUISE automatically stopped and prompted the subject to attract his or her attention. After such an interruption, she always asked if the participant wanted to continue the interaction, at which point he or she could decide to stop the test.

During the interaction, two distractions were voluntarily introduced at fixed moments. The first one was performed by the experimenter in the room at the beginning of the third introduction utterance: he asked the participant if the sound was loud enough. The second distraction was introduced by another experimenter, opening the door and asking the participant if everything was fine, during the fifth question. All test sessions were filmed. The room setup and camera placement are shown on Figure 5.5.

The idea behind the first conversation script was to have the ECA ask questions about itself to obtain the participants' opinions. The script was changed for the second phase of the experiment, based on the feedback of the expert participants of the first phase.

In both phases, after the test, each participant was asked to fill out a questionnaire about LOUISE, their experience when interacting with it and ECAs in general, in a semi-structured interview. The first part, about participants' impressions of their interaction with LOUISE, was composed of 5 questions. The questions were about the clarity of the ECA's speech, how nice they thought the character was, how interesting the experience was, how well the character reacted to their behavior and the duration of the interaction. The last of these 5 questions was only asked to the group of professionals. The details and results of the second part of the questionnaire, composed of general questions about ECAs, are presented in Section 5.3.1. The full questionnaires are given in Appendix C.

Participants

The group in the first experiment was composed of 14 assistive technology professionals, 10 females and 4 males, aged between 23 and 63 years (mean = 37.6, standard deviation = 8.8). They were physicians (geriatricians, to be more specific), psychologists and ergonomists who worked in the field of assistive technologies.

Participants of the older adults group, in the second experiment, were recruited at the Broca Hospital's memory clinic, a consultation center for memory and cognitive issues, visited by over 1000 patients every year. Participation was purely voluntary and was not retributed. All participants gave a written informed consent. 8 older adults, 2 males and 6 females, took part in the experiment; they were aged from 63 to 91 (mean = 76.1, standard deviation = 6.9) and 6 of them had cognitive impairment, either MCI or Alzheimer's disease. In that case, their MMSE scores ranged from 17 to 29 (mean = 22.7, standard deviation = 3.3). Note that, even though in some cases (one, in this experiment), the MMSE score is higher than the standard detection threshold of 26, people may still have cognitive impairment. This is due to the lack of sensitivity of the MMSE, mentioned in Section 2.1.2. Our data on cognitive impairment diagnosis comes from extensive cognitive assessment, conducted as part of the memory consultations at our hospital.

Data analysis

The data samples produced by the Kinect and used in the attention estimation method were recorded during the interaction, as well as the color image with an overlay showing the tracking information (face keypoints and skeleton). These videos were then annotated by three experts, asked for each recorded instant (one per second) to judge if the user was paying attention to the ECA (i.e., looking at the screen) or not. Two of the annotators worked independently. The third annotator had to arbitrate when the two annotators did not agree. The annotations were then compared to the decisions taken by the attention estimator, using only the third annotator's data.

To evaluate our attention estimation method, we computed the correct detection rates by comparing the decisions taken by the algorithm during the experiment with the human annotations. This was done for each data sample, on all the data at once and per subject. We also computed a receiver operating characteristics (ROC) curve for our classifier, using 21 values for the decision threshold, from 0 to 10, with a step of 0.5 (see Figure 5.6). The area under the ROC curve (AURC) was also computed.

In addition, we computed correlations to see which of the three features used in the attention estimation method are the most relevant, using point-biserial correlation coefficients [194] between the ground-truth human annotations of subjects' attention and each of our three features, as well as with the aggregation of all three features in a single attention value. We chose this correlation coefficient because it is adapted to check for correlations between qualitative and numerical data.

Lastly, we did a group comparison with the data from the two experiments that had the same protocol and only differed by the content of the questions. This was done using the Wilcoxon rank sum test for equal medians, which is adapted to small group sizes [221]. Most computations were performed using Mathworks Matlab. Only the per-participant performance scores were computed using Microsoft Excel.

5.2.4 Performance of the attention estimator

The results for both groups are presented in Table 5.1. In the older adults group, with one of the participants, the Kinect's body tracking completely failed during the whole test. It tracked a chair in the background, instead of the person. In another test, also in the older adults group, the body tracking was lost for about one minute. These gave attention estimation scores of 24% and 58% of correct decisions, respectively. This is why the results are presented with and without the body tracking errors. This, however, does not account for short but frequent failures of the Kinect's face tracking. Hence, the results presented here are the ones of the complete system, tested in conditions close to ecological, and not solely of our algorithm in ideal conditions, which would do better with more reliable input data.

Overall, our results are encouraging and show that our simple, cheap and fast attention estimation method is quite reliable and achieved up to 95% of correct recognitions, with an average of 84%, with the hysteresis threshold values given in Section 5.2.2. Regarding the areas under the ROC curves, they show that our detections are better than random decisions. The ROC curves, presented on Figure 5.6, show that our algorithm achieves a better

Table 5.1: Performances of the attention estimator in percentage of correct decisions.

Data set	**Perf.**	**AURC**	**samples removed**
Experts per participant	64 to 94%		none
Experts per participant average	89% ± 6%		none
Experts all samples	89%	0.70	none
Experts without tracking errors	89%	0.72	0.4%
Older adults per participant	24 to 95%		none
Older adults per participant average	75% ± 16%		none
Older adults average without tracking errors	81% ± 7%		1 participant
Older adults all samples	76%	0.64	none
Older adults all samples without tracking errors	80%	0.77	8.3%
Total per participant average	84%		none
Total all samples	84%	0.68	none
Total without tracking errors	86%	0.74	3.3%

sensitivity than specificity, which means it produces more false negatives than false positives. This indicates that the hysteresis threshold decision is a good idea, as it allows to reduce the number of reactions to the detections of false losses of attention (which, by the way, tend to occur when face tracking fails).

Regarding the correlations between our indicators (called features) and their relation to attention, the most correlated feature is the yaw angle (rotation of the head around the vertical axis), with a value of 0.57 for the point-biserial correlation coefficient on the whole data set. The other indicators are also relevant, as the significance tests of point biserial are positive. The correlation coefficients for the shoulder's orientation and pitch angle are 0.33 and 0.35, respectively.

Lastly, the Wilcoxon rank-sum test shows that there is no statistical

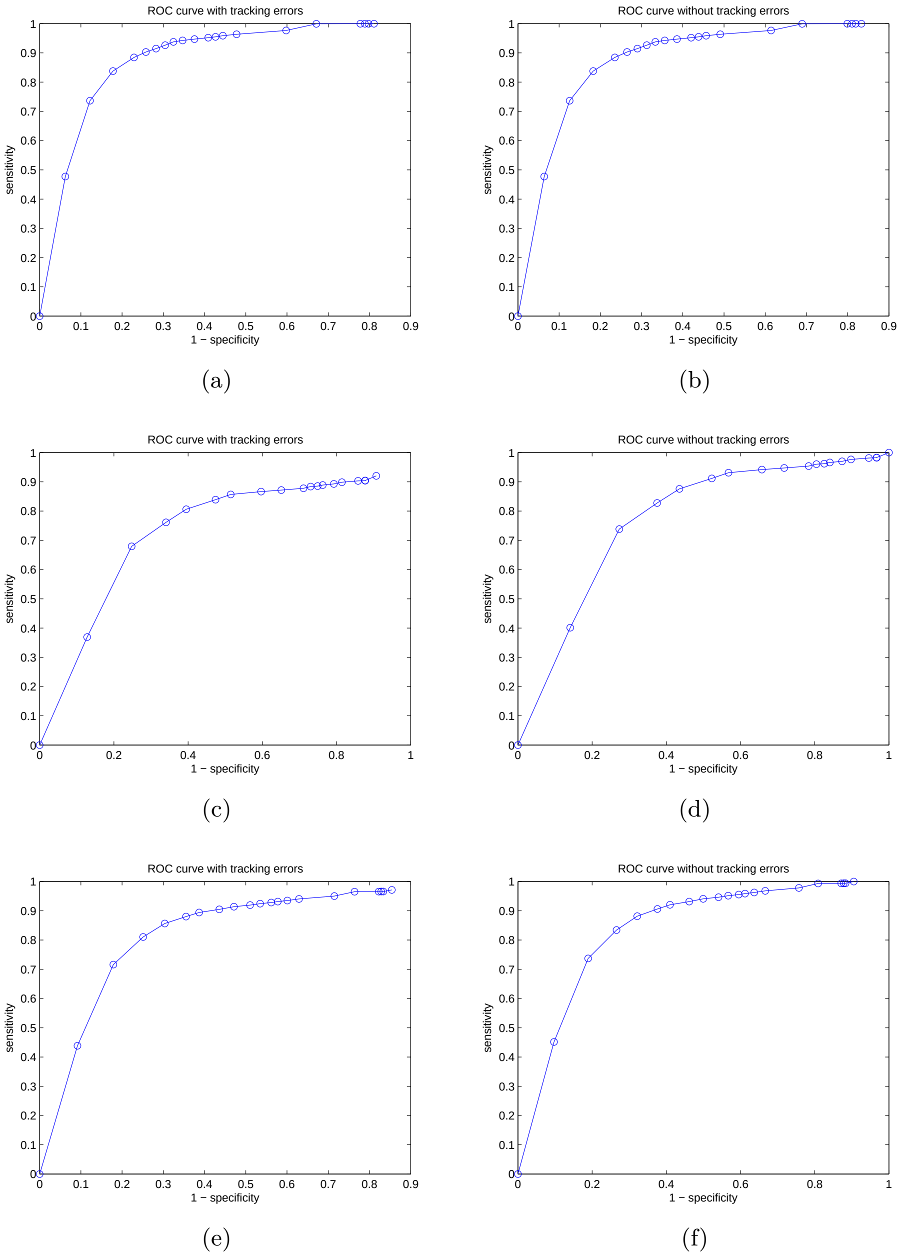


Figure 5.6: Receiver Operating Characteristic (ROC) curves for the attention estimator. (a) Experts. (b) Experts without body tracking error. (c) older adults. (d) older adults without body tracking error. (e) All data. (f) All data without body tracking error.

differences in medians between our two groups when the failed test value is taken off the data set (p = 0.08). This suggests that the results of our attention estimation method are not expected to vary with age. Hence, it does not seem necessary to tailor the algorithm for older adults.

5.2.5 Qualitative results

In the first group, with assistive technology professionals, all participants interacted naturally with the ECA. Most displayed high levels of attention. For some users, the experimenter introducing the second distraction had to insist to distract the participant. This shows that the ECA and the context of the experiment were significantly engaging.

In the second group, composed of older adults, only 6 out of 8 participants successfully interacted with LOUISE. One had very poor hearing and did not understand well what the character was saying. However, when he did understand or was helped, he could interact successfully. The other participant who could not interact successfully throughout the scenario lost track of the context and refused to continue after the first distraction; her cognitive impairment was the highest of the cohort (MMSE = 17).

For the first question about understandability, 10 out of 14 expert participants thought it was clear enough, 2 did not answer and 2 said no. One of them commented that it would turn out to be probably too complicated for cognitively impaired older adults. Regarding likability, 5 participants appreciated the character a lot, 6 thought it was alright and 2 said it was not good enough. Most of the ones who said it was alright or not good enough commented that the lack of facial expression and blinking made them feel uneasy. This is well summed up by what one of the participants said: “the voice is nearly likable but the frozen and cold face leaves me with an impression of quite unpleasant strangeness”. In addition, one participant advised us to add directed gaze to show engagement on the character’s side. Finally, several participants thought that LOUISE should smile. As far as the third question about their experience is concerned, all participants felt the interaction interesting (6) or very interesting (7). No additional comments were given. Question 4, regarding how the character adapted its behavior to their’s, got more divergent opinions; 6 of them judged it reacted very well, while 3 said it was fine and 4 considered it a bit poor. Furthermore, 7 participants thought the duration of the interaction was adapted, 4 said it was too short and the last 2 didn’t have an opinion. Overall, the feedbacks of the professionals were positive, except about the lack of expressiveness of LOUISE. Indeed, the character’s idle animation had it turn its head slightly from right to left and back, which gave the impression that she was not looking at the user; in addition, it did not blink nor did it perform facial expressions or gestures.

In the older adults groups, feedbacks about clarity were mostly positive;

the participants rather liked the character but also complained that it lacked expressiveness and was not smiling; most participants thought their experience was interesting; the feedbacks about how well the character reacted to their behavior, however, were more mitigated.

Lastly, in the older adults group, out of a total of 19 distractions over all the experiment sessions, the experimenter only had to ask the participant to look at the screen once. This shows that the attention recapture strategy is quite effective. However, most of the time, the subject did not get distracted long enough for the character to say one of the prompting phrases (one or more prompting phrases were triggered only 6 times in 19 distractions), and directly asked the after prompting question. Contrary to what we expected, we did not observe any situation in which a patient with cognitive impairment looked away from the screen because he or she lost track of what he or she was doing. The only self-induced distractions were related to hearing and understanding issues. This could be explained by the fact that the experiments were conducted in a controlled environment, with no other sources of distractions than the ones we have voluntarily introduced. In real use conditions, there may be multiple sources of distractions, such as people walking by in the corridors, screaming patients, etc. This is why we think that this observation does not mean that this feature is not useful.

5.2.6 Anthropological analysis of the interactions[2]

To gain insights on the conversation management issues for the future automation of LOUISE and provide evidence-based suggestions for improving the human-computer interaction design, anthropological expertise was recruited. The goal was to analyze the interactions between the participants in the older adults group and LOUISE during the experiment carried out to evaluate the attention estimator presented in Section 5.2.2.

Method and Hypotheses

The empirical material consisted in the video recordings of the experimentation. To guide our analysis, take into account the context of the experiment, and follow best practices of ethnographic research [93, 150], the data was complemented with interviews and a data session with LOUISE's leading designers. The video material was coded by two independent annotators in order to perform a quantitative conversation analysis (CA) using the ELAN software to supplement and test the qualitative analysis [188].

The data and interviews highlighted two targets of general interest: to improve the ECA's overall interactional adaptability to people with dementia (PWD); and to cope with the structural asymmetry of the interactional

[2]Thanks to Giovanni Carletti and Yann Laurent for their help in writing this section.

setting (programmed "one-off volley" questions/answers vs. improvised conversation; institution vs. individual; enabler vs. enabled). Five hypotheses emerged from our qualitative analysis: H1) PWD proportionally utter more words than healthy older adults, seen as control subjects (CS); H2) PWD develop more topic expansion; H3) silences are longer before answers to open questions than before answers to polar questions; H4) the social setting of the experiment significantly modifies the HCI; H5) PWD speak more with the experimenter than CS.

The following indicators, averaged over participants, were computed to assess the validity of each hypothesis (hypothesis, PWD/CS): word count (H1, 60.7/26.5); number of introductions of new information (H2, 6.0/2.5); silence's length, in seconds, per question type, open (H3, 1.0/0.5) or polar (H3, 1.4/0.9); and number of utterances shared by the patient and the physician, namely physician turns (H4 and H5, 9.0/7.5) and patient-to-physician (H4 and H5, 6.3/5.5).

Results

The following result analysis is based on the ethnographic data indicators and CA. In this small sample, PWD uttered more than the healthy older adults, producing spontaneous topical development (H1 and H2). The average conversational time was around two minutes and a half. Contrary to our expectations, silences were surprisingly shorter when answering open questions compared to polar questions (H3). Silences shorter than 300 ms were considered part of a speech turn. Interestingly, qualitative evidence suggest that the silences and pauses could be related with the stage of the disease. The average silence was shorter for CS, with a mean of 0.83, than for PWD (mean 1.32). Further research on this topic is needed, due to the small number of participants.

Both annotators describe the protocol as a multi-party interaction setting. In fact, there are at least three participants to the conversation: physician, ECA, patient. Sometimes, PWD are accompanied by relatives who spontaneously produce some turns. In this respect, sequential analysis cannot be performed by treating the situation as a general face-to-computer interaction. The physician's tokens were 24.5 out of 63.5 produced by PWD and 140 for the ECA. Every time the experimenter spoke, people modified their behavior (H4). Lastly, patients addressed more utterances to the physician than CS, but the sample is too small to judge about the significance of this result.

Discussion

The linguistic analysis of the tested scenario highlighted that almost all questions present a wide focus design (see Appendix B.1 for the list of questions prepared for use by LOUISE). This could explain the absence of statistically significant differences in silence length before open questions (H3). Wide-focus questions are cognitively heavier than contrasted and narrow questions (e.g., do you like music or theater more? [214]) and are used with preference for contiguity in natural talk [176]. Moreover, wide-focus questions are more open to interpretation and may trigger longer answers, mitigation or contextualization reactions. This is why narrow or contrasted questions should be privileged for the interactions between ECAs and PWD.

In natural interaction speech, the sequential embedding of topic shifts has the function of indicating some linguistic context in order to notify the conversation partner(s) of a change of topic [128, 108]. That is to say in social conversations, such as small talks, conversational partners tend to expand the topic to give the other participant(s) elements to react to, either to keep talking about the same topic or to direct the exchange towards a new topic that they would like to discuss next. To make the interaction between people and ECA more natural, it would be worth applying this implicit social rule to introduce new topics. That being said, it seems relevant to ask if it is clinically desirable that PWD are put in a position to speak fluently with ECAs.

Qualitative interaction analysis showed that the healthy older adults displayed some typical adaptation behavior towards the ECA (anthropomorphic robot voice imitation); PWD seemed to adapt too, but in a distinctively polite manner. Further comparison and context-aware scenarios may help analysis and positioning in answering the following question: should machines allow and promote this behavior or should they stimulate different, clinically relevant reactions? This question only accounts for one part of the interactional deal, i.e., the stimulating machine. Although the tested conversational scenario followed a principle of equilibrium-keeping, it was built as a series of questions designed to address the professionals' issues without focusing on the conversational nature of the experimental task. The numerous superposed utterances show that categorically splitting conversation and interaction is not relevant. In fact words' meaning is strictly dependent on their use in discourse [61]. To take into account this coupling, one could ask if it is preferable to limit the topic shifts and concentrate on fluent topic design.

Finally, if machines are needed to engage the user in a focused interaction, it could be useful to embed social features, such as bits of casual conversation (small talks), in the protocols' design. In addition, it could be worth integrating conversational and interactional analysis findings, in order to make the conversations more symmetrical and account for social variations.

5.2.7 Discussion and impact on the next design iteration

Thanks to these first experiments, we could validate the performance of our simple and cost-effective attention estimation method. It was evaluated that it recognizes over 80% of the attention situations correctly. Also note that this result is quantified on a per-sample basis, which means that the system may react correctly, with a short delay (the Kinect sensor generates about 30 samples per second). However, we could not eliminate all false detections of inattention, which may lead to unwanted interruptions of the conversation. In addition, the results indicate that there are no significant differences in performance of our algorithm between the two groups, which suggests that a tailored attention estimator for older adults with cognitive impairment is not necessary, though these results were produced with few subjects and should be further validated in a larger study. When correct attention estimations were obtained, our attention recapture strategy turned out to be effective. However, contrary to what we expected, we did not observe any self-induced distraction. This may be due to the fact that most of our subjects still had a relatively high level of cognitive functioning and should be confirmed in other experiments.

In the first phase of this experiment, with the group of professionals, we were warned that the questions in the interaction scenario were too complex for our target users. This allowed us to simplify the questions in the dialogs, for them to be easy to understand for older adults with cognitive impairment. The details of the utterances in the two interaction scenarios are presented in Appendix B.1.

Thanks to the feedbacks given by the patients and the assistive technology professionals who tried interacting with LOUISE, we were able to improve the system. The main issue revealed by this early evaluation is the need for the character to be expressive and pleasant, by smiling for instance. Furthermore, people were disturbed by the fact that the character did not direct its gaze towards them. For these reasons, the requirements for the animation module in our second prototype are the following:

- the character should be able to perform facial expressions;
- the character's gaze should be directed in a meaningful way;
- it should be possible for the virtual human to perform gestures;
- LOUISE should be able to blink.

We thus decided to switch development platform and use a state-of-the-art BML realizer for the animation of the character. We decided to use SmartBody [196], as it is a quite comprehensive tool for virtual human social behavior generation and it is distributed under an open-source license, which makes it flexible enough to fit our needs perfectly.

Conducting this first set of experiments and the anthropological analysis of the interaction videos allowed us to identify several key aspects for the automation of the system's dialog management. Firstly, our observations led us to think that we should address the issue of "context reminding" after distractions, to help people keep track of what they were doing. At a more fundamental level, following the guidelines obtained after our anthropological analysis, it appears that asking only narrow or contrasted questions should lead to shorter answers, which are easier to manage for automatic speech recognition (ASR) and reduce cognitive load on patients. However, since PWD tend to provide longer answers than healthy people, a keyword-spotting ASR solution should be preferred.

Lastly, it led us to reflect about our testing practices: the presence of the experimenter in the room and the fact that he or she addresses the subjects directly yield to a three-party interaction, whereas the system is intended for an interaction only involving the user and the ECA. Furthermore, our analysis revealed that the subjects emitted few criticisms about the system, contrary to what we expected. In addition, all participants but the one with the strongest cognitive impairment focused very hard on the task. In our future experiments, we should thus try to move the experimenter out of the way, for the interaction to happen only between the participant and the ECA. We think the day hospital environment in which the tests were performed yielded what we call a "white-blouse effect", which prevented people from expressing their criticism and made them try harder to focus than they normally would. Indeed, what we call the "white-blouse effect" is the consequence of the position of authority of physicians (or experimenters, in our case) over the patients. As a results, patients usually do as they are told, ask few questions and do not contradict them. In a hospital, people wearing a white blouse, which is normally the physicians' attribute, are therefore seen as authority figures, by extension. Experiments in a more casual context, and with more participants, are thus warranted to increase ecological validity.

5.3 Design insights from questionnaires and focus groups

One of the secondary goals of this case study is to obtain information about how older adults perceive ECAs and what they expect from them, in terms of appearance, deployment device, functionalities and acceptance. To this aim, on the occasion of two public-relation events, we showed demonstrations of our first prototype of LOUISE to the public and asked people to fill in short questionnaires, so they could give us their opinions (Section 5.3.1). In addition, the same questions were also asked to the participants of the WoZ study presented above. A similar approach was used by Yaghoubzadeh *et* *al.* [229], who conducted interviews and focus groups to inform the design an ECA used as a scheduling assistant for older adults and younger adults with cognitive limitations.

After addressing all the criticisms made to us in a new preliminary design for the second version of LOUISE (see Section 5.4), which was intended to be fully automatic, we conducted a focus group with older adults (Section 5.3.2). In addition, we asked physicians and psychologists working at the Broca hospital for recommendations, in informal contexts, throughout the design process, and in a staff meeting, after giving them a presentation and showing them a demonstration (Section 5.3.2).

5.3.1 Questionnaires

On two occasions, an open-house event at the Broca Hospital and an exhibition of assistive technologies organized by Old'up, an association of older adults for active aging, we showed demonstrations of the first prototype of LOUISE, described earlier in Section 5.2.1. We then asked visitors to fill in a short questionnaire, to get their opinion about ECAs. In addition, as mentioned earlier, the same questions were asked to the participants of our WoZ study.

Our questionnaire was composed of 2 yes/no questions and 5 multiple choice questions.

1. What appearance would you prefer for the ECA?
2. Would you like to be able to personalize the ECA according to your tastes?
3. What parts of the ECA would you like to personalize?
4. What do you think an ECA would be useful for?
5. On which display(s) would you like to visualize the ECA?
6. The perception system of the ECA includes a motion sensor, a video camera and a microphone. Would you accept such a system to be installed in your home, knowing that no data would be transferred elsewhere?
7. Would you like the system to be always on, so it is able to react quickly when you need it?

The questionnaire was answered by a total of 37 people, including 7 people between 65 and 79 years old and 2 over 80 years old. Most visitors who answered the questionnaires at the open-house event of the Broca hospital were care staff members (physicians, physiotherapists, nurses, etc.) or members of patients' families. For the multiple choice answers, except the last one (about the system being always on), people could tick more than one box. The full questionnaire is presented in Appendix C and the full results are presented in Table 5.2.

Regarding the character's embodiment, the respondents' preference is clearly for it to look like a young woman (19 people or 51%). The other options were not chosen much; the second most frequent answer was "no preference" (7 people, 19%). In older adults' answers, the preference also goes to a young woman (5/9 people) or no preference (3/9). Interestingly, for some of the people who said that they had no particular preference, the most important is for the character to be pleasant and smiling. A person also suggested that the embodiment should depend on the application.

Answering the second question, most people (32, or 86%) who took the questionnaire reported that they would like to be able to personalize the character. One person said that there is no need for personalization if the character is pleasant. Among the personalization features, the ones that were selected the most are the character's voice (29/37, or 78%) and clothing (18/37, or 49%). The other personalization features we suggested, haircut, face and personality, also got good scores (from 40 to 43%). Some respondents suggested to personalize facial expressions or eyes.

Regarding the applications of ECAs, the virtual assistant got the most positive answers (28/37, or 76%), that is to say an ECA that helps managing one's schedule and providing reminders of appointments, medication intake, relatives' birthdays, etc. The second most frequently ticked box was the virtual butler that helps controlling connected home appliances, such as electric shutters (23/36, or 62%). In addition, more than half (21/37, or 56%) of the people who answered our questionnaire thought that an ECA would be useful to guide them through complex tasks. On the contrary, fewer people (13/37 or 35%) said that an ECA as a virtual coach would be useful. Lastly, a few other use scenarios were suggested by our respondents: two people said it would be useful for teaching foreign languages; one person suggested it could sing songs to older adults; and another person proposed to use ECAs to give directions in public buildings.

The fourth question about the preferred displays got mitigated answers and there is no clear majority for one type of display over another. The only thing that clearly appeared is that few respondents (8/37, or 21%) would like the ECA to be shown on a dedicated display. This suggests that an ECA application should be deployed on a device that people already own. Possibilities of answers included computer, tablet, smartphone, tv or dedicated screen. In addition, two people answered "nowhere, I do not want to use this technology".

For acceptability, older adults seemed more concerned about having this kind of system at home than younger adults: only 4 out of 9 older adults said they would accept to have this system installed in their home whereas 20 out of 28 younger adults would accept it. Lastly, only a small proportion

Table 5.2: Full results of the questionnaire about ECAs. 37 people answered and 9 of them were older adults (65 and older).

Topic	Answer	Total (%)	Older Adults (%)
Gender	Female	26 (70%)	5 (55%)
	Male	11 (30%)	4 (45%)
Appearance	Young woman	19 (51%)	5 (55%)
	No opinion	7 (19%)	3 (33%)
	Animal	4 (10%)	0
	Older woman	3 (8%)	0
	Older man	2 (5%)	0
	Young man	2 (5%)	0
	Little girl	1 (3%)	1 (12%)
	Little boy	0	0
	Other	1 (3%)	0
Embodiment Personalization	Yes	32 (86%)	7 (78%)
Personalization features	Voice	29 (78%)	7 (78%)
	Clothes	18 (49%)	5 (55%)
	Hair	16 (43%)	4 (45%)
	Personality	16 (43%)	3 (33%)
	Face	15 (41%)	3 (33%)
	Other	6 (16%)	1 (12%)
Applications	Assistant	28 (76%)	7 (78%)
	Butler	23 (62%)	6 (67%)
	Task helper	21 (57%)	4 (45%)
	Coach	13 (35%)	1 (12%)
	Other	4 (10%)	0
	Nothing	3 (8%)	1 (12%)
Display	Computer	21 (57%)	5 (55%)
	TV	20 (54%)	4 (45%)
	Tablet	19 (51%)	1 (12%)
	Phone	16 (43%)	0
	Dedicated disp.	8 (22%)	1 (12%)
	Other	3 (8%)	0
	Nowhere	2 (5%)	1 (12%)
Camera acceptance	Yes	24 (64%)	4 (45%)
Activation	By user only	15 (41%)	5 (55%)
	Programmed	14 (38%)	1 (12%)
	Always on	5 (14%)	2 (22%)
	Do not know	3 (8%)	1 (12%)

of people said they would want the system to be always on (5/37, or 14%); some said it could turn on automatically at fixed hours (14/37, or 38%); and a small majority said they would like the system to be activated only when they decide it (15/37, or 40%). A few people (3/37) did not answer; one person argued that it depends on the user's cognitive abilities. Indeed, a person living with Alzheimer's disease is likely to forget to turn it on, which would make the ECA useless.

5.3.2 Focus group

After the recommendations of professionals and older adults about the character's animation being too poor, we implemented a new front-end for the ECA with rich animation capabilities. The behavior generation of the virtual human character was controlled in BML, using SmartBody and the Microsoft Speech API's voice synthesizer, and the Panda 3D[3] game engine was used for rendering and sound playback. The character models were created using Autodesk Character Generator[4], which allows to quickly and easily create human character models with textures and rigs, ready for use in a game engine. In addition, the system was automated by using speech recognition and a dialog scenario, implemented in the form of a state machine. The layout of the application is depicted in Figure 5.7.

Participants

The focus group was organized in the *Café multimédia* [228], a group activity to inform older adults about information and communication technologies (ICTs), to foster e-inclusion. 9 older adults, 7 females and 2 males, between 67 and 89 years old participated in the focus group (mean = 75.5, standard deviation = 8.3) and most of them had been involved in the *Café multimédia* program for over a year, which made them better informed than average older adults about ICTs. In addition, all participants were retired; most of them had a higher education level than average for their generation; three of them had MCI; and two of them had a relative living with MCI or dementia.

Organization of the session

The session was animated by two developers of LOUISE and two psychologists, in charge of the *Café multimédia* program. To fuel the discussion and give participants a good understanding of the technology, a short general presentation about ECAs, what they are, how they work, their potential applications and their advantages and drawbacks compared to robots, was

[3] Panda 3D is an open-source game engine, distributed under the very liberal BSD-3-Clause license. It was originally created by Disney and it is now maintained at Carnegie Mellon University. It is available at https://www.panda3d.org/.

[4] https://charactergenerator.autodesk.com/

Figure 5.7: Layout of the LOUISE application with the SmartBody-controlled animation.

made. Then, a demonstration of the new LOUISE ECA was shown and two people could interact with it, in the presence of the group. The demonstration was a simple use case, validated by physicians, that consists in reminding users to drink water, having him or her choose between still water and sparkling water and guiding him or her through the task of preparing a glass of water. This scenario was hard-coded into a preliminary version of the system presented in Section 5.4. The discussion was then oriented around four main subjects: feedbacks on demonstration of the system, embodiment, applications and acceptance.

Outcome

The exchanges with the group lasted for about an hour and a half and all four subjects were covered. After the demonstration of the system, we had positive feedbacks from the group of older adults. They agreed on stating that LOUISE's speech was clear, fluid and easy to understand. They also said that the layout of the application (see Figure 5.7), with a camera placed so that the hands of the character are visible, is well adapted, as "hand gestures are very important", the participants said, especially for showing directions. Furthermore, the participants said that the fact that the ECA has the initiative of the dialog is well adapted for older adults with cognitive impairment. However, they thought that the quality of the speech recognition was poor and that the character was not expressive enough. The later comment is most likely due to the fact that the BML behaviors used in the demonstration scenario did not contain enough facial expressions. The participants also proposed some improvements: they thought that the character should smile more, be pleasant, lively (including its voice) and fun. In addition, among the possibilities of ECAs that we listed to them, they thought that an evolution of the relationship between the ECA and its user would be the most interesting. This discussion led the participants to propose the idea of introducing the evolution of LOUISE's functionalities to adapt to the user's declining cognitive capabilities. They stated that the fun aspect of an ECA could be a good starting point to install it in a patients' home, at a point when they do not really need it yet, so that they can get used to it, and then, as the illness progresses, the ECA could propose more and more assistive functionalities.

Regarding the functionalities of an assistive ECA for older adults, the participants thought that it should centralize various cognitive compensation and entertainment applications. More specifically, they listed the following use cases as being the most interesting : provide general information such as date, time, weather forecast or news; help manage the user's schedule, by providing reminders of events, appointments and medication intake; serving as an interface to make and receive calls in a video chat application for the patient and allowing relatives to access the video link, to check on their loved ones, especially when patients are not answering calls; showing adapted physical exercises; and offering games, so the ECA system can be entertaining to be accepted before it actually becomes useful.

For the embodiment, the participants proposed that the ECA could look like a humanoid robot. They argued that “a robot should look like a robot”; 6 out of 9 participants agreed on this idea. It is worth noting that some of the participants had been consulted about assistive robots on several occasions and therefore had already made an opinion on this matter. We also think this would be a good idea, since it could lower expectations on the ECA’s capabilities, compared to a human-looking embodiment, thus avoiding the “uncanny valley” effect (see Section 3.1.1) and avoid potential misjudgment for people with dementia. Other propositions were made by some participants: a cartoon-like human appearance and well-known cartoon characters; but they did not meet as much consensus as the humanoid robot looks. In addition, the participants said that it is important that the ECA can be personalized, to adapt it to each person’s tastes.

Lastly, regarding acceptability, the participants did not see a problem with the fact that the system includes a video camera. The main idea that came out of the exchange is that it would be acceptable if it is useful and, as already mentioned above, making a fun ECA, which includes entertainment applications, would be a way for it to be accepted before it becomes useful, as the cognitive impairment worsens.

5.3.3 Discussion

Thanks to our questionnaires and the focus group, we could gather information from a total of 46 people, including 18 older adults (people over 65 years old) and 14 assistive technology professionals.

We have identified that the character should look like a young woman, as also reported by Morandell *et al.* [138], or a humanoid robot, and that most people would like to be able to personalize the appearance of the ECA. In addition, it seems very important for the character to be lively, expressive and pleasant. Our system should therefore allow to easily accommodate several character models, so people can choose the one they prefer, and have rich animation capabilities, especially for the character’s face. In addition, we strongly agree that the robot-like appearance would be among the best options, as it wouldn’t lead to too high expectations of the ECA’s capabilities and it would be an implicit reminder for people with dementia that they are not addressing a real person. In addition, Bergmann, Eyssel and Kopp [20] have observed, in a study comparing a robot character with a realistic boy character, that a humanoid robot-looking ECA is perceived as warm and competent as a human-looking one; the robot-looking ECA even had slightly better evaluations. We would add that, with a character model and animations that are anthropomorphic enough, it is likely that none of the advantages of ECAs identified in the literature (see Section 2.2.1) would be lost. This development phase is however too early to focus on such fine details. Since it is easy to create human-looking characters of young women, as asked by most people who gave us their opinion on this matter, thanks to character creation tools, such as Autodesk Character Generator[5] or Adobe Fuse CC[6], we thus leave the issue of robot-looking ECAs for future work.

Regarding the applications, the virtual personal assistant that reminds patients of events, appointments and medication intakes and provides them with general information is the most popular. Other applications could include step-by-step task guidance or simplified access to communications, through video chat and emails. In addition, the majority of people would want the ECA to centralize assistive services and/or help control connected home appliances. Furthermore, physicians who were consulted less formally, in staff meetings, thought that the applications we proposed were relevant and also suggested that ECAs should address screams, a frequent behavioral disorder in patients with dementia, and that they should read and teach poems for cognitive stimulation. For these reasons, we think that LOUISE should be a flexible tool that allows to implement various assistive scenarios. This is why we decided to create an interaction manager that can handle XML dialog descriptions, as presented in Section 5.4.3.

Regarding actual use cases, our work focused on adapted interaction management for older adults with cognitive impairment on two elementary interaction functions, which intervene in most assistive scenarios: choice making and guiding through tasks step-by-step. This matter is covered in Chapter 6, as well as Chapter 7, where another system within which LOUISE technology could be applied is presented.

For the deployment platform, the information we have collected suggests that an ECA system should be adaptable to a device the user already owns or, at least, is not only dedicated to the ECA. However, this falls out of the scope of this research.

Lastly, regarding acceptance, some people have privacy concerns about the perception system but we think these would be overcome if the application is useful enough. The best option may be to have a configuration parameter to choose the activation settings (always on, activation hours, or full control), depending on the application, the user's preferences, and his or her needs, given his or her level of cognitive impairment.

5.4 Phase 2: Automating LOUISE

The WoZ study and information collection through questionnaires and focus group correspond to the conception phase. Thanks to that information we

[5] See footnote 4.

[6] http://www.adobe.com/fr/products/fuse.html

could move on to the development phase, which consists in putting together a fully functional prototype and refining it through iterations. In this section, we present the design we propose for the final version of LOUISE, our assistive ECA and its functionalities. The focus group presented in Section 5.3.2 is also part of the development phase. In addition, we conducted a usability study, which is reported in Chapter 6.

The goal here is to build the prototype and test it. Our main contributions are to propose a specific dialog management approach for older adults with cognitive impairment, based on a finite state machine, and to propose a scenario description in XML to create such dialogs. It is important to note that the XML description of scenarios is closely linked to the structure of conversation management and that the way dialogs are managed depends just as much on the structure of the scenarios as on the dialog manager. This allows us to study how the interaction with older adults with cognitive impairment unfolds and produce guidelines on how to build interactions for our target public. Please note that the system presented in this section is the result of iterations in the use cases presented in Chapter 6.

5.4.1 System overview

The system architecture of LOUISE is similar to the classical structure of ECAs, described in Section 4.2. It is depicted on Figure 5.8. The user behavior data is extracted thanks to the Kinect sensor and used as input for a multimodal behavior analysis module; the results of the analysis are passed on to the interaction manager, which uses an XML dialog description, tagged "scenario" on Figure 5.8. Once the behavior that should be produced by the character is determined by the interaction manager, its BML description is sent to the behavior realizer, composed of a behavior controller, a voice synthesizer and a game engine. The behavior analysis and interaction manager modules are implemented in the same program and communicate with the behavior realizer by exchanging messages through the Apache ActiveMQ[7] middleware, a centralized message broker.

Our system implements the following key features:

- attention estimation, as in the early prototype;
- user speech turn detection;
- automatic speech recognition;
- interaction management allowing for prompting in case of inattention, context reminders when the user is paying attention again and handling of wrong answers or speech recognition errors;
- XML dialog description;

[7] http://activemq.apache.org/

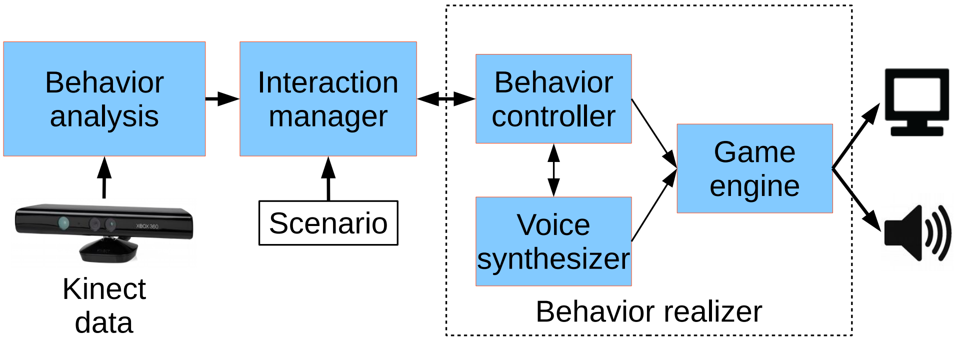


Figure 5.8: Architecture of the LOUISE ECA system.

- BML behavior realization allowing for gestures, facial expressions, head movements, gaze direction, eye saccades and blinks;
- display of images for concept illustration and example videos for step-by-step task guidance.

The last feature is meant to provide information redundancy: instead of just mentioning an object, or describing an action verbally, a picture of the object or a video of the action to perform can be presented to the user. This is meant to compensate for agnosia and aphasia disorders observed in dementia, as recommended by Lapointe *et al.* [111].

5.4.2 Behavior analysis

The behavior analysis module comprises four sources of information:

- user body tracking, in the form of a skeleton;
- user face tracking;
- speech signal;
- sound source localization.

User body and face tracking are used to estimate the user's attention, as described in Section 5.2.2; the speech signal is used for automatic speech recognition (ASR); and the sound source localization is combined with the user body tracking to detect the user's speech turns.

To perform ASR, we used the Microsoft Speech API speech recognition engine and the open-source CMU Sphinx ASR engine. The reason why we chose these tools is that speech recognition is there processed locally; it therefore does not require an access to the Internet. In addition, we thought that users' privacy would be better ensured with these engines than with a cloud-based solution. The ASR is performed given reduced grammar sets, which only include typical words or phrases for greetings, positive answer and negative answers.

To detect users' speech turns, we implemented a very simple method that consists in comparing the sound source angle estimation, given only when the energy of the signal is sufficient, with the angular position of the tracked user. The sound source angle estimation is provided by the Kinect SDK with a confidence indicator γ, between 0 and 1. To determine if the user is speaking or not, the confidence has to be greater than 0.3 and the difference between the estimated sound source angle and the user's position has to be less than a tolerance threshold τ, which is a linear function of the confidence value, given by Equation 5.7,

$$\tau = -10 \times \gamma + 13 \tag{5.7}$$

where τ is the tolerance threshold, and γ is the confidence value. For typical confidence values, this corresponds to an angular difference of 5° for a high confidence of 0.8 and 10° for the minimum acceptable confidence of 0.3.

Also note that we have tried using the movements of the lips, observed thanks to face tracking data, whether by detecting opening of the mouth or the variations of mouth openings, but the data produced by the Kinect was not reliable enough to obtain any robustness.

5.4.3 Interaction manager

Our interaction manager consists in a large finite state machine (FSM) with five orthogonal regions. Each of these region is responsible for an internal representation of an aspect of dialog management:

- state of the dialog;
- state of the speech turn;
- state of the user;
- timer 1;
- timer 2.

This section contains several figures of simplified representations of parts of the FSM. Their nomenclature is based on the standard Unified Modeling Language (UML) representation for FSMs. In these UML diagrams, a black circle represents the entry point; a box represents a state; a large box with small boxes inside represents a sub-machine; an arrow represents a transition; the words next to an arrow represent the event that triggers the corresponding transition and the words between square brackets stand for guards, that is to say conditions of transition.

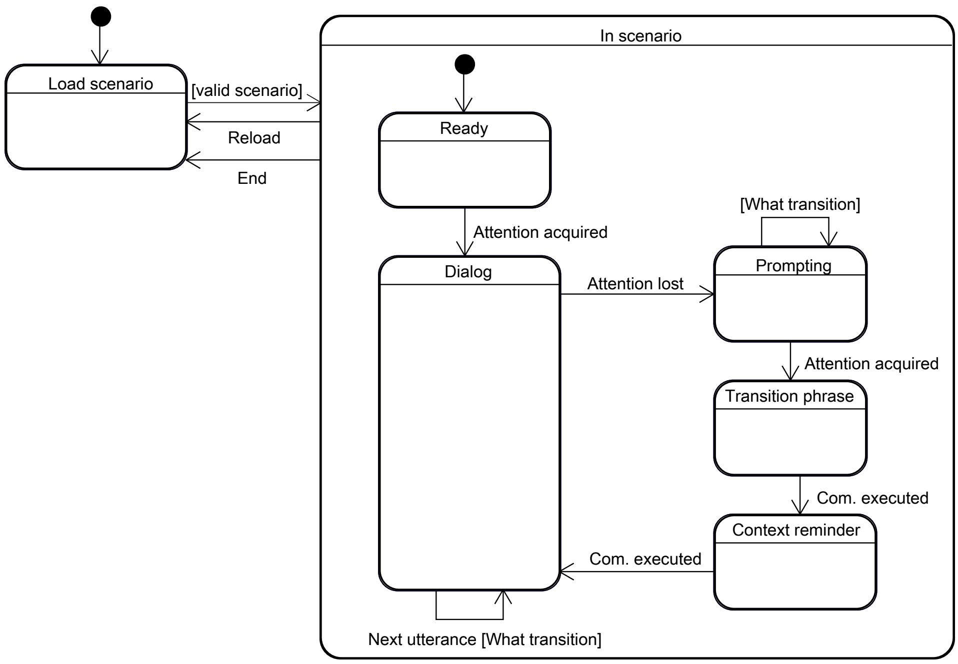


Figure 5.9: Dialog state representation in LOUISE's interaction manager.

