



**HAL**  
open science

**Un système semi-autonome pour la création de corpus  
d'interaction humain-machine en réalité virtuelle  
Application au système Acorformed pour la formation  
des médecins à l'annonce d'évènements indésirable  
graves**

Magalie Ochs, Philippe Blache, Grégoire de Montcheuil, Jean-Marie Pergandi,  
Jorane Saubesty, Daniel Francon, Daniel Mestre

► **To cite this version:**

Magalie Ochs, Philippe Blache, Grégoire de Montcheuil, Jean-Marie Pergandi, Jorane Saubesty, et al.. Un système semi-autonome pour la création de corpus d'interaction humain-machine en réalité virtuelle Application au système Acorformed pour la formation des médecins à l'annonce d'évènements indésirable graves. WACAI-2018, May 2018, Porquerolles, France. hal-01907603

**HAL Id: hal-01907603**

**<https://hal.science/hal-01907603>**

Submitted on 29 Oct 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Un système semi-autonome pour la création de corpus d'interaction humain-machine en réalité virtuelle

Application au système Acorformed pour la formation des médecins à l'annonce d'évènements indésirable graves

Magalie Ochs<sup>1</sup>, Philippe Blache<sup>2</sup>, Grégoire de Montcheuil<sup>2,3,4</sup>, Jean-Marie Pergandi<sup>3</sup>,  
Jorane Saubesty<sup>2</sup>, Daniel Francon<sup>5</sup>, et Daniel Mestre<sup>3</sup>  
Aix-Marseille Université CNRS ENSAM, Université de Toulon, <sup>1</sup>LIS UMR 7020, <sup>2</sup>LPL UMR 7309, <sup>3</sup>ISM  
UMR7287 ; <sup>4</sup>Boréal Innovation, <sup>5</sup>Institut Paoli-Calmettes (IPC), Marseille, France

## ABSTRACT

Dans cet article, nous introduisons une méthodologie particulière fondée sur l'analyse séquentielle de plusieurs corpus : à partir d'un corpus d'interaction humain-humain pour construire un système semi-autonome ayant pour objectif de recueillir un nouveau corpus d'interaction humain-machine, une étape avant le développement d'un système entièrement autonome construit sur la base de l'analyse des corpus collectés. La méthodologie présentée est illustrée dans le contexte du développement d'une plateforme pour la formation en réalité virtuelle des médecins à l'annonce d'évènements indésirables graves.

## Keywords

Patient virtuel ; réalité virtuelle ; corpus ; domaine de la santé ; formation par la simulation

## 1. INTRODUCTION

Cet article présente une *méthodologie fondée sur des corpus* élaborée pour concevoir un *environnement de réalité virtuelle* visant à former des médecins à l'annonce de *mauvaises nouvelles à un patient virtuel*. De nombreux travaux ont montré que les médecins doivent être formés non seulement pour effectuer des actes médicaux ou chirurgicaux, mais aussi pour développer des compétences relationnelles pour la communication avec les patients [3, 11, 15]. Parmi toutes les mauvaises nouvelles possibles, les médecins peuvent être confrontés à la situation complexe d'annoncer un dommage associé à un soin, qui peut engager leur responsabilité : situation médicale imprévisible, dysfonctionnement, erreur médicale, etc. La façon dont *les médecins livrent des mauvaises nouvelles liés aux dommages associés aux soins* a un impact significatif sur le processus thérapeutique : évolution de la maladie, respect des recommandations de trai-

tement, possibilités de litige [1]. Cependant, les cliniciens expérimentés et les étudiants en médecine considèrent cette tâche comme difficile, intimidante et stressante.