Dialog state representation

The dialog state region, depicted on Figure 5.9, allows to load a scenario file and keep track of the current state of the dialog. This component's behavior depends on the scenario description. A dialog "tree" is represented as a set of utterances, with transitions between utterances that can be conditional or not. Each utterance corresponds to a speech turn (or part of a speech turn) of the ECA. Transitions are then performed without condition if the utterance is a statement, or only a part of a speech turn, quickly followed by another utterance; or with conditions if the utterance is a question. In the later case, the next utterance is selected based on the user's answer. This region of the FSM is also responsible for performing interruptions when the user stops paying attention, attention recapture prompting, outputting transition sentences and handling context reminders.

As depicted on Figure 5.9, the dialog starts when attention is acquired by a transition from the *Ready* state to the *Dialog* state. When an utterance has been completed, a *Next utterance* event is triggered by the speech turn manager described below, possibly containing user answer data. Based on the conversation tree in memory, the type of the current utterance (statement or question) and the answer data, the next utterance is selected and sent to the speech turn manager.

When the user is inattentive for more than 2 seconds, an *Attention lost*

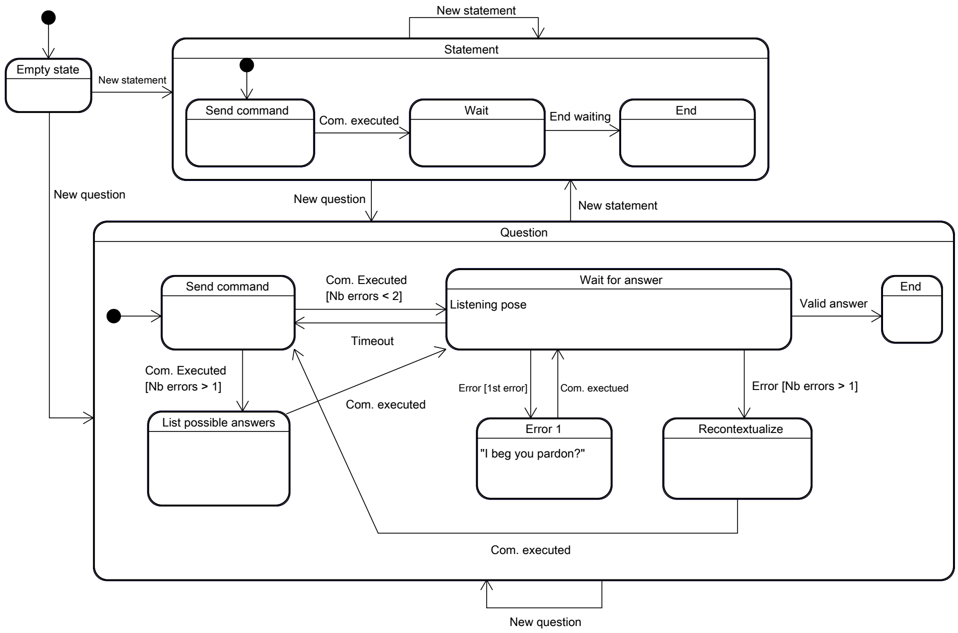


Figure 5.10: State of speech turns representation in LOUISE's interaction manager.

event is sent by the user state representation region. This triggers a state transition between the *Dialog* state and the *Prompting* state. The FSM remains in this state and sends prompting utterances sequentially (several different prompting utterances can be chained), until the user pays attention again and an *Attention acquired* event is received. The FSM then changes its state to the *Transition phrase* state, then to the *Context reminder* state, then back to the *Dialog* state. The last two transitions occur when the corresponding utterances are done being spoken by the behavior realizer.

Speech turn handler

LOUISE's interaction manager can handle two types of utterances: statements and questions. In the case of statements, LOUISE will keep the speech turn in the next utterance. In the case of questions, LOUISE will give the next speech turn to the user. Utterances are therefore managed in two submachines: *statement* and *question*. The corresponding region of the global state machine is depicted in Figure 5.10.

The utterance handler of LOUISE is composed of an *empty* state, which is the default initial state and the two submachines mentioned above. Transitions occur when a new utterance is sent by the dialog manager and the target submachine simply depends on the type of utterance, that is to say whether it is a question or a statement (attention prompts are considered to be statements).

The *statement* submachine is composed of three states: *send command*, *wait* and *end*. The *send command* state performs the action of sending the BML behavior instruction to produce to the behavior realizer. Once the behavior has been executed by the virtual character, a feedback message is received and a *com. executed* event is fired, which triggers a transition to the *wait* state. This state allows to control the time between the current statement and the next utterance. Once the waiting time is up, signaled by the *end waiting* event, the *statement* submachine transitions to the *end* state and the dialog manager can send the next utterance.

The *question* submachine has the same basic functioning but, instead of waiting for a time delay before the next utterance can be sent, it waits for an answer from the user. In addition, it handles incorrect answers and/or speech recognizer errors. When the state machine is in the *wait for answer* state, the character adopts a listening posture, as depicted on Figure 5.11. If an error occurs for the first time for the current question, it goes to the *error* state, in which an error resolution behavior is sent to the behavior realizer. Then, once the behavior has been executed, the state machine goes back to the *wait for answer* state and the user has a second chance to answer. If a second error occurs, a context reminder is sent in the *recontextualize* state; then the question is asked again and the possible answers are listed in the *list possible answers* state, before going back to the *wait for answer* state. Lastly, if after a given time the user does not answer, the question is asked again. This sequence is repeated until a valid answer is provided and can potentially loop forever, as there is no escape case.

User state representation

Our interaction manager includes an internal representation of the user's state, depicted on Figure 5.12. It is composed of a *no user* state and a *user detected* submachine. The transition between *no user*, which is the entry state, and *user detected* occurs when a user is detected by the Kinect sensor. The opposite transition occurs when user tracking is lost.

The *user detected* submachine is composed of three orthogonal regions. The first region represents the attentional state of the user, which is updated on every data frame by the attention estimation component of the behavior analysis module. It is also possible to disable attention estimation, by transitioning to the *in task* state. This feature is useful for step-by-step task guidance applications: when LOUISE asks the user to perform an action with an object, it is likely, even desired, that he or she directs his or her attention towards the object, instead of the ECA's display. In this case, we would not want LOUISE to perform attention recapture prompts. This is why we added the *in task* state. The second and third regions are used to manage speech turns and allow to represent if the user is currently speaking

Figure 5.11: LOUISE in the listening posture.

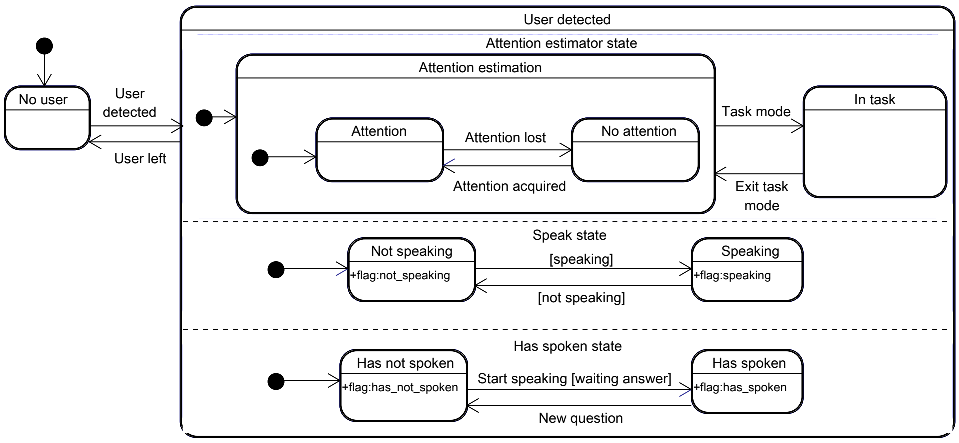


Figure 5.12: User state representation in LOUISE's interaction manager. The *+flag* labels indicate that when the state that contain them is active, the corresponding flag has to be raised to notify the other regions of the FSM of the current state of the user.

or not and if he or she already has spoken, or not. These parts of the machine are reset to *not_speaking* and *has_not_spoken* after each user speech turn, that is to say every time a situation requiring the user to provide an answer has just been resolved.

5.4.4 Behavior realizer

The behavior realizer component in LOUISE allows to interpret and execute Behavior Markup Language (BML) commands, which performs all communication functions allowed by this behavior description domain-specific language (DSL): speaking with synchronized lip animations, gesturing, showing facial expressions, performing head movements, directing the character's gaze, blinking, and performing eye saccades. In addition, it allows to display images and a virtual television screen, on which example videos of expected actions can be shown.

Our behavior realizer is built on top of SmartBody [196], a state-of-the-art BML realizer presented earlier in Section 4.4.1. It can be seen as an animation controller, or behavior controller, as it is labeled on Figure 5.8. Its role is to interpret BML commands and transform them into character animations and sounds. When a BML command is received, it sends the speech contents to the voice synthesizer, which in turn synthesizes the voice samples and returns the phoneme scheduling. Given the schedule for the phonemes, the animation is computed and executed. The game engine is used for rendering the character and playing the synthesized sounds. The integration of SmartBody in the game engine is done by creating a SmartBody scene, which contains the animation skeleton, or "rig"; creating a scene in the game engine with the full model of the character (mesh, rig and textures); and copying the state of all joints of the skeleton in the SmartBody scene over to the ones in the game engine scene. The game engine that we used is Panda 3D[8].

We chose SmartBody for five main reasons: it comes with a large database of animations that allow to have the virtual human character perform many communication gestures; it allows to "retarget" animations, that is to say adapting a key frame animation created for a given character model to another model, even with a different skeleton topology; some of the behaviors, namely gaze directions, head movements and eye saccades are performed procedurally, which allows to use them with any character model; it includes a wrapper for the Apache ActiveMQ[9] middleware, called Virtual Human Message, which makes it ready for simple external communications with other modules of an ECA application; and it is distributed under an open-source GNU license. Using this tool makes it easy to add as many

[8]See footnote 3
[9]See footnote 7

Figure 5.13: Rendering of the Charlotte embodiment.

character models as desired. The animation database can also be extended if one wants to use gestures that are not already available.

We programmed our application so that any model created with Autodesk Character Generator[10] can be added with little work. The full process of creating and adding a new model should take less than an hour. Adding other models, with a different skeleton topology, is also possible but takes more time, as it requires creating a few animations for facial expressions and lip animations, writing a script to specify the skeleton joint mapping for animation retargeting and creating a file to list the joints of the face. It is also worth noting that SmartBody already comes with a few character models and natively supports models created with Mixamo, now called Adobe Fuse CC[11], another popular virtual human character creation tool. For our experiments, we had two different female character models available, so participants had a choice of embodiment. The first one, called Louise, is depicted on Figures 5.7 and 5.11; the second one, called Charlotte, is shown on Figure 5.13.

For implementation, we used the Python API of both SmartBody and Panda 3D. Note that this integration was already done and distributed with

[10]See footnote 4
[11]See footnote 6

the SmartBody software, but we had to rewrite it because it is obsolete. This method was chosen because it allowed to easily create methods that can be called in BML events (see Section 4.6 for details about BML). We used this feature to perform several additional actions: notifying the interaction manager when a behavior is completed, by sending a Virtual Human Message, which in turn fires a *Com. executed* event upon reception; displaying and removing images; displaying and removing the virtual television set; start and stop playing videos on the virtual television set; and having the character speak louder for one utterance. Figure 5.14 shows the layout of the application when displaying images and videos.

5.4.5 Scenario description language

Scenarios are described thanks to XML files. The scripting is done using a custom XML specification. This description language is inspired by the finite state machine implementation of dialog trees. Utterances are seen as states and have transitions to other utterances, with or without condition, whether it is a statement or a question. In this section, we explain how it is structured.

A scenario is described in a

```
<scenario> </scenario>
```

markup. The structure of the scenario has to be written between the start tag and the end tag.

A scenario is composed of utterances. There are several types of utterances: *question*, *statement*, *prompt* and *transition*. The last two correspond to attention prompting and transition phrases spoken before context reminders, right after the user is detected as paying attention again.

An utterance is defined thanks to a double-tag markup:

```
<utterance id="name" type="chosen_type" wait="time_in_seconds"
mode="mode">
... content ...
</utterance> .
```

The parameters *id* and *type* are mandatory; the *wait* and *mode* parameters are optional. Here are the details of each parameter's influence.

- The *id* will allow to give each utterance a name to refer to it in the scenario description. Please note that ids may be any sequence of letters and numbers and signs, but may not contain whitespaces.

- The *type* parameter allows specifying which kind of utterance is being described: "statement" is for declarative sentences; "question" is for questions; "interruption" is for attention prompting; and "welcome-back" is for transition between attention prompting and context reminder. For the latest case, only one utterance may be specified, and

(a)

(b)

Figure 5.14: (a) The Charlotte character showing a bottle of sparkling water. (b) The Charlotte character showing an example video of taking a water bottle's screw cap off.

its id should be "wb". All other types allow as many entries as one likes.

- The *wait* parameter allows to control timing. In the case of an utterance of type "question", it allows controlling the time the person has to answer, before the question is repeated. For other utterances, it allows controlling the time before the transition to the next behavior (except for "welcomeback"). If this parameter is not set, default values are used (6 seconds to answer a question, 1.5 second for "statement" and 2 seconds for others).

- The *mode* parameter is used to specify a special functioning mode. So far, it only has one valid value: "task". This allows switching off the attention estimator while the user is asked to perform a task action. If this parameter is omitted, or if an invalid value is given, the attention estimation will be activated by default.

Each utterance has to contain at least a command and a transition (even the last utterance; in this case just use its own id). The command is specified thanks to a double tag markup as follows:

```
<command> content in BML </command> .
```

Between the two tags, the behavior corresponding to this step in the interaction scenario has to be specified in BML and SSML. For instance:

```
<command>
    <speech>
        Hello! <break time="1s"/> My name is John.
    </speech>
</command> .
```

A transition is specified as

```
<transition condition="answer">
    id_of_the_next_utterance
</transition> .
```

The *condition* parameter is used for questions and allows selecting the next utterance, based on the answer provided by the user. So far, the exact word specified in the condition parameter has to be recognized by the ASR module for the transition to occur. This parameter is optional. In the case of "statement" utterances, there is only one transition, without condition. For instance:

```
<transition>next_utterance</transition> .
```

For "question" utterances, there needs to be as many transitions as valid answers. The *condition* parameter is used in this case only. For instance:

```
<transition condition="yes">
    utterance_if_yes
</transition>
<transition condition="no">
    utterance_if_no
</transition> .
```

Note that it is possible to have only one transition, with a condition. The dialog will not move on for as long as the condition is not met.

An utterance may also contain an optional *recontextualisation* field. This allows to define context reminder behaviors that are relevant to the current topic of the conversation. It is also a BML command, which will be used in the context reminder phase. That is to say, the BML instructions in this field are executed by the character after a loss of the user's attention. This field is used just like the *command* field:

```
<recontextualisation>
    BML instructions
</recontextualisation> .
```

If this field is omitted for a given utterance, the recontextualisation of the previous utterance will be used. One has to make sure to define at least one, in the first utterance. However, prompting utterances (of type "interruption") and transition after prompting, (type "welcomeback") do not require the *recontextualisation* field.

The scenario should include at least one utterance of type "question" or "statement", one of type "interruption" and one and only one of type "welcomeback". It may contain several prompting utterances that will be looped in the order defined thanks to the transition field, until user attention is detected again. The first utterance to start the interaction should always have "start" as its id. Similarly, the last utterance should always have id="end". If one wants a scenario to have several ends, an utterance for which the *id* is "end", containing an empty command field, should be defined and the final utterances for each branch of the dialog tree should transition to it.

Special utterances ("interruption" and "welcomeback") are considered in a different context and *ids* may be reused. In fact, the first prompting utterance to be executed needs to have "start" as its *id*. The last utterance does not require to have "end" as its *id* but needs to have "start" as its only transition.

This XML description of interaction scenarios can be seen as the basis for a more evolved DSL that could be called Assistive Interaction Scenario Markup Language or AISML. The XML schema specification for the current version is given below.

```
<?xml version="1.0" encoding="UTF-8" ?>
```

```
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema">

    <xsd:complexType name="BLM">
        <xsd:sequence>
            <xsd:any minOccurs="0" maxOccurs="unbounded"/>
        </xsd:sequence>
    </xsd:complexType>

    <xsd:simpleType name="uttType">
        <xsd:restriction base="xsd:string">
            <xsd:enumeration value="statement"/>
            <xsd:enumeration value="question"/>
            <xsd:enumeration value="interruption"/>
            <xsd:enumeration value="welcomeback"/>
        </xsd:restriction>
    </xsd:simpleType>

    <xsd:element name="transition">
        <xsd:complexType>
            <xsd:simpleContent>
                <xsd:extension base="xsd:IDREF">
                    <xsd:attribut name="condition"
                    type="xsd:string"/>
                </xsd:extension>
            </xsd:simpleContent>
        </xsd:complexType>
    </xsd:element>

    <xsd:element name="utterance">
        <xsd:complexType>
            <xsd:sequence>
                <xsd:element name="command" type="BLM"/>
                <xsd:element ref="transition" minOccurs="1"
                maxOccurs="unbounded"/>
                <xsd:element name="recontextualization"
                type="BLM" minOccurs="0" maxOccurs="1"/>
            </xsd:sequence>
            <xsd:attribut name="id" type="ID" use="required"/>
            <xsd:attribut name="type" type="uttType"
            use="required"/>
            <xsd:attribut name="wait"
            type="xsd:nonNegativeInteger"/>
            <xsd:attribut name="mode" type="xsd:string"/>
```

```
        </xsd:complexType>
    </xsd:element>

    <xsd:element name="scenario">
        <xsd:complexType>
            <xsd:sequence>
                <xsd:element ref="utterance" minOccurs="4"
                maxOccurs="unbounded"/>
            </xsd:sequence>
        </xsd:complexType>
    </xsd:element>
</xsd:schema>
```

Note that there is a small inconsistency between this XML schema definition and the explanations given about the formalism: we have reused ids, which is not correct as the xsd:ID type imposes that each ID is unique. The “start” and “end” ids should be reserved for statement and question utterances and another set of mandatory ids, such as “start_prompt” and “end_prompt”, should be defined for interruption utterances.

To finish this section, let us consider an example of the basic content of a scenario file.

```
<scenario>

    <utterance id="start" type="statement">
        <command>
            <speech id="sp" type="application/ssml">
                Hello!
            </speech>
            <head id="hd" start="sp:end" type="NOD"
            amount="0.5"/>
            <event start="hd:end" message="triggerEoBEvent()"/>
        </command>
        <transition> Ready? </transition>
        <recontextualisation>
            <speech id="sp" type="application/ssml">
                I was saying.
            </speech>
            <event start="sp:end" message="triggerEoBEvent()"/>
        </recontextualisation>
    </utterance>

    <utterance id="Ready?" type="question" wait="5">
        <command>
            <speech id="sp" type="application/ssml">
```

```
            Are you ready?
        </speech>
        <event start="sp:end" message="triggerEoBEvent()"/>
    </command>
    <transition condition="yes"> Cool! </transition>
    <transition condition="no"> ComeBackLater </transition>
</utterance>

<utterance id="Cool!" type="statement" mode="task">
    <command>
        <speech id="sp" type="application/ssml">
            Cool! Let us get started!
        </speech>
        <event start="sp:end" message="triggerEoBEvent()"/>
    </command>
    <transition> end </transition>
</utterance>

<utterance id="ComeBackLater" type="statement" wait="3">
    <command>
        <speech id="sp" type="application/ssml">
            OK. Come back when you are ready.
        </speech>
        <event start="sp:end" message="triggerEoBEvent()"/>
    </command>
    <transition> end </transition>
</utterance>

<utterance id="end" type="statement">
    <command>
        <speech id="sp" type="application/ssml">
            Bye bye!
        </speech>
        <event start="sp:end" message="triggerEoBEvent()"/>
    </command>
    <transition> end </transition>
</utterance>

<utterance id="start" type="interruption">
    <command>
        <speech id="sp" type="application/ssml">
            Hey! Are you there?
        </speech>
        <event start="sp:end" message="triggerEoBEvent()"/>
```

```
        </command>
        <transition> start </transition>
    </utterance>

    <utterance id="wb" type="welcomeback">
        <command>
            <speech id="sp" type="application/ssml">
                Welcome back.
            </speech>
            <event start="sp:end" message="triggerEoBEvent()"/>
        </command>
        <transition> end </transition>
    </utterance>

</scenario>
```

This short scenario starts by greeting the user with “Hello!” and a polite head nod. This first utterance also adds “I was saying” as a context reminder sentence. As no other *recontextualization* field is used, this will be the context reminder sentence throughout the scenario. There is only one attention prompting sentence, “Hey! Are you there?”, and one transition sentence “Welcome back.”. Therefore, in case of inattention of the user, the ECA will say “Hey! Are you there?”, wait for two seconds, the default wait time, as no value has been specified for the optional *wait* parameter, until the user pays attention again. Then, it will always say “Welcome back. I was saying...” as both transition and context reminder. If the user is detected as paying attention to the ECA, it will go on and ask the question “Are you ready?”. The user will then have 5 seconds to answer. If there is no answer, the character will repeat the question. If the user answers “yes”, LOUISE will say “Cool! Let us get started.”, and then end the conversation by saying “Bye bye!”; if the user says “no”, LOUISE will say “OK. Come back when you are ready.” and end the conversation the same way. In case of answers other than “yes” or “no”, possibly due to an ASR error, the error recovery strategy, not present in the excerpt, will be used: the character will first say “I beg you pardon?”, then the user will be able to answer again for 5 seconds. If a wrong answer or ASR failure occurs again, LOUISE will say the context reminder (not very useful in this example), then ask the question again and list possible answers by saying “You can answer this question by yes or by no.”. If more errors occur, it will keep following this context reminder – repeat the question – list possible answers scheme.

Chapter 6

Assisted task management using LOUISE

Résumé en français. *Le prototype de LOUISE, entièrement automatique, que nous proposons est un logiciel d'ACA polyvalent qui repose sur la description d'arbres de dialogue en XML. Par conséquent, la gestion d'interaction, qui dépend des applications, est autant déterminée par le contenu de l'arbre de dialogue que par les éléments compilés du programme. Le gestionnaire d'interaction décrit dans la section 5.4.3 peut par conséquent être vu comme une structure sur laquelle construire des scénarios d'assistance. A ce jour, il y a peu de connaissances sur l'interaction entre les personnes âgées atteintes de troubles cognitifs et les ACAs. Il est donc nécessaire de mener des expériences, à l'aide du prototype que nous avons créé, avec des patients ayant un diagnostique de TCL ou de démence, dans des conditions écologiques.*

Dans ce chapitre, nous présentons notre étude sur la gestion spécifique des interactions entre les personnes âgées atteintes de troubles cognitifs et LOUISE, fondée sur quatre cas d'application simples mais crédibles. Dans la section 6.1, nous présentons les buts de cette étude. Dans la section 6.2, nous identifions quatre aspects de la conversation qui interviendraient dans la plupart des scénarii d'assistance dans lesquels LOUISE pourrait servir d'interface Homme-machine et proposons des structures conversationnelles pour les gérer. Après quoi, dans la section 6.3, nous présentons quelques scénarii d'utilisation de LOUISE sur lesquels nous nous sommes appuyés pour tester notre prototype, ainsi que nos structures conversationnelles. Enfin, nous rapportons nos expériences auprès des personnes âgées dans les sections 6.4 et 6.5, puis nous discutons nos choix de conception de structures conversationnelles à la lumière des observations réalisées lors de cette phase expérimentale dans la section 6.6.

Le but est d'évaluer notre prototype de LOUISE et de le raffiner, ainsi que d'explorer les spécificité de la gestion de scenarii d'assistance aux personnes âgées atteintes de démence. Pour cela, nous avons proposé des stratégies

et structures conversationnelles concernant quatre aspects des conversations possible : l'initiative du dialogue ; poser des questions et réagir aux réponses ; proposer un choix parmi plusieurs items ; et guider la réalisation d'une tâche pas-à-pas.

Nous avons ensuite élaboré quatre scénarii simples, mais plausibles, d'utilisation de LOUISE, implémentant ces stratégies et structures conversationnelles : rappel de s'hydrater et guidage dans la tâche de se préparer un verre d'eau ; choix du menu d'un repas ; rappel de prise de médicaments et guidage dans la tâche d'ingestion ; rappel et guidage de prise de tension.

Nous avons ensuite évalué ces scénarii auprès de 14 patients de l'hôpital Broca ayant des troubles cognitifs légers à sévères. Les participants ont majoritairement exprimé un haut niveau de satisfaction en termes de charme, de clarté, et de simplicité. De plus, nous avons obtenu de bons résultats d'utilisabilité sur le scénario de choix du menu. Nous avons également pu identifier que la stratégie de guidage dans une tâche pas-à-pas devrait être différente selon le niveau de fonctionnement cognitif de la personne. Cependant, nous avons identifié une limitation majeure de notre système, émanant de la qualité de la reconnaissance vocale utilisée et de l'inadaptation d'une solution nécessitant de ne répondre que par "oui" ou par "non" pour ce public. Enfin, nous avons identifié que le capteur Kinect pour Xbox 360 n'est pas non plus complètement adapté, dans la mesure où il ne fonctionne pas toujours correctement lorsque la personne est placée à moins d'un mètre, ce qui est difficile à éviter dans ce cadre d'utilisation.

The fully automatic prototype that we propose for LOUISE is a general-purpose ECA software that relies on dialog tree descriptions in XML. Therefore, the interaction management part, which is application-dependent, is determined just as much by the contents of the dialog tree as by the compiled software components. The interaction manager described in Section 5.4.3 can thus be seen as a framework to build assistive scenarios. So far, little is known about the interaction between older adults with cognitive impairment and ECAs. It is therefore necessary to experiment with the prototype we have created, in ecological conditions, with patients with MCI or dementia.

In this chapter, we present our study about the specific management of interactions between older adults with cognitive impairment and LOUISE, based on four believable yet simple use cases. In Section 6.1, we present the goals of this study. In Section 6.2, we identify four aspects of conversation that will intervene in most assistive scenarios in which LOUISE could be included as a user interface, and propose conversational structures to handle those. After that, in Section 6.3, we present some use scenarios for LOUISE that we relied on to test both our prototype and conversational structures. Lastly, we report on our experiments with older adults in Sections 6.4 and 6.5, then discuss our choices of design and conversational structures in the light of the observations made during this experimental phase in Section 6.6.

6.1 Goals

The main goal of this study is to complete the development phase of the LOUISE ECA. This means two things: evaluating and refining the system as a whole, particularly our model of interaction management; and proposing, testing and refining conversational structures for goal-oriented dialogs between LOUISE and elders with cognitive impairment. All of this is done through "usability" testing, in ecological conditions, based on "believable" use cases. To this aim, we created several scenarios containing the basic actions that will very frequently be used in assistive applications in which LOUISE is susceptible to be included: asking questions and reacting to answers, having the user make a choice and providing step-by-step instructions to complete a task. Addressing these issues is conducted in an iterative way, as hypotheses about effective interaction management strategies need to be refined and validated.

The secondary goal is to test the system with people with varying levels of cognitive functioning, to identify the maximum level of cognitive impairment that LOUISE, and likely most tailored ECA systems, can handle.

One of the things that is perfectly normal in this approach, but may seem surprising, is that changes to the system are susceptible to be made in the middle of the evaluation study. This may seem like an unorthodox practice, but it is done deliberately in our case. The reasons for this are that testing is very time-consuming and recruiting patients with dementia or cognitive impairment is not easy. Indeed, as seen in Chapter 2, studies involving this kind of public in evaluation of ECAs rarely have more than 15 participants. As a result, when a few tests reveal clear flaws, we do not want to waste time on further testing before improvements are made.

6.2 Structures of conversation

For each of the basic actions that compose the scenarios, we propose a specific conversational structure. These are based on the observations we have made in the anthropological analysis of interactions between elders with cognitive impairment and the Wizard of Oz version of LOUISE, reported in Section 5.2.6, and the knowledge about the symptoms of dementia, presented in Section 2.1.

6.2.1 Dialog initiative

In the Wizard of Oz experiment, none of the participants complained that the ECA always had the dialog initiative, neither did they try to take the initiative themselves. In addition, in the focus group, the participants clearly stated that LOUISE should initiate the dialog. This is why we built our system with the assumption that LOUISE will always keep the initiative and did not try to accommodate for mixed-initiative dialogs. This assumption is also in line with the knowledge of dementia symptoms: people with dementia seldom take actions spontaneously, because of apathy and prospective memory disorders. In this regard, managing the interaction between an ECA and an older adult with cognitive impairment is easier than dialog management for cognitively healthy people, which would require handling mixed initiative dialogs.

6.2.2 Asking questions and reacting to answers

Our interaction manager could very well handle open questions with multiple possible answers, though choosing the appropriate response could be challenging and would likely require some language processing of the user's answers, to interpret his or her intents, in the behavior analysis stage. However, with our target user group, we have identified that narrow or contrasted questions, such as yes/no questions, should be privileged. This is also in line with the recommendations of Zajicek [234], who suggested that menus in speech-based interfaces for older adults should be limited to 3 items at a time. Hence, our conversational trees, which in fact are more like conversational graphs, as there can be branches that reunite at some points, will normally have a maximum of two branches at each node.

6.2.3 Multiple choices

To perform multiple choices, all options are presented first, then a yes/no question is asked for each item. Each time an item is mentioned, a picture illustrating the concept is displayed, as shown on Figure 5.14. When an image is displayed, the character points at it and directs its gaze towards it, by turning its eyes and head, so its body posture is still facing the user. Then, when the image disappears, the character turns its head and eyes towards the user again. This is meant to direct the user's attention towards the object of interest when necessary. This specific choice of gaze direction and posture behaviors is inspired by the findings of Pejsa *et al.* [154], who studied the influence of gaze and postures on the user's quality of attention and reported that the best compromise is for the ECA to fully look towards the user when speaking (affiliative strategy) and fully looking towards the object of interest when making mentions of it (referential strategy). This also resembles the model of engagement behaviors for collaborative tasks involving a human and a robot proposed by Rich *et al.* [168].

If no item is selected, there are two options: choosing none or starting over at the first option. After one option has been selected, an explicit confirmation is asked. When a choice is confirmed, an acknowledgment behavior is executed and the dialog can go on. This is similar to the approach for confirmations recommended by Yaghoubzadeh *et al.* [229], who stated that it is best to confirm each bit of information individually.

Here is a simple example dialog with only two options, illustrating this conversational structure.

```
ECA: ‘‘We have still water or sparkling water.’’
ECA: ‘‘Would you like still water?’’
User: ‘‘No.’’
ECA: ‘‘Would you like sparkling water?’’
User: ‘‘Yes.’’
ECA: ‘‘You want sparkling water, is that correct?’’
User: ‘‘Yes.’’
ECA: ‘‘OK. The sparkling water is in the green bottle on your
left-hand side.’’
```

The typical conversation graph we used for having the user choose one option in several possibilities is depicted on Figure 6.1. The “None?” option represents the case in which it is allowed not to choose an item. In that case, the conversation will jump back to the first item proposition only after the question of picking no item is asked, instead of doing it right after the n^{th} option was refused.

6.2.4 Step-by-step task instructions

For task instructions, each step is managed in the following order:

1. the ECA explains the action to perform, while pointing and gazing at the virtual television set, on which the example video has started looping;

2. it lets the user watch the example video as it loops on the virtual TV screen for a few seconds and keeps gazing at the screen during that time – this is when the *wait* parameter for statement utterances is the most useful (see Section 5.4.5);

3. LOUISE directs its gaze towards the user and instructs him or her to perform the action (the attention estimation is disabled and the video is still looping on the virtual screen);

4. LOUISE waits for a few seconds, thanks to the *wait* parameter in statement utterances;

5. the attention estimator is enabled again, the video is stopped and the ECA asks if the user has finished performing the action;

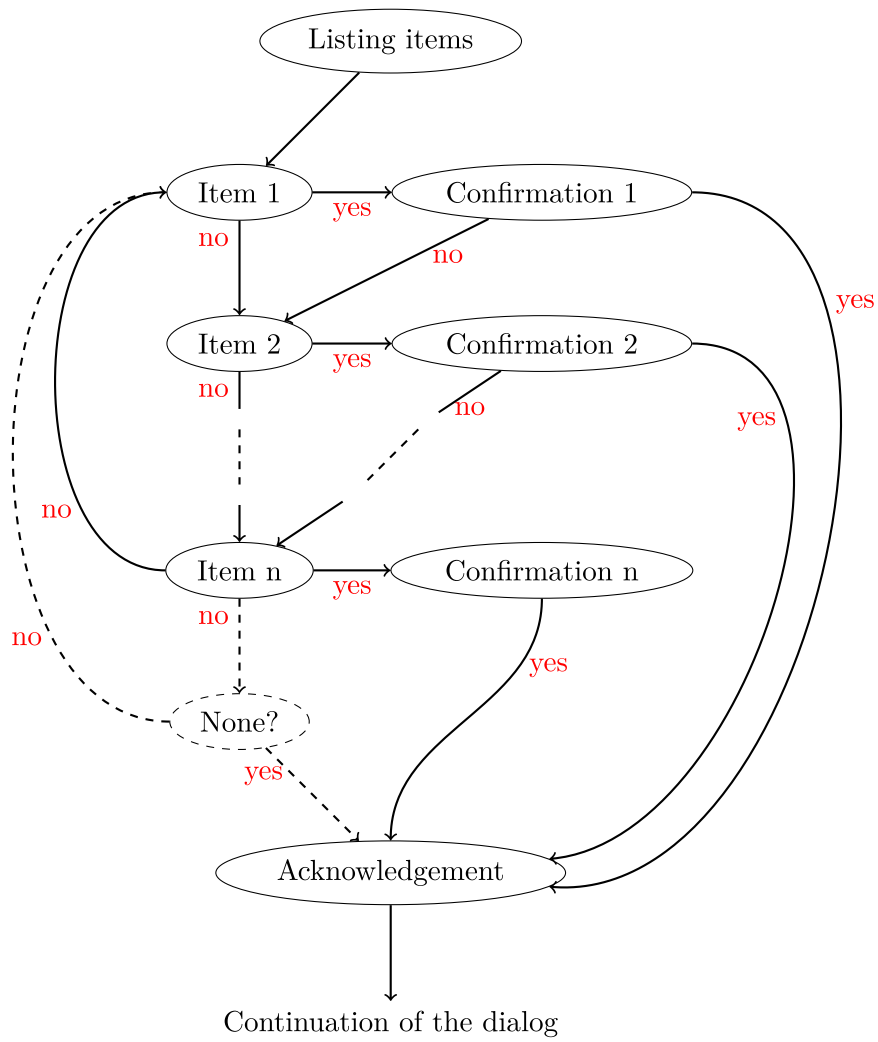


Figure 6.1: Conversation graph for selecting items with multiple choices.

6. if the answer is "yes", the task is assumed completed and the dialog can go on to the next instruction, or to the conclusion;
7. if the answer is "no", the dialog manager goes back to the first step and the instructions are given again;
8. when the task is complete, the ECA congratulates the user and the video is stopped;
9. in case of attention loss, the video is stopped and only gets restarted after the context reminder is performed.

Note that the gaze and posture management approaches are the same as when still images are shown, except that the character keeps looking at the screen when the user is supposed to watch the example video. In addition, the fact that LOUISE gazes at the user again when it asks him or her to perform an action is supposed to make the user think that it is checking for the execution of the action.

Here is an example of dialog, managed using this structure.

```
(attention estimation on)
ECA: ‘‘To take the pill in the box, you have to place your
hand on the hole and turn the box upside-down.’’
(five seconds delay, video looping, ECA gazing at the TV)
(ECA gazing at the user)
ECA:‘‘Now, take the pill in the box.’’
(attention estimation off)
(five seconds delay, video looping)
(attention estimation on, stop video)
ECA: ‘‘Have you taken the pill in the box?’’
User: ‘‘No.’’
ECA: ‘‘To take the pill in the box, you have to place your
hand on the hole and turn the box upside-down.’’
(five seconds delay, video looping)
ECA:‘‘Now, take the pill in the box.’’
(attention estimation off)
(five seconds delay, video looping)
(attention estimation on)
ECA: ‘‘Have you taken the pill in the box?’’
User: ‘‘Yes.’’
ECA: ‘‘Well done!’’
```

6.3 LOUISE use scenarios

For this experiment to have some ecological validity, we wanted to use believable use scenarios for LOUISE. Throughout the study, we have used four different scenarios, which were validated by physicians or in previous exchanges with older adults:

1. reminding to drink water and guiding through the task of preparing a glass of water;

2. reminding medication intake and guiding through the task of taking pills;

3. guiding through the task of measuring one’s blood pressure; and

4. choosing the composition of a meal.

The blood pressure scenario was created to replace the medication intake for testing with the most strongly impaired patients, as the physicians of the hospital were afraid that someone would choke when taking the Tic-Tac candies, used as fake pills (in the Broca hospital's services, pills are administered to the patients with moderate to severe dementia by mashing them down and mixing them with compote). The water-drinking scenario could be used both at home or in institutions; the medication intake and blood pressure scenarios mostly target patients still living in their homes; and the menu scenario mostly targets institutionalized patients.

In the remainder of this section, we discuss the usefulness of each scenario and provide details about how we structured them. The XML description of the hydration reminder scenario in its final version, that is to say after they were improved through testing, is given in Appendix B.2 as a complete example, which includes all types of conversation structures we have investigated.

6.3.1 Drinking water

It is common knowledge that good hydration is essential to health. Older adults in general tend not to drink enough water, partly because they do not feel thirst (and hunger) as much as their younger counterparts. In addition, this issue becomes critical when the weather gets very hot and older adults may die from dehydration. In France, on several occasions, heat waves caused many casualties in older adults and, since then, the government has prepared emergency plans to address such situations. This use case for LOUISE could be particularly helpful in that context. In dementia, apathy may cause a lack of energy and motivation to initiate actions; as a result, people may ever feel thirsty but not do anything about it. Furthermore, a lot of dementia patients have prospective memory disorders, which may lead them not to act or make them unable to complete the necessary tasks. This is why we proposed to use an ECA to remind people to drink water and guide them through the necessary sequence of actions. The system could potentially include an intelligent cup, such as the Vessyl smart cup[1], which could not only allow verify that some liquid has been poured into the cup, but also detect what kind of liquid it is or measure the liquid's temperature, to warn the user if the beverage is too hot, for instance.

The scenario unrolls as follows:

1. the character smiles, then greets the user with a verbal phrase and a polite nod;
2. the user has to greet back to start the interaction;
3. LOUISE introduces itself;

[1] http://myvessyl.com/

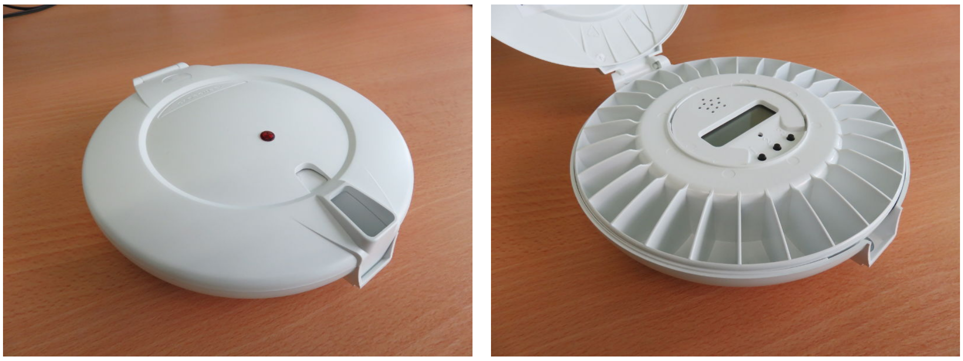
4. the ECA asks the user if he or she is thirsty;
5. if not, the character reminds how important it is to drink water frequently, then asks if the person wants to drink;
6. if the user refuses to drink, the ECA jumps to the conclusion, as we would not want the ECA to force people to drink, for ethical reasons;
7. when the user has agreed to drink, the ECA has him or her choose between still water and sparkling water;
8. once the choice is made, the user is asked if he or she needs help to prepare the glass of water;
9. if the answer is negative, the ECA jumps to the conclusion;
10. if the answer is positive, the ECA guides the user through the task in three steps – taking the screw cap off the bottle, pouring the water in the glass and putting the cap back on;
11. the interaction is concluded by saying that the user can come back if he or she wants to drink;
12. the dialog ends with the ECA saying goodbye.

6.3.2 Taking pills

The medication time reminder is one of the most common applications. There are electronic pill dispensers on the market that ring when it is time to take medicines and keep ringing until the pills are retrieved from the box. Some advanced models can even send an SMS when the pills have not been retrieved at medication time. The issue we address in this use case is that people may not remember what the ring is about, what to do to retrieve the pills in the box, or may retrieve the pills without actually taking them afterwards. Instead, we propose to use an ECA to do the reminder and explain how to take the pills in the box. In addition, an ECA could also remind the person what the pills are for, which may increase the odds that the person actually takes his or her medicines. Furthermore, if the pill box is connected to the ECA, it could provide some feedback about the actions of the user on the device to the ECA. Another possibility is to perform activity recognition through video analysis. These strategies could provide some confidence about whether the patient has taken the pills or not. In case of doubt, a caregiver could be informed.

We used the Careousel Mk3 pill dispenser[2], presented on Figure 6.2. The round box has a hole on top, from which the pills are retrieved by putting

[2] http://www.pivotell.co.uk



(a) (b)

Figure 6.2: Careousel Mk3 pill dispenser. (a) Closed. (b) Open.

one's hand on it and turning the box upside-down. The hole is too small to put one's finger inside to take the pills and, once prepared by a caregiver, the box can be locked. Inside the device is a half torus-shaped piece with slots to prepare the doses for each medication intake. When it is medication time, the box rings, the LED on top of it blinks, and the half torus rotates by one slot. It does not stop ringing and blinking until it is turned upside down. We designed our scenario as if the ECA could replace the alarm, then guide the user through the task of taking pills. We did not actually connect the box to the ECA, but it would be easy to do so, and our scenario would not have to be modified much.