La formation des professionnels de la santé pour l'annonce d'évènements indésirables graves est aujourd'hui recommandée par plusieurs agences nationales (e.g. La Haute Autorité de la Santé, HAS). Ces formations sont organisées sous la forme d'ateliers au cours desquels les médecins annoncent des mauvaises nouvelles aux acteurs jouant le rôle de patients. Cette solution est complexe à mettre en œuvre : elle nécessite plusieurs personnes, elle est coûteuse et prend beaucoup de temps (chaque séance de 30 mn nécessite une heure de préparation). Notre projet <sup>1</sup> vise à développer *un outil de formation en réalité virtuelle* avec un agent conversationnel animé jouant le rôle d'un patient virtuel.

Les approches classiquement utilisées pour développer un système interactif incluant un agent conversationnel animé consistent à analyser automatiquement ou manuellement *un corpus* annoté d'interactions humain-humain [7]. Ensuite, un prototype entièrement autonome est généralement développé et évalué à travers des études perceptives. Dans l'approche présentée ici, la méthodologie est fondée sur plusieurs corpus de différents types, permettant d'acquérir des données et des informations spécifiques sur les modules et fonctionnalités requis. Plus précisément, la procédure commence classiquement avec la collecte de données naturelles (enregistrements d'interactions humain-humain), mais nous introduisons une nouvelle étape fondée sur l'acquisition de données interactionnelles humain-machine. Dans ce cadre est développé un agent conversationnel mettant en œuvre à la fois des modules automatiques et manuels, permettant de simuler une interaction humain-machine entièrement automatisée, sans avoir besoin de développer l'ensemble du système (en particulier le module de compréhension).

Cette méthodologie présente plusieurs avantages. Les corpus d'interaction humain-humain et humain-machine correspondent au même contexte. Leur étude permet de préciser les caractéristiques comportementales verbales ou non verbales que l'utilisateur peut adopter face à un humain virtuel en comparaison aux interactions interpersonnelles. En conséquence, le prototype correspond complètement au comportement observé des utilisateurs lors d'une interaction humain-

1. ACORFORMed, [12], <http://www.lpl-aix.fr/~acorformed/>.

machine. De plus, l'utilisation d'un système semi-autonome permet de s'abstraire des modules critiques, difficiles à développer et cruciaux pour une interaction réussie, par exemple le système de reconnaissance de la parole, qui peut fortement détériorer l'interaction en cas de bugs répétés. Ces modules sont intégrés dans une troisième étape sans impact sur l'évaluation du premier prototype. Dans le reste de cet article, nous illustrons une application de cette méthodologie avec le développement d'un prototype pour la formation des médecins pour l'annonce d'événements indésirables graves aux patients.

L'article est organisé de la manière suivante. Dans la première section, nous présentons la méthodologie globale fondée sur différents corpus. Dans la section 3, nous présentons le corpus multimodal d'interaction humain-humain utilisé pour développer le premier prototype semi-autonome. La section 4 est consacrée à la présentation de la plateforme semi-autonome et des différents outils développés pour collecter un corpus annoté automatiquement d'interaction humain-machine. Dans la section 5, nous décrivons le corpus humain-machine et la manière dont il est utilisé pour développer le prototype final.

## 2. UNE MÉTHODOLOGIE FONDÉE SUR L'ANALYSE DE PLUSIEURS CORPUS

La méthodologie utilisée pour développer la plateforme de formation en réalité virtuelle comprend plusieurs étapes, chacune fondée sur la création et l'analyse d'un corpus (interaction humain-humain ou humain-machine) dans le contexte de l'annonce de mauvaises nouvelles. Ces étapes peuvent être spécifiées comme suit (Figure 1) :

1. Analyse d'un corpus d'interaction humain-humain pour la spécification de l'organisation du discours multimodal, la définition des différents modules comportementaux, le développement d'un premier prototype (agent semi-autonome).
2. Le prototype est ensuite utilisé pour collecter un nouveau corpus d'interaction humain-machine ;
3. Le nouveau corpus humain-machine est analysé pour spécifier les modules manquants et améliorer les différentes fonctionnalités afin de développer la version autonome du système ;
4. Évaluation du système, classiquement avec des études perceptives.

Pour la première étape, le corpus initial d'interaction médecin-patient (enregistré lors des séances d'entraînement) a été annoté manuellement (Section 3). Les résultats de l'analyse ont été utilisés pour développer une première version de la plateforme de formation en réalité virtuelle. Cette plateforme, telle que décrite dans la section 4, est semi-autonome : certains modules de l'architecture sont simulés par un expérimentateur. Cette plateforme semi-autonome a été utilisée pour collecter un nouveau corpus d'interaction humain-machine, en considérant différents dispositifs d'interaction (PC, casque de réalité virtuelle et salle de réalité virtuelle). Les corpus collectés sont utilisés pour développer les modules manquants et ainsi améliorer le prototype initial dans la perspective d'une plateforme de formation en réalité virtuelle entièrement autonome.

## 3. ANALYSE D'UN CORPUS MULTIMODAL HUMAIN-HUMAIN

La modélisation du patient virtuel repose sur un corpus audiovisuel d'interactions entre médecins et acteurs jouant le rôle de patients (appelés "patients standardisés") lors de véritables formations dans des établissements médicaux français (il n'est pas possible, pour des raisons éthiques, d'enregistrer de vraies situations d'annonce de mauvaises nouvelles). L'utilisation de "patients standardisés" dans la formation médicale est une pratique courante. Les acteurs sont soigneusement formés (dans notre projet, les acteurs sont aussi des infirmiers) et suivent des scénarios préétablis définis par des experts pour jouer les réactions des patients les plus fréquemment observés. En moyenne, une consultation simulée dure 9 minutes. Le corpus collecté est composé de 13 vidéos d'interaction patient-médecin (chaque vidéo implique un médecin différent et/ou un couple acteur-patient différent), avec différents scénarios<sup>2</sup>. Le corpus initial a été annoté de manière semi-manuelle sur une durée totale de 119 minutes.

Différents outils ont été utilisés pour annoter le corpus. Tout d'abord, le corpus a été segmenté automatiquement en utilisant SPPAS [5] et transcrit manuellement en utilisant Praat [6]. Les comportements non-verbaux des médecins et des patients ont été annotés manuellement en utilisant ELAN [17]. Différents gestes des médecins et des patients ont été annotés : mouvements de la tête, changements de posture, direction du regard, mouvements des sourcils, gestes de la main et sourires. Trois experts ont annoté chacun un tiers du corpus. Afin de valider l'annotation, 5% du corpus a été annoté par un autre annotateur. L'accord inter-annotateur, en calculant le Kappa de Cohen, était satisfaisant ( $k = 0,63$ ). Le corpus est décrit plus en détails dans [14].

Le corpus annoté a été analysé à deux fins différentes :

- pour construire le *modèle de dialogue du patient virtuel* : le modèle de dialogue du patient virtuel est fondé sur la notion de "terrain commun" ("*common ground*") [8, 18], *i.e.* un modèle de situation représenté par différentes variables qui est mis à jour en fonction de l'échange d'informations entre les interlocuteurs. Les variables décrivant le modèle de situation (par exemple la cause du dommage), propres à l'annonce de mauvaises nouvelles, ont été définies à partir de l'analyse manuelle du corpus transcrit et étant donnés les objectifs pédagogiques en termes de dialogue [12]. Nous avons utilisé pour l'implémentation le système de dialogue OpenDial [10]. Dans cette approche, le modèle sélectionne automatiquement la réaction verbale du patient en fonction de l'énoncé verbal reconnu du médecin ainsi que suivant l'historique du dialogue (les informations déjà fournies par le médecin étant encodées dans le *terrain commun*). Le modèle de dialogue est décrit plus en détail dans [12] ;
- concevoir les *comportements non-verbaux du patient virtuel* : le corpus a été utilisé pour enrichir la bibliothèque de comportements non-verbaux du patient virtuel avec des gestes spécifiques pour l'annonce de mauvaises nouvelles. Les gestes récurrents identifiés dans le corpus sont ceux utilisés pour indiquer la douleur. Au total, nous avons enrichi la bibliothèque de comportements non-verbaux du personnage virtuel avec 16 nouveaux gestes spécifiques à ce contexte.