The scenario is organized as follows:

1. LOUISE smiles and greets the user;
2. the user has to greet LOUISE back;
3. the character introduces itself;
4. LOUISE says it is time to take medication and informs that it will show what to do on the virtual television set;
5. the ECA guides the person through the task of taking pills in 5 steps – preparing a glass of water, as in the water scenario (3 steps), retrieving the Tic-Tac candy (replacing the pills) from the box and swallowing the pill with the water – each step is managed as described in Section 6.2;
6. at the last step, if the user says he or she has not taken the pills, LOUISE takes a worried expression and asks if someone should be called;

7. if yes, the interaction is concluded by saying it has called someone for help and the dialog manager jumps to the last utterance;
8. if no, the dialog manager goes back to the last step of pill intake instruction (putting the pills in one's mouth and swallowing them with the water);
9. if the person says he or she has taken the pills, LOUISE congratulates him or her and asks if he or she has drunk the whole glass;
10. if not, the user is advised to finish the glass, because it is important to stay hydrated;
11. otherwise, LOUISE congratulates the user;
12. LOUISE asks the user if he or she wants it to come back when it is medication time again;
13. once the user as answered, an acknowledgment behavior is executed;
14. the ECA says goodbye and the interaction ends.

6.3.3 Measuring blood pressure

Blood pressure is an important physiological measure that many people have to keep track of. There are some automatic blood pressure monitors people can use at home, such as the Microlife BP A100 Plus[3] we used in this study, shown on Figure 6.3. There also are wireless blood pressure monitors, such as the one created by Withings[4], that connect to a smartphone application, through Bluetooth or WiFi. The idea behind this scenario is to address the same issues as in the pill dispenser scenario, but for blood pressure.

The scenario, meant to replace the medication intake scenario for hospitalized people who are at risk of choking when swallowing the pills, follows the same order. The only differences are that people can ask for help at the point when they are supposed to put the armband on, instead of at the last action, as it is the most critical one, and that the task is longer and decomposed in more steps.

6.3.4 Choosing the menu of a meal

As mentioned earlier, this scenario addresses a use case that applies to people in hospitals or specialized care institutions. Some hospitals, but mostly nursing homes, let patients or residents choose the menu for their meals. However, the care staff does not always have the time to ask each person

[3] http://www.microlife.com/
[4] http://www.withings.com/us/en/products/blood-pressure-monitor

Figure 6.3: Microlife automatic blood pressure monitor

what he or she wants to eat. This particularly problematic with people with dementia, as they may take a very long time to decide and are likely to be indisposed at the time when the caregiver comes. These issues would be easily addressed by using an ECA to have residents choose their meals, as it can take as much time as necessary and come back later if the person is not ready. The introduction of ECAs in institutions could therefore, among other things, allow to give people a choice they no longer have. While this could be an improvement of their quality of life, it may also motivate them to eat, as people with dementia tend to refuse to eat and older adults in general do not feel hunger as much as younger people.

In our test scenario, the user is asked to pick a starter among three options (ground carrots, tomato salad or boiled eggs with mayonnaise), a main course among four options (pork with lentils, spaghetti carbonara, salmon with spinach or steak with fries) and a desert between two options (fruit or yogurt). It is possible to refuse to take a starter or a desert but it is mandatory to choose a main course. The choice for each part of the meal is presented as described in Section 6.2. The dialog is structured as follows:

1. the character smiles and greets the user with a polite head nod;

2. the user has to greet back;

3. LOUISE introduces itself and says what the purpose of its visit is;
4. the user is asked to choose a starter;
5. the user is asked to choose a main course;
6. the user is asked to choose a desert;
7. LOUISE ends the interaction.

6.4 Experiment Protocol

The participants who tested the family of scenarios described in Section 6.3 were seated at a table in front of a screen on which the ECA was displayed; the Kinect sensor was placed on top of the screen, in the middle. All of the necessary objects for the experiment (a glass, a bottle of still water, a bottle of sparkling water and the pill dispenser or the blood pressure monitor, depending on the case) were placed on the table, as depicted on Figure 6.4. The software ran on a Toshiba laptop with an Intel i7 quad core CPU, an Intel HD Graphics GPU and 8GB of RAM; the Kinect sensor we used was the Kinect for Xbox 360; and the character was displayed on a DELL 22-inches display. Participants were informed of what was going to happen, that they had to greet the ECA back when it greets them and that they had to answer questions by “yes” or “no”. If the participant agreed, a video recording of the session was done. We used two cameras: one to film the person and one to film the table and the ECA. In addition, the LOUISE application was instrumented to record the input data and the evolution of the dialogs, including the states of the interaction manager FSM.

Participants had to interact with LOUISE through 3 scenarios in a row, always in the same order: reminder to drink water, meal menu choice and medication intake reminder or blood pressure measurement, depending on the participants, since we were not allowed to have hospitalized patients do the medication intake scenario. Before the interaction, they were asked to choose their favorite embodiment between the Louise and Charlotte characters. During the tests, the Microsoft Speech API speech recognizer was used, except for the last test, in which we tried a could-based speech recognizer (Microsoft Bing Speech API[5]).

Once the test, which took about 20 minutes, was over, we conducted individual semi-structured interviews with each participant and had them fill a questionnaire consisting in 5 Likert scale (1 to 4) questions regarding their impressions of the ECA during the test and 7 general questions about ECAs. To build the questionnaire, inspiration was taken from the assistive social agent technology acceptance model proposed by Heerink *et al.* [86],

[5] https://www.microsoft.com/cognitive-services/en-us/speech-api

Figure 6.4: Setup of the table for the experiment.

which is specifically designed for older adults. The Likert scale questions were about how pleasant they thought the character was, how much they liked its appearance, how clear the character's speech was, how clear the instructions it gave were and if the conversation's pace was too fast. The general questions were about the possibility of personalization, the applications of ECAs, how the character should address them (using their first name of their last name, calling them "*vous*" or "*tu*" – a very frequent issue in French), the type of display they would like the ECA to be visualized on, their acceptance of the perception system and the activation parameters (always on, activation times or user-activation only). The questionnaire is shown in Appendix C.

6.4.1 Participants

To be included in the study, participants had to be over 65, have a diagnosis of cognitive impairment and give written informed consent. People with severe auditory impairment, unable to speak or with severe visual disability were excluded. 14 participants, 11 females and 3 males, enrolled in the study. They were between 71 and 89 years old (mean = 78.8, standard deviation = 5.8) and their MMSE scores ranged from 8 to 30 out of 30 (mean = 23.8, standard deviation = 4.9). All were diagnosed with Mild Cognitive Impairment (MCI) [157] (9/14) in at least one cognitive domain (usually attention for participants who achieved a perfect score at the MMSE) or Alzheimer's disease (5/14) at the Broca hospital's memory consultation center by a multidisciplinary team composed of neuropsychologists, geriatricians, neurologists and psychiatrists, through comprehensive medical and neuropsychological assessment.

6.4.2 Data analysis

11 participants, out of 14, agreed to the video recording of the experiment. The videos, for a total of about 3.5 hours, were annotated using Noldus The Observer XT 11.5[6], a specialized software for event logging on videos. We counted the following events:

- valid answers;
- incorrect answers;
- no answer;
- answer attempts when the ECA is not listening;
- system not reacting to an answer;
- speech recognizer error;
- correct actions;
- incorrect actions;
- appropriate timing of the actions;
- attention estimation (AE) error;
- reaction of the attention estimator when participants performed an action at an inappropriate time;
- looking for help;
- intervention of the experimenter.

6.5 Results

Overall, 13/14 participants were capable of interacting with LOUISE, leaving out the participant with the most severe cognitive impairment (male, age = 73, MMSE = 8/30) who seemed to be intimidated by the character. Furthermore, 11/14 participants completed 3 scenarios each: in one test, the third scenario (taking pills) was interrupted early by the experimenter because the attention estimator was failing; and one of the participants could not do the blood pressure scenario because she could not move her only arm. 3 participants, hospitalized patients who were not allowed to do the pill dispenser scenario (because of the risk of choking on the candy), did the

[6] http://www.noldus.com/

blood pressure scenario; the other 9 did the pill dispenser scenario. In addition, only 3 participants went through the whole water reminder scenario (choosing between mineral and sparkling water and being guided through the process of serving themselves a glass), as most participants told the ECA that they were not thirsty. All participants with MCI (9/14) successfully accomplished the pill task. Only one of the participants with Alzheimer's disease (female, age = 83, MMSE = 26) successfully accomplished the much more complex blood pressure task without help.

6.5.1 Quantitative results and observations

Table 6.1 presents, for each evaluation indicator we have used, the minimum and maximum totals for one participant, the per participant average and standard deviation, the number of participants with whom the corresponding situation occurred and the total number of occurrences (or global percentage, when the values are percentages). This data was gathered thanks to video annotations and experiments' logs. It only includes data for 9 participants: only 10 of the participants who agreed to be filmed did interact with LOUISE; and the last test, in which the cloud-based speech recognizer was used, was actually conducted in WoZ mode (the system allowed the experimenter to press keys to notify answers in case of malfunction), since the speech recognition did not work because of a configuration error. Also note that the attention estimation data was not included for Participant 11, who could not be tracked by the Kinect sensor as she had only one arm that she could not move.

The first issue we observed is that some of the test participants tended to forget that they could only answer by "yes" or "no" and answered incorrectly, particularly the participants with the most severe memory impairment, scoring less than 20/30 at the MMSE, who account for 29 of the 56 incorrect answers (52%). This confirms the observations revealed in the anthropological analysis we conducted in the first round of experiments, reported in Section 5.2.6, that people with cognitive impairment tend to provide more elaborate answers and interact in a "social" way. Given this result, our error management strategy that lists the possible answers proved useful, as participants could correct their mistakes in all cases, especially after the system was changed for the valid answer reminder to be performed after two errors instead of three. Although the cloud-based speech recognizer did not work for Participant 14, considering that it was supposed to spot the "yes" and "no" keywords in the recognized sentence, all of her answers were considered as correct in the video annotation.

Regarding speech recognition, the performance of the ASR we have used are far from satisfying, with a word error rate of more than 20%. In addition, the system sometimes did not react to the participant's answers, when they did not speak loud enough (9.9% of all answers) or reacted to the ECA's

Table 6.1: Quantitative results for 9 participants (7 females, 2 males), aged between 72 and 89 (mean = 79, standard deviation = 6) with MMSE scores comprised between 14 and 30 (mean = 25.2, standard deviation = 4.3), obtained through video annotation and experiment logs. The “Nb. part.” column indicates the number of participants for whom the situation occurred.

Indicator	Min	Max	Mean ± Std. dev.	Nb. part.	Total
Questions	22	103	38.2 ± 15.0	9	344
Actions	0	13	4.8 ± 2.9	8	43
Participants					
Valid answers	21	55	28.7 ± 7.6	9	258
Invalid answers	0	20	6.2 ± 4.7	8	56
% invalid ans.	0%	28.2%	14.4 ± 8.6%		16.3%
No answer	0	28	3.3 ± 5.5	3	30
% no answer	0%	27.2%	3.6 ± 5.2%		8.7%
Answers when not listening	0	13	6.7 ± 3.8	8	60
Correct actions	0	8	4.1 ± 2.1	8	37
Incorrect actions	0	5	0.7 ± 1.0	2	6
% incorrect act.	0%	38.5%	7.3 ± 11.0%		14.0%
Wrong action timing	0	4	1.1 ± 1.3	4	10
System					
No reaction to answer	0	12	3.4 ± 3.3	6	31
% no reaction	0%	21.0%	8.2 ± 6.2%		9.9%
Speech reco. error	2	16	6.6 ± 3.2	9	59
Word err. rate	6.9%	33.3%	19.7 ± 6.2%		20.8%
AE errors	0	11	2.6 ± 2.3	6	23
AE unwanted reactions	0	3	0.5 ± 0.9	2	5
Requests for help	0	6	2.0 ± 1.3	7	18
Experimenter interventions	0	11	5.9 ± 5.2	8	23

own speech (the exact number of times could not be computed, since the logs are not reliable enough due to the amount of speech recognition errors). These three factors combined were quite confusing for the participants.

Regarding the attention estimator, it sometimes made errors that led to unwanted attention prompting (on 23 occasions). Most of these errors are due to losses of face tracking caused by participants being too close to the sensor. Indeed, they were seated close to the screen whereas the theoretical minimum distance for the Kinect sensor for Xbox 360 to work properly is about 80 cm and this distance is closer to 1 m in practice. This was particularly problematic when users came closer to the sensor to reach for objects on the table. We therefore mounted the Kinect slightly behind and above the screen. This positioning, as well as the hunched posture of some participants, caused the pitch angle measure to be incorrect and the attention estimator produced more errors than in the first round of experiments. Attention estimator errors also negatively impacted the user experience and caused confusion. In addition, attention management does not seem sufficient to keep people with dementia engaged in the interaction, since participants 10 and 14, who respectively scored 14 and 22 at the MMSE, sometimes did not answer the ECA although they were looking at it (Patient 10 did not answer on 28 occasions). However, when participants did get distracted, which happened once with Participant 2 and 3 times with Participant 10, the ECA could successfully recapture their attention.

Given the structure of dialog management for step-by-step task instructions presented in Section 6.2, 3 of the first 4 participants, who all had MCI, tended to perform the action while the video instruction was being shown (9/10 actions were performed too early) and the fourth one did not because she had asked the experimenter if she had to perform the action right away. This caused the attention estimator to react at an inappropriate moment on 5 occasions. The structure of interaction management for tasks was thus changed to allow the user to perform the task right away, as detailed in Section 6.6. This was successful, since the issue did not occur for the other participants with MCI. However, this only seemed suitable for one of the participants with Alzheimer's disease, who had the highest MMSE in the Alzheimer's group. Attention prompting was also triggered on 3 occasions by the last participant while she was performing actions of the blood pressure task, because she went past the time allowed (5 or 10 seconds, depending on the action).

On many occasions (60 in total), the participants tried to talk to the ECA while it was not listening to them. In particular, in the menu choice scenario, we observed that several participants wanted to choose a dish right after the presentation of the available options. As this was not allowed by our scenario, it caused them some frustration. In addition, the explicit confirmation after each choice annoyed some participants, who complained that the ECA was repeating itself a lot. Participants with MCI also tended

Table 6.2: Ratings of LOUISE in terms of pleasantness, appearance, clarity of speech, ease of following its instructions, and pace of the conversation on a scale from 1 (very negative opinion) to 4 (very positive opinion).

Question	Mean	Std Dev.	Min	Max	4	3	2	1
Pleasantness	3.38	0.43	2	4	7	4	2	0
Appearance	3.31	0.43	3	4	5	8	0	0
Clarity	3.54	0.50	3	4	7	6	0	0
Instructions	3.38	0.47	3	4	5	8	0	0
Pace	3.46	0.66	1	4	8	4	0	1

to anticipate their answer to the confirmation question or after two speech recognition errors, when the ECA kept the speech turn for a longer time than usual to remind them of the valid answers. Another issue we encountered is that 2 participants with Alzheimer's disease tended to state that they had correctly performed an action when they had not, on a total of 8 occasions (this includes Participant 14, who is not counted in Table 6.1). In addition, the "welcome back" transition sentence after a loss of attention turned out to be confusing for several participants. This was also due to the phrase we used ("*vous revoilà*", which translates "you are back"), which did not seem very natural.

Lastly, most of the requests for help and interventions of the experimenter were caused by system malfunctions (attention estimator errors or speech recognition error). The few exceptions had to do with participants formulating incorrect answers but not getting reminded of the correct answers soon enough (this is why we changed the system to perform the valid answers reminder after 2 errors instead of 3), helping participants with the lowest MMSE perform a difficult action (putting on the blood pressure monitor's armband), so they could go on to complete the task, and to remind people that they had to answer the ECA when they were not responding.

6.5.2 Feedback

The feedbacks in the questionnaires were mostly positive. The results to the Likert questions in the questionnaire are given in Table 6.2. This includes answers from 13 participants out of 14, as the one who did not interact with LOUISE was not asked to fill in the questionnaire.

Regarding pleasantness, most participants had a positive or very positive impression of LOUISE and only two participants out of 13 had a negative opinion. Regarding the appearance, all participants expressed a positive or very positive opinion. In addition, all participants thought LOUISE's speech was clear or very clear, and that the instructions were easy or very easy to follow. Lastly, regarding the pace of the conversation, only one participant, with a low MMSE score (17/30), said that the conversation was too fast; all others said it was not too fast or not fast at all. However, several participants (5, all with MCI) complained that the system was too slow. Overall, the feedbacks from the participants are very positive, which is encouraging, as people belonging to this age group are usually considered to be reluctant to use new technologies – this suggests that our participants may perceive LOUISE as being a medium with inherent user-friendliness. Some results, however, regarding the ease of following the ECA's task instructions are surprising, as they do not match our observations: two participants obviously had some difficulties following the instructions but said it was easy. These participants had low MMSE scores (14/30 and 22/30). This may be due to the fact that they did not realize that what they were doing was incorrect or did not want to admit it. It could also be linked to the white-blouse effect.

The feedback given in the semi-structured interviews were not as positive. As already mentioned, 5 participants complained that the system was too slow to speak and to respond to their answers. One of the participants said it was because the ECA's utterances were too long. Below are some translations of verbatims on this matter:

`Participant 1: ‘‘She should ask the questions faster.’’`

`Participant 2: ‘‘It is a bit slow... I put myself in the shoes of people who want to pick their menu and they will likely want it to go faster.’’`

`Participant 6: ‘‘It is this waiting time [that is too long].’’`

`Participant 7: ‘‘The fluidity should be improved.’’`

`Participant 8: ‘‘If it is for a person in full awareness, it may be a bit slow.’’`

Participants were not sure when the time was appropriate to give their answers, which partly explained the many times when they attempted to answer whereas the system was not ready to listen to them. 2 participants complained about it:

`Participant 1: ‘‘We should be able to answer as if we had the person directly in front of us.’’`

`Participant 7: ‘‘Should not it be explained that one has to answer when Louise, in this case, puts her hand to her ear?’’`

A participant made a comment about her expectations of the ECA's capabilities being too high because of its appearance:

```
Participant 1: ''I have the impression that it's not virtual.
I have the impression that I have a person in front of me
who will answer.''
```

This may partly explain why several participants tended to formulate invalid answers, thus not adapting to the fact that they are talking to a computer, and tried to answer when the ECA was not listening.

Only 2 participants emitted negative comments about the system's malfunctions:

```
Participant 2: ''It can annoy people when it is not working.''

Participant 3: ''For someone who cannot take initiatives
anymore, I forsee some difficulties''.
```

Regarding the system's usefulness, only 1 participants formulated comments spontaneously. She did not think it would be useful for her:

```
Participant 3: ''I do not see the interest for someone like me,
who still have their head on their shoulders.''
```

However, she mitigated her answer by saying that it could be useful it it were meant for cognitive stimulation exercises:

```
Participant 3: ''If it were requiring attention, concentration,
reponsiveness, it would probably seem more effective.''
```

2 other participants said it could be useful in some situations:

```
Participant 4: ''My very old parents were not capable to
prepare their pill dispenser anymore. [...] I think that
for pill dispensers... there is room for improvement! You
are on to something.''

Participant 8: ''I think that for people who lost their memory,
it is very interesting.''
```

Lastly, participants 2 and 8 spontaneously formulated positive comments about the system, regarding the presentation of the task instructions and the system's ease of use:

```
Participant 2: ''Yes, I think it is good to have [both visual
and auditory instructions].''

Participant 8: ''Well, it's not too complicated.''

Participant 8: ''It's clear, it's easy.''
```

6.5.3 General questions about ECAs

As mentioned above, in the questionnaires filled after the tests, we also asked more general questions about ECAs. The results are shown in Table 6.3.

Regarding the possibility of personalizing the character's appearance, more than half of the participants (7/13) said they would like to be able to do that. This is a smaller proportion than in the previous phases of the study.

Regarding the applications of ECAs they would be interested in, 6 participants selected the virtual assistant (providing information and doing reminders); 6 were interested in the coaching application; 7 said they would like step-by-step task instructions; 7 thought that it would be useful for controlling home appliances; and 2 said it was not useful for anything. Compared to the previous phase, reported in Section 5.3.1, there are more mitigated results, as there is no application that clearly got more votes than the others. This may also be due to the size of the sample.

Regarding the type of display, similar results as the ones observed in the previous phase were obtained: only one participant wanted a dedicated display for the ECA; 7 said they would like it to be displayed on a television set, which is a larger proportion than in the first study; 7 participants selected the tablet; only three selected the smartphone; and one said "nowhere, I do not want to use this technology". Again, most people would like an assistive ECA to be added on a device that they already own or is not only dedicated for this use.

Regarding how the ECA should address people, 9 participants said it should use their first name; 3 would prefer it to call them Mr or Mrs X; and one participant did not have an opinion. For the very frequent issue in French language regarding the use of "*tu*" (informal) or "*vous*" (formal), each possibility was selected by 6 participants and one did not know. This means that these options should be configuration parameters, as this kind of preferences depend on each user and may influence their adherence to the system. For instance, people who wants the ECA to address them in a very polite manner may reject the ECA, thinking it is disrespectful, if it calls them by their first name and uses the "*tu*" pronoun. In addition, it might change over time, as this also depends on the kind of relationship that exists between people and their ECA. For information, the guidelines for positive treatment of people with dementia impose the use of "*vous*", without exceptions. This should therefore be the default setting. However, we do not see why the ECA could not call its user"*tu*" if that is what he or she wants.

Lastly, regarding the acceptability of the perception system, few people (5) said they would want it to be installed in their homes, and only 3 would want it to be always on and one person did not know. By comparison, 7 people would not accept it and 8 said it should only be activated when

Table 6.3: Full results of the questionnaire about ECAs in the usability study. 13 participants answered.

Topic	Type	Answer	Total (%)
Embodiment personalization	single answer	Yes	7 (54%)
Applications	multiple answers	Task helper	7 (54%)
		Butler	7 (54%)
		Assistant	6 (46%)
		Coach	6 (46%)
		Nothing	2 (15%)
		Other	0
Addressing the user	single answer	*Tu*	9 (69%)
		Vous	3 (23%)
		Do not know	1 (8%)
Naming the user	single answer	First name	6 (46%)
		Last name	6 (46%)
		Do not know	1 (8%)
Display	multiple answers	TV	7 (54%)
		Tablet	7 (54%)
		Computer	6 (46%)
		Phone	3 (23%)
		Dedicated screen	1 (8%)
		Nowhere	1 (8%)
Camera acceptance	single answer	No	7 (54%)
		Yes	5 (38%)
		Do not know	1 (8%)
Activation	single answer	By user only	8 (61%)
		Always on	3 (23%)
		Programmed	1 (8%)
		Do not know	1 (8%)

they decide it. It is however difficult to say if they fully understood the implications of the question, especially the ones with the strongest cognitive impairment.

6.6 Discussion

Performing LOUISE usability testing allowed us to debunk several issues and sometimes fix them, or at least get ideas for solutions. In addition, the fact that we allowed some changes between tests to be performed enabled us to iterate quickly and evaluate our solutions right away.

Globally, the fact that all but one participants were able to interact with the ECA and that inability to do so was more linked to system malfunctions or limitations (i.e. the speech recognizer only allowing to answer by "yes" or "no") than to the person's capabilities suggest that LOUISE, after some improvements, could be suitable for people with MCI or moderately severe dementia (down to 10/30 at the MMSE) who are still capable of speaking, provided that they do not have severe hearing impairment. However, it will probably not be suitable for guiding through a task, at least, not in its current state, that is to say without activity recognition capabilities. It would likely be the most useful as a cognitive prosthesis for people with dementia, particularly to address memory loss by performing reminders (date, time, place where they are, appointments, medication intakes, etc.). However, it could also find useful applications for people with MCI, such as cognitive stimulation exercises, eased access to video chat to remain socially active or entertainment. According to our participants in the focus group, finding such applications for people with MCI would be interesting, as it would be a way for them to start using the system and get used to it before it becomes truly useful, as they progressively lose their cognitive abilities.

The most limiting factor is the quality of automatic speech-to-text transcription, especially for elderly voice [232]. In fact, Aman *et al.* [8] have demonstrated that performance in automatic speech recognition dropped significantly when users were disabled older adults, which are the ones who need assistive technologies the most. In addition, the ECA tended to react to its own speech, which is a problem we put a lot of effort in trying to fix but could not eliminate completely. This last issue proved to be a difficult point as the speech recognizer we used, the Microsoft Speech API, did not allow to turn it on and off as desired. We thus used our speech turn management protocol to make the system more robust, by ignoring speech recognition results outside of user speech turns, but, because of delays in the system, the problem still occurred in some cases. In addition, this caused some of the participants' answers to be ignored by the system. Another reason why this flaw is difficult to fix is that, if the delay between the end of the ECA's speech turns and the moments when it starts reacting to speech recognition data is too long, we may miss the user's answer. These malfunctions are very damaging to the quality of experience and also causes misunderstanding and confusion for our users with cognitive impairment. However, the technologies on which ECAs are built have made progress in the past decade and keep improving fast. We thus believe that these technical limitations will soon be overcome. In our system, we have implemented a new speech recognition module at the end of the testing phase, using the recently launched Microsoft Bing cloud-based speech recognition API. It seems more reliable than the one performed locally and it handles word-spotting. In addition, it can be turned on and off as desired, which also fixes the problem of undesirable recognition of the ECA's own speech. Unfortunately, we have not evaluated it with patients yet. Using this new tool could change our error management strategy, since it may enable the system to distinguish between recognition errors, which should make the ECA ask the user to repeat his or her answer, and incoherent answers, which should trigger a context reminder utterance. Note that initially, it had one more step, that allowed for a second error before the ECA does a context reminder and lists the possible answers. But this was too long and caused participants frustration, instead of quickly telling them which answers the system expects. Hence, we decided to remove it to produce the design presented in Section 5.4.3.

Related issues are the facts that people were not sure when to answer and that their expectations of the ECA's understanding capabilities were too high. As a result, they sometimes tried to formulate answers when the system was not listening. They also were not always sure if the system had heard their answer when it took some time to react and repeated their answer. It therefore appears that the listening pose (with hand on ear) is not sufficient and the system's feedback notification capabilities should be augmented with other visual cues to indicate more clearly when the person should speak, when their answer has been heard and when it is being processed. This would likely diminish frustration. In addition, the character's appearance should be changed for it to look like a humanoid robot, as proposed by older adults in our focus group (see Section 5.3.2); this would likely lower people's expectations and be a constant reminder that they are talking to a machine and should adapt their speech to that context. Furthermore, reminding people of the valid answers when the ECA does not understand them proved useful during our tests. However, it originally took too long for the ECA to perform that reminder, as it waited for the third speech recognition error, and reducing that number to 2 errors improved this feature's efficacy.

Regarding the Kinect's placement and functioning range, the best solution would be to use another sensor, which works at a closer range. The Kinect sensor for Xbox One[7] could be a good solution. Interestingly, its API

[7] http://www.xbox.com/en-US/xbox-one/accessories/kinect

contains attention estimation functionalities, which deserve to be compared to our own method. It also allows lip-reading and mouth movements detection, which could be useful to improve user speech turn detection. However, using this device is very costly in terms of computing power. A cheaper solution may be to use a simple camera with a powerful and well-optimized face-tracker such as OpenFace[8].

The system's attention management proved useful in a few cases and is likely to be even more useful in a less controlled environment. However, in some cases, people with Alzheimer's disease did not answer the ECA's questions, although they were looking at it. This suggests that a strategy to handle this kind of situation should be proposed. For instance, the ECA could prompt the person and call him or her by his or her name (first or last, depending on their preferences) to make clear that it is addressing him or her.

To address the issue linked to task instructions, we re-arranged the scenarios for the dialog to be structured slightly differently. Instead of explaining and showing what action should be performed first and then asking the person to perform the task, we let the user do the task right after it is explained. We then ask if it is completed, and only explicitly ask the person to act if the answer is negative. This also changes the policy of attention estimation deactivation and gaze direction. The resulting dialog structure is shown in the following example:

```
(attention estimator off)
ECA: ‘‘To take the pill in the box, you have to place your
hand on the hole and turn the box upside-down.’’
(ten seconds delay, video looping, ECA gazing at the TV)
(attention estimator on, ECA gazing at the user)
ECA: ‘‘Have you taken the pill in the box?’’
User: ‘‘No.’’
(attention estimator off, ECA gazing at the user)
ECA: ‘‘Go on, take the pill in the box.’’
(ten seconds delay, video looping, ECA gazing at the user)
(attention estimator on, ECA gazing at the user)
ECA: ‘‘Have you taken the pill in the box?’’
User: ‘‘Yes.’’
ECA: ‘‘Well done!’’.
```

Note that, if the answer is still "no" after the prompt to perform the action, the ECA does not go back to the instructions, but keeps repeating the prompt and asking if the action has been completed. This strategy worked better with the participants with MCI, but the participants with dementia struggled when they had to look at the example and perform the actions

[8] https://github.com/TadasBaltrusaitis/OpenFace

at the same time. This is due to their reduced attentional capabilities, as it requires switching attentional focus rapidly and several times. As a result, they performed incorrect actions and thought they were correct. This suggests that there should be a different mode for task guidance that depends on the person's level of cognitive impairment. We think that the first way of managing step-by-step task instructions we have described in Section 5.4.3 would be well adapted to the patients with severe cognitive impairment, but it should be more explicit that they have to concentrate on the instruction first and wait for the ECA to tell them to perform the action. In addition, the instruction sentences have to be carefully written, not to make people think that they have to perform the action right away. Taking advice from linguistic anthropology experts could be very helpful in that department.

Following the steps of a task is more difficult for people with dementia than for people with MCI and, in our experiment, this situation was worsened by the fact that the blood-pressure measurement task is more complex than the pill dispenser task. However, this allowed us to observe that guiding patients through the task step-by-step and asking them for confirmation at each step is not sufficient for them to perform it autonomously, as soon as it is a bit complex. In that regard, some very encouraging results were obtained by Hoey *et al.* [89], who worked on automated task assistance for patients with dementia using activity recognition. Given the results of their work and our own observations, we think that having people with dementia do things they can no longer do alone, thanks to an assistive ECA system, seems achievable, at least in some cases, but it would require activity recognition and/or feedback from the objects involved in the task. This would also allow to compute a confidence indicator for task completion, which could be harnessed by a caregiver to help only when necessary, as also suggested by Hoey *et al.* [89]. Furthermore, training users to use the ECA could increase the success rate.

Regarding item selection in a list, the way the interaction is managed by the ECA worked quite well, since all 13 participants who did interact with the ECA were able to complete the scenario. It could be improved by adding a step when people can say what they want, if they know, before reviewing the items one by one. To introduce more flexibility in dialog management, some inspiration could be taken from the work of Yaghoubzadeh *et al.* [230], who have obtained some good results in using a conversation manager that allows for mixed-initiative interaction in a scheduling assistant ECA, though their system was only tested with a few older adults without cognitive impairment. However, it would be difficult to let the person interrupt the ECA to say what he or she wants. This remains a challenging issue to address because, as we have mentioned above, the ECA will react to its own speech. It may be possible to do it by using high-end directional microphones or head-mounted ones, but the former is too expensive and the later is not practical, especially for older adults with cognitive impairment, and possibly reduced mobility. Another solution could consist in using advanced speaker identification techniques. A simple improvement, however, would be to remove the confirmation step and confirm all chosen items at the end of the dialog, checking them one by one, as recommended in [229], to avoid annoying repetitions.

Lastly, after checking with two psychologists, we decided to remove the "welcome back" transition phrase because they said that it sounded very unnatural and it was redundant with the recontextualization phrase. In addition, when the inattention time was too short, the ECA spoke that sentence directly, without doing the prompting, which was confusing for our test participants.

Chapter 7

Virtual Promenade, virtual reality exposure therapy for post-fall syndrome treatment in older adults

Résumé en français. *Les jeux vidéo modernes comportent de nombreux humains, humanoïdes ou personnages anthropomorphes qui jouent des rôles variés dans le scénario : alliés, ennemies, adversaires et, le plus important, avatars. Dans les premiers cas, les humains virtuels sont utilisés comme non-soi, parfois appelés personnages non-jouables. Dans le cas des avatars, les humains virtuels sont utilisés comme soi, c'est-à-dire qu'ils sont le moyen par lequel le joueur peut accéder au monde virtuel et interagir avec les objets et personnages qui le peuplent. Dans les chapitres 5 et 6, nous avons vu avec LOUISE une étude de cas des humains virtuels comme non-soi, dans laquelle ils sont utilisés pour représenter un agent conversationnel. Dans ce chapitre, nous explorons l'utilisation des humains virtuels comme soi, sous la forme d'avatars, pour mettre les personnes âgées atteintes d'un syndrome post-chute (SPC) dans une situation de marche, qui leur cause de l'anxiété, grâce à la réalité virtuelle. Nous discutons brièvement la possibilité de combiner les deux approches, soi et non-soi, des humains virtuels dans la section 7.7.3.*

A cet effet, nous avons développé un système, appelé Virtual Promenade, qui combine un jeu vidéo de marche virtuelle et un fauteuil haptique dont l'assise bouge en reproduisant le mouvement des hanches de la marche humaine. Ce système est prévu pour être déployé dans des environnements de soin (hôpitaux et centres de rééducation) et utilisé par des professionnels de santé avec leurs patients. Un étude préliminaire a été menée pour s'assurer de la faisabilité de l'utilisation d'un jeu vidéo et du fauteuil haptique par des patients hospitalisés dans un service de rééducation et que le siège mobile ne cause pas de douleurs, ni d'autres effets secondaires indésirables, tels que

la nausée, aux patients [124, 218]. Aucun obstacle majeur à la faisabilité, ni aucune source d'inquiétude en termes de sécurité ou d'effets indésirables n'ont été trouvés.

*Dans ce chapitre, nous commençons par présenter les buts de cette recherche et l'approche de conception adoptée pour cette étude de cas. Nous présentons ensuite le système que nous proposons dans la section 7.2. Puis, nous rapportons le processus de conception participative, incluant des tests utilisateur (section 7.3), des groupes de discussion (*focus group*) avec des kinésithérapeutes et des psychomotriciens (section 7.4), et le suivi de professionnels sur une journée de travail ordinaire (section 7.5). Nous poursuivons cette présentation de la phase de conception par une discussion, dans la section 7.6. En dernier lieu, nous présentons une évaluation du prototype en conditions écologiques avec des patients de l'hôpital Broca, dans la section 7.7.*

Le SPC ressemblant, sur plusieurs aspects, au syndrome de stress post-traumatique (SSPT) [43, 27, 96], nous proposons de le traiter par la réalité virtuelle, cette technique ayant fait ses preuves dans le cadre des SSPT. Le but principal de cette étude de cas est de concevoir et évaluer le système évoqué plus haut, en étroite collaboration avec plusieurs parties prenantes, en particulier un médecin, les patients, les psychomotriciens et les kinésithérapeutes. Les objectifs secondaires sont d'identifier les aspects de la conception qui sont importants aux yeux des personnes âgées afin qu'elles trouvent le dispositif plaisant, de mieux comprendre comme elles abordent le fait d'incarner un avatar, de déterminer le mode de commande de l'avatar le plus adapté et d'identifier d'éventuels obstacles de faisabilité ou d'acceptabilité.

Nous avons commencé par assembler le dispositif et créer une ébauche rudimentaire de jeu, dans lequel le joueur contrôle un avatar dans un environnement urbain. Nous avons ensuite co-conçu le jeu avec des personnes âgées en bonne santé en alternant des séances de test informelles et des phases d'implémentation. Cela nous a conduit à ajouter deux environnements virtuels (forêt et parc automnal), jugés très plaisants par les personnes âgées, qui n'ont que très peu apprécié l'environnement urbain initialement proposé. Nous avons également ajouté plusieurs modèles de personnages pour servir d'avatar et créé des didacticiels expliquant comment jouer. Enfin, nous avons expérimenté plusieurs manettes de jeu et notre choix s'est arrêté sur une manette de Nintendo 64.

Après avoir validé cette première version du jeu lors d'une séance de test avec 9 personnes âgées et fait quelques ajustements, nous l'avons présentée aux kinésithérapeutes de l'hôpital Broca, ainsi qu'aux psychomotriciens. Nous avons également passé une journée dans le service de rééducation pour se faire une meilleure idée de la réalité du terrain. Nous avons pu, grâce à leurs retours et à nos observations, faire de nouveaux ajustements et mieux comprendre comment le dispositif pourrait s'intégrer dans leurs pratiques de soin et quelles sont les contraintes du terrain.

Enfin, nous avons mené une évaluation du dispositif avec 8 patients hospitalisés dans le service de rééducation de l'hôpital Broca, qui avaient tous des troubles cognitifs légers à modérés et plus de 75 ans. Nous avons observé que les commandes du jeu étaient peu adaptées à ce public, contrairement aux personnes âgées en bonne santé. Nous avons donc modifié le dispositif au cours de la phase d'évaluation, principalement. Principalement en changeant de manette (notre choix s'est porté sur un joystick type simulateur de vol qui nécessite peu de force pour l'actionner) et en créant un mode "facile", dans lequel le contrôle de l'avatar est simplifié : le joueur ne contrôle que la vitesse de déplacement et l'avatar se dirige tout seul dans l'environnement. Avec ces modifications, tous les sujets ont pu utiliser le dispositif avec succès et ont rapporté un haut niveau de satisfaction en termes de facilité d'utilisation et d'appréciation de l'activité. En revanche, le peu de stress observé chez les participants pendant les séances suggère que le jeu n'est pas assez immersif et/ou ne propose pas de situations suffisamment stressantes.

MODERN VIDEO GAMES include many human, humanoid or anthropomorphous characters that play various roles in the storyline: allies, enemies, competitors and, most important, avatars. In the former cases, virtual humans are used as non-self, sometimes called non-playable characters (NPCs). In the case of avatars, virtual humans are used as self, that is to say they are the means by which the player can access the virtual world and interact with the objects and other characters that populate it. In Chapters 5 and 6, we have seen with LOUISE a case study of virtual humans as non-self, in which they are used to represent a conversational agent. In this chapter, we explore the use of virtual humans as self, in the form of avatars, to put older adults with post-fall syndrome (PFS) in situations of walking, which causes them anxiety, through virtual reality. We briefly discuss the intriguing possibility of merging self and non-self virtual humans in Section 7.7.3.

To this aim, we developed a system, called Virtual Promenade, which includes a virtual strolling video game and a haptic chair with a moving seat that reproduces the movements of the hips in human walk. This system is meant to be deployed in care environments (hospitals or rehabilitation centers) and used by health-care professionals with patients. A preliminary study had been conducted to check for the feasibility of using a video game and the haptic chair with patients hospitalized in a rehabilitation service and that the moving seat system does not cause pain or unwanted side effects, such as nausea, to patients [124, 218]. No major obstacle to feasibility nor source of concern about safety or undesirable effects were found.

In this chapter, we start by presenting the research goals and design approach for this case study. We then present the system we propose in Section 7.2. After that, we report the living lab participatory design process of the system, which included user testing (Section 7.3), focus groups with physiotherapists and psychomotricians (Section 7.4), and following professionals on a typical day at work (Section 7.5). We end this reporting of the conception phase with a discussion in Section 7.6. Lastly, we present an evaluation of the prototype in ecological conditions with patients of the Broca hospital in Section 7.7.

7.1 Goals and design approach

The fear of falling associated to PFS presents similarities with post-traumatic stress disorder (PTSD), according to the definition given in the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM V) [9]: PTSD occurs when one is the victim of a traumatic event and his or her response includes intense fear, helplessness or feelings of horror. These symptoms are also observed in the psychological consequences of falls, which suggests that they can be seen as traumatic events [43, 27, 96].

As the psychological symptoms of PFS can be compared to PTSD, a worthwhile approach would be to explore the potential of Virtual Reality Exposure (VRE) treatment, since it has been shown to be a useful therapeutic tool to improve motor rehabilitation in combat-related PTSD [28]. VRE has also been shown to yield interesting therapeutic results for the motor rehabilitation of patients who had a stroke [112] or present walking disorders caused by chronic diseases, such as Parkinson's disease or multiple sclerosis [90, 15]. Several experiments have also been conducted with VRE to treat various psychological disorders, such as acrophobia, arachnophobia, aviophobia, claustrophobia or social phobias [76]. This is why we propose to use a 3D virtual strolling serious game, combined with a haptic chair, to increase immersion, to help treat PFS in older adults.

The main goal of this research is to create a tool to provide virtual reality therapy for post-fall syndrome (PFS) in older adults that is adapted for use by the target public and to the constraints of the rigid care environments. This system is meant to be used by care professionals such as physiotherapists or psychomotricians and has to comply with the following requirements:

- the activity has to be performed while seating, so that patients who cannot stand up, for physical or psychological reasons, are able to participate;
- the activity should be adapted for use by most patients in geriatric rehabilitation care, who potentially have other pathologies than PFS, such as frailty and/or cognitive impairment;
- the system should be safe, easy to deploy, quick to operate and require little or no training for professionals to make their patients use it;

- the technology has to be accepted by both patients and professionals.

As a secondary goal, we would like to identify what design aspects are important to older adults, so they find the activity pleasant and stay motivated to engage in it regularly, for an extended time period. In addition, we would like to obtain information about how older adults apprehend the activity, more specifically how they relate to their avatar in the game. Furthermore, we are interested in investigating the suitability of game controllers for our target public and determine what the most adapted way to control an in-game avatar is. Lastly, we would like to assess the feasibility of deploying such a system in a hospital environment, through our own experience, and the acceptance of this technology for patients and professionals, through testing and consultation.

To tackle these challenges, as for LOUISE, we adopted a living lab participatory design process to build the system. For this project, our methodology resembles the design-with-users approach [21], design-with-stakeholders in our case, as we jointly undertook the conception of the system with a physician and closely involved healthy older adults, as proxies to our end users, end users themselves (patients of the Broca hospital in geriatric rehabilitation care), physiotherapists and psychomotricians. Compared to the LOUISE project, reported in Chapter 5, the system is less complex, which allows for faster iterations on the prototype to work in close collaboration with stakeholders and better reflect their recommendations and requests. In this case study, as in the LOUISE project, we focus on the conception and development phases of the living lab product life cycle (see Section 3.2.2).

To avoid unnecessary work and reduce costs (which is essential for future deployment), we used commercially available elements, so we could focus solely on the game and movement sensory feedback design. More specifically, we have repurposed an existing haptic chair and experimented with game controllers available on the market. To foster easy implementation, yielding fast prototyping and short iteration cycles, we chose to use the Unity game engine[1]. This tool combines the advantages of being quite comprehensive, user-friendly, cost-effective and of having multi-platform release capabilities. In addition, we have used ready-for-use graphical and animation materials, as well as free sounds.

7.2 System description

The system we propose, called Virtual Promenade, is composed of a game controller, a virtual strolling game, a haptic chair, a programmable power source (connected to the motor, to control its input voltage) and a power controller program that receives commands from the game through TCP

[1] http://unity3d.com/

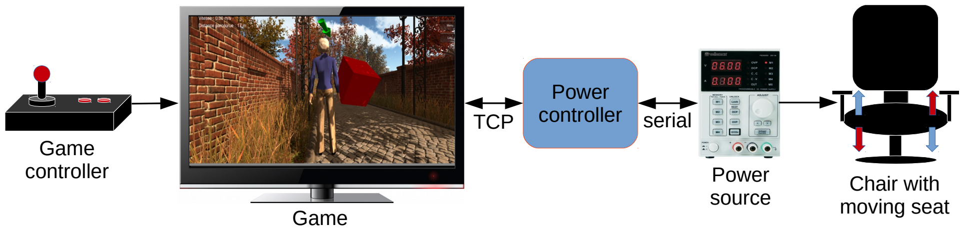


Figure 7.1: The *Virtual Promenade* system.

Figure 7.2: Picture of the *Virtual Promenade* system setup.

and sends commands to the power source through a serial-over-USB protocol. This setup is represented in Figure 7.1 and a picture of the system is presented in Figure 7.2.

7.2.1 VP game

Our initial design, before we improved it and added contents through participatory design increments, called for a virtual strolling game set in a virtual outdoor environment, a model of a modern city, in which the player controlled a virtual character to stroll freely in a third-person view mode. We chose a city environment for three reasons: cities are particularly stressful environments; the users asked for an outdoor environment in the preliminary study; and, according to the experts, the fear of falling mostly arises when older adults go outside, whereas they usually feel safe at home. A screen capture of the environment is presented in Figure 7.3. This approach, which consists in building a minimal prototype and starting to co-design the game

Figure 7.3: Screen capture of the early version of the virtual strolling game.

with end users right away is somewhat similar to the one used by Uzor *et al.* [204] (see Section 2.2.3 for more detail).

Throughout the design, our character models were created using Autodesk Character Generator, as for the embodiments in the automated prototype of LOUISE (see Chapter 5). The walk animations and the environment models were downloaded for free or purchased on the Unity Asset Store[2], a community-based content sharing platform provided by Unity. Lastly, the sound effects were downloaded on free sound effect exchange platforms.

7.2.2 Backwell moving seat

The moving seat system was created by the Backwell company[3]. It was intended to prevent lumbago back pain in workers who sit at their desk all day, by mobilizing the upper-body muscles through the movement of the seat, activated periodically, say 5 minutes every half-hour. The chair has a mechanical system moved by a single motor to mimic the movement of the hips during walking. The motor and actuators are placed under the chair's seat. An animation can be seen on the manufacturer's website[4]. We chose this chair because it is an affordable and easy-to-repurpose solution, as speed can easily be controlled by varying the motor's input voltage.

Combining this chair with our virtual strolling program has two interests: firstly, it may enhance the virtual strolling experience by adding proprioceptive sensations to the visual and auditory sensations of the game; secondly, the walk-like movement stimulates the motor scheme (brain activities involved in motor control). To connect the game and the chair, the motor's

[2] http://www.assetstore.unity3d.com/
[3] http://back-well.com
[4] See footnote 3

Figure 7.4: Pictures of the game controllers we experimented with.

input voltage is regulated thanks to a Velleman PS3005D programmable laboratory power source. The goal is to synchronize the speed of the chair's seat movement with the player's avatar in the virtual strolling game. This is achieved through TCP client-server communication between the game and the power control program, which, in turn, sends voltage control commands to the power source through serial-over-USB communication.

7.2.3 Game controllers

We deliberately chose to use a classic game controller for human-machine interaction rather than an interaction device that would require the user to perform physical tasks. Our aim was to find a game controller that is easy to use by older adults, who are usually not familiar with this type of device. We thus experimented with several, commercially available game controllers:

- a Logitech flight simulator-like joystick;
- a Mayflash arcade game controller;
- a Sony Playstation controller;
- a Nintendo 64 controller.

In the preliminary feasibility study [124, 218], a Thrustmaster USB Joystick (a flight simulator-like joystick) was used and it was observed that it yielded good usability for a simple task, where the user could only go forward and backward in a corridor. However, this model was too light and made for right-handed people only. Therefore, we first experimented with a Logitech flight simulator controller that is heavier than the Thrustmaster one and made for use by both left-handed and right-handed people. The controllers we experimented with are shown on Figure 7.4.

7.3 Playtesting

We conducted playtesting sessions with 7 healthy older adults, as a proxy to our target users, and one patient who was hospitalized after a fall. The participants were recruited on a voluntary basis. All participants were women over 80 years old. Some are "expert users" who frequently participate in research activities of our living lab. Among the 8 women who participated in the trials, one did not want to try the system herself and only gave her opinion based on observing two other participants play.

We had three formal trial sessions and a few informal trials with one or several participants at a time, after they had participated in another activity at the living lab. This design phase led us to add several environments and character models, as well as a tutorial level, to the game. In the last trial session, participants had to fill a satisfaction questionnaire with 15 questions in blocks of 5. The first block was about the tutorial and the ease of use of the controls; the second and third blocks were about the environments and the feeling of immersion.

After having implemented a first version of the game thanks to the feedbacks, we validated the design with a group of 9 healthy older adults, 7 women and 2 men, during a *Café Multimédia*[5] activity [228]. Participants were between 67 and 91 years old, about 75 on average, and two of them had MCI.

In the remainder of this section, we summarize our observations and highlight the design choices we have made in this first design phase. Then, we discuss the final design of the game. Lastly, we report on the validation.

7.3.1 Virtual environments, avatars and view

As mentioned earlier in Section 7.2.1, we started off with only one environment, a model of a Japanese city, as a city environment was recommended by the experts and we had noted in the feasibility study that people asked for an outdoor environment. Only one of the 4 participants that tried the city environment liked it. The others said that it felt cold and unwelcoming. The participant who liked this environment said it was only because, as an urban resident, she could relate to this environment. We thus added two other environments: an autumn forest and a park full of trees and flowers. All of the 5 participants who tried the game with these environments were very satisfied with them and said they were realistic enough.

Regarding the avatars, we only had one character model at first, a thin lady with gray short hair. Participants in the first trial complained that they did not identify with this character because its body shape looked nothing like a real elderly person. We thus added 7 new character models. Our models include 5 female characters and 3 male. We made them so that

[5] See Section 3.2.3

their body shape was more realistic, according to what our participants recommended. All subjects seemed satisfied with the choice of character models. One person even said she identified well with the character she chose.

Lastly, we experimented with two views: first-person and third-person. In the first trial, we only had the third-person view. The subjects complained that the camera was too close to the character, which made it harder for them to see where they were going. We changed the camera's position to be higher and further from the character. We did not have other complaints in the other trials. We also added a first-person view. We observed that people performed better in controlling their avatar in first-person view, but 4 out of 5 subjects said that they preferred the third-person view because it was more reassuring. We think that keeping both options is interesting because the third-person view is reassuring and the first-person view yields better immersion, and may elicit increased anxiety, which could be useful in the later stages of VRE therapy.

7.3.2 Game controllers

The Logitech flight-simulator joystick did not turn out to be adapted in the first trial. The subjects said it was too hard to maneuver, and it required a table to put it on, which was not practical. Indeed, the addition of two control axes and the use of a heavier joystick (for more stability) made the use of this type of controller more difficult. In addition, in the preliminary phase, the game and the haptic chair were mostly tested separately, which prevented from identifying the issue regarding joystick placement.

The Sony Playstation game controller was too difficult to use as it required fine motor control of the thumbs and older adults are not used at all to that kind of device. Walking in a strait line turned out to be particularly difficult for our participants.

The arcade controller was quite promising and had two main advantages: it can be put on the user's lap and it is big enough so the user can rest his or her arm on it. In addition it has big buttons that users successfully used to navigate the instructions in the tutorial. However, it has a major flaw: the directional stick is made with all-or-nothing switches, which doesn't allow the player to control the speed of the avatar.

Lastly, the Nintendo 64 controller was used successfully. It is lightweight and practical, as the player can hold the handle in the middle with one hand and use the analog stick with the other hand (which the participants did), by grabbing the stick between the thumb and the index finger. The analog stick also has a guide that helps going straight. However, this device is obsolete and can have the same drawbacks as the Sony Playstation controller, though it is a bit easier to use, as it still requires fine motor control.

This issue is very important. One of the subjects even reported spontaneously that controlling the character being too difficult for her hindered her experience as she could not concentrate on the walking sensation. Facing this issue, a good practice is to support several controllers for adaptability to as many patients as possible.

7.3.3 Haptic chair and immersion

The feedback on the immersion was very positive: only one participant out of eight said she did not think the haptic chair enhanced the experience. All other participants said they were feeling a good walking sensation. However, two subjects thought that it would be better if their feet were moved as well and one said that the walking sensation was reduced at low speed. This means that the synchronization between the chair's mechanism and the movement of the on-screen character may require some fine tunning.

7.3.4 Tutorial

The tutorial was only introduced in the third session of trials. All 4 participants went through it successfully. However, they required help because they kept moving the character while trying to read the following instructions and could not do both at the same time. This issue could be solved by freezing the avatar when the instructions text is displayed so they cannot move their avatar while reading.

7.3.5 Discussion of design choices

In our test sessions, we observed that older adults are sensitive to the visual layout of the game and we identified it as an important motivational factor. We observed that the combination of visual, auditory and proprioceptive sensations yields a good level of immersion, especially when using first-person view. However, we could not find the ideal controller for older adults, but we observed that the best one so far was the Nintendo 64 controller.

In the game we implemented through this design phase, the player controls a human avatar, whose appearance can be selected among 8 models, and has to stroll through a virtual environment. The strolling can be either free, guided through the instructions of a therapist, or through objects (large blinking cubes) that the player has to collect. Having three different tasks is meant to allow for variations and gradually increasing difficulty, to keep patients motivated throughout the virtual reality therapy program. When collectable objects are involved, a 3D green arrow at the top of the view helps players find the objects by pointing in the direction of the next object to collect. Once the player has picked-up a cube (no explicit action required), the next cube appears. If the player follows the path correctly, the next cube appears in sight from the location of the one he or she has

(a)

(b)

(c)

Figure 7.5: Screen captures of the virtual strolling game. (a) Tutorial. (b) Forest. (c) Park.

just collected. Lastly, the view can be selected among first-person view and third-person view at any time.

The game starts with an interactive tutorial and has two levels with increasing difficulty:

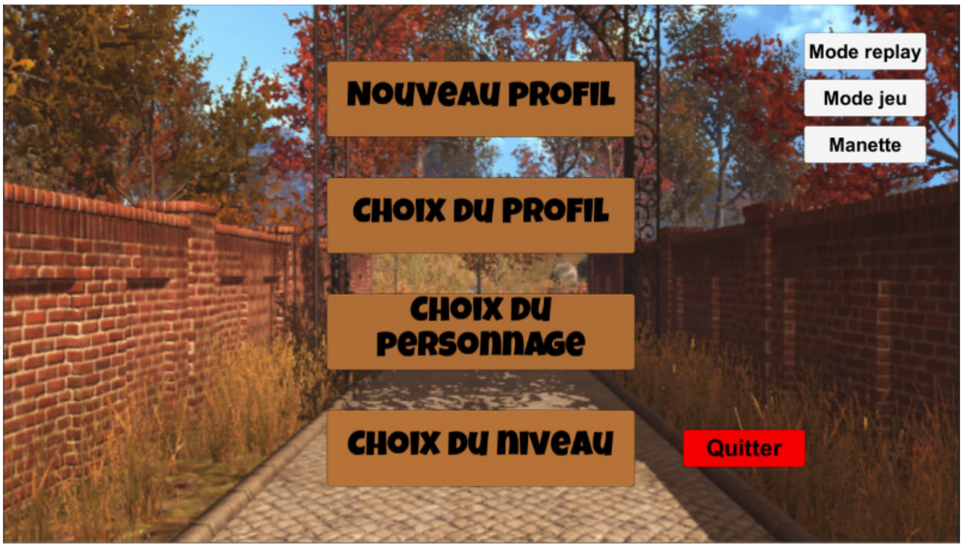
- The tutorial takes place in an outdoor space with grass and a few trees (to give visual orientation cues). It explains how to control the character, the mechanics of the game when pick-up objects are involved and how to switch between first- and third-person views. The instructions are given in text bubbles.
- The first level takes place in an autumn forest path, in which the player can stroll freely. The purpose of this level is mostly to give the player some time to get acquainted with the controls without goal constraints.
- The second level is a park. This time, the player has to stay on a paved path with turns and collect the cubes.

Screen captures of the tutorial and game levels are shown on Figure 7.5. The game only uses four controls: the joystick (or left analog stick on the Sony controller), used to move the character; one button to display the next instruction; another button to go back to the previous instruction; and a button to toggle between first-person view and third-person view. We also added a menu for character and level selection, shown in Figure 7.6, which is navigated with the computer's mouse and meant to be operated by the therapist. For the therapist to have full control on the session, it is possible to switch level, character model and view at any time. In addition, there are duplicated controls, so the therapist can change the view and control the instructions' display (next or previous) from the keyboard.

Lastly, regarding sound design, the avatar produces footstep sounds as it walks, which are different depending on the environment (there are dead leaves all over the ground in the forest and pavement in the park). In the background, we put nature sounds with signing birds in the forest and park environments and a relaxing music in the tutorial.

7.3.6 Validation of the first version

All 9 participants of the *Café Multimédia* event tried the Virtual Promenade system in two groups. They played the game with the Nintendo 64 controller, one by one, while the others were watching, and we have shown them all three levels. All participants enjoyed playing the game. They gave positive feedbacks on the environments. Some preferred the forest and others liked the park more. They could all successfully use the game controller and we have observed that it only took them a few minutes to get used to it. The only two negative comments were about the renderings of the character



(a)

(b)

Figure 7.6: Screen capture of the game's menus. (a) Main menu. (b) Character selection menu

(a) (b)

Figure 7.7: Renderings of the Monique character model's face. (a) Before replacement. (b) After replacement.

models' faces in the character selection menu, which, they said, looked like "ex-convicts", and the rotation speed of the avatar. To address this issue, we have replaced the renderings they did not like with ones of better quality, as illustrated on Figure 7.7, and decreased the avatar's rotation speed parameter.

7.4 Focus groups and co-design with care professionals

Once the first version of the game had been fully implemented, we showed it to the physician who co-designed it with us. He was satisfied with the design but asked that we add a very simple task to start the game, in which the player only has to go forward. We thus decided to split the forest level in two parts: in the first part, the player can only go forward and has to walk towards a big rock. Then, in the second part, he or she has to go around the rock, after which he or she can stroll freely, until the avatar has walked 200 meters. Since there were three levels, each corresponding to a more difficult task than the previous one, we also split the tutorial in three parts, each providing the necessary instructions for the following level. As a result, the game is organized as follows:

1. in a first tutorial, the player learns how to go forward;
2. the player has to make his or her avatar walk to a big rock in the forest environment;

3. the second tutorial explains how to turn and go backward;
4. the player has to go around the rock and walk for 200 meters;
5. in the third tutorial, the player learns to follow the path by picking up the cubes;
6. the player has to pick all cubes up in the park environment.

After that, we organized two focus groups: one with physiotherapists and one with psychomotricians[6]. Each session took the form of a meeting and was organized with the corresponding team of the Broca hospital. All the members of both teams came, including the manager of the physiotherapists' team. They were given a short presentation about the purpose of the project and the prototype. Then, all those who wanted could try out the system. Since there were only three of them, all psychomotricians tried it out.

7.4.1 Feedback from the physiotherapists

Four members of the team and the manager tried the system. Several participants thought that the project was interesting; some of them were even quite enthusiastic. They said that it could be helpful for some patients with fear of falling and agreed that it would ease the rehabilitation work if the virtual reality therapy could reduce patients' anxiety. Indeed, with some patients, they struggle to get them to get up from their wheelchair, even when they are helped by the therapist and have a bar to catch on. They also recommended to test the system with patients recently admitted in the rehabilitation care service, as they would be the primary beneficiaries of this new therapeutic tool.

Regarding the integration of this new activity in their care practices, they said that, in current practices, it is the work of the psychomotricians to help patients with fear of falling reduce their anxiety. This is why we decided to have a second focus group with these professionals.

Lastly, regarding the game, they noticed several minor flaws: the ground in the forest environment is very uneven, which may be too stressful for patients with fear of falling, given that the forest levels are supposed to be the easy ones that players would start with; the footstep sounds are too loud; and the cubes are opaque, which does not make one want to go through them. They added that walking through cubes seemed like a strange idea and proposed that we replace the cubes with other objects to make the task more meaningful. In addition, they emitted doubts about the ergonomics of the Nintendo 64 game controller and said that it may be too difficult for patients with dementia to play the game. Lastly, they suggested that we

[6]Physiotherapists mostly focus on balance, muscle strength and coordination, whereas psychomotricians address sensory-motor and psycho-motor aspects of rehabilitation.

used a means of interaction with the game that would require the patient to be more active, such as moving the arms or the feet, as people do when walking, to make the character go forward.

7.4.2 Feedback from the psychomotricians

Following the advice from the physiotherapists' team, we invited the psychomotricians of the Broca hospital in the living lab to also have their opinion on our system. Before the meeting, we had made some changes to the game, as detailed below. There are only three members in the team and they all tried the system.

The psychomotricians did not have any comments about the game or the system itself. They were less enthusiastic than their colleagues of the physical therapy service but did not emit negative comments either. They reported that they did not use technology in their practices, except for a music player, and that they did not really see how such a system could be included in their care practices. They said that it should rather be used in the physical therapy service. However, a member of the team suggested that it could be helpful at the end of their work with the patients to motivate them to walk again and her colleagues said they agreed with her.

7.4.3 Discussion and improvements

Between the two sessions we fixed the minor issues that the physiotherapists had reported by leveling the ground of the forest environment, lowering the volume of the footstep sounds, and making the cubes translucent. We also made sure to keep support for all the game controllers we had experimented with so we could choose the one that is the more adapted for each patient.

The exchanges with physiotherapists were more fruitful than with psychomotricians. This is likely due to the fact that there were more people attending the meeting. Regarding the use of the system, the physiotherapists seemed the more enthusiastic, compared to the psychomotricians. This may be due to the fact that physiotherapists have (recently) started using computer-based gamified activities in their practices, as we will see in Section 7.5, whereas psychomotricians do not use them at all. However, we have shown the system to another psychomotrician, who was an intern in our living lab, and she was really enthusiastic. When we told her about the lack of enthusiasm of her colleagues, she said they were "old school" (although they were quite young). That being said, we think that this system is at the frontier between the two specialties. In addition, we do not think that any team would start using the system before we can answer all their questions regarding its usability with real patients, especially those with dementia, and show evidence of therapeutic effects. This brings us back to the theory of technology acceptance mentioned in Section 1.2.2: the main predictor is the (perceived) usefulness of the system.

7.5 Shadowing physiotherapists

To gain some insights about the deployment of the Virtual Promenade system, to better understand the current care practices in a geriatric physical therapy service and to develop a closer collaboration with the physiotherapists of the Broca hospital, two members of our team each spent one day in their service. More specifically, this was done for three purposes: witnessing first-hand the rehabilitation activities offered by professionals at the hospital to better understand the nature and organization of their tasks, in order to project how our system could fit in the physical therapy practices; seeing for ourselves the space and equipment dedicated to rehabilitation care to better identify the constraints of our target deployment field; and acquiring some practical knowledge about physical therapy to better understand physiotherapists' professional interests and concerns and facilitate future collaborations with them.

During this day, we observed, asked questions, took notes and even participated in the activities of the service, when possible. For instance, when physiotherapists ask a patient with little muscle strength in his or her legs to walk, they have to almost carry him or her, which mobilizes two professionals, and a third one has to follow them with the patient's wheelchair, so they can stop as soon as the person refuses to go further. We mostly helped out by pushing wheelchairs around, which was appreciated.

First of all, we have observed that, most of the time, each physiotherapist takes care of two patients (sometimes three) at the same time. While he or she has one patient do exercises, the other patient(s) rest. Secondly, we noticed that they have very limited space to work and must have the patients walk and do exercises with obstacles in the hospital's corridors. Thirdly, we have seen that they sometimes have patients perform seated exercises, to work on muscle strength with patients who cannot stand up because they still have not fully recovered from fractures. This is due to the fact that they try to start the rehabilitation work as soon as possible, since long periods of bed rest cause muscular atrophy and psychomotor regression.

Lastly, we have seen that the service was recently equipped with a device for gamified balance exercises. This system consisted in a posturographic platform connected to a computer and a large display. This system, called Gymplate, is made by Techno Concept[7]. The patients stand on the platform and perform balance and posture exercises by playing games in which they control an avatar or steer an airplane by moving their center of pressure. The system is very quick to set for the therapist, who only has to click a few buttons. We interviewed patients, who reported that they enjoyed this

[7] http://www.technoconcept.fr/

activity. We noticed that patients were motivated by the fact that they could see their scores and their progression curves over the sessions immediately after they had played.

7.6 Discussion

Through this participatory design phase, we could produce a fully working prototype that proved easy to use and enjoyable for older adults. We identified that the aesthetics of the game was an important factor for them, which led us to focus on including pleasant environments. This turned out to be more important to elders' quality of experience than the synchronization of the movements of the seat with the speed of the player's avatar in the game. This is a good example of how by starting testing with a minimal prototype, we could quickly orientate the design and avoid unnecessary developments. This is in line with the work done by Uzor *et al.* [204] who could also save development time to produce a game that older adults enjoyed, by involving them in the very early phases of the design and letting them drive the design.

Once we had developed a more mature prototype, we could refine the design and get insights for deployment by consulting the professionals who take care of patients with PFS and performing shadowing in the physical therapy service. Doing so, we have mostly identified that to convince professionals of using our system in their practices, it will be necessary to validate the usability of the system with patients who suffer from PFS, particularly those with cognitive impairment, and to conduct clinical trials to check for evidence of therapeutic effects. The later is out of the scope of this research, but an evaluation of the system in ecological conditions, with hospitalized patients, is the logical next step of our design process, which we report in the following section.

7.7 Pilot study in ecological settings

To validate the usability and acceptance of our system, we conducted a pilot study in ecological conditions, that is to say with patients hospitalized in rehabilitation care or long term care for having fear of falling and doing rehabilitation. In addition, we wanted to look for potential therapeutic effects. The tests were conducted at the Broca hospital.

7.7.1 Protocol

We recruited patients who were over 75 and who had fear of falling. We did not include patients who had severe dementia (MMSE < 10), suffered from an acute pathology or were receiving palliative care. The patients

Figure 7.8: Setup for the experiments of the Virtual Promenade pilot study.

were indicated to us by physicians or physiotherapists. All participants gave written informed consent and were not retributed.

The intervention consisted in having patients sit on the haptic chair and play the virtual strolling game for two or three 15 to 20 minutes individual sessions, over two to three weeks. Participants were placed about 3 meters from a large display (42-inch plasma screen or video-projected image) on which the game was displayed. During the sessions, after choosing their avatar, they had to play all three levels of the game. In the first session, they learned how to play by going through all three parts of the tutorial and played in third-person view; in the second session, the tutorials were skipped and participants played in first-person mode. If a third session was done, they played in first-person view in the park and in third-person view in the forest. During the tests, we used the Nintendo 64 controller but quickly switched to the Thrustmaster USB joystick, used in the feasibility study, and visible on Figure 7.8; we will explain why in Section 7.7.2. When participants agreed to, a video recording of the session was taken. The experiment's setup is depicted on Figure 7.8.

At the end of each session, we conducted a directed interview with the participant. It consisted in discussing 18 questions with the participants to debrief the session. The questions were about:

- the number of avatar models to choose from;
- how important the possibility of choosing one's avatar is;

- how satisfied participants were of their avatar's appearance;
- how realistic they thought the walk animations of the avatar were;
- how much they physically identified with the avatar;
- how much they felt like they walked via the avatar;
- how much the seat's movements gave them the impression of walking;
- how pleasant the seat's movements were;
- how much they liked the environment;
- how easy the controls were;
- how helpful the tutorial was for them;
- how legible the text instructions were;
- how much they enjoyed playing the game;
- if they felt more immersed in first-person view than in third-person view;
- how much their fear of falling diminished between the beginning and the end of the sessions;
- and if they would like to play again in another sessions.

The first and last questions were yes/no questions, as well as the one about the quality of immersion depending on the view. All other questions were 1 to 4 Likert scale questions. The full questionnaire is given in Appendix C. To build this questionnaire, we took inspiration from the work of Heerink *et al.* [86], who proposed and evaluated a specific technology acceptance model for the use of assistive social agents by older adults.

To gain insights about potential therapeutic effects, patients were evaluated along the Fall Efficacy Scale (FES) [198] and PTSD checklist scale (PCS) [26], to quantify their fear of falling and their anxiety disorders due to fall, before and after the intervention. In addition, we asked them to fill in a Katz's autonomy scale [101] questionnaire, which informs on the level of autonomy of each patient.

7.7.2 Results

10 participants, 9 females and 1 male, were recruited. Only 7 people completed the study by doing at least two sessions, and one only did one session. The participants dropped from the study due to sickness, depression and/or refusal to participate in the tests. We therefore only report the data for the

8 participants who had at least one session. They were between 75 and 99 years old (mean = 88.9, standard deviation = 5.9) and their MMSE scores ranged from 27 to 12 (mean = 20.9, standard deviation = 3.7). Two had dementia; one had Parkinson's disease; and the other ones had MCI. Also note that, for one participant, we only had access to his MoCA score (MoCA = 12/30), which was converted to an estimation of the corresponding MMSE score (MMSE = 19/30) thanks to the conversion table by Roalf *et al.* [173], validated by Falkowski *et al.* [65]. The participants all had some level of dependency, as they scored between 3.5 and 5.5 out of 6. In addition, the highest scores are likely to be overrated, as these people had dementia and intense fear of falling and said they could do things on their own that they obviously could not.

Usability

With the three first participants, we tried to use the Nintendo 64 controller, but quickly noticed a major usability flaw: they had poor fine motor control in the hand, which made the use of this controller very difficult. It was manageable for the task in which they only had to go forward, but they struggled as soon as they had to worry about the direction. In addition, due to cognitive impairment, Participants 1 and 2, who scored 24 and 23 at the MMSE, respectively, were too slow to start moving forward as soon as the avatar was oriented towards the direction they wanted to go to and they just kept rotating the avatar in place until the experimenter in the room intervened, not to let them stuck for too long. They also had difficulties finding the cubes in the park environment and the experimenter had to intervene orally, to give them instructions, or physically, to put them in the right direction, on several occasions. Participant 3, who had a higher MMSE (27/30), did slightly better and could play on her own, with a little difficulty in the cube pickup task. This resulted in dreadful effects on immersion, as participants could not focus on the display nor on the seat's movements as they were too busy struggling with the controls. For Participants 1 and 2 to be able to enjoy the experience, the experimenter ended up controlling the avatar for them in the park environment. On the videos of the sessions, we observed a much higher level of engagement in that situation than when they were struggling with the controls.

To address the issue, we had Participants 2 and 3 play with the arcade controller in their second sessions. However, it only addressed the motor control issue but not the cognitive impairment. In addition, as mentioned earlier, this controller presents the drawback of not allowing for speed control.

We then decided to go back to the Thrustmaster USB Joystick, which was evaluated positively in the preliminary study, and created an "easy" mode, in which the avatar follows a predefined path and patients only have to worry about speed. To address the issues of the joystick being too light and of its placement on a table, it was taped to a table with wheels like the ones all patients have in their hospital rooms, as can be seen on Figure 7.8. These tables combine all the advantages of taking little space, being easy to move and being adjustable in height. Lastly, the fact that this joystick is made for right-handed people was not much of an issue, as in the generation of our target users, left-handed people were forced to write with their right hand and most of them became ambidextrous. In addition, in the "easy" mode, the hand with which the game is controlled does not matter much, as it only requires pushing the joystick. Nevertheless, we maintained support for all the game controllers we have used in the project, to be able to adapt the system to each participant as much as possible.

After we made these changes, we kept this setup for all sessions, including a third session with Patient 3. Patients 1 and 2 did not try this solution. Thanks to these changes, all participants could successfully play the game and we observed high levels of engagement on the videos. Three patients needed a little help: two were losing track of what they were doing and had to be reminded to keep playing by the experimenter and the patient with Parkinson's disease was tired and stopped on several occasions. All of them played only the easy mode, as all of them had moderate cognitive impairment, except one, who had Parkinson's disease.

During the experiment, we observed another important design flaw, regarding the safety of the armchair with the moving seat: the top part of the armrests can rotate by about 10 degrees inward or outward. The problem resides in the fact that patients with fear of falling always grab the armrests and push on them to sit down or get up. This may be risky and makes the transfer of the patients from their wheelchairs to the haptic chair more difficult. Another issue with the chair we used in the experiments is that it has wheels, which we had to replace with fixed footings for the experiments. This is not practical and increases the installation time of the system. This could be fixed easily by using wheels that can be blocked thanks to a small brake on each wheel.

Lastly, we observed that participants did not really want to read the written instructions in the tutorials and did not understand well, or had a hard time remembering, how to see the next or previous instruction. The fastest method was to give them verbal instructions and take their hand to put it on the controller and show them the effects when pushing or pulling the handle of the joystick. As a result, the tutorial level was only an occasion for them to try out the controls, before having to perform the actual tasks.

Feedbacks of the participants

Only 7 participants answered the questionnaire. They mostly reported that they enjoyed the activity: when answering the 15$^{\text{th}}$ question, they gave an average score of 3 out of 4, and only two participants gave a grade of 2, only in one of the two sessions they had. After the first session, 6 participants said they wanted to play again in another session; after the second sessions, this number dropped to 3. This means that there might be a lack a variety in the activity. Regarding the pleasantness of the movements of the chair, opinions were mixed: 3 participants said it was “not really pleasant” (Likert grade of 2 out of 4); the others said it was pleasant (3 out of 4). On the contrary, participants gave very positive feedbacks about the avatars’ appearances, walk animations and the environments (3.3/4, 3.2/4 and 3.7/4 on average, respectively); none of them had a negative opinion about these features and all of them were satisfied with the number of possible choices for the avatar’s embodiment, which most of them thought it was important that they could choose (3/4 on average). However, only one participant reported that she identified with the character she had chosen, which is not an indicator of satisfaction but rather an information on how participants related with their avatars. Aside from those, one patient had a negative comment about the cube-picking task and said it bothered her to have to walk through them. She suggested that we made them smaller and agreed that it would be better to do a more meaningful activity.

Regarding the self-reported ergonomics of the system, all said that the written instructions were legible or very legible (3.4/4 Likert score on average). In addition, they said the controller was easy or very easy to use, except the participant with Parkinson’s disease, who said it was rather difficult (3.1/4 on average), although she only played in the “easy mode”. Regarding the instructions given in the tutorial, all said they were helpful (3/4). However, this mostly referred to the verbal instructions they had received and the fact that they could try the controls before having to perform the tasks.

Lastly, regarding immersion and impressions of walking when doing the activity, participants had mixed opinions: 2 participants said that they did not feel like walking at all; 2 said “a little” or “not really”, depending on the sessions; and 3 said “a little” in both sessions. The impression of walking due to the movement of the chair was a bit more positive: 3 said “a little” in both sessions; 2 said the same thing in only one session; and 2 said “not really” in both sessions. However, in one play session, a participant reported that she did not notice that the seat was moving. Furthermore, 4 participants said the first-person view was more immersive than the third-person view; 2 said it was not; and one did not answer, as she did not do the second session and had not tried it. We also have observed that, in third-person view, participants tended to refer to their avatar as “he” or “she”. Our male participant even said: “He is not in a hurry, this one. He doesn’t get paid on a hourly basis.”, about his avatar.

Effects on fear of falling

For starters, the answers we obtained in the FES did not reflect the actual fear of falling for at least four patients: one did not care and answered randomly; the patient with the lowest MMSE (12/30) had a high FES score (45/64) but walked, without even using a stick, from her room to the lab, located in the hospital's garden, and back; and two other had low scores because they did not answer around half of the questions about how worried at the idea of falling they would be while performing activities they do not currently perform, as they are hospitalized. The necessity for patients to project themselves in activities they no longer perform to answer related questions was an issue in all cases. This raises doubts about the validity of this test when it is taken by older adults with cognitive impairment. At the end of the day, the most reliable evaluation of fear of falling was the one reported by physiotherapists or observed when transferring patients from their wheelchairs to the experiment's chair. We would say that, in 8 participants, 2 had an intense fear of falling and required the help of a physiotherapist for the transfers; 4 had high to moderate fear of falling; 1 was a little anxious; and 1 had no fear at all.

The PCS seems reliable, even though one patient did not fall but had a fracture due to osteoporosis and another patient did not take the test seriously. We are thus left with only 5 patients who answered the PCS before and after the intervention, as the sixth one did not finish the study. For those patients, the score is not significantly different before and after the intervention.

Lastly, regarding the self-reported diminution of fear of falling between the beginning and the end of the session, in the first session, 4 patients said "not at all" and 3 said "not really"; in the second session, 3 patients out of 6 said "not at all" again, one said "not really" and 2 said "a little".

7.7.3 Discussion

Regarding usability, after the introduction of an "easy" mode, all patients could successfully play the game. Regarding acceptance, participants' feedback about the activity was mostly positive, and very few negative comments were made by people who have tried the system. Regarding the potential of the system for therapeutic outcomes, given the small size of the sample and the unreliability of the data, no conclusion can be made. In addition, making any sort of conclusion would require the study to include a control group. However, we have identified that using the FES as a metric for fear of falling in older adults with cognitive impairment does not seem to be a reliable indicator of success.

This preliminary study allowed us to debunk issues that we had not anticipated and to improve the game. By using the Thrustmaster joystick, which does not require much strength or fine motor control to operate, and by introducing an easy mode, in which the avatar steers itself, the system can likely be used by most older adults, even with moderately severe cognitive impairment or moderate Parkinson's disease. The easy mode that we have introduced also saves time for the therapists that would manage the sessions, as fewer instructions have to be given.

Regarding the tutorial, we have identified that, for this type of patients, the best is for the therapist to provide the instructions. However, having a level in which patients can try out the controls seems like a good idea. If the system is meant for a patient to play on his or her own, it may be useful to include an ECA such as LOUISE to fulfill two functions: explain how to play, in lieu of the tutorial, and perform prompting and context reminders when people lose track of what they are doing or get distracted, as we have observed with two participants.

There are no clear indication that the system actually helps reducing fear of falling. During the sessions, only one patient expressed anxiety, but it was quite intense, as she urinated in her pants. However, her high stress level was also caused by the transfer from her wheelchair to the experiment's chair, which was very difficult in her case, so much that she needed several minutes to recover before we could start the session. This may show that the game is not immersive enough, as the sense of presence in virtual reality exposure therapy is usually characterized by in-session anxiety [162]. This hypothesis is also supported by the mitigated answers we obtained about the feelings of walking. This aspect should be given much importance in future work, as presence is also linked to the efficacy of virtual reality treatments [162, 70].

Lastly, the cube-picking task does not seem to be interesting enough. Cubes should probably be replaced with more meaningful objects, as it was already suggested by the physiotherapists. Interestingly, Uzor *et al.* [204] have reported a similar observation: in a game that consists in picking up objects falling from the back of a truck by adjusting the flying altitude of a pigeon, controlled through body movements, older adults involved in the participatory design asked that these objects be meaningful (mails or fruits). In our case, we could make the task in the Park environment more meaningful by replacing cubes with flowers that have to be collected to complete a bouquet or toys lost by a child. In any case, this choice of object should be validated by consulting some older adults to ask them what kind of objects they would like to collect in the game.

Chapter 8

Conclusion

In this thesis, we have reported our work on the use of virtual human technology to improve older adults' care. In our approach, we considered that virtual humans can play two distinct roles in elders' care scenarios: be the avatar of the person in the virtual world (as self) or represent an artificial agent that lives in that virtual world (as non-self). This led us to conducting two case-studies: (1) exploring the use of ECAs, virtual humans capable of verbal and nonverbal communication with their users, to improve the accessibility of information and communication technologies, which would be useful for cognitive compensation and stimulation, for older adults with cognitive impairment; and (2) investigating the use of avatars in virtual environments to expose older adults with PFS to virtual walking situations, in order to reduce their fear of falling. More specifically, we have designed developed and tested an ECA, called LOUISE, to serve as user interface in assistive systems for cognition and a virtual reality treatment system, composed of a serious game and a haptic chair. Throughout this research work, we followed a living lab participatory design methodology, for which we defined adapted steps to best fit our target audience.

In this final chapter, we start by presenting conclusions for each of our two case studies in Section 8.1. Then, we present a general conclusion for this work in Section 8.2. After that, we propose ideas for future work in Section 8.3. Lastly, we end this thesis by discussing the long-term ethical implications of virtual human technology in Section 8.4.

8.1 Case studies for virtual humans

In this section, we summarize and discuss our work and main findings in each case study.

8.1.1 LOUISE

People with dementia require help to perform activities of daily living. In that regard, ECAs have been shown to be a promising solution (see Section 2.2.1) to assist them, so they remain autonomous for as long as possible, and have the advantage of being cheap. This is why our colleague, Dr. Pino, a psychology researcher at the Broca hospital, initiated this project. In the case study of the LOUISE (LOvely User Interface for Servicing Elders) project, reported in Chapters 5 and 6, we have conducted the conception and development phases of an ECA, seen as a toolbox, to be integrated in cognitive assistive technologies. This work was oriented around the specificity of the interaction between the ECA and older adults with cognitive impairment, particularly for managing attentional disorders, disorientation, reduced short-term memory and aphasia. Throughout the study, we have also collected information about the desires and expectations of the stakeholders regarding this type of interface, as well as the usefulness of some of the assistive scenarios in which it could be integrated.

WoZ study

We started off by conducting a WoZ study with an elementary semi-automated prototype for LOUISE, featuring an attention management strategy based on estimating the user's attention in real time, thanks to a new algorithm that we have proposed, and having the ECA prompt the user in case of loss of attention, until his or her attention was directed towards the ECA again. We have tested this system with 14 assistive technology professionals and 9 older adults, most of whom had mild to moderate cognitive impairment, and performed an anthropological analysis of the video recordings of the tests with older adults. We have found the following results:

1. people with cognitive impairment are capable of interacting with an ECA;

2. older adults mostly reacted positively to the ECA;

3. elders with dementia tended to interact with the ECA in a more social way than cognitively-healthy older adults, as they uttered more words and did more topic development;

4. our attention recapture strategy was effective but not sufficient, as in the case of the most severely impaired subject it did not allow to carry on the interaction because of the subject's disorientation;

5. our simple, cheap and fast attention estimation system was over 80% accurate;

6. the performance of the attention estimator did not seem to depend on age, which suggests that it is not necessary to create a tailored method for older adults.

Questionnaires and focus group

To collect information about the design requirements, we have asked their opinions to the participants of our experiments and to the public, on the occasions of demonstrations given in two public events. In total, 37 people answered our questionnaire. Then, we have conducted a focus group with 9 older adults, 3 of whom had MCI. We found, that:

1. people think ECAs would be the most useful as personal assistants, that is to say, providing help in managing one's schedule, reminding medication intakes, giving access to video chat, reading emails, etc.;
2. other applications reported as useful included guiding step-by-step through a complex task, serving as an interface to control home appliances or centralizing cognitive compensation and stimulation applications;
3. physicians we have consulted informally in staff meetings proposed that ECAs address behavioral disorders, such as screams, and read and teach poems to patients, for cognitive stimulation;
4. people would not want the ECA to be installed on a dedicated device but rather on a device they already own, or that has other purposes, such as a computer, a tablet or a television set, for the most popular options;
5. ECAs should be pleasant or even fun, expressive and smiley;
6. people want the ECA to look like a young woman, or a humanoid robot;
7. most people would want to personalize the embodiment of an assistive ECA.

Assisted task management

Given this information and the results of our Woz study, we put together a fully-functional automatic prototype of LOUISE, which includes:

- rich animation capabilities, thanks to the use of the state-of-the-art BML realizer SmartBody;
- display of images and example videos;

- easy addition of character models for the embodiment;
- a custom-built interaction manager module that allows to describe interaction scenarios in a dedicated XML syntax.

The tailored interaction management we propose for older adults with cognitive impairment allows to manage attention, as in the first prototype, perform context reminders after the user gets distracted and specify the possible answers for a given question if people do not answer adequately. We used this system to explore the specificities of dialog management for elders with dementia. We created interaction scenarios, based on realistic use cases, to study two tasks that would intervene in most applications of ECAs as user interfaces for assistive technologies – choice with multiple options and step-by-step task guidance – for which we proposed a dialog structure. We then conducted a usability study to refine and validate our system with 14 older adults with mild to moderately severe cognitive impairment. We have found the following issues and made some adjustments:

1. all participants with mild to moderate cognitive impairment were capable of interacting with LOUISE (MMSE > 10);
2. participants gave very positive feedbacks (over 3.3/4 on average on the Likert scale) about the pleasantness of the ECA, its appearance, the clarity of its synthetic speech and the pace of the conversation;
3. the ECA can have the dialog's initiative, but people with MCI sometimes want to take the initiative;
4. our strategy for choice-making is effective, since all participants could make choices and small changes to improve it have been identified;
5. our strategy for step-by-step task instruction has been improved for people with MCI, but the changes do not suit people with more severe cognitive impairment;
6. the minimum distance for the Kinect sensor for Xbox360 to track users' bodies and faces is too large for the system to work reliably in real use condition;
7. the performance of off-line speech recognizers is not good enough;
8. the quality of speech recognition is critical, as people get confused when it makes mistakes;
9. our observations confirm that people with cognitive impairment tend to elaborate answers in a social way, as seen in the WoZ experiment.

Thanks to our living lab participatory method, we could create an ECA that older adults enjoyed using, debunk several usability issues and fix some of them. We have stated in Chapter 2 that, in some respects, the implementation of such a system could be simpler when dealing with older adults with cognitive impairment than with healthy people, particularly regarding rendering quality and interaction management. Our observations in this case study tend to confirm these assumptions as people were satisfied with the visuals of the application, though the animations could be better and the rendering quality offered by Panda 3D, the game engine we have used, is below average. Regarding the simplicity of interaction management, a simple system can cover the needs in most cases, but some more advanced possibilities should be explored. In addition, the use cases that we have tested are quite simple and we think that, with a few improvements, especially for step-by-step task guidance, the ECA could already be useful for people with moderate dementia. More generally this suggest that even ECAs with little intelligence could provide useful services to people with dementia, as they require help in performing simple activities.

8.1.2 Virtual Promenade

As PFS, a combination of psychological and psycho-motor symptoms resulting from a fall, resembles PTSD in several ways, we proposed to use virtual reality therapy for older patients with fear of falling, as suggested by Dr. Bloch, a geriatrician of the Broca hospital, who started the project [124]. In our second case study, we have conducted the participatory design of a system, called Virtual Promenade, which includes a serious game and a haptic chair with a moving seat that imitates the movements of the hips in human walk. The serious game we have developed is a virtual strolling game, in which patients control a virtual human avatar they have to make walk through several virtual environments. The haptic chair is connected to the game, so that the speed of movement of the seat corresponds to the walking speed of the players' avatar.

Playtesting

We started by putting together the prototype with an elementary version of the game and conducted playtesting of the game with older adults. Based on the feedbacks, opinions and wishes expressed by the participants, we could build a first version of the game.