---

2. Le corpus est sur Ortolang

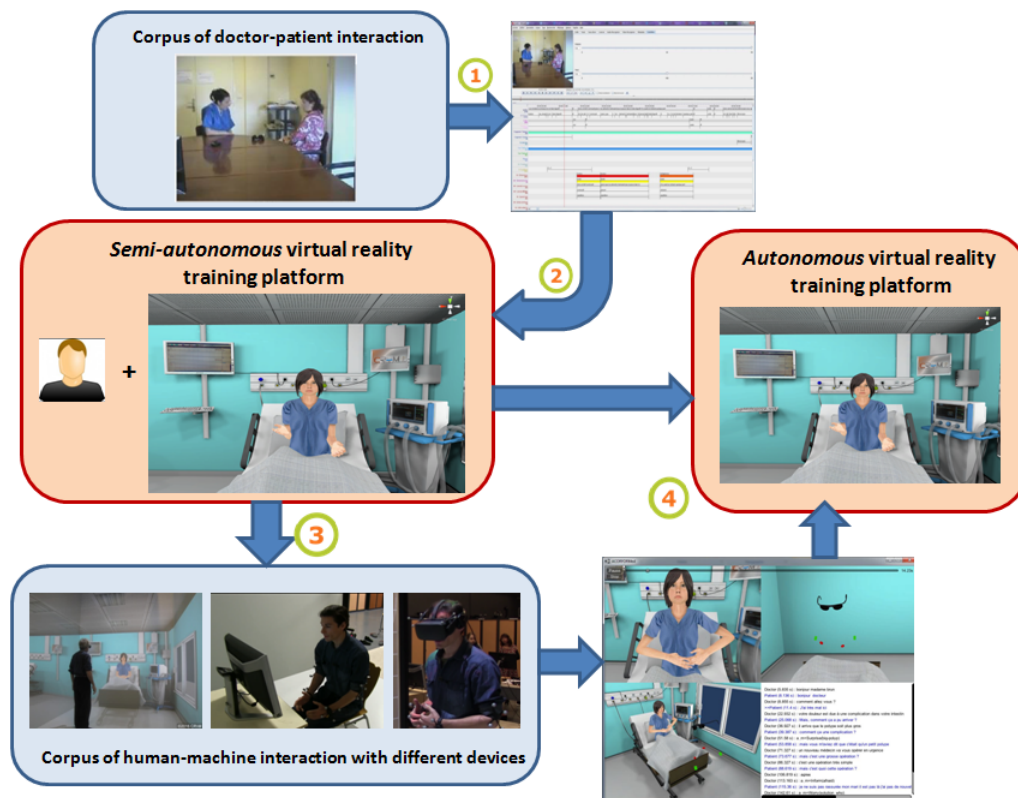


Figure 1: Méthodologie

Le modèle de dialogue ainsi que la bibliothèque de comportements non-verbaux du patient virtuel ont été utilisés pour mettre en œuvre un premier prototype de la plateforme de formation en réalité virtuelle, comme décrit dans la section suivante.

#### 4. UNE PLATEFORME SEMI-AUTONOME DE FORMATION EN RÉALITÉ VIRTUELLE

Afin d'évaluer l'architecture générale de la plateforme de formation et afin de collecter un corpus d'interaction humain-machine spécifique à notre projet, nous avons développé une plateforme semi-autonome dont l'architecture est décrite à travers la figure 2.