Thanks to this co-design process with older adults and their caregivers, we could create a virtual strolling game that elders were able to play and enjoyed playing. Doing so, we have made the following observations:

1. older adults wanted the virtual environments to be pleasant, rather than realistic, and this was identified as a motivational factor;

2. older adults wanted to relate to their avatar's body shape;

3. a flight simulator-like joystick would not be adapted for use by older adults if it requires strength to maneuverer, as it was the case with the Logitech joystick we used in our trials;

4. the Nintendo 64 game controller seemed well adapted for use by healthy older adults, though it may require fine motor control.

Focus groups and shadowing with professionals

After this first phase, we conducted two focus groups with professionals that intervene in rehabilitation care: physiotherapists and psychomotricians. After meeting each group, we implemented the changes to the application that they asked for. Lastly, we have followed physiotherapists in a typical day at work (shadowing) to better understand their activities and the target field of deployment of our system.

The consultation and shadowing of professionals allowed to refine the design, better understand how they would relate to the introduction of a system like Virtual Promenade in their workplace, how they think it would fit in their care practices and get insights about the constraints of the field, which are mostly space, operation time, safety and autonomy of the patients to play the game. This last point is particularly interesting, as we have observed that physiotherapists usually handle two patients at a time, sometimes even three: a system they only have to start and can let the patient use alone, while they take care of other patients, would be the ideal solution for it to be integrated smoothly in rehabilitation care practices.

Pilot study

Lastly, we tested the system in ecological conditions with patients hospitalized in rehabilitation care. In 10 participants, only 8 completed at least one trial of the system. We observed that the Nintendo 64 controller, which seemed adapted for use by healthy older adults, was not adapted for patients with poor fine motor control. In addition, having to actually control the character to visit virtual environments was too difficult for older adults with moderate cognitive impairment. As a result, we have replaced the Nintendo 64 game controller with a Thrustmaster flight simulator-like joystick that does not require as much strength to operate as the Logitech one used in the first phase of the study, attached it to a hospital bed-side table because it was too light and, most importantly, introduced an "easy" mode in the game, in which the player's avatar follows a predefined path, so the player only has to worry about speed. In addition, we maintained support for the other game controllers we have experimented with throughout the project. These changes allowed older adults with moderately severe dementia or Parkinson's disease to successfully use the system. Moreover, we have found that:

1. participants were satisfied with the usability of the game;
2. older adults enjoyed playing the game and were satisfied with the environments, the graphics and animations;
3. they did not really want to use the written instructions and the experimenter had to intervene to explain participants how to play;
4. most participants did not have the impression to walk when performing the activity;
5. using the system only triggered anxiety feelings in one participant, but this may also be due to the stress of the transfer from her wheelchair to the system's chair;
6. we did not observe any effects on fear of falling;
7. the FES did not seem to be a reliable success indicator as the scores we obtained did not reflect well the intensity of fear of falling we have observed, particularly for people with cognitive impairment, who had much difficulties projecting in activities they no longer perform, answered randomly or said what they thought we wanted them to say ;
8. adjustable armrests cause a safety issue, as people push on the armrests for seating down and getting up.

We do not have not enough data to conclude whether or not using this system will have any therapeutic effects, as this would require to create a therapy program and conduct a randomized controlled clinical trial. However, our intuition is that the system may not be immersive enough for patients to relate the activity offered by the Virtual Promenade system to actual walking situations. This can be due to several factors: the use of a third-person view seemed to be less effective than first-person view, as people reported that they felt more like walking in the second session in first-person view; the chosen environments (forest and park) may not allow people to relate enough to their daily lives; the activity may not resemble walking enough; the movements of the chair's seat may be too subtle; and the fact that they chose an avatar may create a distance between the users and what happens in the virtual world. We thus propose improvements in Section 8.3.3 to address these shortcomings and make the system more immersive.

8.2 Living lab-based design for older adults

Overall, in both case studies, the use of a living lab participatory design process allowed to create virtual human-based prototypes that older adults enjoyed using and to achieve good usability, particularly in the Virtual Promenade project. In addition, testing the system in ecological conditions, with the target users, allowed to shed light on several shortcomings in our designs, which could sometimes be fixed right away, thanks to the insights obtained by working in the field with the target population, or by involving care professionals in the design process. For instance, in the Virtual Promenade project, testing with hospitalized patients revealed issues of usability and safety that could not have been observed with healthy older adults. Furthermore, the adaptations we made to implement the living lab method also proved useful: the use of off-the-shelf elements allowed to iterate fast on the prototypes' design and keep costs low; involving healthy older adults as proxies to our target users allowed to obtain useful information; so did involving a wider audience, by doing demonstrations to the public; and shadowing sessions allowed to create a relationship of trust with the care professionals in the field, improve communication with them, and get design insights. This last point stresses out the importance of acculturation in the context of multi-disciplinary work: when having to collaborate with professionals of other disciplines, it is critical to understand their activities, concerns and interests, as well as assimilating the vocabulary of their profession to communicate effectively. This is a long but necessary process, as it allows to get past purely theoretical knowledge of care practices to grasp the reality of the field, which is one of the founding principles of the living lab approach.

During these case studies, we have observed that most older adults are sensitive to the aesthetics of virtual human characters or virtual environments. However, graphic realism, quality of animation and quality of rendering did not seem to be as important to them as aesthetics. This is in line with what we have stated in Section 2.1.5: as older adults have (for now) little experience of AAA games, they have lower expectations in terms of graphics quality and realism. If we had to keep only one qualifier of what older adults want a virtual human-based application to look and feel like, it would be "pleasant". Indeed, they mostly asked that the characters in the LOUISE ECA system smiled and gave an impression of being nice to them. Similarly, for the environments in the Virtual Promenade, they thought the first city environment we had presented to them to be cold and unwelcoming, and asked for nicer ones, with flowers and trees. Of course, this notion is quite subjective and depends on people's personal tastes. This is why in applications such as LOUISE, which have only one or a few users, it seems important that the software allows for some personalization features of the visuals. For applications that are meant to be used by many different users, such as our virtual strolling game in the Virtual Promenade system, it seems best to find visuals that most people would appreciate, even though there will always be people that will not like them.

Working with older adults, we have observed that a lot of them are not particularly reluctant to new technologies. However, due to the complexity of our systems, it was sometimes difficult to explain to them what it was that we wanted them to try out and how our system could help them in a way they would understand, especially when they had cognitive impairment. The most difficult public was hospitalized patients, who tend to be depressed, tired by rehabilitation activities and sick. When recruiting subjects for the pilot study in the Virtual Promenade project, more than half of the patients we have asked to participate refused to do so. A lot of them said that all they wanted was to get out of the hospital; some said they did not want to be "lab rats"; some were concerned that it would tire them or that it would interfere with their rehabilitation program; some were afraid that they would not be able to use the system; but only a few said that they were not interested because they did not like computer-based technology.

Interestingly, as we were wearing white blouses when going in the hospital's services to recruit patients for our studies, we were sometimes assimilated to care professionals, and some patients who did not want to try our systems made false excuses, instead of simply rejecting our offer. One patient even argued that the very physician who told us he would be a good candidate for our study recommended him to wait before he enrolled, as it was too early after his fracture! People were probably afraid to disappoint the care staff by refusing to participate to the study. This points out to some ethical questions about how hospitalized patients should be recruited for our studies. Though we recruited patients ourselves, the fact that we wore a white blouse may have been misleading to some people, even though we clearly stated that we were computer scientists and wore a badge that said so. It is also very important to insist on the fact that they are completely free to refuse to participate or interrupt their participation at any time, without any consequences, which we did. We think that perceived usefulness of the system is again a very important factor for people to participate in our research, as people who wanted to try the Virtual Promenade seemed to be mostly interested in the potential therapeutic effects of the system, though we did inform them that we could not tell yet if it had any positive effects or not.

We found that, in most cases, older adults with cognitive impairment are perfectly capable of giving feedbacks about their experience, regarding how much they appreciated the system and what they liked or did not like in it. However, it was sometimes difficult to obtain information from the most strongly impaired patients, particularly when it required them to imagine situations and project into them or understand some complex questions. In addition, having collected this information, it may be difficult to know what importance should be given to some parts of it, as it sometimes feels like they tell the experimenter what they think he or she wants to hear, instead of what they really think, or answer randomly when they do not have an opinion or do not understand the question. For instance, when asked about the usefulness of the applications we had thought about for LOUISE, their acceptance of its perception system or, in the Virtual Promenade project, if the first-person view was more immersive than the third-person view, they often did not really know what to answer. Getting honest answers, sometimes simply getting an answer, requires time, patience, reformulating the question in ways people understand, showing them the objects or elements the question is about if they forgot or if they do not understand, and discussing their answers with them to make sure that there is no misunderstanding. In that regard, involving healthy older adults in the process is particularly useful, as it is much easier to collect their opinions on complex matters. In addition, though we think it is very important that people with dementia get to be part of the decisions that impact their way of life, it is often the caregivers and family members who take the decisions of buying assistive technology for them, sometimes without consulting them. It is thus mandatory to collect family caregivers' feedback. However, it is also critical to make sure that patients, who are the end users of the system, will actually be able to use it and will enjoy the system, as otherwise, it will quickly be put away. This is why both patients and caregivers have to be involved in the design process, and their opinions and wishes should be given equal importance.

Lastly, based on our experience in conducting living lab participatory design of virtual humans for older people, we share our lessons learned by proposing recommendations that we detail in the box below.

Recommendations for living lab participatory design

Start small; test early To avoid unnecessary developments and quickly converge towards a satisfying solution, it is best to start with a prototype that is just evolved enough for participants involved in the design to get a good sense of what it can do. Then, testing it early allows to quickly debunk the biggest issues and get insights about the most important features and what aspects matter the most for stakeholders. For instance, when developing Virtual Promenade, a lot of effort was put to fine-tune the synchronization of the seat's movements to those of the in-game avatar; and yet, what turned out to matter the most was the pleasantness of the environments. This recommendation is in line with the "lean" approaches and the Agile development movement.

Testing is expensive; make the most out of it The most precious information is usually obtained in test sessions. It is important to prepare well for the sessions, by anticipating questions on specific elements of the design. It is also a good practice to collect as much information as possible in every test. In addition, if after a few tests it seems obvious that some shortcomings in the design are so crippling that participants see only that, we found it best to fix the issue before doing any further testing. Otherwise, odds are that testing will not allow to debunk other issues. Given how expensive recruiting hospital participants and testing are, carrying on testing a system without correcting such flaws would quite likely be a waste of time and resources, even though making changes makes it more difficult to interpret data at the end of the experiment.

Choose versatile development tools In the course of participatory design, there usually are a lot of changes, big or small, to make. Choosing a versatile development tool allows gaining a lot of time and fixing issues quickly when they arise, sometimes in a day or two, between two test sessions. For instance, in the Virtual Promenade project, using Unity and ready-for-use elements, downloaded for free or bought for cheap on the Unity Asset Store, allowed to iterate fast. Similarly, in the LOUISE project, the fact that we used XML dialog tree descriptions, which are very flexible, allowed us to quickly make changes and test various dialog management approaches or create a new scenario, e.g. when physicians forbade that patients with dementia take the fake pills.

Make things easy to test By this, we mean that the system should be easy to install and to operate. In research, it is not rare that little work is dedicated to packaging software, as there are little usability requirements for experimenters to operate the system. In living lab development however, given how frequent and intensive testing is, it is worth spending a little more time to make the system fast to install and easy to test. There are at least three reasons for this: reduced installation and operation time will save researchers a lot of time and energy, when it adds up at the end of the study; if a system is too long to install and hard to operate, it is likely that it will not do well in the field, as in hospitals and care institutions, due to staff restriction issues, these criteria are particularly critical; if the system fails when participants are already there for the test, unless it is quick to restart, some tests will have to be canceled, which should be avoided, given the cost of recruiting subjects. This is also an opportunity for researchers to be their own subjects to assess how easy the system is to deploy.

Stay aware Software technologies evolve very fast. It is likely that during the course of a project, some new, more powerful or more reliable tool becomes available. It is therefore necessary to stay aware of the novelties, as a software update may improve your system a lot. For instance, while we conducted the LOUISE project, several new cloud-based speech recognition services that could have improved our system became available. We did implement a speech recognizer using one of those services, but have not had a chance to test it with patients.

Go see for yourself When designing any system, it is very helpful to go out in the field to observe, get a sense of the reality of the products' deployment environment and get to know the people who will use it. This helps to come up with designs that have more chances of being truly adapted. Though it does not replace evaluation in ecological conditions, it is likely that it will reduce the number of necessary iterations to reach a well-adapted design.

Go the extra mile To maximize the number of tests and contributions of stakeholders to the design process, it is important to be flexible, and make things easy for stakeholders to contribute. For instance, hospital staff have little time to spend on participation in a living lab co-creation process. It is therefore necessary to adapt to their schedule, make good use of their time and go meet them where they are: for instance, by taking a few minutes to present a project and asking targeted questions in a staff meeting. This principle should also be applied for other stakeholders: some hospitalized patients would only have agreed to test the LOUISE prototype if we brought it to their room, which we unfortunately could not do for logistic reasons, but we estimate that if we did, we would have doubled the number of participants from the long-term care service, who are underrepresented in the study. This would have represented a 20% increase of our sample size.

Adapt your discourse This recommendation applies to both living lab participatory design and multidisciplinary work in general. When involving various stakeholders in the design, it is necessary to adapt one's discourse to the audience, given their field of expertise, education level and cognitive abilities. It is key that people who get involved in the process understand what you tell them, which is not always easy. In addition, it is necessary to be able to understand each stakeholder's point of view and what their points of concern are, to ask the right questions to the right people.

8.3 Future work

The two projects conducted in this thesis have already yielded encouraging results. However, more research is necessary on several aspects. This is why we propose ideas for future improvements of each of our prototypes, including adding a virtual therapist in the Virtual Promenade system, by integrating LOUISE in it.

8.3.1 Conducting more validations studies

The main limitation of this work is the small number of subjects included in both projects. This is due to the difficulty of recruiting patients to participate in our studies and the high costs of testing, in terms of time, at least. Another limitation is the lack of comparability of the results in a given study, as the system is changed between tests. While this is a deliberate choice that has clearly shown to be a useful practice in living lab participatory design, some extra validation experiments with stabilized systems and at a larger scale than the ones reported in this thesis, should be conducted.

8.3.2 LOUISE

So far, our work in the LOUISE project has focused on attention and interaction management. However, the conversation management we have proposed is quite elementary and must be improved, especially to account for the interpersonal variability of dementia and its evolution in time for each patient. Moreover, besides making the system more robust, it could be worth looking at adding extra sources of information, such as emotion detection or data from external devices, connected to the Internet of things (pill dispenser, refrigerator, etc.), which could add value to the system, particularly for people with moderate to severe dementia, who have difficulties expressing themselves and performing simple tasks. In addition, the older participants in the focus group we conducted for the LOUISE project have expressed the wish for a very pleasant or even fun interface, which could improve likability and acceptance. Furthermore, as LOUISE is a multi-purpose tool, it could be used in larger systems, such as assistive homes, used in mobile robots and in other research works. Lastly, it is worth considering how the availability of better off-the-shelf software and hardware may improve the system, with a little refactoring work.

Adaptable interaction manager

As we have seen in Chapter 6, it appears that the interaction scheme that suits older adults with MCI may not work for people with more severe cognitive impairment and *vice versa*. In addition, in the cases of severe cognitive impairment, an assistive ECA may have to deal with out-of-scope demands linked to repetitive questioning. Indeed, people with dementia are highly susceptible to getting obsessed by some issues they are having, such as when their relatives will be visiting them or when they are getting discharged from the hospital. Although some frequently asked questions can probably be identified and put in a database, this may highly depend on each patient. This point, as well as our observations and the information we have collected, suggest that a highly important issue to address in the context of ECAs in cognitive assistive systems is personalization. Indeed, we think the system should be able to be tailored according to at least three factors: the level of cognitive impairment, which will influence the way the interaction is managed; the person's personal tastes and topics of interest, which will influence the contents and embodiment; and other personal attributes, such as the user's name, preferences for how the ECA should address him or her, the names of relatives, the reason for hospitalization, etc. The most important thing to focus on in future research is therefore to come up with parametric approaches to adapt the interaction manager in LOUISE to account for personalization and varying levels of cognitive impairment. The ideal solution would be for the system to self-configure, based on previous interactions, to adapt to the needs of each person, which may vary across time, as the dementia symptoms get worse.

In addition, more flexibility should be added to the system, to account for at least two things: personal preferences in how the ECA addresses its user, that is to say being able to use the person's name or surname in the interaction and use either "*tu*" or "*vous*", depending on the user's wishes; and allow for user initiative in the dialog. The former could be done thanks to approaches to dialog descriptions such as in Artificial Intelligence Markup Language [10], in which it is possible to have XML markups that are to be substituted with words or phrases that are either defined *a priori* (the user's name, for instance) or obtained during the conversation. This could also allow to improve our dialog tree description formalism by limiting redundancy when several utterances only differ by a few words, to choose between still water and sparkling water, for instance; or make the formulation of the utterances more natural, by allowing to replace some words with synonyms. These evolutions could be addressed by improving the XML language, called AISML, we have created for goal-oriented dialog description. To allow for mixed-initiative dialog, the interaction manager should be changed more deeply and rely on a frame-based or information state dialog management approach (see Section 4.5).

More comprehensive user behavior analysis

Improving the user behavior analysis of LOUISE would allow to better manage interactions by giving the ECA more social awareness. The additional analysis components that could be used include, but are not limited to: emotion analysis, which could allow to sense the user's affective state and adapt to it; prosody analysis, which would allow to produce appropriate active listening behavior; activity recognition, which would ease guiding users step by step through a task and provide some confidence indicators of correct task completion; and feedback from connected devices or programs, which would serve similar purposes as activity recognition but is either complementary or a more reliable alternative, depending on the applications.

Fun interface

Some work could be conducted to experiment with various embodiments, animations, voices and sound effects to make the interface more fun or more likable than it currently is. For instance, cartoon-like characters with stereotypical behaviors could amuse older adults. This could also make the ECA easier to understand, as exaggerated movements and facial expressions may be better perceived by people with dementia. In addition, a good choice of embodiment may remind people with cognitive impairment that they are not dealing with a real person and thus avoid "uncanny valley" effects and false affordances, by reducing expectations.

While health and wellbeing are very serious matters, gamification and fun can be harnessed to make things easier and more pleasant. This was experimented a lot and yielded good results to increase patients' motivation to perform physical therapy exercises (see Section 2.2.3). It could therefore be interesting to investigate the effects it would have on people with dementia's adhesion to applications for cognitive compensation.

Use of LOUISE or elements from it in other works and systems

During the course of this thesis, LOUISE or assets created for the project have been used in other works. The first one is a research on the perception by older adults of emotions expressed by a robot with an articulated neck, for which we provided renderings of the Louise character (see Chapter 5) performing facial expressions [13]. The robot's head was composed of three actuators and a tablet, on which faces were displayed. The experiment, partly conducted in the *Café Multimédia* activity, consisted in having older adults recognize emotions produced by the robot in four different conditions and ask them what condition they liked the most. The conditions that were compared included: smiley-like face, smiley-like face with head movements, Louise face, and Louise face with head movements. In some cases, participants recognized better the expressions performed with the Louise face than with the smiley-like face. In addition, about half of the participants reported that they preferred that the robot had a human-like face rather than a smiley-like face. This suggests that using LOUISE on a mobile platform, which would make it somewhere in between an ECA and a robot could be an interesting path to explore. For instance, the Buddy robot by Blue Frog Robotics[1] has a tablet in its head and is close to being an ECA on wheels, with a few extra actuators to move its head.

The second use of LOUISE in another work that has already started is a WoZ version of our second prototype for anthropology work: we have taken out the behavior analysis stage and replaced it with keyboard inputs, so the experimenter can have full control on the interaction and simulate full understanding of natural speech. The structures of dialogs are also described using AISML, with the only difference that conditions for transitions are the letters corresponding to the keys to press to trigger them, instead of words or phrases that a user has to say. Our anthropologist colleague [35] thinks that LOUISE is an interesting tool to study natural interaction with dementia patients, as it allows to perform fully reproducible verbal and nonverbal behaviors, which reduces sources of bias.

Lastly, as LOUISE is a multi-purpose tool, it would be easy to integrate it in many other applications. In Chapter 6, we have proposed four use cases for LOUISE in some of which the ECA could be connected to external devices, such as an electronic pill dispenser. To work with these additional sources of information that, as mentioned above, would allow to obtain confidence information about correct task execution by the user, a few type of messages sent from the device to the ECA could be sufficient to cover almost all cases: "action in progress", "action done" and "user inactive". Each of these messages would require to contain at least two information items: an action name and a confidence indicator. In some cases, there would also be a need to send information from the ECA to the device, such as notifying it that the user is expected to perform Action X.

Better off-the-shelf hardware and software

During the course of this work, some new APIs for cloud-based ASR have been released. We have used one (provided by Microsoft[2]) to implement a new speech recognition component in LOUISE, as the speech recognizer used in the experiments reported in Chapter 6 was not reliable enough, but were unable to test it thoroughly, though we think it will likely improve the overall functioning of the system. Similarly, a newer version of the Kinect sensor than the one used in this work is now available, along with new and more reliable behavior analysis functionalities. As mentioned in the recommendations, computer hardware and software evolve very fast,

[1] http://www.bluefrogrobotics.com/en/home/
[2] http://www.microsoft.com/cognitive-services/en-us/speech-api

and new exciting tools are often released, especially in the past few years, since software giants, such as Google or Microsoft, and a myriad of smaller companies have started to offer cloud-based “artificial intelligence” services. As a result, given that ECAs rely on these technologies, systems such as LOUISE will be constantly improved through hardware and software updates, while keeping costs low. Until recently, this was limited to ASR and user tracking, but, nowadays, countless open-source or commercial solutions for applications such as emotion recognition from facial expression or voice tone analysis are already available, and a lot more are about to be. This is why it is more important than ever to keep track of the novelties in the field, because the piece of software that will take ECAs to a whole new level might be right around the corner.

8.3.3 Virtual Promenade

Thanks to participatory design, we could produce a design for the Virtual Promenade system which will likely be usable by almost all elderly patients with PFS. Nevertheless, it could be improved in several aspects: creating more content, with more meaningful tasks and more stressful situations, such as city environments with pedestrians and cars, slippery floors or stairs; exploring other means of interaction and richer haptic sensations to reinforce the impression of walking while seating; and using 3D displays or virtual reality helmets to increase immersion.

Adding more game mechanics and content

We have seen in our design phases that people asked for more meaningful tasks than picking up virtual cubes. There are a number of ways it could be done: the player could have to pick up more meaningful objects, such as flowers for a bouquet; the task could consist in following a dog for it not to get lost; it could also consist in taking a child to a certain place, such as a toy store or the zoo. Doing so will likely enhance engagement in the activity, which may lead to better immersion in the virtual environments. In addition, it would be more motivating if the game included scores, feedbacks and rewards.

As the activity is meant to be performed frequently, over several weeks, it is necessary to add more content than there already is, otherwise patients may quickly get bored, even if they have dementia, as, though they may not remember the activity, their procedural memory, which still works, makes them more expert and the tasks become unchallenging. New environments with tasks that get more challenging than in the current levels and/or correspond to more stressful situations are required to keep people motivated and progressively have them stroll in environments that get more and more stressful, as their fear of falling decreases.

Body and haptics interaction with the game

To reinforce the impression of walking, we would like to explore the benefits of having players perform walk-like movements with their arms and/or their feet to make their avatar move. This however is more challenging than using a simple game controller, especially for synchronizing the movements of the seat with the ones of the patient. In addition, it is not guaranteed that it will actually increase immersion. For instance, in a study by Nabiyouni *et al.* [144], the authors compared three locomotion techniques for virtual reality applications: "non-natural", using a game controller; "semi-natural", using a human-size hamster ball called Virtusphere, that allows to walk while staying in place; and "natural", where people were actually walking in a room equipped with motion capture sensors. It turned out that the semi-natural condition offered less fidelity and decreased performance, compared to the non-natural condition, which may decrease the quality of immersion. However, in our case, this means of interaction could be beneficial, as it would require patients to do some physical exercise.

Lastly, people with PFS often present a postural disorder called "retropulsion" or "backward desequilibrium", which is a known symptom the psychomotor disadaptation syndrome [141]. Therefore, a simple modification that could improve the therapeutic qualities of the system would be to add a pressure sensor on the backrest of the chair and prevent the avatar from moving if the patient rests his or her back against it. This would force patients to hold their back, which would stimulate their back muscles, and adopt a position that may cause them to be anxious, as patients with PFS often lean backward when sitting, stiffly pushing their back against the backrest, out of fear of falling forward.

3D displays and head-mounted displays

To increase immersion, it would be worth experimenting with 3D displays, such as 3D television sets. The ideal would be to use one that does not require spectators to wear polarized glasses, as this element would cause hygiene-related issues for use in care environments, but this feature is only available on high-end products and this would multiply the cost of the system by a factor of at least 4. In addition, due to visual impairment, such as cataract, which is very frequent in older adults, it is likely that a fair proportion of patients would not be able to see the stereoscopic effect, no matter what technology of 3D display is used.

A cheaper solution would be to use a virtual reality (VR) helmet. However, we have crossed-out that option *a priori*, as it is not likely to be suitable for this application, for use in a medical context. There are several reasons for this: VR helmets cause cyber-sickness to many people, even young ones, mostly due to latency; they may be too heavy for frail older adults; their use in a medical context would cause hygiene-related issues; and, to the best of our knowledge, there is very little available data about how people with dementia would react to immersion with a helmet [122], except that older adults do not particularly experience more cyber-sickness than younger adults [235]. This technology being quite young, some products that address most of these issues may become available in a near future, as well as some literature about the use of VR helmets with older adults with dementia. Then, it would be interesting to test VR helmets for the Virtual Promenade system.

8.3.4 Combining both approaches of virtual humans

The two approaches to virtual humans that we have considered as being distinct are in fact complementary and have to be merged. Indeed, as we have seen in Section 7.7.3, patients who participated in the Virtual Promenade tests did not really want to read the written instructions and had difficulties navigating them (displaying the next or previous instruction). To explain how to play, give instructions for tasks and motivate patients; it would be worth considering including the LOUISE ECA in the virtual strolling game of Virtual Promenade. In addition, we have not observed many situations in which older adults with cognitive impairment lose track of what they are doing and stop doing the activity when we tested the use case scenarios we have proposed for LOUISE, whereas it happened several times with two patients who participated in the pilot study of the Virtual Promenade system. This suggests that the attention management capabilities of LOUISE would have more value when the ECA is helping patients perform a task that requires sustaining one's attention for several minutes than in face-to-face conversation. In our Virtual Promenade system, LOUISE could be embodied as a virtual physiotherapist or virtual physician character. Its feature of step-by-step task guidance with example videos would be very useful to show how to handle the controls. In addition, the ECA could ask people what tasks they want to do and manage the session. This would allow rehabilitation professionals to let their patients play the game more autonomously; they could then take care of other patients during that time, while keeping an eye on them. Doing this would create a very interesting mixed-interaction paradigm, making use of both verbal behaviors, nonverbal behaviors and actions on the game controller. However, though some degree of automation saves time for the therapist, having patients playing the game in full autonomy is not desirable, as it would lower the acceptance of the system by professionals, out of concern of seeing their jobs replaced by machines. In addition, this would require a tremendous amount of development time and testing, which would make the system too expensive.

8.4 An ethical epilogue

The introduction of smartphones and the availability of fast Internet connection everywhere and at any time has started to transform our way of life. This allows to carry in our pockets small devices from which we can access a myriad of very useful (or very useless) services, which are often free of charge for the end user. By doing so, we knowingly accept to trade some of our privacy for practical applications, such as GPS guidance or easily finding restaurants, bars or shops in our vicinity. In addition, via the free and freemium business models, service providers collect a lot of private information that they sometimes sell to third-parties. Some of these services, as well as some specific ones, could also be useful for older adults with dementia, sometimes more critical to their safety than for healthy people. But would/do they understand the implications of using smartphone-based applications or systems such as emergency call necklaces with GPS tracking in terms of invasion of their privacy?

In addition, it is often family members that take the decision of buying assistive technologies for their loved ones. As their main concern is usually the safety of their relatives, some are willing to consider very intrusive applications or systems, such as accessing the video produced by a webcam placed in the patient's home to check on them or using monitoring systems that keep track of the person's activities (does he or she eat well, drink well, take his or her medicines, go to the bathroom, etc?). We think that it is therefore very important that people with dementia be informed of the implications of using such systems and consulted before these are imposed to them, at least as long as they are able to choose for themselves. In a way, equipping older adults with assistive technologies is like moving them to a "virtual nursing home": when they are not involved in the decision, some dreadful results are to be expected. When people are moved to a specialized institution without being consulted or against their will, they may get very depressed. With assistive technologies, odds are that the person will put the device away and refuse to use it or even (attempt to) destroy it. We believe that participatory design of assistive products and services allows older adults with dementia to be part of the conversation from the start, so that their concerns and wishes can be addressed and taken into account upstream.

During the design phases of LOUISE, we have asked people if they would want the ECA system to be always on. It turns out that the majority of older adults who gave us their opinion would not. However, for some of them, when the illness reaches a moderate to severe stage, the system would be useless if it is only activated when they turn it on, as it is likely that they would forget to do so. It is therefore important to consider this aspect with care and to consult people to find the best compromise between the quality and nature of the services offered by ECAs and the respect of people's choices. For older adults with dementia, who may not figure out how to turn the system off when it bothers them, it could be useful to come up with solutions to give them some control. For instance, it could be useful to detect anger and verbal inputs such as "stop bugging me" or similar phrases using coarse language that disinhibited people are susceptible to formulate.

Lastly, as we have mentioned in Section 3.1.4, there are some well-known manipulation techniques, that are usually conveyed verbally, and that can be used to influence people's actions [97]. Thanks to these techniques, and to the perceived trustworthiness of ECAs, these artificial agents, as they evolve from a technological point of view, will soon be capable of manipulating patients with dementia. While this may sound awful, in dementia care, it can prove useful, even beneficial, for some patients: in situations of care refusal for instance, an ECA could influence people so they feed, hydrate, take their medicines, etc. It could also be used to influence people so they adopt a healthier lifestyle. For instance, in the Swedish science-fiction television show *Äkta människor* (Real Humans), an elderly person is given a new geriatric care robot, which looks and behaves like a real person, to replace the one he had, because it had malfunctions. This new robot, unlike its predecessor, is programmed to do and make the person do what is considered as being good for him: it refuses to prepare its favorite dishes and serves him healthier food instead; it controls the amount of wine he drinks and restricts him when it thinks he had enough; and it even mixes sleeping pills with his food, so he goes to sleep early. Needless to say, the character does not appreciate his new robot and wishes he could have the old one back. However, the situation improves when his family members come visit him and ask the robot to be less strict. Manipulation is definitely a slippery path to get on, even with good intentions. It is therefore necessary to ask questions about how much manipulation is desirable, where the ethical boundary that should not be crossed is, and whether an assistive robot or ECA should do what people want it to do or what is good for them. It is therefore necessary to establish best-practices recommendations for the use of social agents in elders' care, which could be done thanks to a living lab approach.

Conclusion

Dans cette thèse, nous avons rapporté nos travaux sur l'utilisation de la technologie des humains virtuels pour améliorer le soin aux personnes âgées. Dans notre approche, nous avons considéré que les humains virtuels peuvent jouer deux rôles distincts dans les scénarii de soin aux personnes âgées : représenter une intelligence artificielle qui habite ce monde virtuel (comme non-soi) ou être l'avatar de la personne dans le monde virtuel (comme soi). Cela nous a conduit à mener deux études de cas : (1) explorer l'utilisation des ACA, des humains virtuels capables de communication verbale et non-verbale avec leurs utilisateurs, dans le but d'améliorer l'accessibilité des technologies de l'information et la communication pour les personnes âgées ayant des troubles cognitifs ; et (2) étudier l'utilisation d'avatars dans des environnements virtuels pour exposer les personnes âgées atteintes de SPC à des situations de marche, afin de réduire leur peur de chuter. Plus spécifiquement, nous avons conçu, développé et testé un ACA, appelé LOUISE, dont la fonction est de servir d'interface utilisateur dans les systèmes d'assistance cognitive, et un système de thérapie en réalité virtuelle, composé d'un jeu sérieux et d'un fauteuil haptique. Tout au long de ces travaux de recherche, nous avons suivi une méthodologie de conception participative living lab, pour laquelle nous avons défini des étapes adaptés pour s'accommoder au mieux à notre public cible.

Dans ce dernier chapitre, nous commençons par présenter les conclusions de chacune de nos deux études de cas. Puis, nous présentons une conclusion générale de ces travaux. Ensuite, nous proposons des idées pour des travaux futures. En dernier lieu, nous terminons cette thèse par une discussion sur les implications éthiques à long terme de la technologie des humains virtuels.

Études de cas sur les humains virtuels

Dans cette section, nous résumons et discutons nos travaux et principaux résultats, pour chaque étude de cas.

LOUISE

*Les personnes vivant avec une démence ont besoin d'aide pour accomplir les activités de la vie quotidienne. De ce point de vue, il a été montré que les ACA sont une solution prometteuse (voir section 2.2.1) pour les assister, afin qu'elles restent autonomes aussi longtemps que possible, et ont l'avantage d'être peu onéreux. C'est pourquoi notre collègue, Dr. Pino, chercheur en psychologie à l'hôpital Broca, a initié le projet consistant à en valider la pertinence in situ. Dans l'étude de cas du projet LOUISE (*LOvely User Interface for Servicing Elders*), rapportée dans les chapitres 5 et 6, nous avons mené les phases de conception et de développement d'un ACA, vu comme une boite à outils, destiné à être intégré dans les technologies d'assistance cognitive. Ces travaux ont été centrés sur les spécificité de l'interaction entre l'ACA et les personnes âgées ayant des troubles cognitifs, particulièrement pour gérer les troubles attentionnels, la désorientation, la mémoire à court-terme réduite et l'aphasie. Au cours de cette étude, nous avons également collecté des informations concernant les souhaits et attentes des parties prenantes pour ce type d'interface, ainsi que sur l'utilité de certains des scénarii d'assistance dans lesquels des ACA seraient intégrés.*

Étude en "magicien d'Oz"

Nous avons commencé par mener une étude en "Magicien d'Oz" avec un prototype semi-automatique élémentaire de LOUISE, incluant une stratégie de gestion de l'attention fondée sur la mesure du niveau d'attention de l'utilisateur en temps réel, grâce à un nouvel algorithme que nous avons proposé, et des rappels attentionnels effectués par l'ACA en cas de perte d'attention, jusqu'à ce qu'il ou elle dirige à nouveau son attention sur l'ACA. Nous avons testé ce système avec 14 professionnels des technologies d'assistance et 9 personnes âgées, dont la plupart avaient des troubles cognitifs légers à modérés, et effectué une analyse anthropologique des enregistrements vidéo pris lors des tests avec les personnes âgées. Nous avons fait les observations suivantes :

1. *les personnes ayant des troubles cognitifs sont capables d'interagir avec un ACA ;*

2. *les personnes âgées ont majoritairement réagi positivement à l'ACA ;*

3. *les personnes âgées atteintes de démence ont tendance à interagir de manière plus sociale que les personnes âgées sans troubles cognitifs, puisqu'elles ont prononcé plus de mots et fait plus de développements thématiques ;*

4. *notre stratégie de re-capture de l'attention est efficace mais pas suffisante, puisque, dans le cas du sujet le plus fortement handicapé, cela*

n'a pas permis de poursuivre l'interaction, car le sujet était désorienté ;

5. *notre système d'estimation d'attention simple, rapide et bon marché est précis à plus de 80% ;*

6. *les performances de notre estimateur d'attention semblent indépendantes de l'âge, ce qui suggère qu'il n'est pas nécessaire de créer une méthode sur mesure pour les personnes âgées.*

Questionnaires et groupes de discussion

*Afin de collecter des informations sur les spécifications de conception, nous avons demandé leur avis aux participants de nos expériences et au public, lors de démonstrations faites à l'occasion de deux évènements publiques. Au total, 37 personnes ont répondu à notre questionnaire. Ensuite, nous avons organisé un groupe de discussion (*focus group*) avec 9 personnes âgées, dont 3 avaient des troubles cognitifs légers. Nous avons trouvé que :*

1. *les personnes interrogées pensent que les ACA seraient le plus utile en tant qu'assistants personnels, c'est-à-dire pour aider à gérer un emploi du temps, rappeler les prises de médicaments, donner accès à la visioconférence, lire des courriels, etc. ;*

2. *les autres applications jugées utiles incluent le guidage pas-à-pas dans des tâches complexes, la commande d'appareils ménagers ou la centralisation de services de compensation et de stimulation cognitive ;*

3. *les médecins que nous avons consultés de manière informelle, lors de réunions d'équipe, ont proposé que les ACA soient utilisés dans la gestion des troubles psycho-comportementaux, tels que les cris, et lisent et apprennent des poèmes aux patients, pour les stimuler ;*

4. *les personnes interrogées ne voudraient pas que des ACA soient installés sur un appareil dédié, mais préfèreraient qu'ils soient utilisés sur un appareil qu'ils possèdent déjà ou qui ne serve pas qu'à cela, tels qu'un ordinateur, une tablette ou un téléviseur, pour les options les plus plébiscitées ;*

5. *les ACA doivent être plaisants, voire amusants, expressifs et souriants ;*

6. *les personnes interrogées voudraient que l'ACA ait l'apparence d'une jeune femme ou d'un robot humanoïde ;*

7. *la plupart des répondants voudraient qu'il soit possible de personnaliser l'apparence de l'ACA.*

Gestion de tâches assistée

Étant donné ces informations et les résultats de notre étude en magicien d'Oz, nous avons assemblé un prototype entièrement automatisé de LOUISE qui inclut :

- *des capacités d'animation riches, grâce à l'utilisation de SmartBody, un réalisateur BML au niveau de l'état de l'art ;*
- *l'affichage d'images et de vidéos d'exemples ;*
- *l'ajout facile de modèles de personnages ;*
- *un module de gestion d'interaction ad hoc permettant de décrire les scénarii d'interaction selon une syntaxe XML dédiée.*

La gestion d'interaction sur mesure pour les personnes âgées atteintes de troubles cognitifs que nous proposons permet de gérer l'attention, d'effectuer des rappels de contexte à la suite d'un moment de distraction de l'utilisateur et de préciser à l'utilisateur les réponses possibles à une question donnée, si les personnes ne répondent pas de manière adéquate. Nous avons utilisé ce système pour explorer les spécificité de la gestion de dialogue pour les personnes âgées démentes. Nous avons créé des scénarii d'interaction, fondés sur des cas d'usage réalistes, afin d'étudier deux tâches intervenant dans la plupart des applications – choix entre plusieurs items et guidage pas-à-pas – pour lesquelles nous avons proposé des structures de dialogue. Nous avons ensuite mené une étude d'utilisabilité, afin de raffiner et valider notre système, avec 14 personnes âgées ayant des troubles cognitifs légers à sévères. Nous avons observé les éléments suivants et avons fait des ajustements :

1. *tous les participants ayant des troubles légers à modérés ont été capables d'interagir avec LOUISE (MMSE > 10) ;*
2. *les participants ont fait des retours très positifs (plus de 3,3/4 de moyenne sur l'échelle d'appréciation) concernant le charme de l'ACA LOUISE, son apparence, la clarté de sa voix de synthèse et le rythme de la conversation ;*
3. *l'ACA peut avoir l'initiative du dialogue, mais les personnes avec TCL souhaitent parfois prendre l'initiative ;*
4. *notre stratégie de choix d'item est efficace, puisque tous les participants ont pu faire des choix et, de plus, des petits changements à faire pour l'améliorer ont pu être identifiés ;*
5. *notre stratégie pour le guidage de tâches pas-à-pas a été améliorée pour les personnes avec TCL, mais ces changements ne conviennent pas aux personnes ayant des troubles cognitifs plus sévères ;*

6. *la distance minimum nécessaire au capteur Kinect pour Xbox 360 pour suivre les corps et visages des utilisateurs est trop grande pour que le système fonctionne de manière fiable dans des conditions réelles d'utilisation ;*

7. *les performances de la reconnaissance vocale hors-ligne ne sont pas satisfaisantes ;*

8. *la qualité de la reconnaissance vocale est critique, car les erreurs sont sources de confusion pour les personnes ;*

9. *nos observations confirment que les personnes ayant des troubles cognitifs ont tendance à faire des réponses plus élaborées, de manière sociale, comme vu dans l'étude en magicien d'Oz.*

Grâce à notre méthode de conception participative living lab, nous avons pu créer un ACA que les personnes âgées ont apprécié utiliser, dévoiler plusieurs problèmes d'utilisabilité et en résoudre certains. Nous avons affirmé dans le chapitre 2 que, sous certains aspects, l'implémentation d'un tel système pourrait être plus simple lorsqu'il s'adresse à des personnes âgées atteintes de troubles cognitifs que pour des personnes jeunes et sans handicap, particulièrement concernant la qualité de rendu graphique et la gestion d'interaction. Nos observations au cours de cette étude de cas tendent à confirmer ces hypothèses, car nos participants étaient satisfaits du visuel de l'application, bien que les animations auraient pu être de meilleure qualité et que la qualité de rendu offerte par Panda 3D, le moteur de jeu que nous avons utilisé, soit en dessous de la moyenne. En ce qui concerne la simplicité de la gestion d'interaction, un système simple peut répondre aux besoins dans la plupart des cas, mais des possibilités plus évoluées restent à explorer. De plus, les cas d'usage que nous avons testé sont très simples et nous pensons que, avec quelques améliorations, particulièrement pour le guidage de tâches pas-à-pas, l'ACA pourrait déjà se montrer utile pour les personnes ayant une démence modérée. Plus généralement, cela suggère que même un ACA à l'intelligence limitée pourrait fournir des services utiles aux personnes démentes, puisqu'elles ont besoin d'aide pour faire des activités simples.

Virtual Promenade

Comme le SPC, une combinaison de symptômes psychologiques et psychomoteurs résultant d'une chute, ressemble au SSPT sur plusieurs aspects, nous avons proposé d'utiliser la thérapie en réalité virtuelle pour les patients âgés ayant peur de chuter, comme suggéré par le Dr. Bloch, un gériatre de l'hôpital Broca, à l'origine du projet [124]. Dans notre seconde étude de cas, nous avons mené la conception participative d'un système, appelé Virtual Promenade*, alliant un jeu sérieux à un fauteuil haptique dont l'assise bouge* *en imitant le mouvement des hanches de la marche humaine. Le jeu sérieux que nous avons développé est un jeu de promenade virtuelle, dans lequel les patients contrôlent un avatar humain à travers plusieurs environnements virtuels. Le fauteuil haptique est connecté au jeu, afin que la vitesse de mouvement du siège corresponde à la vitesse de marche de l'avatar du joueur.*

Séances de test

Nous avons commencé par assembler un prototype comportant une version élémentaire du jeu et avons organisé des séances de test du jeu avec des personnes âgées. En se basant sur les retours, opinions et souhaits des participants, nous avons pu réaliser une première version du jeu.

Grâce à ce processus de co-conception avec les personnes âgées et leurs aidants, nous avons pu créer un jeu de promenade virtuelle auquel les personnes âgées peuvent jouer et qu'elles apprécient. Ce faisant, nous avons fait les observations suivantes :

1. *les personnes âgées ont souhaité que les environnements virtuels soient plaisants, plutôt que réalistes, et cela a été identifié comme un facteur de motivation ;*
2. *les personnes âgées souhaitaient pouvoir s'identifier un minimum à la forme corporelle de leur avatar ;*
3. *un joystick de simulateur de vol n'est pas adapté pour les personnes âgées s'il faut de la force pour le manœuvrer, comme cela fut le cas avec le joystick Logitech utilisé dans nos tests ;*
4. *la manette de jeu de Nintendo 64 a semblé bien adaptée pour les personnes âgées en bonne santé, bien que son utilisation nécessite une motricité fine.*

Groupes d'échange et suivi de professionnels

*Après cette première phase, nous avons mené deux groupes d'échange (*focus group*) avec des professionnels de la rééducation : kinésithérapeutes et psychomotriciens. Après avoir rencontré chaque groupe, nous avons implémenté les changements demandés. Enfin, nous avons suivi des kinésithérapeutes dans une journée de travail ordinaire, afin de mieux comprendre leur activité et le terrain de déploiement ciblé.*

La consultation et le suivi de professionnels ont permis de raffiner la conception, de mieux comprendre comment ils perçoivent l'introduction d'un système tel que Virtual Promenade sur leur lieu de travail, comment ils pensent que cela pourrait s'intégrer à leurs pratiques de soin et d'obtenir des informations sur les contraintes du terrain, qui sont principalement l'espace disponible, la rapidité de mise en œuvre, la sécurité et l'autonomie *des patients pour jouer au jeu. Ce dernier élément est particulièrement intéressant, car nous avons observé que les kinésithérapeutes s'occupent généralement de deux patients à la fois, voire trois : un système qu'ils n'ont qu'à démarrer et qu'ils peuvent laisser le patient utiliser seul, pendant qu'ils s'occupent d'un autre, serait la solution idéale pour son intégration en douceur dans les pratiques de rééducation.*

Étude pilote

Enfin, nous avons testé le système dans des conditions écologiques, avec des patients hospitalisés en service de rééducation. Sur les 10 participants recrutés, seuls 8 ont essayé le système pour au moins une séance. Nous avons observé que la manette de jeu de Nintendo 64, qui semblait adaptée lorsque nous avons testé avec des personnes âgées en bonne santé, ne convenait pas pour les patients avec une faible motricité fine. De plus, contrôler le personnage pour visiter des environnements virtuels s'est avéré trop difficile pour les personnes âgées ayant des troubles cognitifs modérés. Par conséquent, nous avons remplacé la manette de Nintendo 64 par un joystick de simulateur de vol Thrustmaster, plus souple que le modèle Logitech utilisé dans la première phase de l'étude, que nous avons collé sur une table de chevet d'hôpital, parce qu'il était trop léger, et, plus important, nous avons créé un mode "facile" dans lequel l'avatar du joueur suit un chemin prédéfini, de manière à ce que le joueur n'ait à se préoccuper que de la vitesse de marche. De plus, nous avons maintenu la compatibilité avec les autres contrôleurs de jeu testés au cours de l'étude. Ces changements ont permis à des personnes âgées atteints de troubles cognitifs modérément sévères ou de la maladie de Parkinson d'utiliser le système avec succès. De plus, nous avons trouvé que :

1. *les participants étaient satisfaits de l'utilisabilité du jeu ;*
2. *les personnes âgées ont apprécié le jeu et étaient satisfaites des environnements, du rendu graphique et des animations ;*
3. *les participants ne voulaient pas vraiment lire les instructions écrites et l'expérimentateur a du intervenir pour leur expliquer comment jouer ;*
4. *la plupart des participants n'ont pas eu l'impression de marcher pendant l'activité ;*
5. *l'utilisation du système n'a causé de l'anxiété que dans un seul cas, mais cela peut également être du au stress suscité par le transfert entre le fauteuil roulant et le fauteuil expérimental ;*
6. *nous n'avons pas observé d'effets sur la peur de chuter ;*
7. *Le FES n'a pas semblé être un indicateur de succès fiable car les scores obtenus ne reflètent pas correctement l'intensité de la peur de chuter*

observée, particulièrement chez les patients ayant des troubles cognitifs, puisqu'ils ont eu du mal à se projeter dans des activités qu'ils ne réalisent plus, qu'ils ont répondu au hasard, ou qu'ils ont dit ce qu'ils pensaient que nous voulions entendre ;

8. *les accoudoirs réglables posent un problème de sécurité, car les personnes s'appuient dessus lors des transfert, et ils risquent de se dérober sous eux.*

Nous n'avons pas suffisamment de données pour nous prononcer sur les éventuels effets thérapeutiques du dispositif, car cela nécessite de créer un programme thérapeutique et de mener des essais contrôlés randomisés. Cependant, notre intuition est que le système n'est probablement pas assez immersif pour que les personnes assimilent l'activité proposée par Virtual Promenade à des situations de marche réelles. Cela peut être lié à plusieurs facteurs : l'utilisation de la vue à la troisième personne semble moins efficace que celle de la vue à la première personne, car les participants ont rapporté une sensation de marche plus grande dans la deuxième séance avec la vue à la première personne ; les environnements retenus (forêt et parc) ne permettent peut-être pas aux personnes de faire le lien avec des situations de leur vie quotidienne ; il se peut que l'activité ne rappelle pas suffisamment la marche ; les mouvements de l'assise sont possiblement trop subtils ; et le fait qu'ils choisissent un avatar est susceptible de créer une distance entre l'utilisateur et ce qui se passe dans le monde virtuel. Nous proposons donc des améliorations dans les travaux futurs pour combler ces lacunes et rendre l'expérience plus immersive.

Conception living lab pour les personnes âgées

D'une manière générale, dans les deux études de cas, le suivi d'un processus de conception participative living lab a permis de créer des prototypes, fondés sur les humains virtuels, que les personnes âgées ont apprécié et de garantir une bonne utilisabilité, particulièrement dans le cas de Virtual Promenade. De plus, les tests des systèmes en conditions écologiques avec les utilisateurs ciblés ont permis de mettre en lumière plusieurs lacunes de conception, qui ont parfois pu être résolues tout de suite, grâce aux connaissances acquises en travaillant sur le terrain ou en impliquant les professionnels de santé dans le processus de conception. Par exemple, dans le projet Virtual Promenade, les tests avec des patients hospitalisés ont permis de révéler des problèmes d'utilisabilité et de sécurité qui n'auraient pas pu être observés avec les personnes âgées en bonne santé. D'autre part, les adaptations que nous avons faites pour implémenter la méthode living lab se sont également avérées utiles : l'utilisation d'éléments "pris sur l'étagère" a permis des itérations rapides dans la conception des prototypes, tout en *limitant les coûts ; l'implication des personnes âgées en bonne santé comme "proxy" des utilisateurs ciblés a permis d'obtenir des informations utiles, tout comme l'implication d'un public élargi, lors de démonstration publiques ; et le suivi de professionnels sur le terrain a permis d'établir avec eux une relation de confiance, d'améliorer la communication, et d'obtenir des informations pertinentes pour la conception. Ce dernier point souligne l'importance de l'acculturation dans le contexte des travaux pluridisciplinaires : lorsqu'il est question de collaborer avec des professionnels d'autres disciplines, il est d'une importance critique de comprendre leur activité, leurs préoccupations et intérêts, et d'assimiler le vocabulaire de leur profession pour communiquer efficacement. Cela est un processus long, mais nécessaire, car cela permet de passer outre la connaissance purement théorique des pratiques de soin, pour saisir la réalité du terrain, ce qui est l'un des principes fondateurs de l'approche living lab.*

Au cours de ces études de cas, nous avons observé que la plupart des personnes âgées sont sensibles à l'esthétique des personnages humains et environnements virtuels. En revanche, le réalisme graphique, la qualité de l'animation et du rendu n'ont pas semblé aussi importants. Cela est cohérent avec ce que nous avons affirmé dans la section 2.1.5 : comme (pour le moment) les personnes âgées sont peu familières des jeux vidéos AAA, leurs attentes en termes de qualité graphique et de réalisme sont peu élevées. Si nous devions ne conserver qu'un seul qualificatif de ce que les personnes âgées souhaitent pour une application fondée sur les humains virtuels, ce serait "plaisant". En effet, elles ont majoritairement demandé que les personnages de l'ACA LOUISE soient souriants et sympathiques. De manière similaire, elles ont trouvé que le premier environnement urbain proposé dans le projet Virtual Promenade était froid et peu accueillant, et elles ont demandé des lieux plus agréables, avec des arbres et des fleurs. Bien entendu, cette notion est assez subjective et dépend des goûts personnels de chacun. C'est pourquoi dans les applications telles que LOUISE, qui ont un seul ou quelques utilisateurs, il semble important que le logiciel permette la personnalisation de l'aspect visuel. Pour des applications qui sont supposées être utilisées par des personnes plus nombreuses, comme c'est le cas du jeu de marche virtuelle dans le système Virtual Promenade, la meilleure approche est de proposer des visuels que la plupart des personnes apprécieront, même s'il s'en trouvera toujours qui ne les aimeront pas.

En travaillant au contact des personnes âgées, nous avons observé que peu d'entre elles sont particulièrement réticentes aux nouvelles technologies. Cependant, étant donné la complexité de nos systèmes, il s'est parfois avéré difficile de leur expliquer, de manière compréhensible pour elles, ce que nous voulions leur faire essayer et comment cela pourrait leur être utile, en particulier lorsqu'elles étaient atteintes de troubles cognitifs. Le public le plus difficile est celui des patients hospitalisés, qui ont tendance à être déprimés, fatigués par les activités de rééducation et malades. Lorsque nous recrutions *des sujets pour l'étude pilote de Virtual Promenade, plus de la moitié des patients auxquels nous avons proposé de participer ont refusé. La plupart d'entre eux n'étaient intéressés que par leur date de décharge de l'hôpital ; quelques-uns nous ont dit qu'ils ne voulaient pas être des "cobayes" ; certains avaient peur que cela les fatigue ; d'autres étaient inquiets de ne pas réussir à utiliser le système ; mais seulement quelques personnes ont dit qu'elles n'étaient pas intéressées parce qu'elles n'aiment pas les technologies informatiques.*

Il est intéressant de noter que, comme nous portions des blouses blanches, nous étions parfois assimilé à des soignants, et certains patients qui ne souhaitaient pas participer à nos études ont cherché de fausses excuses, plutôt que de simplement décliner notre proposition. Par exemple, un patient nous a dit que le médecin lui avait recommandé d'attendre que sa fracture soit mieux remise avant de participer à l'étude, alors que c'était justement ce médecin qui nous l'avait indiqué comme étant un bon candidat. Ces personnes étaient probablement inquiètes de décevoir l'équipe soignante en refusant de participer à l'étude. Cela soulève des questions éthiques sur la manière dont les patients hospitalisés devraient être recrutés pour les études. Même si nous avons recruté les patients nous-même, le fait que nous portions des blouses peut avoir été la cause de malentendus, bien que nous ayons clairement expliqué que nous étions informaticiens et que nous portions un badge qui indiquait cela. Il est également important d'insister sur le fait que les patients sont tout à fait libres de refuser de participer aux études, ou d'interrompre leur participation à tout moment, sans aucune conséquence, ce que nous avons fait. Nous pensons que, là encore, la perception d'utilité est un facteur déterminant pour les personnes qui participent à nos recherches, puisque les sujets qui ont souhaité essayer Virtual Promenade ont semblé être particulièrement intéressés par les potentiels effets thérapeutiques, bien que nous les ayons clairement informé que nous n'avions aucunes certitudes à ce sujet.

Il nous a semblé que, dans la plupart des cas, les personnes âgées ayant des troubles cognitifs sont parfaitement capables de donner des retours concernant ce qu'ils ont apprécié ou pas lors du test. Cependant, il a été parfois difficile d'obtenir des informations des patients les plus atteints, en particulier lorsque les questions étaient complexes ou nécessitaient qu'ils se projettent dans une situation particulière. De plus, ayant collecté ces informations, il se peut qu'il soit difficile de juger de l'importance à y accorder, car il semble que, parfois, les participants disent à l'expérimentateur ce qu'ils pensent qu'il ou elle souhaite entendre, au lieu de ce qu'ils pensent réellement, ou répondent au hasard, lorsqu'ils n'ont pas d'opinion ou ne comprennent pas la question. Par exemple, lorsque nous les avons interrogé sur l'utilité des applications envisagées pour LOUISE, leur acceptation de son système de perception ou, dans le projet Virtual Promenade, si la vue à la première personne était plus immersive que celle à la troisième personne, *la plupart ne savaient pas vraiment quoi répondre. Obtenir des réponses honnêtes, parfois obtenir des réponses tout court, requiert du temps, de la patience, de reformuler la question de manière à ce que la personne comprenne, de remontrer les éléments qui font l'objet de la question, au cas où ils auraient oublié ou n'auraient pas compris, et de discuter avec eux leurs réponses pour éviter les malentendus. Sur cet aspect, impliquer des personnes âgées sans troubles s'avère particulièrement utile, car il est bien plus facile de recueillir leurs opinions sur des sujets complexes. De plus, bien que nous pensons qu'il est très important que les personnes démentes soient associées de manière directe aux décisions qui impactent leur mode de vie, ce sont bien souvent les aidants et les proches qui prennent les décisions d'achat de technologies d'assistance, parfois sans les consulter. Il est donc indispensable de collecter également les retours des aidants familiaux. Cependant, il est également critique de s'assurer que les patients, qui sont les utilisateurs finaux du système, vont être en mesure de l'utiliser et vont l'apprécier, puisque, sinon, il sera vite abandonné. C'est pourquoi à la fois les patients et leurs aidants doivent prendre part au processus de conception, et leurs opinions et souhaits doivent avoir le même impact.*

Enfin, en se basant sur l'expérience que nous avons acquise en menant des travaux de conception participative living lab, nous partageons les principaux enseignements à retenir en proposant des recommandations, détaillées dans l'encadré du chapitre 8.

Travaux futurs

Les deux projets menés au cours de cette thèse ont déjà produit des résultats encourageants. Cependant, des travaux de recherche supplémentaires sont nécessaires sur plusieurs aspects. C'est pourquoi nous proposons des idées d'amélioration futures de chacun de nos prototypes, notamment l'ajout d'un thérapeute virtuel dans Virtual Promenade, en y intégrant LOUISE.

Mener de nouvelles études de validation

La principale limite de ces travaux est le petit nombre de sujets inclus dans chaque projet. Cela est lié aux difficultés de recrutement de patients pour nos études et des coûts élevés des campagnes de test, en termes de temps, au moins. Une autre limitation est le peu de comparabilité des résultats au cours d'une étude donnée, puisque le système évolue au cours de l'évaluation. Bien que cela soit un choix délibéré qui s'est clairement montré utile dans le cadre d'une conception participative living lab, des expériences de validation supplémentaires, avec des prototypes stabilisés et à une plus grande échelle, sont nécessaires.

LOUISE

Jusqu'ici, nos travaux dans le projet LOUISE se sont concentrés sur la gestion de l'interaction et de l'attention. Cependant, le mode de gestion de l'interaction que nous avons proposé est relativement élémentaire et doit être amélioré, en particulier pour prendre en compte la variabilité interpersonnelle de la démence et son évolution au cours du temps. De plus, outre l'amélioration de la robustesse du système, il pourrait être intéressant d'ajouter d'autres sources d'information, comme la détection d'émotion ou des données provenant d'appareils externes, connectés à l'Internet des objets (pilulier, réfrigérateur, etc.), ce qui pourrait ajouter de la valeur au système, en particulier pour les personnes atteintes de démence modérée à sévère, qui ont les plus grandes difficultés pour accomplir des tâches simples et pour s'exprimer. Aussi, les participants âgés du groupe d'échange mené dans le cadre du projet LOUISE ont exprimé le souhait d'une interface plaisante, voire amusante, ce qui pourrait augmenter l'appréciation et l'acceptabilité du dispositif. Puisque LOUISE est un outil générique, il pourrait également être utilisé dans des systèmes plus conséquents, telles que des maisons intelligentes, embarqué sur des robots mobiles et utilisé dans d'autres travaux de recherche. Enfin, il est utile de garder en tête que, moyennant de menus travaux, le système pourrait être facilement amélioré, au fur et à mesure de l'évolution des composants logiciel et matériel dont il dépend.

Gestionnaire d'interaction adaptable

Comme nous l'avons vu dans le chapitre 6, il apparait que le schéma interactionnel qui convient aux personnes âgées avec des TCL peut ne pas fonctionner pour des personnes qui souffrent de troubles cognitifs plus sévères et vice versa. De plus, dans les cas de troubles cognitifs sévères, un ACA d'assistance risque d'être confronté à des demandes hors-sujet, liées au questionnement répétitif. En effet, les personnes démentes sont fortement susceptibles d'être obsédées par des soucis personnels, telles que la prochaine visite de leurs proches ou la date de leur décharge de l'hôpital. Bien que certaines des questions les plus fréquentes puissent probablement être identifiées, cela reste dépendant de chaque patient. Cet élément, ainsi que nos observations et les informations que nous avons collectées, suggèrent que la personnalisation est une problématique d'importance majeure dans le cadre de l'utilisation des ACA dans les systèmes d'assistance cognitive. En effet, nous pensons que le système doit pouvoir s'adapter à au moins trois facteurs : la sévérité des troubles cognitifs, qui impacte la manière dont les interactions sont gérées ; les goûts et centres d'intérêt de la personne, qui doivent influencer les contenus et l'apparence ; et d'autres attributs personnels, tels que le nom de l'utilisateur, ses préférences sur la manière dont l'ACA s'adresse à lui, les noms de ses proches, la raison de son hospitalisation, etc. Par *conséquent, le point le plus important auquel les recherches futures devront s'intéresser est celui d'une approche paramétrique pour adapter le gestionnaire d'interaction de LOUISE pour qu'il soit personnalisable et qu'il puisse composer avec des niveaux variables de troubles cognitifs. La solution idéale serait que le système s'auto-configure, en se fondant sur les interactions précédentes, pour s'adapter aux besoins de chaque personne, qui varient au cours du temps, au fur et à mesure que les symptômes de la démence empirent.*

*De surcroît, une plus grande flexibilité devrait être donnée au système, pour prendre en compte au moins deux paramètres : les préférences de l'utilisateur sur la manière dont l'ACA s'adresse à lui, c'est-à-dire s'il préfère être appelé par son nom ou son prénom et s'il préfère être tutoyé ou vouvoyé ; et pour permettre que l'utilisateur prenne l'initiative de la conversation. Le premier élément pourrait être réalisé grâce à l'utilisation d'approches de description de dialogue telles que l'*Artificial Intelligence Markup Language *[10], dans lequel il est possible de mettre des balises XML auxquelles seront substitués des mots ou morceaux de phrases qui sont soit définis à priori (le nom de l'utilisateur, par exemple) ou obtenus au cours de l'interaction. Cela pourrait également permettre d'améliorer la description des arbres de dialogue en limitant la redondance lorsque plusieurs utterances ne diffèrent que par quelques mots, comme, par exemple, lorsqu'on demande de choisir entre de l'eau plate et de l'eau gazeuse ; ou pour rendre la formulation des utterances plus naturelle, en permettant de remplacer certains mots par des synonymes. Ces évolutions pourraient être faite en améliorant le langage XML, appelé AISML, dont nous avons jeté les bases, qui sert à décrire des dialogues à but dirigé. Pour permettre des dialogues à initiative mixte, le gestionnaire d'interaction devra être modifié plus profondément, et reposer sur des approches de gestion de dialogue de type* "frame-based" *ou* "information state" *(voir section 4.5).*

Analyse comportementale de l'utilisateur plus complète

L'amélioration de l'analyse du comportement de l'utilisateur de LOUISE permettrait de mieux gérer les interactions en donnant à l'ACA une plus grande conscience sociale. Les composants d'analyse comportementale additionnels pourraient inclure, entre autres : l'analyse des émotions, qui permettrait de percevoir l'état affectif de l'utilisateur et de s'y adapter ; l'analyse de la prosodie, qui permettrait de produire des comportements d'écoute active appropriés ; la reconnaissance d'activité, qui faciliterait le guidage pas-à-pas de l'utilisateur au cours d'une tâche et permettrait d'obtenir un niveau de confiance sur l'accomplissement correct de la tâche ; et des connexions avec des appareils ou programmes externes, qui serviraient des objectifs similaires à la reconnaissance d'activité de manière complémentaire, ou plus fiable, suivant les applications.

Vers une interface amusante

Des travaux pourraient être menés pour tester plusieurs apparences, animations, voix et effets sonores qui rendraient l'interface plus amusante ou plus plaisante qu'actuellement. Par exemple, des personnages type dessin animé, dont les comportements sont stéréotypés, pourraient amuser les personnes âgées. Cela pourrait également rendre l'ACA plus facile à comprendre, car des mouvements et expressions faciales exagérés peuvent être plus facilement perçus par les personnes démentes. De plus, un bon choix d'apparence pourrait rappeler aux personnes ayant des troubles cognitifs qu'elles n'ont pas affaire à une vraie personne et de cette manière éviter des effets d'"uncanny valley*" ou de fausses affordances, en réduisant les attentes.*

Bien que la santé et le bien-être soient des sujets très sérieux, la ludification et le divertissement peuvent être exploités pour rendre les choses plus faciles et plus plaisantes. À ce sujet, de nombreuses expériences ont été menées et de bons résultats ont été observés pour augmenter la motivation des patients à réaliser des exercices de rééducation (voir section 2.2.3). Il pourrait donc être intéressant d'étudier les effets que cela aurait sur l'adhésion des personnes démentes aux dispositifs de compensation cognitive.

Utilisation de LOUISE ou de ses éléments dans d'autres travaux

*Au cours de cette thèse, LOUISE ou du matériel créé pour le projet ont été utilisés dans d'autres travaux. Le premier est une recherche sur la perception par les personnes âgées des émotions exprimées par un robot doté d'un cou articulé, pour lequel nous avons fourni des rendus du personnage Louise (voir chapitre 5) réalisant des expressions faciales [13]. La tête du robot se composait de trois actionneurs et d'une tablette, sur laquelle les visages étaient affichés. L'expérience, en partie menée dans le Café multimédia, consistait à demander à des personnes âgées de reconnaitre les émotions exprimées par le robot dans quatre conditions différentes et d'indiquer quelle condition elles avaient préférée. Les conditions comparées étaient les suivantes : visage schématique (*smiley*), visage schématique avec mouvement de la tête du robot, visage de Louise et visage de Louise avec mouvements de la tête. Dans certains cas, les participants ont mieux reconnu les expressions du visage de Louise que celles du visage schématique. De plus, environ la moitié des participants ont préféré qu'ils préfèreraient que le robot ait un visage humanoïde (Louise) plutôt qu'un visage schématique. Cela suggère que l'utilisation de LOUISE sur une plate-forme mobile, qui la placerait quelque part entre l'ACA et le robot, serait une piste intéressante à explorer. Par exemple, la tête du robot Buddy de Blue Frog Robotics[3] est doté d'un tablette, et il se rapproche d'un ACA sur roues, avec quelques actionneurs*

[3] http://www.bluefrogrobotics.com/en/home/

supplémentaires lui permettant de bouger sa tête.

La seconde utilisation de LOUISE dans d'autres travaux, qui a déjà commencé, consiste à utiliser une version "magicien d'Oz" de notre deuxième prototype pour des études anthropologiques : nous avons remplacé le module d'analyse comportementale par des entrées au clavier, afin que l'expérimentateur puisse avoir le contrôle complet de l'interaction et simuler la pleine compréhension du langage naturel. Les structures des dialogues sont décrites en AISML, avec pour seule différence que les conditions de transition sont les lettres correspondantes aux touches du clavier sur lesquelles il faut appuyer pour les déclencher, au lieu de mots que l'utilisateur doit prononcer. Notre collègue anthropologue [35] pense que LOUISE est un outil intéressant pour étudier l'interaction naturelle avec les personnes démentes, car cela permet de produire des comportements verbaux et non-verbaux complètement reproductibles, ce qui réduit les sources de biais.

Enfin, LOUISE étant un outil générique, il serait facile de l'intégrer dans de nombreuses autres applications. Dans le chapitre 6, nous avons proposé quatre cas d'usage de LOUISE, dont certains pourraient nécessiter la connexion de LOUISE à des appareils externes, tels qu'un pilulier électronique. Pour travailler avec ces sources d'information complémentaires, qui, comme mentionné ci-dessus, permettraient d'obtenir un indice de confiance sur l'exécution correcte d'une tâche, quelques types de messages, envoyés à l'ACA par l'appareil, pourraient suffire à couvrir la plupart des cas : "action en cours", "action effectuée" et "utilisateur inactif". Chacun de ces messages devrait comprendre au moins deux champs d'information : un identifiant d'action et un indice de confiance. Dans certains cas, il faudrait également pouvoir envoyer des informations de l'ACA vers l'appareil, une notification que l'utilisateur est supposé réaliser une action donnée, par exemple.

Disponibilité de meilleurs composants logiciel et matériel

Au cours de cette thèse, de nouvelles API de reconnaissance vocale sur le nuage ont été lancées. Nous avons utilisé l'une d'entre elles (proposée par Microsoft[4]) pour implémenter un nouveau composant de reconnaissance vocale pour LOUISE, étant donné que ceux utilisés lors des tests décrits dans le chapitre 6 n'étaient pas suffisamment fiables. Cependant nous n'avons pas eu l'occasion de le tester avec les patients, mais nous pensons qu'il y a de bonnes chances que cela améliore le fonctionnement global du système. D'une manière similaire, une nouvelle version du capteur Kinect est désormais disponible, accompagnée de fonctionnalités d'analyse du comportement étendues et plus fiables que dans la précédente version. Le matériel et les logiciels informatiques évoluent très rapidement, et de nouveaux outils sont

[4] http://www.microsoft.com/cognitive-services/en-us/speech-api

lancés régulièrement, tout particulièrement ces dernières années, depuis que les géants du logiciel, tels que Google ou Microsoft, ainsi qu'une myriade de petites entreprises spécialisées, ont lancé des offres d'"intelligence artificielle" sur le nuage. Il en résulte que, comme les ACA dépendent de ces technologies, les systèmes tels que LOUISE s'améliorent constamment, au fil des mises-à-jour logicielles et matérielles, tout en limitant les coûts. Jusque récemment, cela se limitait à la reconnaissance vocale et au suivi des utilisateurs, mais, aujourd'hui, d'innombrables solutions, libres ou commerciales, pour des applications de reconnaissance d'émotion ou d'analyse prosodique sont disponibles, ou le seront très prochainement. C'est pourquoi il est plus important que jamais de se tenir au courant des nouveautés dans ce domaine, car la brique logicielle qui fera passer les ACA sur un tout autre plan est peut-être toute proche.

Virtual Promenade

Grâce à la conception participative, nous avons pu concevoir le système Virtual Promenade de sorte qu'il soit utilisable par pratiquement toutes les personnes âgées souffrant d'un SPC. Néanmoins, il pourrait être amélioré sur plusieurs aspects : ajout de nouveaux contenus comportant des tâches ayant plus de sens et des situations plus stressantes, tels qu'un environnement urbain peuplé de piétons et de voitures, un sol glissant ou des escaliers ; exploration d'autres modes d'interaction et de sensations haptiques plus riches pour renforcer l'impression de marche en étant assis ; et l'utilisation d'écrans 3D ou de casques de réalité virtuelle pour augmenter l'immersion.

Ajouts de contenus et de nouvelles mécaniques de jeu

Nous avons vu dans nos phases de conception que les participants ont demandé des tâches ayant plus de sens que celle consistant à collecter des cubes virtuels. Il y a de nombreuses manières de faire cela : le joueur pourrait ramasser des objets plus symboliques, comme des fleurs pour un bouquet ; la tâche pourrait consister à suivre un chien, afin de ne pas le perdre ; cela pourrait également consister à emmener un enfant dans un lieu donné, comme un magasin de jouets, ou un zoo. Cela pourrait stimuler l'engagement dans l'activité, ce qui pourrait conduire à une meilleure immersion dans les environnements virtuels. De plus, cela serait plus motivant si le jeu incluait des scores, des retours et des récompenses.

Comme l'activité est destiné à être réalisée fréquemment, pendant plusieurs semaines, il est également nécessaire de prévoir davantage de contenus, autrement les patients risquent de s'ennuyer rapidement, même s'ils sont atteints de démence, car, s'ils ne se rappellent pas de l'activité, leur mémoire procédurale, qui fonctionne toujours correctement, les rend plus experts, et les tâches peuvent devenir trop faciles. De nouveaux environ-

nements et des tâches de difficulté croissante et/ou qui correspondent à des situations stressantes sont nécessaires pour que les patients restent motivés et qu'on puisse les faire se promener dans des environnements de plus en stressants, au fur et à mesure que leur peur de chuter diminue.

Interaction corporelle et haptique

Pour renforcer l'impression de marcher, nous voudrions explorer les avantages de faire faire aux joueurs des mouvements de marche avec les bras et/ou les jambes pour faire bouger leurs avatars. Cela est cependant plus difficile que d'utiliser une simple manette de jeu, particulièrement pour synchroniser les mouvements de l'assise avec ceux du patient. De plus, il n'y a aucune garantie que cela améliorerait réellement l'immersion. Par exemple, dans une étude menée par Nabiyouni et al. [144], les auteurs ont comparé trois modalités de locomotion pour les applications virtuelles : "non-naturelle", via l'utilisation d'une manette de jeu ; "semi-naturelle", en utilisant une boule de hamster à taille humaine appelée Virtusphere, qui permet de marcher en faisant du sur-place ; et "naturelle", dans laquelle les sujets marchent réellement, dans une salle équipée de capteurs de mouvement. Il s'avère que la condition semi-naturelle offre moins de fidélité et mène à des performances moindres que la condition non-naturelle, qui peut diminuer le niveau d'immersion. Cependant, dans notre cas, ce mode d'interaction pourrait être bénéfique, puisque les patients feraient de l'activité physique.

Enfin, les personnes atteintes de SPC présentent souvent des troubles posturaux appelés "rétro-pulsion" ou "déséquilibre arrière", connus pour figurer parmi les symptômes du syndrome de désadaptation psycho-motrice [141]. Par conséquent, une modification simple du système, qui pourrait en améliorer les qualités thérapeutiques, consisterait à ajouter un capteur de pression sur le dossier du fauteuil et à empêcher tout déplacement de l'avatar tant que le patient s'appuie sur le dossier. Cela forcerait les patients à tenir leur dos, ce qui stimulerait les muscles posturaux, et à adopter une position susceptible de provoquer de l'anxiété, de peur de tomber en avant.

Écrans 3D et casques de réalité virtuelle

Afin d'augmenter l'immersion, cela vaudrait la peine d'étudier l'utilisation d'écrans 3D, tels que des téléviseurs stéréoscopiques. L'idéal serait d'utiliser un modèle avec lequel les spectateurs n'ont pas besoin de porter de lunettes polarisées, car cela pourrait causer des problèmes d'hygiène dans les environnements de soin, mais ces caractéristiques ne sont présentes que sur des modèles haut-de-gamme, ce qui multiplierait le coût total du système par 4, au minimum. De plus, à cause de déficiences visuelles, telles que la cataracte, observée fréquemment chez les personnes âgées, il est probable qu'une proportion significative des patients ne puissent pas voir l'effet *stéréoscopique, quelle que soit la technologie d'affichage utilisée.*

Une solution plus économique consisterait à utiliser un casque de réalité virtuelle. Cependant, nous avons éliminé cette option à priori, car il y a peu de chances que cela convienne à cette application, pour une utilisation dans un contexte médical. Il y a plusieurs raisons à cela : les casques de réalité virtuelle peuvent causer la cyber-nausée à de nombreuses personnes, même jeunes, principalement à cause de la latence ; ils risquent d'être trop lourds pour les personnes âgées fragiles ; leur utilisation dans un contexte médical peut causer des problèmes d'hygiène ; et, à notre connaissance, il y a très peu de données disponibles sur la manière dont les personnes démentes pourraient réagir à l'immersion avec casque [122], sinon que les personnes âgées n'ont pas particulièrement plus de problèmes de cyber-nausée que les personnes jeunes [235]. Cette technologie étant relativement jeune, des produits qui éliminent la plupart de ces problèmes pourraient être disponibles dans un futur proche, ainsi que de la littérature sur l'utilisation de ces casques de réalité virtuelle avec les personnes âgées atteintes de démence. A ce moment, il serait intéressant de tester ces casques pour le système Virtual Promenade.

Combiner les deux approches des humains virtuels

Les deux approches des humains virtuels que nous avons considérées comme distinctes sont en fait complémentaires et doivent être fusionnées. En effet, comme nous l'avons vu dans la section 7.7.3, les patients ayant participé aux tests de Virtual Promenade ne voulaient pas vraiment lire les instructions écrites et ont eu des difficultés à les faire passer (afficher l'instruction suivante ou précédente). Pour expliquer comment jouer, donner les consignes des tâches et motiver les patients, cela vaudrait la peine de considérer l'inclusion de l'ACA LOUISE dans le jeu de marche virtuelle de Virtual Promenade. De plus, nous n'avons pas observé beaucoup de situations dans lesquelles les personnes âgées atteintes de troubles cognitifs perdent le fil de ce qu'elles font et cessent de le faire lors de nos tests des scénarii d'usage de LOUISE, alors que cela s'est produit plusieurs fois, avec deux des patients, au cours de l'étude pilote de Virtual Promenade. Cela suggère que les capacités d'estimation d'attention de LOUISE auraient plus de valeur lorsque l'ACA aide les patients à réaliser une tâche qui requiert une attention soutenue pendant plusieurs minutes que pour la conversation face-à-face. Dans Virtual Promenade, LOUISE pourrait prendre l'apparence d'un kinésithérapeute ou d'un médecin virtuel. Sa fonction de guidage pas-à-pas dans une tâche avec des exemples vidéo pourrait être très utile pour montrer comment contrôler l'avatar du joueur. De plus, l'ACA pourrait demander aux patients quelles tâches ils veulent réaliser et gérer le déroulement de la séance. De cette manière, les rééducateurs pourraient laisser leurs patients utiliser le dispositif de manière autonome ; ils pourraient ainsi s'occuper d'autres patients pendant ce temps là, tout en gardant un œil sur eux. Faire *cela créerait un paradigme d'interaction mixte, utilisant à la fois la communication verbale, non-verbale et les actions sur la manette de jeu. Cependant, bien qu'un certain degré d'automatisation fait gagner du temps aux professionnels du soin, il n'est pas souhaitable que les patients jouent en toute autonomie, car cela risque de diminuer l'acceptation du dispositif par les professionnels, qui pourraient craindre qu'on les remplace par des machines. De plus, cela requerrait un très important travail de développement et de tests, ce qui rendrait le système trop onéreux.*

Épilogue éthique

L'arrivée des smartphones *et l'accès à une connexion internet haut-débit partout et n'importe quand ont enclenché une transformation de nos modes de vie. Cela permet de transporter dans nos poches de petits appareils depuis lesquels nous accédons à une myriade de services très utiles (ou très inutiles), souvent gratuits pour les utilisateurs finaux. En faisant cela, nous acceptons sciemment de troquer une partie de notre intimité contre des applications pratiques, tels que le guidage par GPS ou la possibilité de trouver facilement restaurants, bars, ou boutiques à proximité de là où nous nous trouvons. De plus, à travers les modèles commerciaux gratuits ou "*freemium*", les fournisseurs de services collectent de nombreuses données personnelles qu'ils vendent parfois à des tierces parties. Certains de ces services, et quelques services spécifiques, pourraient également être utiles pour les personnes âgées démentes, et parfois même plus critiques pour leur sécurité que pour les personnes sans troubles. Mais comprendraient/comprennent-elles ce qu'implique l'utilisation d'applications mobiles ou de systèmes tels que les colliers d'appel d'urgence munis de puces GPS en termes de violation de leur intimité ?*

De plus, ce sont souvent les membres de la famille qui prennent les décisions d'achat de technologies d'assistance pour leurs proches. Étant donné que leur principale préoccupation est généralement la sécurité de la personne, certains sont prêts à considérer l'utilisation de systèmes très intrusifs, comme l'accès au flux vidéo d'une caméra placée au domicile du patient ou des systèmes de suivi qui enregistrent l'activité de la personne (s'alimente-t-elle, s'hydrate-t-elle, prend-t-elle ses médicaments, va-t-elle à la selle, etc ?). Nous pensons donc qu'il est extrêmement important que les personnes démentes soient informées des implications de l'utilisation de tels systèmes et qu'elles soient consultées avant que cela leur soit imposé, du moins tant qu'elles sont capables de choisir. D'une certaine manière, équiper une personne âgée de technologies d'assistance est comparable à les placer dans une "maison de retraite virtuelle" : lorsqu'elles ne prennent pas part à la décision, des conséquences délétères sont à prévoir. Lorsque les personnes sont placées dans des institutions spécialisées sans avoir été préalablement con-

sultées ou contre leur gré, elles peuvent tomber dans une grave dépression. Avec les technologies d'assistance, il est probable que la personne les mette de côté et refuse de les utiliser, voire qu'elle les détruise (ou tente de le faire). Nous pensons que la conception participative de produits et services d'assistance permet aux personnes âgées démentes de prendre part à la conversation dès le début, de manière à ce que leurs préoccupations et leurs souhaits puissent être pris en compte en amont.

Durant les phases de conception de LOUISE, nous avons demandé aux participants s'ils voudraient que le système soit allumé en permanence. Il apparait que la plupart des personnes âgées qui nous ont donné leur avis ne le souhaitent pas. Cependant, pour certaines d'entre elles, lorsque la maladie entre dans un stade modéré à sévère, le système ne serait d'aucune utilité s'il n'est activé que lorsqu'elles l'allument, car il est fort probable qu'elles oublient de le faire. C'est pourquoi il est nécessaire de considérer ces aspects avec soin et de consulter les personnes pour trouver le meilleur compromis entre la qualité et la nature des services proposés par les ACA et le respect du choix des personnes. Pour les personnes âgées atteintes de troubles cognitifs, qui n'arriveraient pas à arrêter le système lorsqu'il les importune, il pourrait être utile de trouver des solutions pour leur donner un certain niveau de contrôle. Par exemple, cela serait utile de détecter l'agacement ou des entrées verbales comme "laisse moi tranquille" ou d'autres expressions plus grossières que des personnes désinhibées sont susceptibles d'employer.

*Enfin, comme nous l'avons mentionné dans la section 3.1.4, il existe des techniques de manipulation bien connues, généralement verbales, qui peuvent servir à influencer les actions des personnes [97]. Grâce à ces techniques, ainsi qu'à la confiance qu'inspirent les ACA, ces agents artificiels, au fil de leurs évolutions techniques, seront bientôt capables de manipuler les personnes démentes. Bien que cela puisse paraître effrayant, dans la prise en charge de la démence, cela pourrait se révéler utile, voire bénéfique, pour certains patients : dans les situations de refus des soins, par exemple, un ACA pourrait influencer les personnes, afin qu'elle s'alimentent, qu'elles s'hydratent, qu'elles prennent leurs médicaments, etc. Cela pourrait également servir à pousser les personnes à adopter un mode de vie plus sain. Par exemple, dans la série télévisée de science-fiction suédoise Äkta människor (*Real Humans*), un homme âgé se voit offrir un nouveau robot gériatrique qui a l'apparence et le comportement d'une vraie personne, pour remplacer celui qu'il possédait, car ce dernier ne fonctionnait plus correctement. Ce nouveau robot, contrairement à son prédécesseur, est programmé pour faire ce qui est considéré comme bon pour lui : il refuse de lui préparer ses plats favoris et lui en sert des plus sains ; il contrôle la quantité de vin qu'il boit et le restreint quand il juge qu'il en a eu assez ; et il va jusqu'à introduire des somnifères dans sa nourriture pour qu'il se couche tôt. Il va sans dire que le personnage n'apprécie pas vraiment son nouveau robot et voudrait récupérer l'ancien. Cependant, la situation s'améliore lorsque ses* *proches viennent lui rendre visite et demande au robot d'être moins strict. La manipulation est assurément un chemin hasardeux sur lequel s'aventurer, même avec de bonnes intentions. Il est par conséquent nécessaire de se poser la question du niveau de manipulation souhaitable, des limites éthiques à ne pas franchir et si un ACA ou un robot d'assistance doit faire ce que veulent ses bénéficiaires ou ce qui est pour leur bien. Il est donc nécessaire d'établir des recommandations sur les bonnes pratiques à appliquer pour l'utilisation d'agents sociaux dans le soin aux personnes âgées, ce qui pourrait être réalisé grâce à l'approche living lab.*

Bibliography

[1] Abramowitz, J. S., Deacon, B. J., and Whiteside, S. P. *Exposure therapy for anxiety: Principles and practice.* Guilford Press, 2012.

[2] Acampora, G., Cook, D., Rashidi, P., and Vasilakos, A. A survey on ambient intelligence in healthcare. *Proceedings of the IEEE 101*, 12 (December 2013), 2470–2494.

[3] Agmon, M., Perry, C. K., Phelan, E., Demiris, G., and Nguyen, H. Q. A pilot study of wii fit exergames to improve balance in older adults. *Journal of Geriatric Physical Therapy 34*, 4 (2011), 161–167.

[4] Agogué, M., Comtet, G., Le Masson, P., Menudet, J.-F., and Picard, R. Managing innovative design within the health ecosystem: the living lab as an architect of the unknown. *Revue Management et Avenir Santé 1*, 1 (Oct 2013), 17–32.

[5] Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Gamst, A., Holtzman, D. M., Jagust, W. J., Petersen, R. C., et al. The diagnosis of mild cognitive impairment due to alzheimer's disease: Recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. *Alzheimer's & dementia 7*, 3 (2011), 270–279.

[6] Alm, N., Astell, A., Gowans, G., Ellis, M., Dye, R., Vaughan, P., and Riley, P. Cognitive prostheses: Findings from attempts to model some aspects of cognition. In *Human Centered Design*, M. Kurosu, Ed., no. 6776 in Lecture Notes in Computer Science. Springer Berlin Heidelberg, January 2011, pp. 275–284.

[7] Alm, N., Dye, R., Gowans, G., Campbell, J., Astell, A., and Ellis, M. A communication support system for older people with dementia. *Computer*, 5 (2007), 35–41.

[8] Aman, F., Vacher, M., Rossato, S., and Portet, F. Speech recognition of aged voice in the aal context: Detection of distress sentences. In *Speech Technology and Human-Computer Dialogue (SpeD), 2013 7th Conference on* (2013), IEEE, pp. 1–8.

[9] American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (DSM-5®).* American Psychiatric Pub, 2013.

[10] Artificial Intelligence Foundation. Aiml: Artificial intelligence markup language. [Online] http://www.alicebot.org/aiml.html.

[11] Asteriadis, S., Karpouzis, K., and Kollias, S. The importance of eye gaze and head pose to estimating levels of attention. In *Games and Virtual Worlds for Serious Applications (VS-GAMES), 2011 Third International Conference on* (2011), IEEE, pp. 186–191.