La plateforme est *semi-autonome* car certains modules du système sont automatiques (par exemple la génération de dialogue) et d'autres sont manuels. En particulier, les modules de reconnaissance et de compréhension de la parole sont simulés par un humain : la production verbale du médecin est interprétée en temps réel par l'opérateur qui sélectionne le signal d'entrée adéquat à transmettre au système de dialogue. En effet, ces modules peuvent être particulièrement critiques en cas d'échec et endommager alors fortement l'interaction. Ils représentent en outre une des parties les plus difficiles du système à développer. Le remplacement du module par l'opérateur permet une reconnaissance et une compréhension parfaite de la parole. Ceci permet de contrôler complètement les paramètres correspondants et de se concentrer sur l'évaluation des autres modules, tels que la

supervision du dialogue et le comportement non-verbal du patient virtuel. De plus, cela rend possible l'évaluation de l'interaction globale (par exemple le sentiment de présence, la satisfaction de l'utilisateur, la perception de la crédibilité de l'agent) et cela sans le développement complet du système à travers une évaluation du système semi-autonome.

Une interface spécifique a été conçue à cet effet pour permettre à l'expérimentateur de sélectionner les phrases aussi proches que possible de celles qui ont été dites par les médecins (Figure 3).

L'interface contient 136 phrases prototypiques organisées en différentes phases de dialogue : les salutations, les sentiments du patient, la description du problème chirurgical, la description de la remédiation. Ces phrases ont été définies sur la base de l'analyse du corpus transcrit d'interaction médecin-patient (Section 3). Chaque phrase prototypique correspond une famille d'énoncés possibles, tels qu'identifiés dans le corpus. Les phrases sont codées dans un fichier XML. Les raccourcis clavier sont associés à chaque phrase/motif et peuvent être configurés pour être facilement sélectionnés par l'expérimentateur. Plusieurs pré-tests ont été construits pour tester l'interface et former l'expérimentateur. Notez qu'à la différence d'un "Magicien d'Oz", l'expérimentateur ne sélectionne pas la réaction du patient virtuel mais envoie seulement au modèle de dialogue la phrase du médecin reconnue.

Le système de dialogue génère ensuite une séquence d'instructions, envoyée à un système d'animation d'ACA, le système VIB [13]. Ce système calcule les paramètres d'animation (*Facial Animation Parameters* - FAP - et *Behavioral*

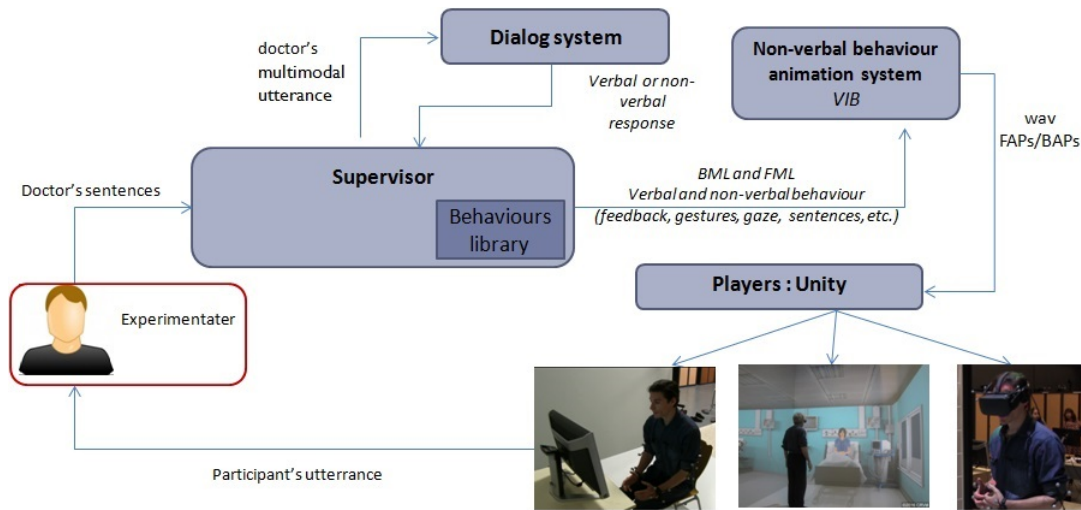


Figure 2: Architecture générale de la plateforme de formation en réalité virtuelle