[12] Asteriadis, S., Karpouzis, K., and Kollias, S. Visual focus of attention in non-calibrated environments using gaze estimation. *International Journal of Computer Vision* (Dec. 2013), 1–24.

[13] Badeig, F., Wargnier, P., Pino, M., De Olivera Lopes, P., Grange, E., Crowley, J. L., Rigaud, A.-S., and Vaufreydaz, D. Impact of head motion on the assistive robot expressiveness – evaluation with elderly persons. In *1st International Workshop on Affective Computing for Social Robotics (ACSR)* (Aug 2016), IEEE.

[14] Banala, S. K., Agrawal, S. K., and Scholz, J. P. Active leg exoskeleton (alex) for gait rehabilitation of motor-impaired patients. In *Rehabilitation Robotics, 2007. ICORR 2007. IEEE 10th International Conference on* (2007), IEEE, pp. 401–407.

[15] Baram, Y., and Miller, A. Virtual reality cues for improvement of gait in patients with multiple sclerosis. *Neurology 66*, 2 (2006), 178–181.

[16] Bateni, H. Changes in balance in older adults based on use of physical therapy vs the wii fit gaming system: a preliminary study. *Physiotherapy 98*, 3 (2012), 211–216.

[17] Benkaouar, W., and Vaufreydaz, D. Multi-sensors engagement detection with a robot companion in a home environment. In *Workshop on Assistance and Service robotics in a human environment at IEEE International Conference on Intelligent Robots and Systems (IROS2012)* (2012), pp. 45–52.

[18] Benkaouar, W., and Vaufreydaz, D. Multi-sensors engagement detection with a robot companion in a home environment. In *Workshop on Assistance and Service robotics in a human environment at IEEE International Conference on Intelligent Robots and Systems (IROS2012)* (Oct 2012), p. 45–52.

[19] Benveniste, S. *Incremental Design of Therapeutic Music Games Theory and Application to the Treatment of Behavioral Disorders and Alzheimer's Disease*. PhD thesis, MINES ParisTech, 2010.

[20] Bergmann, K., Eyssel, F., and Kopp, S. A second chance to make a first impression? how appearance and nonverbal behavior affect perceived warmth and competence of virtual agents over time. In *International Conference on Intelligent Virtual Agents* (2012), Springer, pp. 126–138.

[21] Bergvall-Kåreborn, B., Howcroft, D., Ståhlbröst, A., and Wikman, A. M. Participation in living lab: Designing systems with users. In *Human Benefit Through the Diffusion of Information Systems Design Science Research*. Springer, 2010, pp. 317–326.

[22] Bergvall-Kåreborn, B., and Ståhlbröst, A. Living lab: an open and citizen-centric approach for innovation. *International Journal of Innovation and Regional Development 1*, 4 (Jan. 2009), 356–370.

[23] Bhala, R. P., O'Donnell, J., and Thoppil, E. Ptophobia: Phobic fear of falling and its clinical management. *Physical therapy 62*, 2 (1982), 187–190.

[24] Bickmore, T., Schulman, D., and Yin, L. Maintaining engagement in long-term interventions with relational agents. *Applied Artificial Intelligence 24*, 6 (2010), 648–666.

[25] Bickmore, T. W., Silliman, R. A., Nelson, K., Cheng, D. M., Winter, M., Henault, L., and Paasche-Orlow, M. K. A randomized controlled trial of an automated exercise coach for older adults. *Journal of the American Geriatrics Society 61*, 10 (2013), 1676–1683.

[26] Blanchard, E. B., Jones-Alexander, J., Buckley, T. C., and Forneris, C. A. Psychometric properties of the ptsd checklist (pcl). *Behaviour research and therapy 34*, 8 (1996), 669–673.

[27] Bloch, F., Blandin, M., Ranerison, R., Claessens, Y., Rigaud, A., and Kemoun, G. Anxiety after a fall in elderly subjects and subsequent risk of developing post traumatic stress disorder at two months. a pilot study. *The journal of nutrition, health & aging 18*, 3 (2014), 303–306.

[28] Bloch, F., Rigaud, A.-S., and Kemoun, G. Virtual reality exposure therapy in posttraumatic stress disorder: A brief review to open new opportunities for post-fall syndrome in elderly subjects. *European Geriatric Medicine 4*, 6 (2013), 427–430.

[29] Boger, J., Poupart, P., Hoey, J., Boutilier, C., Fernie, G., and Mihailidis, A. A decision-theoretic approach to task assistance for persons with dementia. In *Proceedings of the 19th international joint conference on Artificial intelligence* (2005), Morgan Kaufmann Publishers Inc., pp. 1293–1299.

[30] Bondi, M. W., and Kaszniak, A. W. Implicit and explicit memory in alzheimer's disease and parkinson's disease. *Journal of Clinical and Experimental Neuropsychology 13*, 2 (1991), 339–358.

[31] Brandherm, B., Prendinger, H., and Ishizuka, M. Interest estimation based on dynamic bayesian networks for visual attentive presentation agents. In *Proceedings of the 9th international conference on Multimodal interfaces* (2007), ACM, pp. 346–349.

[32] Bundesen, C., and Habekost, T. Principles of visual attention: Linking mind and brain.

[33] Cafaro, A., Vilhjálmsson, H. H., Bickmore, T., Heylen, D., and Pelachaud, C. Representing communicative functions in saiba with a unified function markup language. In *Intelligent Virtual Agents* (2014), Springer, pp. 81–94.

[34] Campbell, A. J., Borrie, M. J., and Spears, G. F. Risk factors for falls in a community-based prospective study of people 70 years and older. *Journal of gerontology 44*, 5 (1989), M112–M117.

[35] Carletti, G., 2016. Private communications.

[36] Carrasco, E., Epelde, G., Moreno, A., Ortiz, A., Garcia, I., Buiza, C., Urdaneta, E., Etxaniz, A., González, M. F., and Arruti, A. Natural interaction between avatars and persons with alzheimer's disease. In *Computers Helping People with Special Needs*, K. Miesenberger, J. Klaus, W. Zagler, and A. Karshmer, Eds. Springer Berlin Heidelberg, January 2008, pp. 38–45.

[37] Čereković, A., and Pandžić, I. S. Multimodal behavior realization for embodied conversational agents. *Multimedia Tools and Applications 54*, 1 (2011), 143–164.

[38] Chen, C.-C. Improvement in the physiological function and standing stability based on kinect multimedia for older people. *Journal of Physical Therapy Science 28*, 4 (2016), 1343–1348.

[39] Chen, J., Chen, D., Li, X., and Zhang, K. Towards improving social communication skills with multimodal sensory information. *IEEE Transactions on Industrial Informatics 10*, 1 (Feb 2014), 323–330.

[40] Chen, K., and Chan, A. A review of technology acceptance by older adults. *Gerontechnology 10*, 1 (2011), 1–12.

[41] Cherry, E. C. Some experiments on the recognition of speech, with one and with two ears. *The Journal of the acoustical society of America 25*, 5 (1953), 975–979.

[42] Chittaro, L., Ieronutti, L., and Rigutti, S. Supporting presentation techniques based on virtual humans in educational virtual worlds. In *Cyberworlds, 2005. International Conference on* (2005), IEEE, pp. 8–pp.

[43] Chung, M. C., McKee, K. J., Austin, C., Barkby, H., Brown, H., Cash, S., Ellingford, J., Hanger, L., and Pais, T. Post-traumatic stress disorder in older people after a fall. *International journal of geriatric psychiatry 24*, 9 (2009), 955–964.

[44] Cohen-Mansfield, J., Dakheel-Ali, M., and Marx, M. S. Engagement in persons with dementia: the concept and its measurement. *The American journal of geriatric psychiatry 17*, 4 (2009), 299–307.

[45] Cook, A. M., and Polgar, J. M. *Assistive technologies: Principles and practice.* Elsevier Health Sciences, 2014.

[46] Coon, W., Rich, C., and Sidner, C. L. Activity planning for long-term relationships. In *Intelligent Virtual Agents* (Edinburgh, UK, Aug. 2013), Springer, p. 425.

[47] Costa Jr, P. T., and McCrae, R. R. Four ways five factors are basic. *Personality and Individual Differences 13*, 6 (June 1992), 653–665.

[48] Courgeon, M. *Marc : modèles informatiques des émotions et de leurs expressions faciales pour l'interaction Homme-machine affective temps réel.* PhD thesis, Université Paris Sud - Paris XI, Nov. 2011.

[49] Courgeon, M., and Clavel, C. MARC: a framework that features emotion models for facial animation during human-computer interaction. *Journal on Multimodal User Interfaces 7*, 4 (Dec. 2013), 311–319.

[50] Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski, J., Chowdhury, N., and Baker, A. Optimizing inhibitory learning during exposure therapy. *Behaviour research and therapy 46*, 1 (2008), 5–27.

[51] Cristancho-Lacroix, V., Wrobel, J., Cantegreil-Kallen, I., Dub, T., Rouquette, A., and Rigaud, A.-S. A web-based psychoeducational program for informal caregivers of patients with alzheimer's disease: A pilot randomized controlled trial. *Journal of Medical Internet Research 17*, 5 (2015).

[52] Davis, F. D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly* (1989), 319–340.

[53] Deci, E. L., and Ryan, R. M. *Intrinsic motivation.* Wiley Online Library, 1975.

[54] Deci, E. L., and Ryan, R. M. *Intrinsic Motivation and Self-determination in Human Behavior.* Plenum, 1985.

[55] Deci, E. L., and Ryan, R. M. The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological inquiry 11*, 4 (2000), 227–268.

[56] Denis, G., and Jouvelot, P. Motivation-driven educational game design: applying best practices to music education. In *Proceedings of the 2005 ACM SIGCHI International Conference on Advances in computer entertainment technology* (2005), ACM, pp. 462–465.

[57] Dubois, B., Feldman, H. H., Jacova, C., DeKosky, S. T., Barberger-Gateau, P., Cummings, J., Delacourte, A., Galasko, D., Gauthier, S., Jicha, G., et al. Research criteria for the diagnosis of alzheimer's disease: revising the nincds–adrda criteria. *The Lancet Neurology 6*, 8 (2007), 734–746.

[58] Ekman, P., and Friesen, W. V. The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica 1*, 1 (1969), 49–98.

[59] Ekman, P., and Friesen, W. V. Unmasking the face: A guide to recognizing emotions from facial cues, 1975.

[60] Ekman, P., and Friesen, W. V. *Manual for the facial action coding system.* Consulting Psychologists Press, 1978.

[61] Encrevé, P., and De Fornel, M. Le sens en pratique [construction de la référence et structure sociale de l'interaction dans le couple question/réponse]. *Actes de la recherche en sciences sociales 46*, 1 (1983), 3–30.

[62] European Commission. Growing the european silver economy. [Online] http://ec.europa.eu/research/innovation-union/pdf/active-healthy-ageing/silvereco.pdf, Feb. 2015.

[63] Eysenck, H. J. *The measurement of personality*, vol. xviii. University Park Press, Baltimore, MD, US, 1976.

[64] Fagel, S., Hilbert, A., Mor, M., and Mayer, C. The virtual counselor automated character animation for ambient assisted living. In *Proceedings of 6th International Conference on Advances in Computer-Human Interactions (ACHI'13)* (2013), IARIA, pp. 184–187.

[65] Falkowski, J. A., Hynan, L. S., Krishnan, K., Carter, K., Lacritz, L., Weiner, M., Rossetti, H., and Cullum, C. M. Conversion of moca to mmse scores. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring 1*, 1 (2015), 125.

[66] Faur, C., Clavel, C., Pesty, S., and Martin, J.-C. PERSEED: A self-based model of personality for virtual agents inspired by socio-cognitive theories. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII)* (Sept. 2013), pp. 467–472.

[67] Fisk, A. D., Rogers, W. A., Charness, N., Czaja, S. J., and Sharit, J. *Designing for Older Adults: Principles and Creative Human Factors Approaches*, second ed. Human Factors & Aging Series. CRC Press, 2009.

[68] Folstein, M. F., Folstein, S. E., and McHugh, P. R. "mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *Journal of psychiatric research 12*, 3 (1975), 189–198.

[69] Forum des Living Labs en Santé et Autonomie. La co-construction. [Online] http://www.forumllsa.org/co-working, 2016.

[70] Fox, J., Christy, K. R., and Vang, M. H. The experience of presence in persuasive virtual environments. In *Interacting with presence: HCI and the sense of presence in computer-mediated environments*, G. Riva, J. Waterworth, and D. Murray, Eds. DeGruyter Open, 2014, pp. 164–178.

[71] Frisoli, A., Borelli, L., Montagner, A., Marcheschi, S., Procopio, C., Salsedo, F., Bergamasco, M., Carboncini, M. C., Tolaini, M., and Rossi, B. Arm rehabilitation with a robotic exoskeleleton in virtual reality. In *Rehabilitation Robotics, 2007. ICORR 2007. IEEE 10th International Conference on* (2007), IEEE, pp. 631–642.

[72] Garcia, J. A., Pisan, Y., Tan, C. T., and Navarro, K. F. *Entertainment Computing – ICEC 2014: 13th International Conference, ICEC 2014, Sydney, Australia, October 1-3, 2014. Proceedings.* Springer Berlin Heidelberg, Berlin, Heidelberg, 2014, ch. Assessing the Kinect's Capabilities to Perform a Time-Based Clinical Test for Fall Risk Assessment in Older People, pp. 100–107.

[73] Gaudet, M., Tavernier, B., Mourey, F., Tavernier, C., Richard, D., and Marot, J. Le syndrome de régression psychomotrice du vieillard. *Médecine et hygiène 44*, 1656 (1986), 1332–1336.

[74] Gaver, W. W. Technology affordances. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (1991), ACM, pp. 79–84.

[75] Gerardi, M., Cukor, J., Difede, J., Rizzo, A., and Rothbaum, B. O. Virtual reality exposure therapy for post-traumatic stress disorder and other anxiety disorders. *Current Psychiatry Reports 12*, 4 (2010), 298–305.

[76] Gerardi, M., Cukor, J., Difede, J., Rizzo, A., and Rothbaum, B. O. Virtual reality exposure therapy for post-traumatic stress disorder and other anxiety disorders. *Current psychiatry reports 12*, 4 (2010), 298–305.

[77] Gibson, J. J. The ecological approach to visual perception.

[78] Gillespie, A., Best, C., and O'Neill, B. Cognitive function and assistive technology for cognition: a systematic review. *Journal of the International Neuropsychological Society: JINS 18*, 1 (January 2012), 1–19. PMID: 22152338.

[79] Green, C. S., and Bavelier, D. Action video game training for cognitive enhancement. *Current Opinion in Behavioral Sciences 4* (2015), 103–108.

[80] Griol, D., Iglesias, J. A., Ledezma, A., and Sanchis, A. A dialog management methodology based on evolving fuzzy-rule-based (frb) classifiers. In *Evolving and Adaptive Intelligent Systems (EAIS), 2014 IEEE Conference on* (2014), IEEE, pp. 1–8.

[81] Gunes, H., Schuller, B., Pantic, M., and Cowie, R. Emotion representation, analysis and synthesis in continuous space: A survey. In *Automatic Face Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on* (March 2011), pp. 827–834.

[82] Hamada, T., Kuwahara, N., Morimoto, K., Yasuda, K., Akira, U., and Abe, S. *Universal Access in Human-Computer Interaction. Addressing Diversity: 5th International Conference, UAHCI 2009, Held as Part of HCI International 2009, San Diego, CA, USA, July 19-24, 2009. Proceedings, Part I.* Springer Berlin Heidelberg, Berlin, Heidelberg, 2009, ch. Preliminary Study on Remote Assistance for People with Dementia at Home by Using Multi-media Contents, pp. 236–244.

[83] Hamm, J., Money, A. G., Atwal, A., and Paraskevopoulos, I. Fall prevention intervention technologies: A conceptual framework and survey of the state of the art. *Journal of Biomedical Informatics 59* (2016), 319 – 345.

[84] Han, J., Shao, L., Xu, D., and Shotton, J. Enhanced computer vision with microsoft kinect sensor: A review. *Cybernetics, IEEE Transactions on 43*, 5 (2013), 1318–1334.

[85] Hartholt, A., Traum, D., Marsella, S. C., Shapiro, A., Stratou, G., Leuski, A., Morency, L.-P., and Gratch, J. All Together Now: Introducing the Virtual Human Toolkit. In *13th International Conference on Intelligent Virtual Agents* (Edinburgh, UK, Aug. 2013).

[86] Heerink, M., Kröse, B., Evers, V., and Wielinga, B. Assessing acceptance of assistive social agent technology by older adults: the almere model. *International journal of social robotics 2*, 4 (2010), 361–375.

[87] Heloir, A., and Gibet, S. A qualitative and quantitative characterisation of style in sign language gestures. In *Gesture-Based Human-Computer Interaction and Simulation*, M. S. Dias, S. Gibet, M. M. Wanderley, and R. Bastos, Eds., no. 5085 in Lecture Notes in Computer Science. Springer Berlin Heidelberg, Jan. 2009, pp. 122–133.

[88] Heylen, D., Kopp, S., Marsella, S. C., Pelachaud, C., and Vilhjálmsson, H. The next step towards a function markup language. In *Intelligent Virtual Agents* (2008), Springer, pp. 270–280.

[89] Hoey, J., Poupart, P., von Bertoldi, A., Craig, T., Boutilier, C., and Mihailidis, A. Automated handwashing assistance for persons with dementia using video and a partially observable markov decision process. *Computer Vision and Image Understanding 114*, 5 (2010), 503–519.

[90] Holden, M. K. Virtual environments for motor rehabilitation: review. *Cyberpsychology & behavior 8*, 3 (2005), 187–211.

[91] Huang, C. M., and Mutlu, B. Robot behavior toolkit: Generating effective social behaviors for robots. In *Human-Robot Interaction (HRI), 2012 7th ACM/IEEE International Conference on* (March 2012), pp. 25–32.

[92] Huang, H.-H., Matsushita, H., Kawagoe, K., Sakai, Y., Nonaka, Y., Nakano, Y., and Yasuda, K. Toward a memory assistant companion for the individuals with mild memory impairment. In *Cognitive Informatics & Cognitive Computing (ICCI* CC), 2012 IEEE 11th International Conference on* (2012), IEEE, pp. 295–299.

[93] Hymes, D. *Foundations in sociolinguistics: An ethnographic approach*. Psychology Press, 2001.

[94] International Telecommunications Union. ICT facts and figures. [Online] http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2015.pdf, May 2015.

[95] Ishii, R., Nakano, Y. I., and Nishida, T. Gaze awareness in conversational agents: Estimating a user's conversational engagement from eye gaze. *ACM Transactions on Interactive Intelligent Systems 3*, 2 (July 2013), 11:1–11:25.

[96] Jayasinghe, N., Sparks, M. A., Kato, K., Wyka, K., Wilbur, K., Chiaramonte, G., Barie, P. S., Lachs, M. S., O'Dell, M., Evans, A., et al. Posttraumatic stress symptoms in older adults hospitalized for fall injury. *General hospital psychiatry 36*, 6 (2014), 669–673.

[97] Joule, R.-V., Girandola, F., and Bernard, F. How can people be induced to willingly change their behavior? the path from persuasive communication to binding communication. *Social and Personality Psychology Compass 1*, 1 (2007), 493–505.

[98] Judge, K. S., Camp, C. J., and Orsulic-Jeras, S. Use of montessori-based activities for clients with dementia in adult day care: Effects on engagement. *American Journal of Alzheimer's Disease and Other Dementias 15*, 1 (2000), 42–46.

[99] Karg, M., Samadani, A.-A., Gorbet, R., Kühnlenz, K., Hoey, J., and Kulić, D. Body movements for affective expression: A survey of automatic recognition and generation.

[100] Karlin, S. Tactus technology. *Spectrum, IEEE 50*, 4 (2013), 23–23.

[101] Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A., and Jaffe, M. W. Studies of illness in the aged: the index of adl:

a standardized measure of biological and psychosocial function. *Jama 185*, 12 (1963), 914–919.

[102] Kelley, J. F. An iterative design methodology for user-friendly natural language office information applications. *ACM Transactions on Information Systems 2*, 1 (Jan. 1984), 26–41.

[103] Keogh, J. W., Power, N., Wooller, L., Lucas, P., and Whatman, C. Physical and psychosocial function in residential aged-care elders: effect of nintendo wii sports games. *Journal of aging and physical activity* (2014).

[104] Kim, J., Son, J., Ko, N., and Yoon, B. Unsupervised virtual reality-based exercise program improves hip muscle strength and balance control in older adults: a pilot study. *Archives of physical medicine and rehabilitation 94*, 5 (2013), 937–943.

[105] Knopman, D. S., and Nissen, M. J. Implicit learning in patients with probable alzheimer's disease. *Neurology 37*, 5 (1987), 784–784.

[106] Koolagudi, S. G., and Rao, K. S. Emotion recognition from speech: a review. *International Journal of Speech Technology 15*, 2 (2012), 99–117.

[107] Kopp, S., Krenn, B., Marsella, S., Marshall, A. N., Pelachaud, C., Pirker, H., Thórisson, K. R., and Vilhjálmsson, H. Towards a common framework for multimodal generation: The behavior markup language. In *Intelligent Virtual Agents*, J. Gratch, M. Young, R. Aylett, D. Ballin, and P. Olivier, Eds., no. 4133 in Lecture Notes in Computer Science. Springer Berlin Heidelberg, January 2006, pp. 205–217.

[108] Lambrecht, K. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*, vol. 71. Cambridge university press, 1996.

[109] Lan, Y., Harvey, R., and Theobald, B.-J. Insights into machine lip reading. In *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (March 2012), pp. 4825–4828.

[110] Lange, B., Flynn, S., Chang, C.-Y., Liang, W., Si, Y., Nanavati, C., Chieng, C.-L., and Rizzo, A. Development of an interactive stepping game to reduce falls in older adults. *International Journal on Disability and Human Development 10*, 4 (2011), 331–335.

[111] Lapointe, J., Verreault, A., Bouchard, B., Bouchard, J., Potvin, A., and Bouzouane, A. How to maximize the effectiveness of prompts in assistive technologies according to the particular

cognitive profile of people with alzheimer ? *International Journal of Smart Home 7*, 5 (September 2013), 19–38.

[112] Laver, K., George, S., Thomas, S., Deutsch, J., and Crotty, M. Cochrane review: virtual reality for stroke rehabilitation. *European journal of physical and rehabilitation medicine 48*, 3 (2012), 523–530.

[113] Lee, A., and Kawahara, T. Recent development of open-source speech recognition engine julius. In *Proceedings: APSIPA ASC 2009: Asia-Pacific Signal and Information Processing Association, 2009 Annual Summit and Conference* (2009), Asia-Pacific Signal and Information Processing Association, 2009 Annual Summit and Conference, International Organizing Committee, pp. 131–137.

[114] Lee, K.-F., Hon, H.-W., and Reddy, R. An overview of the sphinx speech recognition system. *Acoustics, Speech and Signal Processing, IEEE Transactions on 38*, 1 (1990), 35–45.

[115] Leonhardt, M. D. *Enhancing affective communication in embodied conversational agents through personality-based hidden conversational goals*. PhD thesis, Universidade federal do rio grande do sul, Mar. 2012.

[116] Lewin, K. Action research and minority problems. *Journal of social issues 2*, 4 (1946), 34–46.

[117] Li, L., Xu, Q., and Tan, Y. K. Attention-based addressee selection for service and social robots to interact with multiple persons. In *Proceedings of the Workshop at SIGGRAPH Asia* (New York, NY, USA, Nov 2012), WASA '12, ACM, pp. 131–136.

[118] Lohse, K., Shirzad, N., Verster, A., Hodges, N., and Van der Loos, H. M. Video games and rehabilitation: using design principles to enhance engagement in physical therapy. *Journal of Neurologic Physical Therapy 37*, 4 (2013), 166–175.

[119] Lucca, U., Tettamanti, M., Logroscino, G., Tiraboschi, P., Landi, C., Sacco, L., Garrì, M., Ammesso, S., Bertinotti, C., Biotti, A., et al. Prevalence of dementia in the oldest old: The monzino 80-plus population based study. *Alzheimer's & Dementia 11*, 3 (2015), 258–270.

[120] Luppa, M., Luck, T., Weyerer, S., König, H.-H., Brähler, E., and Riedel-Heller, S. G. Prediction of institutionalization in the elderly. a systematic review. *Age and ageing 39*, 1 (Jan. 2010), 31–38.

[121] Maiden, N., D'Souza, S., Jones, S., Müller, L., Pannese, L., Pitts, K., Prilla, M., Pudney, K., Rose, M., Turner, I., et al. Computing technologies for reflective, creative care of people with dementia. *Communications of the ACM 56*, 11 (2013), 60–67.

[122] Manera, V., Chapoulie, E., Bourgeois, J., Guerchouche, R., David, R., Ondrej, J., Drettakis, G., and Robert, P. A feasibility study with image-based rendered virtual reality in patients with mild cognitive impairment and dementia. *PloS one 11*, 3 (2016), e0151487.

[123] Marcotte, T. D., Scott, J. C., Kamat, R., and Heaton, R. K. Neuropsychology and the prediction of everyday functioning. *Neuropsychology of everyday functioning* (2010), 5–38.

[124] Marivan, K., Boully, C., Benveniste, S., Reingewirtz, S., Rigaud, A.-S., Kemoun, G., and Bloch, F. Rehabilitation of the psychomotor consequences of falling in an elderly population: a pilot study to evaluate feasibility and tolerability of virtual reality training. *Technology and Health Care*, Preprint (2015), 1–7.

[125] Marks, R. Physical activity and hip fracture disability: a review. *Journal of aging research 2011* (2011).

[126] Marsella, S., and Gratch, J. Computationally modeling human emotion. *Communications of the ACM 57*, 12 (Nov. 2014), 56–67.

[127] Martin, R. C. *Agile software development: principles, patterns, and practices*. Prentice Hall PTR, 2003.

[128] Maynard, D. W., and Zimmerman, D. H. Topical talk, ritual and the social organization of relationships. *Social psychology quarterly* (1984), 301–316.

[129] McGurk, H., and MacDonald, J. Hearing lips and seeing voices. *Nature 264*, 5588 (December 1976), 746–748.

[130] McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., and Stadlan, E. M. Clinical diagnosis of alzheimer's disease report of the nincds-adrda work group* under the auspices of department of health and human services task force on alzheimer's disease. *Neurology 34*, 7 (1984), 939–939.

[131] McRorie, M., Sneddon, I., Sevin, E. d., Bevacqua, E., and Pelachaud, C. A model of personality and emotional traits. In *Intelligent Virtual Agents*, Z. Ruttkay, M. Kipp, A. Nijholt, and H. H. Vilhjálmsson, Eds., no. 5773 in Lecture Notes in Computer Science. Springer Berlin Heidelberg, Jan. 2009, pp. 27–33.

[132] McTear, M., Callejas, Z., and Griol, D. *The Conversational Interface*. Springer, 2016.

[133] Microsoft Insights. Attention spans. [Online] https://advertising.microsoft.com/en/WWDocs/User/display/cl/researchreport/31966/en/microsoft-attention-spans-research-report.pdf, 2015.

[134] Mindmakers. BML 1.0 specification. [Online] http://www.mindmakers.org/projects/bml-1-0/wiki.

[135] Mishra, B., Fernandes, S. L., Abhishek, K., Alva, A., Shetty, C., Ajila, C. V., Shetty, D., Rao, H., and Shetty, P. Facial expression recognition using feature based techniques and model based techniques: A survey. In *Electronics and Communication Systems (ICECS), 2015 2nd International Conference on* (2015), IEEE, pp. 589–594.

[136] Moore, D. S., and Ellis, R. Measurement of fall-related psychological constructs among independent-living older adults: a review of the research literature. *Aging and mental health 12*, 6 (2008), 684–699.

[137] Morandell, M., Hochgatterer, A., Fagel, S., and Wassertheurer, S. Avatars in assistive homes for the elderly a user-friendly way of interaction? *HCI and Usability for Education and Work* (2008), 391–402.

[138] Morandell, M. M., Hochgatterer, A., Wöckl, B., Dittenberger, S., and Fagel, S. Avatars@home InterFACEing the smart home for elderly people. In *HCI and Usability for e-Inclusion*, A. Holzinger and K. Miesenberger, Eds., no. 5889 in Lecture Notes in Computer Science. Springer Berlin Heidelberg, Jan. 2009, pp. 353–365.

[139] Mori, M. The uncanny valley. *Energy 7*, 4 (1970), 33–35.

[140] Morris, J. C., Rubin, E. H., Morris, E. J., and Mandel, S. A. Senile dementia of the alzheimer's type: an important risk factor for serious falls. *Journal of gerontology 42*, 4 (1987), 412–417.

[141] Mourey, F., Manckoundia, P., Martin-Arveux, I., Tavernier-Vidal, B., and Pfitzenmeyer, P. Psychomotor disadaptation syndrome. *Geriatrics 59*, 5 (2004).

[142] Murphy, J., and Isaacs, B. The post-fall syndrome. *Gerontology 28*, 4 (1982), 265–270.

[143] Muson, S. A., Poole, E., Perry, D. B., and Peyton, T. Gamification and health. In *The Gameful World*, S. P. Walz and S. Deterding, Eds. The MIT Press, 2014, pp. 597–623.

[144] Nabiyouni, M., Saktheeswaran, A., Bowman, D. A., and Karanth, A. Comparing the performance of natural, semi-natural, and non-natural locomotion techniques in virtual reality. In *3D User Interfaces (3DUI), 2015 IEEE Symposium on* (March 2015), pp. 3–10.

[145] Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., and Chertkow, H. The montreal cognitive assessment, moca: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society 53*, 4 (2005), 695–699.

[146] Nemoto, Y., Egawa, S., Koseki, A., Hattori, S., Ishii, T., and Fujie, P. Power-assisted walking support system for elderly. In *Engineering in Medicine and Biology Society, 1998. Proceedings of the 20th Annual International Conference of the IEEE* (1998), vol. 5, IEEE, pp. 2693–2695.

[147] Norman, D. A. *The design of everyday things: Revised and expanded edition.* Basic books, 2013.

[148] Oakley, A., Dawson, M. F., Holland, J., Arnold, S., Cryer, C., Doyle, Y., Rice, J., Hodgson, C., Sowden, A., Sheldon, T., et al. Preventing falls and subsequent injury in older people. *Quality in Health Care 5*, 4 (1996), 243.

[149] Olson, K. E., O'Brien, M. A., Rogers, W. A., and Charness, N. Diffusion of technology: Frequency of use for younger and older adults. *Ageing international 36*, 1 (2011), 123–145.

[150] O'Reilly, K. *Ethnographic methods.* Routledge, 2012.

[151] Ortiz, A., del Puy Carretero, M., Oyarzun, D., Yanguas, J. J., Buiza, C., Gonzalez, M. F., and Etxeberria, I. Elderly users in ambient intelligence: Does an avatar improve the interaction? In *Universal Access in Ambient Intelligence Environments*, C. Stephanidis and M. Pieper, Eds., no. 4397 in Lecture Notes in Computer Science. Springer Berlin Heidelberg, January 2007, pp. 99–114.

[152] Pak, R., and McLaughlin, A. *Designing Displays for Older Adults.* CRC Press, 2011.

[153] Pecune, F. Toward a computational model of social relations for artificial companions. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction (ACII)* (Sept. 2013), pp. 677–682.

[154] Pejsa, T., Andrist, S., Gleicher, M., and Mutlu, B. Gaze and attention management for embodied conversational agents. *ACM Transactions on Interactive Intelligent Systems (TiiS) 5*, 1 (2015), 3.

[155] Pelachaud, C. Multimodal expressive embodied conversational agents. In *Proceedings of the 13th Annual ACM International Conference on Multimedia* (New York, NY, USA, 2005), MULTIMEDIA '05, ACM, pp. 683–689.

[156] Pentland, A. Social signal processing. *IEEE Signal Processing Magazine 24*, 4 (July 2007), 108–111.

[157] Petersen, R. C. Mild cognitive impairment as a diagnostic entity. *Journal of internal medicine 256*, 3 (2004), 183–194.

[158] Pichierri, G., Murer, K., and de Bruin, E. D. A cognitive-motor intervention using a dance video game to enhance foot placement accuracy and gait under dual task conditions in older adults: a randomized controlled trial. *BMC geriatrics 12*, 1 (2012), 1.

[159] Pino, M. *Human Factors and Psychosocial Challenges in the Design and Use of Assistive Technology for Older Adults with Cognitive Impairment.* PhD thesis, Ecole pratique des hautes études, 2012.

[160] Pino, M., Benveniste, S., Picard, R., and Rigaud, A.-S. User-driven innovation for dementia care in France: The LUSAGE living lab case study. *Interdisciplinary Studies Journal 3*, 4 (2014), 251–268.

[161] Pompeu, J., Arduini, L., Botelho, A., Fonseca, M., Pompeu, S., Torriani-Pasin, C., and Deutsch, J. Feasibility, safety and outcomes of playing kinect adventures!™ for people with parkinson's disease: a pilot study. *Physiotherapy 100*, 2 (2014), 162–168.

[162] Price, M., and Anderson, P. The role of presence in virtual reality exposure therapy. *Journal of anxiety disorders 21*, 5 (2007), 742–751.

[163] Proffitt, R., and Lange, B. User centered design and development of a game for exercise in older adults. *The International Journal of Technology, Knowledge, and Society 8*, 6 (Mar. 2013), 95–112.

[164] Rajman, M. *Speech and language engineering.* EPFL Press, 2007.

[165] Rego, P., Moreira, P. M., and Reis, L. P. Serious games for rehabilitation: A survey and a classification towards a taxonomy. In *Information Systems and Technologies (CISTI), 2010 5th Iberian Conference on* (2010), IEEE, pp. 1–6.

[166] Rendon, A. A., Lohman, E. B., Thorpe, D., Johnson, E. G., Medina, E., and Bradley, B. The effect of virtual reality gaming on dynamic balance in older adults. *Age and ageing 41*, 4 (2012), 549–552.

[167] Rich, C., Ponsler, B., Holroyd, A., and Sidner, C. Recognizing engagement in human-robot interaction. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (2010), pp. 375–382.

[168] Rich, C., Ponsler, B., Holroyd, A., and Sidner, C. L. Recognizing engagement in human-robot interaction. In *Human-Robot Interaction (HRI), 2010 5th ACM/IEEE International Conference on* (2010), IEEE, pp. 375–382.

[169] Rich, C., and Sidner, C. L. Using collaborative discourse theory to partially automate dialogue tree authoring. In *Intelligent Virtual Agents* (2012), Springer, pp. 327–340.

[170] Ries, E. *The lean startup: How today's entrepreneurs use continuous innovation to create radically successful businesses.* Crown Books, 2011.

[171] Rigby, C. S. Gamification and motivation. In *The Gameful World*, S. P. Walz and S. Deterding, Eds. The MIT Press, 2014, pp. 597–623.

[172] Ring, L., Barry, B., Totzke, K., and Bickmore, T. Addressing loneliness and isolation in older adults: Proactive affective agents provide better support. In *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on* (2013), IEEE, pp. 61–66.

[173] Roalf, D. R., Moberg, P. J., Xie, S. X., Wolk, D. A., Moelter, S. T., and Arnold, S. E. Comparative accuracies of two common screening instruments for classification of alzheimer's disease, mild cognitive impairment, and healthy aging. *Alzheimer's & Dementia 9*, 5 (2013), 529–537.

[174] Rousseaux, M., Sève, A., Vallet, M., Pasquier, F., and Mackowiak-Cordoliani, M. A. An analysis of communication in conversation in patients with dementia. *Neuropsychologia 48*, 13 (2010), 3884–3890.

[175] Russell, J. A. A circumplex model of affect. *Journal of Personality and Social Psychology 39*, 6 (1980), 1161–1178.

[176] Sacks, H. On the preferences for agreement and contiguity in sequences in conversation. In *Talk and Social Organisation*, G. Button and J. R. E. Lee, Eds. Multilingual Matters LTD, 1987, pp. 54–69.

[177] Salah, A., Pantic, M., and Vinciarelli, A. Recent developments in social signal processing. In *2011 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (October 2011), pp. 380–385.

[178] Salen, K., and Zimmerman, E. *Rules of play: Game design fundamentals.* MIT press, 2004.

[179] Sanford, J. A. *Universal design as a rehabilitation strategy: design for the ages.* Springer Publishing Company, 2012.

[180] Scherer, S., Marsella, S., Stratou, G., Xu, Y., Morbini, F., Egan, A., and Morency, L.-P. Perception markup language: towards a standardized representation of perceived nonverbal behaviors. In *Intelligent virtual agents* (2012), Springer, pp. 455–463.

[181] Scherer, S., Marsella, S., Stratou, G., Xu, Y., Morbini, F., Egan, A., Rizzo, A., and Morency, L.-P. Perception markup language: Towards a standardized representation of perceived nonverbal behaviors. In *Intelligent Virtual Agents*, Y. Nakano, M. Neff, A. Paiva, and M. Walker, Eds., no. 7502 in Lecture Notes in Computer Science. Springer Berlin Heidelberg, Jan. 2012, pp. 455–463.

[182] Selwyn, N. The information aged: A qualitative study of older adults' use of information and communications technology. *Journal of Aging studies 18*, 4 (2004), 369–384.

[183] Shibata, T. Therapeutic seal robot as biofeedback medical device: Qualitative and quantitative evaluations of robot therapy in dementia care. *Proceedings of the IEEE 100*, 8 (2012), 2527–2538.

[184] Sidner, C., Bickmore, T., Rich, C., Barry, B., Ring, L., Behrooz, M., and Shayganfar, M. Demonstration of an always-on companion for isolated older adults. In *Proceedings of the SIGDIAL 2013 Conference* (Metz, France, August 2013), Association for Computational Linguistics, pp. 148–150.

[185] Sidner, C. L., Lee, C., Kidd, C. D., Lesh, N., and Rich, C. Explorations in engagement for humans and robots. *Artificial Intelligence 166*, 1 (2005), 140–164.

[186] Simons, D. J., and Chabris, C. F. Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception 28*, 9 (1999), 1059–1074.

[187] Skjæret, N., Nawaz, A., Morat, T., Schoene, D., Helbostad, J. L., and Vereijken, B. Exercise and rehabilitation delivered through exergames in older adults: An integrative review of technologies, safety and efficacy. *International journal of medical informatics 85*, 1 (2016), 1–16.

[188] Sloetjes, H., and Wittenburg, P. Annotation by category: ELAN and ISO DCR. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC)* (2008).

[189] Smith, S. T., Sherrington, C., Studenski, S., Schoene, D., and Lord, S. R. A novel dance dance revolution (ddr) system for in-home training of stepping ability: basic parameters of system use by older adults. *British journal of sports medicine 45*, 5 (2011), 441–445.

[190] Stiefelhagen, R., Finke, M., Yang, J., and Waibel, A. From gaze to focus of attention. In *Visual Information and Information Systems* (1999), Springer, pp. 765–772.

[191] Stössel, C., Wandke, H., and Blessing, L. Gestural interfaces for elderly users: help or hindrance? In *Gesture in embodied communication and human-computer interaction*. Springer, 2009, pp. 269–280.

[192] Stricker, R., Hommel, S., Martin, C., and Gross, H.-M. Realtime user attention and emotion estimation on a mobile robot. In *Proceedings of 55th Ilmenau International Scientific Colloquium* (Sep 2010), p. 13–17.

[193] Studenski, S., Perera, S., Hile, E., Keller, V., Spadola-Bogard, J., and Garcia, J. Interactive video dance games for healthy older adults. *The journal of nutrition, health & aging 14*, 10 (2010), 850–852.

[194] Tate, R. F. Correlation between a discrete and a continuous variable. point-biserial correlation. *The Annals of mathematical statistics* (1954), 603–607.

[195] Taylor, P., and Isard, A. Ssml: A speech synthesis markup language. *Speech communication 21*, 1 (1997), 123–133.

[196] Thiebaux, M., Marsella, S., Marshall, A. N., and Kallmann, M. SmartBody: Behavior realization for embodied conversational agents. In *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems* (2008), vol. 1 of *AAMAS '08*, International Foundation for Autonomous Agents and Multiagent Systems, pp. 151–158.

[197] Tinetti, M. E., De Leon, C. F. M., Doucette, J. T., and Baker, D. I. Fear of falling and fall-related efficacy in relationship to functioning among community-living elders. *Journal of gerontology 49*, 3 (1994), M140–M147.

[198] Tinetti, M. E., Richman, D., and Powell, L. Falls efficacy as a measure of fear of falling. *Journal of gerontology 45*, 6 (1990), P239–P243.

[199] Tombaugh, T. N., and McIntyre, N. J. The mini-mental state examination: a comprehensive review. *Journal of the American Geriatrics Society 40*, 9 (1992), 922–935.

[200] Traum, D. R., and Larsson, S. The information state approach to dialogue management. In *Current and new directions in discourse and dialogue*. Springer, 2003, pp. 325–353.

[201] Tsoi, K. K., Hirai, H. W., Chan, J. Y., and Kwok, T. C. Time to treatment initiation in people with alzheimer disease: A meta-analysis of randomized controlled trials. *Journal of the American Medical Directors Association 17*, 1 (2016), 24–30.

[202] Uzor, S., and Baillie, L. Exploring and designing tools to enhance falls rehabilitation in the home. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2013), CHI '13, ACM, pp. 1233–1242.

[203] Uzor, S., and Baillie, L. Investigating the long-term use of exergames in the home with elderly fallers. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems* (New York, NY, USA, 2014), CHI '14, ACM, pp. 2813–2822.

[204] Uzor, S., Baillie, L., and Skelton, D. Senior designers: empowering seniors to design enjoyable falls rehabilitation tools. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2012), ACM, pp. 1179–1188.

[205] Van Doorn, C., Gruber-Baldini, A. L., Zimmerman, S., Richard Hebel, J., Port, C. L., Baumgarten, M., Quinn, C. C., Taler, G., May, C., and Magaziner, J. Dementia as a risk factor for falls and fall injuries among nursing home residents. *Journal of the American Geriatrics Society 51*, 9 (2003), 1213–1218.

[206] van Welbergen, H., Reidsma, D., and Kopp, S. An incremental multimodal realizer for behavior co-articulation and coordination. In *Intelligent virtual agents* (2012), Springer, pp. 175–188.

[207] Van Welbergen, H., Xu, Y., Thiebaux, M., Feng, W.-W., Fu, J., Reidsma, D., and Shapiro, A. Demonstrating and testing the bml compliance of bml realizers. In *Intelligent virtual agents* (2011), Springer, pp. 269–281.

[208] Vardoulakis, L. P., Ring, L., Barry, B., Sidner, C. L., and Bickmore, T. Designing relational agents as long term social companions for older adults. In *Intelligent Virtual Agents*, Y. Nakano, M. Neff, A. Paiva, and M. Walker, Eds., no. 7502 in Lecture Notes in Computer Science. Springer Berlin Heidelberg, January 2012, pp. 289–302.

[209] Venkatesh, V., and Davis, F. D. A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management science 46*, 2 (2000), 186–204.