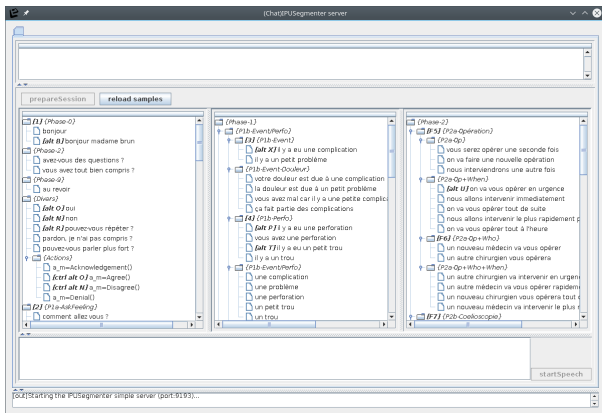


Figure 3: Image d'écran de l'interface de l'expérimentateur pour sélectionner les phrases correspondant aux phrases émises par le médecin

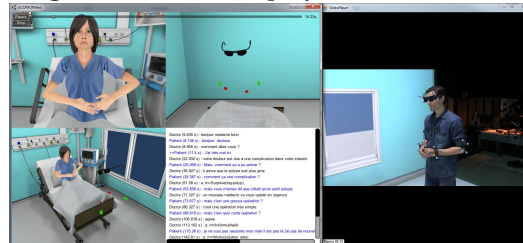
*Animation Parameters* - BAP) pour animer le visage et le corps du patient virtuel. Le résultat est codé en XML et décrit l'intention communicative à effectuer (encodée en FML, *Function Markup Language*) ainsi que les signaux non-verbaux à exprimer (codés en BML, *Behavior Markup Language*). De plus, le système VIB contient un système de synthèse vocale [2] pour générer le discours de manière synchronisée avec le comportement non-verbal (y compris l'animation labiale).

Afin d'expérimenter aussi largement que possible la validité de l'approche, nous avons implémenté le patient virtuel sur différentes plateformes : PC, casque de réalité virtuelle et salle de réalité virtuelle (CAVE). La salle de réalité virtuelle est constituée d'un espace cubique de 3 mètres de profondeur, 3 mètres de large et 4 mètres de haut avec trois écrans verticaux et un écran horizontal (plancher). Cela offre une immersion sensorielle optimale à l'utilisateur. L'environnement a été conçu pour simuler une véritable salle de réveil où l'annonce de mauvaises nouvelles est généralement effectuée.

L'agent virtuel fondé sur la plateforme VIB a été intégré via *Unity*.

Afin d'enregistrer l'interaction et de créer un corpus d'interaction humain-machine dans le contexte de l'annonce d'événements indésirables graves, nous avons mis en place un dispositif spécifique. Premièrement, le médecin est filmé à l'aide d'une caméra. Ses gestes et mouvements de tête sont enregistrés : sa tête (lunettes stéréo), ses coudes et ses poignets sont équipés de cibles suivies. Un microphone haut de gamme enregistre de manière synchrone l'expression verbale du participant. Quant à l'agent virtuel, ses gestes et ses expressions verbales sont enregistrés à partir de *Unity*. La visualisation de l'interaction se fait à travers un lecteur vidéo 3D que nous avons développé (Figure 4). Cet outil rejoue de manière synchrone l'animation et l'expression verbale de l'agent virtuel ainsi que les mouvements et la vidéo du participant.

Figure 4: Lecteur "playback" 3D des interactions



Cet environnement facilite la collecte de corpus d'interaction médecin-patient virtuel afin d'analyser les comportements verbaux et non-verbaux dans différents environnements immersifs. Le corpus collecté est à la base du développement de la plateforme de formation en réalité virtuelle entièrement autonome.