[210] Vichitvanichphong, S., Talaei-Khoei, A., Kerr, D., and Ghapanchi, A. H. Adoption of assistive technologies for aged care: A realist review of recent studies. In *System Sciences (HICSS), 2014 47th Hawaii International Conference on* (2014), IEEE, pp. 2706–2715.

[211] Vilhjálmsson, H., Cantelmo, N., Cassell, J., Chafai, N. E., Kipp, M., Kopp, S., Mancini, M., Marsella, S., Marshall, A. N., Pelachaud, C., et al. The behavior markup language: Recent developments and challenges. In *Intelligent virtual agents* (2007), Springer, pp. 99–111.

[212] Vinciarelli, A., Pantic, M., and Bourlard, H. Social signal processing: Survey of an emerging domain. *Image and Vision Computing 27*, 12 (November 2009), 1743–1759.

[213] Vinciarelli, A., Pantic, M., Heylen, D., Pelachaud, C., Poggi, I., D'Errico, F., and Schroeder, M. Bridging the gap between social animal and unsocial machine: A survey of social signal processing. *IEEE Transactions on Affective Computing 3*, 1 (2012), 69–87.

[214] Vipperla, R., Wolters, M., and Renals, S. Spoken dialogue interfaces for older people. *Advances in home care technologies 1* (2012), 118–137.

[215] Vlaeyen, E., Coussement, J., Leysens, G., Van der Elst, E., Delbaere, K., Cambier, D., Denhaerynck, K., Goemaere, S., Wertelaers, A., Dobbels, F., et al. Characteristics and effectiveness of fall prevention programs in nursing homes: A systematic review and meta-analysis of randomized controlled trials. *Journal of the American Geriatrics Society 63*, 2 (2015), 211–221.

[216] Wagner, J., Lingenfelser, F., Baur, T., Damian, I., Kistler, F., and André, E. The social signal interpretation (ssi) framework: multimodal signal processing and recognition in real-time. In *Proceedings of the 21st ACM international conference on Multimedia* (2013), ACM, pp. 831–834.

[217] Ward, R. E., Leveille, S. G., Beauchamp, M. K., Travison, T., Alexander, N., Jette, A. M., and Bean, J. F. Functional performance as a predictor of injurious falls in older adults. *Journal of the American Geriatrics Society 63*, 2 (2015), 315–320.

[218] Wargnier, P., Phuong, E., Marivan, K., Benveniste, S., Bloch, F., Reingewirtz, S., Kemoun, G., and Rigaud, A.-S. Virtual promenade: A new serious game for the rehabilitation of older adults with post-fall syndrome. In *4th International Conference on Serious Games and Applications for Health (IEEE SeGAH 2016)* (2016).

[219] Welbergen, H., Reidsma, D., Ruttkay, Z. M., and Zwiers, J. Elckerlyc-a bml realizer for continuous, multimodal interaction with a virtual human. *Journal on Multimodal User Interfaces 3*, 4 (2010), 271–284.

[220] Wiemeyer, J., and Kliem, A. Serious games in prevention and rehabilitation—a new panacea for elderly people? *European Review of Aging and Physical Activity 9*, 1 (2012), 41–50.

[221] Wilcoxon, F. Individual comparisons by ranking methods. *Biometrics bulletin* (1945), 80–83.

[222] Wilks, Y., Jasiewicz, J. M., Catizone, R., Galescu, L., Martinez, K. M., and Rugs, D. CALONIS: An artificial companion within a smart home for the care of cognitively impaired patients. In *Smart Homes and Health Telematics*, C. Bodine, S. Helal, T. Gu, and M. Mokhtari, Eds., Lecture Notes in Computer Science. Springer International Publishing, Jan. 2015, pp. 255–260.

[223] Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L.-O., Nordberg, A., Bäckman, L., Albert, M., Almkvist, O., Arai, H., Basun, H., Blennow, K., de Leon, M., DeCarli, C., Erkinjuntti, T., Giacobini, E., Graff, C., Hardy, J., Jack, C., Jorm, A., Ritchie, K., van Duijn, C., Visser, P., and Petersen, R. C. Mild cognitive impairment - beyond controversies, towards a consensus: report of the international working group on mild cognitive impairment. *Journal of internal medicine 256*, 3 (September 2004), 240–246.

[224] Winter, R., and Munn-Giddings, C. *A handbook for action research in health and social care.* Routledge, 2001.

[225] World Health Organization. Towards a common language for functioning, disability and health (icf). [Online] http://www.who.int/classifications/icf/training/icfbeginnersguide.pdf, 2002.

[226] World Health Organization. World report on disability. [Online] http://www.who.int/disabilities/world_report/2011/report.pdf, 2011.

[227] World Health Organization, and Alzheimer's Disease International. *Dementia: a public health priority.* 2012.

[228] Wu, Y.-h., Damnée, S., Kerhervé, H., Ware, C., and Rigaud, A.-S. Bridging the digital divide in older adults: a study from an initiative to inform older adults about new technologies. *Clinical interventions in aging 10* (2015), 193.

[229] Yaghoubzadeh, R., Kramer, M., Pitsch, K., and Kopp, S. Virtual agents as daily assistants for elderly or cognitively impaired people. In *Intelligent Virtual Agents*, R. Aylett, B. Krenn, C. Pelachaud, and H. Shimodaira, Eds. Springer Berlin Heidelberg, January 2013, pp. 79–91.

[230] Yaghoubzadeh, R., Pitsch, K., and Kopp, S. Adaptive grounding and dialogue management for autonomous conversational assistants for elderly users. In *Intelligent Virtual Agents* (2015), Springer, pp. 28–38.

[231] Yasuda, K., Aoe, J.-I., and Fuketa, M. Development of an agent system for conversing with individuals with dementia. In *The 27th Annual Conference of the Japanese Society for Artificial Intelligence* (Jun 2013).

[232] Young, V., and Mihailidis, A. Difficulties in automatic speech recognition of dysarthric speakers and implications for speech-based applications used by the elderly: A literature review. *Assistive Technology 22*, 2 (2010), 99–112.

[233] Young, W., Ferguson, S., Brault, S., and Craig, C. Assessing and training standing balance in older adults: a novel approach using the 'nintendo wii'balance board. *Gait & posture 33*, 2 (2011), 303–305.

[234] Zajicek, M. A methodology for interface design for older adults. In *Enterprise Information Systems Vi.* Springer, 2006, pp. 285–292.

[235] Zakzanis, K. K., Quintin, G., Graham, S. J., and Mraz, R. Age and dementia related differences in spatial navigation within an immersive virtual environment. *Medical Science Monitor 15*, 4 (2009), CR140–CR150.

[236] Zancanaro, M., Gabrielli, S., Jameson, A., Leonardi, C., Not, E., and Pianesi, F. Virtual helper or virtual card player? contrasting responses of older users. In *Your Virtual Butler*. Springer, 2013, pp. 70–78.

Appendices

Appendix A

Publications and communications

A.1 Publications

Here is a list of publications associated with this thesis, ordered by publication dates:

- J. Wrobel, M. Pino, **P. Wargnier** and A.-S. Rigaud “Robots et agents virtuels au service des personnes âgées : une revue de l’actualité en gérontologie”, *NPG Neurologie-Psychiatrie-Gériatrie*, 14(82), 2014, pp 184–193.

- **P. Wargnier**, A. Malaisé, J. Jacquemot, S. Benveniste, P. Jouvelot, M. Pino and A.-S. Rigaud “Towards Attention Monitoring of Older Adults with Cognitive Impairment During Interaction with an Embodied Conversational Agent” In *3rd IEEE VR International Workshop on Virtual and Augmented Assistive Technologies (VAAT)*, Arles, France, March 2015, pp. 23–28.

- **P. Wargnier**, G. Carletti, Y. Laurent-Corniquet, S. Benveniste, P. Jouvelot and A.-S. Rigaud “Field Evaluation with Cognitively Impaired Older Adults of Attention Management in the Embodied Conversational Agent Louise” In *4th IEEE International Conference on Serious Games and Applications for Health (SeGAH)*, Orlando FL., USA, May 2016.

- **P. Wargnier**, E. Phuong, K. Marivan, S. Benveniste, F. Bloch, S. Reingewirtz, G. Kemoun and A.-S. Rigaud “Virtual Promenade: A New Serious Game for the Rehabilitation of Older Adults with Post-fall Syndrome” In *4th IEEE International Conference on Serious Games and Applications for Health (SeGAH)*, Orlando FL., USA, May 2016.

- F. Badeig, **P. Wargnier**, M. Pino, P. De Oliveira Lopes, E. Grange, J. L. Crowley, A.-S. Rigaud and D. Vaufreydaz "Impact of Head Motion on the Assistive Robot Expressiveness – Evaluation with Elderly Persons" In *1st IEEE International Workshop on Affective Computing for Social Robotics (ACSR)*, New York NY., USA, August 2016.

- **P. Wargnier**, P.-E. Fauquet, S. Benveniste, P. Jouvelot, A.-S. Rigaud, G. Kemoun and F. Bloch. "Rehabilitation of the psychomotor consequences of elderly fallers: A pilot study to evaluate the feasibility and tolerability of virtual reality training using a user-centered designed serious game" In *10th World Conference of Gerontechnology (ISG 2016)*, Vol. 15, No. 169, Nice, France, September 2016.

A.2 Communications

Here are the events in which I gave presentations during the preparation of this thesis:

- 2ème congrès européen de stimulation cognitive, Toulouse, France, September 22nd–23rd 2014.

- 5ème journée "primeurs" de la société de gériatrie et de gérontologie d'Île-de-France, Paris, France, January 8th 2015.

- 5ème colloque PARAChute, Evry, France, November 23rd 2016.

Appendix B

Interaction scenarios

B.1 Wizard of Oz study

The contents of the scenario files used in the Wizard of Oz study are presented in Tables B.1 and B.2. The first one corresponds to the assistive technology experts group and the second one corresponds to the older adults group.

B.2 Automated LOUISE study

```
<scenario>
<utterance id="start" type="statement" wait="0">
<command> <saccade mode="talk"/>
<face id="sm" start="0" end="2" type="facs" au="12_left"
amount="1.1"/>
<face start="0" end="2" type="facs" au="12_right"
amount="1.1"/>
<face start="0" end="2" type="facs" au="6" amount="1"/>
<event start="sm:end" message="triggerEoBEvent()"/>
</command>
<transition>Bonjour</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml">Je vous
disais</speech>
<animation name="ChrBrad@Idle01_HoweverLf01"
start="sp:start"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="Bonjour" type="question">
```

Table B.1: Contents of the interaction scenario in the attention estimator validation experiment with assistive technology experts

Utterance	Type
"Hello Sir/Madam. I'm pleased to meet you."	introduction
"My name is Louise. I am a virtual agent. I would like to ask you a few questions. Are you interested?"	introduction + yes/no question
"Very well. I am going to ask you a series of questions about avatar. Shall we start?"	introduction + yes/no question
"What did you know about avatars before meeting me?"	open question
"I can assist you in your daily life. What kind of functionalities would you of a virtual agent?	open question
"Do you think it is important for an avatar to look like a human being?"	yes/no question
"What do you think about my appearance?"	open question
"Do you think I could be useful in your daily life?"	yes/no question
"Would you like an avatar like me to be installed in you home to assist you?"	yes/no question
"I really liked talking to you. What about you? Did you enjoy our conversation?"	yes/no question
"Very well."	acknowledgment
"Hum... I see."	acknowledgment
"Interesting."	acknowledgment
"OK"	acknowledgment
"Thank you for answering my questions."	conclusion
"Goodbye sir/madam!"	conclusion
"Excuse me?"	prompting
"Err... Sir/Madam?"	prompting
"Err... Please?"	prompting
"Would you like to continue?"	after prompting
"I repeat my question."	transition

```
<command>
<speech id="sp" type="application/ssml+xml">Bonjour!</speech>
<head type="NOD" amount=".5"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition condition="bonjour">Smile2</transition>
</utterance>
```

Table B.2: Contents of the interaction scenario in the attention estimator validation experiment with older adults

Utterance	Type
"Hello. (2s pause) My name is Louise"	introduction
"I am a virtual character"	introduction
"What is you name?"	open question
"I am very pleased to meet you. May I ask you a few questions?"	yes/no question
"Had you ever talked to a virtual person?"	yes/no question
"Do you understand well what I say?"	yes/no question
"What is the weather like today?"	open question
"Do you like my voice?"	yes/no question
"Do you like music?"	yes/no question
"As I am virtual, my appearance can be changed. Do you like my hair color?"	yes/no question
"I really liked talking to you. What about you? Did you enjoy our conversation?"	yes/no question
"Very well."	acknowledgment
"Hum... I see."	acknowledgment
"Interesting."	acknowledgment
"OK"	acknowledgment
"I understand"	acknowledgment
"Thanks for talking to me. I was bored being on my own."	conclusion
"Goodbye!"	conclusion
"(clear throat) Are you listening?"	prompting
"Excuse me?"	prompting
"Err... Are you there?"	prompting
"Err... Please?"	prompting
"Err... I am here!"	prompting
"Would you like to continue?"	after prompting
"I repeat my question."	transition

```
<utterance id="Smile2" type="statement" wait="0">
<command>
<saccade mode="talk"/>
<face id="sm" start="0" end="2" type="facs" au="12_left"
amount="1.1"/>
<face start="0" end="2" type="facs" au="12_right"
amount="1.1"/>
<face start="0" end="2" type="facs" au="6" amount="1"/>
```

```
<event start="sm:end" message="triggerEoBEvent()"/>
</command>
<transition>Presentation</transition>
</utterance>

<utterance id="Presentation" type="statement">
<command>
<speech id="sp" type="application/ssml+xml">Je m’appelle Louise.
Je suis un agent virtuel.</speech>
<animation name="ChrBrad@Idle01_MeLf02"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition>Soif</transition>
</utterance>

<utterance id="Soif" type="question">
<command>
<speech id="sp" type="application/ssml+xml"> Avez−vous soif?
</speech>
<animation name="ChrBrad@Idle01_YouLf02"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition condition="oui">EauGazeuse</transition>
<transition condition="non">Recommandation</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml">Je m’appelle Louise.
Je suis la pour vous aider. <break time="0.5s"/> Je vous
demandais</speech>
<animation name="ChrBrad@Idle01_HoweverLf01"
start="sp:start+4"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="Recommandation" type="question">
<command>
<speech id="sp" type="application/ssml+xml">C’est important de
bien s’hydrater.<break time="0.5s"/> Voulez vous boire?</speech>
<animation name="ChrBrad@Idle01_HoweverLf01"
start="sp:start+4"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition condition="oui">EauGazeuse</transition>
<transition condition="non">Conclusion</transition>
```

```
<recontextualisation>
<speech id="sp" type="application/ssml+xml">Je m'appelle Louise.
Je suis la pour vous aider. <break time="0.5s"/> Je vous
disais</speech>
<animation name="ChrBrad@Idle01_HoweverLf01"
start="sp:start+4"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="EauGazeuse" type="question">
<command>
<speech id="sp" type="application/ssml+xml">Voulez-vous de
l'eau gazeuse?</speech>
<animation id="a1" name="ChrBrad@Idle01_IndicateLeftLf01"/>
<gaze start="a1:start" target="left" sbm:joint-range="EYES
NECK"/>
<event start="sp:start" message="displayGasWater()"/>
<gaze start="a1:end" target="front" sbm:joint-range="EYES
NECK"/>
<event start="a1:end+1" message="triggerEoBEvent()"/>
</command>
<transition
condition="oui">ConfirmationEauGazeuse</transition>
<transition condition="non">EauPlate</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml">Vous m'avez dis que
vous aviez soif. <break time="1s"/> Je vous demandais</speech>
<event start="sp:start" message="removeGasWater()"/>
<animation name="ChrBrad@Idle01_HoweverLf01"
start="sp:start+3"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="EauPlate" type="question">
<command>
<speech id="sp" type="application/ssml+xml">Voulez-vous de
l'eau plate?</speech>
<animation id="a1" name="ChrBrad@Idle01_IndicateRightRt01"/>
<gaze start="a1:start" target="right" sbm:joint-range="EYES
NECK"/>
<event start="sp:start" message="removeGasWater()"/>
<event start="sp:start" message="displayMineralWater()"/>
```

```
<gaze start="a1:end" target="front" sbm:joint−range="EYES
NECK"/>
<event start="a1:end+1" message="triggerEoBEvent()"/>
</command>
<transition condition="oui">ConfirmationEauPlate</transition>
<transition condition="non">PasDeChoix</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml">Vous m'avez dis que
vous aviez soif. <break time="1s"/> Je vous demandais.</speech>
<event start="sp:start" message="removeMineralWater()"/>
<animation name="ChrBrad@Idle01_HoweverLf01"
start="sp:start+3"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="PasDeChoix" type="question">
<command>
<speech id="sp" type="application/ssml+xml">Malheureusement,
nous n'avons que de l'eau. En voulez−vous?</speech>
<animation id="a1" ready="sp:start"
name="ChrBrad@Idle01_PleaBt02"/>
<animation id="a2" start="sp:start+3"
name="ChrBrad@Idle01_YouLf01"/>
<event start="sp:start" message="removeMineralWater()"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition condition="oui">EauGazeuse</transition>
<transition condition="non">Conclusion</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml">Vous m'avez dis que
vous aviez soif mais vous n'avez pas voulu d'eau.</speech>
<animation name="ChrBrad@Idle01_HoweverLf01"
start="sp:start+3"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="ConfirmationEauGazeuse" type="question">
<command>
<speech id="sp" type="application/ssml+xml">Vous voulez de l'eau
gazeuse?</speech>
<animation id="a1" start="sp:start+1"
name="ChrBrad@Idle01_IndicateLeftLf01"/>
```

```
<event start="sp:start" message="displayGasWater()"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition condition="oui">VotreDroite</transition>
<transition condition="non">EauPlate</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml">J'ai compris que
vous vouliez de l'eau gazeuse. <break time="1s"/>Je vous
disais.</speech>
<animation name="ChrBrad@Idle01_HoweverLf01"
start="sp:start+3"/>
<event start="sp:start" message="removeGasWater()"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="ConfirmationEauPlate" type="question">
<command>
<speech id="sp" type="application/ssml+xml">Vous voulez de l'eau
plate?</speech>
<animation id="a1" start="sp:start+1"
name="ChrBrad@Idle01_IndicateRightRt01"/>
<event start="sp:start" message="displayMineralWater()"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition condition="oui">VotreGauche</transition>
<transition condition="non">PasDeChoix</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml">J'ai compris que
vous vouliez de l'eau plate. <break time="1s"/>Je vous
disais.</speech>
<animation name="ChrBrad@Idle01_HoweverLf01"
start="sp:start+3"/>
<event start="sp:start" message="removeMineralWater()"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="VotreDroite" type="statement">
<command>
<speech id="sp" type="application/ssml+xml">D'accord. L'eau
gazeuse se trouve sur votre droite.</speech>
<head type="WIGGLE" repeat="2"/>
```

```
<animation id="a1" start="sp:start+1"
name="ChrBrad@Idle01_IndicateLeftLf01"/>
<event start="sp:start" message="removeGasWater()"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition>Transition1</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml">Vous m'avez dis que
vous vouliez de l'eau gazeuse. <break time="0.5s"/>Je vous
disais.</speech>
<animation name="ChrBrad@Idle01_HoweverLf01"
start="sp:start+3"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="VotreGauche" type="statement">
<command>
<speech id="sp" type="application/ssml+xml">D'accord. L'eau
plate se trouve sur votre gauche.</speech>
<head type="WIGGLE" repeat="2"/>
<animation id="a1" start="sp:start+1"
name="ChrBrad@Idle01_IndicateRightRt01"/>
<event start="sp:start" message="removeMineralWater()"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition>Transition2</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml">Vous m'avez dis que
vous vouliez de l'eau plate. <break time="0.5s"/>Je vous
disais.</speech>
<animation name="ChrBrad@Idle01_HoweverLf01"
start="sp:start+3"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="Transition1" type="statement">
<command>
<speech id="sp" type="application/ssml+xml">Je ne peux pas vous
servir mais je peux vous expliquer comment faire pour boire un verre
d'eau.</speech>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
```

```
<transition>Explications1?</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml">Vous m'avez dis que
vous vouliez de l'eau gazeuse. Elle se trouve sur votre droite. <break
time="0.5s"/>Je vous disais.</speech>
<animation name="ChrBrad@Idle01_HoweverLf01"
start="sp:start+3"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="Transition2" type="statement">
<command>
<speech id="sp" type="application/ssml+xml">Je ne peux pas vous
servir mais je peux vous expliquer comment faire pour boire un verre
d'eau.</speech>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition>Explications2?</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml">Vous m'avez dis que
vous vouliez de l'eau plate. Elle se trouve sur votre gauche. <break
time="0.5s"/>Je vous disais.</speech>
<animation name="ChrBrad@Idle01_HoweverLf01"
start="sp:start+3"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="Explications1?" type="question">
<command>
<speech id="sp" type="application/ssml+xml">Souhaitez-vous que
je vous explique comment vous servir de l'eau?</speech>
<animation name="ChrBrad@Idle01_YouLf02"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition condition="oui">DebutExplication</transition>
<transition condition="non">Conclusion</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml">J'ai compris que
vous vouliez de l'eau gazeuse. <break time="1s"/>Je vous
disais.</speech>
<animation name="ChrBrad@Idle01_HoweverLf01"
start="sp:start+3"/>
```

```xml
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="Explications2?" type="question">
<command>
<speech id="sp" type="application/ssml+xml">Souhaitez-vous que
je vous explique comment vous servir de l'eau?</speech>
<animation name="ChrBrad@Idle01_YouLf02"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition condition="oui">DebutExplication</transition>
<transition condition="non">Conclusion</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml">J'ai compris que
vous vouliez de l'eau plate. <break time="1s"/>Je vous
disais.</speech>
<animation name="ChrBrad@Idle01_HoweverLf01"
start="sp:start+3"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="DebutExplication" type="statement">
<command>
<speech id="sp" type="application/ssml+xml">D'accord. <break
time="0.5s"/> Je vais vous expliquer pas a pas. Et vous montrer ce que
vous devez faire sur l'et cran a ma droite.</speech>
<head id="hd" type="WIGGLE" repeat="2"/>
<animation start="sp:start+7" id="a1"
name="ChrBrad@Idle01_IndicateRightRt01"/>
<gaze start="a1:ready" target="right" sbm:joint-range="EYES
NECK"/>
<event start="hd:end" message="addTvToScene()"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition>ExplicationDeboucher</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml"> Vous vouliez que
vous explique comment vous servir un verre d'eau. <break
time="1s"/>Je vous disais.</speech>
<animation name="ChrBrad@Idle01_HoweverLf01"
start="sp:start+3"/>
<gaze target="front" sbm:joint-range="EYES NECK"/>
```

```
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="ExplicationDeboucher" type="statement"
wait="10" mode="task">
<command>
<speech id="sp" type="application/ssml+xml"> Vous devez d'abord
tenir la bouteille a une main et retirer le bouchon avec l'autre
main.</speech>
<animation stroke="sp:end" id="a1"
name="ChrBrad@Idle01_IndicateRightRt01"/>
<gaze start="sp:start" target="right" sbm:joint-range="EYES
NECK"/>
<event start="sp:start" message="playVideo('deboucher.mp4')"/>
<event start="a1:end" message="triggerEoBEvent()"/>
</command>
<transition>ConfirmationAction1</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml"> Je vous montrais
comment retirer le bouchon de la bouteille.</speech>
<gaze start="sp:start" target="front" sbm:joint-range="EYES
NECK"/>
<gaze start="sp:end" target="right" sbm:joint-range="EYES
NECK"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="ConfirmationAction1" type="question" wait="10"
mode="task">
<command>
<speech id="sp" type="application/ssml+xml">Avez-vous retirer le
bouchon?</speech>
<gaze start="sp:start" target="front" sbm:joint-range="EYES
NECK"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition condition="oui">ExplicationVerser</transition>
<transition condition="non">Debouchez</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml"> Vous devez retirer
le bouchon de la bouteille.</speech>
<event start="sp:end" message="triggerEoBEvent()"/>
```

```
</recontextualisation>
</utterance>

<utterance id="Debouchez" type="statement" wait="10"
mode="task">
<command>
<speech id="sp" type="application/ssml+xml"> Allez−y <break
time="1s"/> Retirez le bouchon de la bouteille.</speech>
<animation stroke="sp:start+1" id="a1"
name="ChrBrad@Idle01_YouLf01"/>
<gaze start="sp:start" target="front" sbm:joint−range="EYES
NECK"/>
<event start="a1:end" message="triggerEoBEvent()"/>
</command>
<transition>ConfirmationAction1</transition>
</utterance>

<utterance id="ExplicationVerser" type="statement" wait="10"
mode="task">
<command>
<speech id="sp" type="application/ssml+xml"> Maintenant, vous
pouvez verser l'eau dans le verre.</speech>
<animation stroke="sp:end" id="a1"
name="ChrBrad@Idle01_IndicateRightRt01"/>
<gaze start="sp:start" target="right" sbm:joint−range="EYES
NECK"/>
<event start="sp:start" message="playVideo('verser.mp4')"/>
<event start="a1:end" message="triggerEoBEvent()"/>
</command>
<transition>ConfirmationAction2</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml"> Je vous montrais
comment vous servir un verre d'eau.</speech>
<gaze start="sp:start" target="front" sbm:joint−range="EYES
NECK"/>
<gaze start="sp:end" target="right" sbm:joint−range="EYES
NECK"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="ConfirmationAction2" type="question" wait="10"
mode="task">
<command>
```

```
<speech id="sp" type="application/ssml+xml">Avez-vous verser
l'eau?</speech>
<gaze start="sp:start" target="front" sbm:joint-range="EYES
NECK"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition condition="oui">ExplicationReboucher</transition>
<transition condition="non">Versez</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml"> Vous pouvez verser
l'eau dans le verre.</speech>
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="Versez" type="statement" wait="10"
mode="task">
<command>
<speech id="sp" type="application/ssml+xml"> Allez-y <break
time="1s"/> Versez l'eau dans le verre.</speech>
<animation stroke="sp:start+1" id="a1"
name="ChrBrad@Idle01_YouLf01"/>
<gaze start="sp:start" target="front" sbm:joint-range="EYES
NECK"/>
<event start="a1:end" message="triggerEoBEvent()"/>
</command>
<transition>ConfirmationAction2</transition>
</utterance>

<utterance id="ExplicationReboucher" type="statement"
wait="10" mode="task">
<command>
<speech id="sp" type="application/ssml+xml"> Une fois que vous
avez verser l'eau, n'oubliez pas de reboucher la bouteille. <break
time="40s"/></speech>
<animation stroke="sp:end" id="a1"
name="ChrBrad@Idle01_IndicateRightRt01"/>
<gaze start="sp:start" target="right" sbm:joint-range="EYES
NECK"/>
<event start="sp:start" message="playVideo('reboucher.mp4')"/>
<event start="a1:end" message="triggerEoBEvent()"/>
</command>
<transition>ConfirmationAction3</transition>
<recontextualisation>
```

```
<speech id="sp" type="application/ssml+xml"> Je vous expliquais
comment remettre le bouchon.</speech>
<gaze start="sp:start" target="front" sbm:joint-range="EYES
NECK"/>
<gaze start="sp:end" target="right" sbm:joint-range="EYES
NECK"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="ConfirmationAction3" type="question" wait="10"
mode="task">
<command>
<speech id="sp" type="application/ssml+xml">Avez-vous bien
reboucher la bouteille?</speech>
<gaze start="sp:start" target="front" sbm:joint-range="EYES
NECK"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition condition="oui">FinExplications</transition>
<transition condition="non">Rebouchez</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml"> Vous pouvez
remettre le bouchon.</speech>
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="Rebouchez" type="statement" wait="10"
mode="task">
<command>
<speech id="sp" type="application/ssml+xml"> Allez-y <break
time="1s"/> Rebouchez la bouteille.</speech>
<animation stroke="sp:start+1" id="a1"
name="ChrBrad@Idle01_YouLf01"/>
<gaze start="sp:start" target="front" sbm:joint-range="EYES
NECK"/>
<event start="a1:end" message="triggerEoBEvent()"/>
</command>
<transition>ConfirmationAction3</transition>
</utterance>

<utterance id="FinExplications" type="statement">
<command>
```

```
<speech id="sp" type="application/ssml+xml"> C'est bien. Vous
pouvez maintenant boire votre verre d'eau </speech>
<head type="WIGGLE" repeat="2"/>
<event start="sp:start" message="stopVideo()"/>
<event start="sp:end" message="removeTv()"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition>Conclusion</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml">Vous venez de vous
servir un verre d'eau. Je vous disais</speech>
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="Conclusion" type="statement">
<command>
<speech id="sp" type="application/ssml+xml">Quand vous aurez
soif, revenez me voir!</speech>
<animation name="ChrBrad@Idle01_OfferBoth01"/>
<event start="sp:start" message="removeMineralWater()"/>
<event start="sp:start" message="removeGasWater()"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition>Smile3</transition>
<recontextualisation>
<speech id="sp" type="application/ssml+xml">Nous avions fini de
discuter. <break time="1s"/> Je vous disais.</speech>
<animation name="ChrBrad@Idle01_HoweverLf01"
start="sp:start+3"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</recontextualisation>
</utterance>

<utterance id="Smile3" type="statement">
<command>
<saccade mode="think"/>
<face id="sm" start="0" end="1.5" type="facs" au="12_left"
amount=".9"/>
<face start="0" end="1.5" type="facs" au="12_right"
amount=".9"/>
<face start="0" end="1.5" type="facs" au="6" amount="1"/>
<event start="sm:end" message="triggerEoBEvent()"/>
</command>
```

```
<transition>end</transition>
</utterance>

<utterance id="end" type="statement">
<command>
<speech id="sp" type="application/ssml+xml">A bien tot
!</speech>
<head type="NOD" amount=".4"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition>end</transition>
</utterance>

<utterance id="start" type="interruption">
<command><speech id="sp" type="application/ssml+xml">Coucou!
Je suis la!</speech>
<animation name="Louise@wave" stroke="0.5" relax="2.5"/>
<gaze target="front" sbm:joint-range="EYES NECK"/>
<event start="0" message="removeMineralWater()"/>
<event start="0" message="removeGasWater()"/>
<event start="0" message="setVolumeHigh()"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition>end</transition>
</utterance>

<utterance id="end" type="interruption">
<command><speech id="sp" type="application/ssml+xml">Nous
aitions en train de discuter.</speech>
<gaze target="front" sbm:joint-range="EYES NECK"/>
<event start="0" message="removeMineralWater()"/>
<event start="0" message="removeGasWater()"/>
<event start="0" message="setVolumeHigh()"/>
<event start="sp:end" message="triggerEoBEvent()"/>
</command>
<transition>start</transition>
</utterance>

<utterance id="wb" type="welcomeback">
<command>
<event start="0" message="triggerEoBEvent()"/>
</command>
<transition>1</transition>
</utterance>
```

```
<utterance id="iws" type="welcomeback">
<command>
<event start="0" message="triggerEoBEvent()"/>
</command>
<transition>1</transition>
</utterance>

</scenario>
```

Appendix C

Questionnaires

The first questionnaire was used in the Wizard of Oz study and the same questions were asked to the public in our events, except questions 10 to 14. The second questionnaire was used in the usability study of LOUISE. The third questionnaire was used in the pilot study for Virtual Promenade.

Questionnaire : votre avis sur les agents conversationnels

1. **Sexe :**

   ☐ masculin ☐ féminin

2. **Age : ________**

3. **Concernant l'apparence de l'agent conversationnel, que préfèreriez-vous ?**

   ☐ un jeune homme
   ☐ une jeune femme
   ☐ un petit garçon
   ☐ une petite fille
   ☐ un homme d'âge mur
   ☐ une femme d'âge mur
   ☐ un animal
   ☐ un robot
   ☐ pas de préférence

4. **Aimeriez-vous pouvoir personnaliser le personnage selon vos goûts (changer ses vêtements, sa voix, son apparence, etc.) ?**

   ☐ oui ☐ non

5. **Si vous avez répondu oui à la précédente question, qu'aimeriez-vous pouvoir choisir (cochez autant de cases que vous souhaitez) ?**

   ☐ sa tenue vestimentaire

☐ sa coupe de cheveux

☐ la forme de son visage

☐ sa voix

☐ sa personnalité

☐ autre, veuillez préciser :

____

6. **A quoi pensez-vous qu'un avatar pourrait vous être utile (cochez autant de cases que vous souhaitez) ?**

   ☐ assistant virtuel : vous rappeler vos rendez-vous, de prendre vos médicaments, les anniversaires de vos proches ; vous lire vos courriels etc.

   ☐ vous coacher (pour arrêter de fumer, manger plus équilibré, faire plus d'exercice, etc.)

   ☐ vous guider dans des tâches complexes (cuisiner, bricoler, faire des exercices de relaxation, utiliser un ordinateur, etc.)

   ☐ vous aider à contrôler des appareils connectés (lumières, chauffage, volets, télévision, etc.)

   ☐ rien

   ☐ autre, veuillez préciser :

   ____

   ____

   ____

7. **Sur quel(s) support(s) voudriez-vous visualiser l'avatar (cochez autant de cases que vous souhaitez) ?**

   ☐ sur un ordinateur

☐ sur une tablette tactile

☐ sur un smartphone

☐ sur un téléviseur

☐ sur un/des écran(s) dédié(s) à l'avatar

☐ nulle part, je ne souhaite pas utiliser cette technologie

☐ autre, veuillez préciser :

____

8. **Le système de perception, nécessaire au fonctionnement de l'avatar, comporte un capteur de mouvement, une analyse vidéo et une analyse sonore. Accepteriez-vous qu'un tel système soit installé chez vous, sachant qu'aucune donnée ne serait transmise ?**

☐ oui ☐ non

9. **Voudriez-vous que le système de perception soit activé en permanence afin que l'avatar puisse s'allumer automatiquement et rapidement chaque fois que vous voulez l'utiliser ?**

☐ oui

☐ oui, à condition que je puisse régler les plages horaires d'activation

☐ non ; je préfère qu'il ne soit activé que lorsque je décide de le mettre en marche

10. **Pendant l'expérience, avez-vous trouvé que le personnage était facile à comprendre ?**

☐ oui ☐ non

11. **Le personnage vous as-t-il paru :**

☐ Très sympathique

☐ Assez sympathique

☐ Peu sympathique

☐ Désagréable

12.Avez vous trouvé l'expérience :

☐ Très intéressante

☐ Assez intéressante

☐ Peu intéressante

☐ Pas intéressante

13. Le personnage a-t-il bien réagi à ce que vous disiez ?

☐ Très bien

☐ Assez bien

☐ Pas tellement

☐ Pas du tout

14.Voulez-vous nous faire part de vos remarques, impressions, attentes :

____

____

____

____

____

____

____

____

Questionnaire de satisfaction

1. **Sexe :** ☐ masculin ☐ féminin
2. **Age : ___________________**
3. **Le personnage vous as-t-il paru :**

   ☐ Très sympathique

   ☐ Assez sympathique

   ☐ Peu sympathique

   ☐ Désagréable

4. **Vous avez trouvé l'apparence du personnage :**

   ☐ Très agréable

   ☐ Assez agréable

   ☐ Peu agréable

   ☐ Désagréable

5. **Ce que disais le personnage était-il clair ?**

   ☐ Très clair

   ☐ Assez clair

   ☐ Peu clair

   ☐ Pas du tout clair

6. **Les consignes données par le personnage était-elles facile à suivre ?**

   ☐ Très faciles

   ☐ Assez faciles

   ☐ Assez difficiles

   ☐ Très difficiles

7. **Le rythme de la conversation vous a paru trop rapide ?**

   ☐ Oui, beaucoup

   ☐ Oui, un peu

   ☐ Non, pas tellement

   ☐ Non, pas du tout

8. **Aimeriez-vous pouvoir choisir l’apparence du personnage ?**

   ☐ oui ☐ non

9. **A quoi pensez-vous qu’un avatar pourrait vous être utile (cochez autant de cases que vous souhaitez) ?**

   ☐ assistant virtuel : vous rappeler vos rendez-vous, de prendre vos médicaments, les anniversaires de vos proches ; vous lire vos courriels etc.

   ☐ vous coacher (pour arrêter de fumer, manger plus équilibré, faire plus d’exercice, etc.)

   ☐ vous guider dans des tâches complexes (cuisiner, bricoler, faire des exercices de relaxation, utiliser un ordinateur, etc.)

   ☐ vous aider à contrôler des appareils connectés (lumières, chauffage, volets, télévision, etc.)

   ☐ rien

   ☐ autre, veuillez préciser :

   ____

   ____

   ____

10. Préférez-vous que le personnage vous appelle par votre prénom ou par votre nom ?

☐ Prénom ☐ Nom

11. Préférez-vous qu'il vous tutoie ou qu'il vous vouvoie ?

☐ Tutoiement ☐ Vouvoiement

12. Sur quel(s) support(s) voudriez-vous visualiser le personnage (cochez autant de cases que vous souhaitez) ?

☐ sur un ordinateur

☐ sur une tablette tactile

☐ sur un smartphone

☐ sur un téléviseur

☐ sur un/des écran(s) dédié(s)

☐ nulle part, je ne souhaite pas utiliser cette technologie

13. Le système de perception, nécessaire au fonctionnement de l'avatar, comporte un capteur de mouvement, une analyse vidéo et une analyse sonore. Accepteriez-vous qu'un tel système soit installé chez vous ?

☐ oui ☐ non

14. Voudriez-vous que le système de perception soit activé en permanence afin que l'avatar puisse s'allumer automatiquement et rapidement chaque fois que vous voulez l'utiliser ?

☐ oui

☐ oui, à condition que je puisse régler les plages horaires d'activation

☐ non ; je préfère qu'il ne soit activé que lorsque je décide de le mettre en marche

Questionnaire de satisfaction

1) **Sexe :** ☐ Masculin ☐ Féminin

2) **Age :** _______ ans

3) **Le nombre de personnages proposé est-il satisfaisant ?**

☐ Oui ☐ Non

4) **La possibilité de choisir son personnage est :**

☐ Très Importante

☐ Assez Importante

☐ Peu importante

☐ Pas du tout importante

5) **L'apparence du personnage que vous avez choisi est-elle satisfaisante ?**

☐ Très satisfaisante

☐ Assez satisfaisante

☐ Peu satisfaisante

☐ Insatisfaisante

6) **Trouvez vous que les animations de la marche du personnage sont :**

☐ Très réalistes

☐ Assez réalistes

☐ Peu réalistes

- [ ] Pas du tout réalistes

7) Vous êtes vous identifié(e) physiquement au personnage que vous avez choisi ?

- [ ] Tout à fait
- [ ] Un peu
- [ ] Pas tellement
- [ ] Pas du tout

8) Avez-vous eu l'impression de marcher à travers le personnage ?

- [ ] Oui, Beaucoup
- [ ] Oui, un peu
- [ ] Non, pas tellement
- [ ] Non, pas du tout

9) Les mouvements du fauteuil donnent-ils l'impression de marcher ?

- [ ] Oui, Beaucoup
- [ ] Oui, un peu
- [ ] Non, pas tellement
- [ ] Non, pas du tout

10) Les mouvements du fauteuil sont-ils agréables ?

- [ ] Oui, Beaucoup
- [ ] Oui, un peu

- [ ] Non, pas tellement
- [ ] Non, pas du tout

11) Les différents lieux que l'on vient de vous présenter sont-ils agréables ?

- [ ] Très agréables
- [ ] Assez agréables
- [ ] Peu agréables
- [ ] Épouvantables

12) L'utilisation de la manette est :

- [ ] Très facile
- [ ] Assez facile
- [ ] Plutôt difficile
- [ ] Très difficile

13) Les explications fournies vous ont-elles aidé à contrôler le personnage ?

- [ ] Oui, Beaucoup
- [ ] Oui, un peu
- [ ] Non, pas tellement
- [ ] Non, pas du tout

14) Les instructions présentées à l'écran sont :

- [ ] Très lisibles

- [ ] Assez lisibles
- [ ] Peu lisible
- [ ] Illisible

15) Avez-vous apprécié jouer à ce jeu ?

- [ ] Très satisfait
- [ ] Assez satisfait
- [ ] Peu satisfait
- [ ] Pas du tout

16) Selon vous, la vue à la 1ere personne est plus immersif qu'à la 3ième personne ?

- [ ] Oui
- [ ] Non

17) Votre peur de chuter a t-elle diminué entre le début et la fin de séance ?

- [ ] Oui, Beaucoup
- [ ] Oui, un peu
- [ ] Non, pas tellement
- [ ] Non, pas du tout

18) Souhaiteriez-vous continuer cette activité dans une prochaine séance ?

- [ ] Oui
- [ ] Non

Résumé

Les travaux présentés dans ce manuscrit portent sur l'utilisation des humains virtuels et, plus globalement, des technologies du jeu vidéo pour améliorer le soin aux personnes âgées atteintes de troubles cognitifs. Nos travaux s'articulent autour de deux cas d'utilisation des humains virtuels : comme non-soi et comme soi. Plus spécifiquement, nous avons conçu, implémenté et évalué (1) un agent conversationnel animé, appelé LOUISE (LOvely User Interface for Servicing Elders), ayant pour but de servir d'interface utilisateur accessible dans les dispositifs de compensation cognitives à destination des personnes âgées atteintes de troubles cognitifs, et (2) un dispositif de thérapie psychologique par la réalité virtuelle pour traiter les conséquences des chutes, appelé Promenade virtuelle. Ces deux projets ont été menés suivant des principes de conception participative en living lab, impliquant plusieurs parties prenantes dans le processus de conception (patients, aidants et professionnels de santé). Nous avons produit des prototypes utilisables et appréciés par les personnes âgées qui les ont testés. Nous avons également pu mieux comprendre les spécificités de l'interaction avec un agent conversationnel pour ce public et identifier, en particulier, qu'il est très important à leurs yeux qu'un dispositif mettant en œuvre des humains virtuels soit plaisant.

Mots-Clés

Agents conversationnels animés, jeux sérieux, santé, ergonomie, démence, personnes âgées

Abstract

The work presented in this manuscript investigates the use of virtual huamans, and, more broadly, of video game technologies to improve the care of older adults with cognitive impairment. Our work revolves around two use cases of virtual humans: as non-self and as self. More specifically, we have designed, implemented and evaluated (1) an embodied conversational agent, called LOUISE (LOvely User Interface for Servicing Elders), meant to serve as an accessible user interface in cognitive compensation systems for older adults with cognitive impairment, and (2) a system for virtual reality-based psychological therapy addressing the consequences of falls, called Virtual Promenade. These two projects have been conducted according to the principles of living lab participatory design, involving several stakeholders in the design process (patients, caregivers and healthcare professionals). We were able to create prototypes that the older adults with cognitive impairment who participated in our tests could successfully use and enjoyed using. We were also able to better understand the specificities of interactions between older adults with cognitive impairment and embodied conversational agents and identify, among other findings, that it is very important for a virtual human-based system targetting this public to be pleasant.

Keywords

Embodied conversational agents, serious games, health, ergonomics, dementia, older adults