## 5. DU CORPUS D'INTERACTION HUMAIN-MACHINE À LA PLATE-FORME AUTONOME DE FORMATION

En utilisant le système semi-autonome décrit dans la section précédente, nous avons collecté 108 interactions de participants avec le patient virtuel. Au total, 36 personnes ont participé à l'expérimentation. Dix d'entre elles sont de vrais médecins qui ont déjà une expérience dans l'annonce de mauvaises nouvelles auprès de vrais patients. Chaque participant a interagi avec le patient virtuel trois fois avec trois dispositifs différents : PC, casque de réalité virtuelle et salle de réalité virtuelle. La tâche des participants était d'annoncer une perforation digestive après une endoscopie gastroentérologique en période post-opératoire immédiate<sup>3</sup>. Avant l'interaction, des instructions écrites ont été présentées aux participants : le rôle qu'ils avaient à jouer est le rôle d'un médecin qui venait d'opérer le patient virtuel pour enlever un polype dans l'intestin. Une perforation digestive s'est produite pendant la chirurgie. Ces instructions écrites expliquent précisément les causes du problème, les effets (la douleur) et les solutions proposées (une nouvelle intervention chirurgicale, de toute urgence). Les participants sont invités à lire les instructions plusieurs fois ainsi qu'avant chaque interaction. La compréhension est vérifiée au moyen d'un questionnaire oral. Chaque participant a l'instruction d'annoncer cette situation médicale au patient virtuel trois fois avec trois dispositifs différents : PC, casque de réalité virtuelle et salle de réalité virtuelle. L'ordre des conditions a été contrebalancé.

Le corpus collecté est composé de 108 vidéos (36 par dispositif). La durée totale du corpus est de 5h34 (dont deux heures avec de vrais médecins). En moyenne, une interaction dure 3mn16 (un exemple d'interaction est présenté sur le site ACORFORMed). Notez que grâce aux outils décrits dans la section précédente, une partie du comportement non-verbal du participant peut être annotée automatiquement.

Afin de valider ce premier prototype, nous avons demandé aux participants de remplir différents questionnaires sur leur expérience subjective pour mesurer leur sentiment de présence (avec le *Igroup Presence Questionnaire*, IPQ [16]), le sentiment de co-présence [4], et la perception de la crédibilité du patient virtuel (questions extraites de [9]). Ces évaluations subjectives nous permettent d'étiqueter les vidéos du corpus avec les résultats de ces tests, puis de corrélérer des mesures objectives (par exemple le comportement verbal et non-verbal des participants) à des mesures subjectives (ex. de la crédibilité du patient virtuel).

Nous analysons actuellement le corpus pour le développement de la plateforme de formation entièrement autonome (en particulier le module de compréhension et de génération). Nous avons déjà utilisé le corpus pour entraîner et tester le système de reconnaissance vocale, afin de garantir que le système de reconnaissance vocale puisse reconnaître les participants avec précision. Nous vérifions également que les mots et phrases reconnus activent correctement les règles attendues dans le modèle de dialogue. Ces règles de compréhension sont adaptées en conséquence. De plus, le corpus est également utilisé pour comparer les comportements non-verbaux à travers les différents dispositifs. Par exemple, nous

3. Le scénario a été soigneusement choisi avec les partenaires médicaux du projet pour plusieurs raisons : par exemple le panel de dommages résultants, la difficulté de l'annonce, ses caractéristiques standard d'annonce.

formulons l'hypothèse que les participants utilisent moins de gestes dans la condition de casque de réalité virtuelle puisqu'ils ne voient pas leur corps. Selon les résultats, la reconnaissance automatique des gestes pourrait être adaptée.

## 6. CONCLUSIONS

Le développement d'un agent conversationnel incarné avec des compétences conversationnelles spécifiques est un défi pour la communication humain-machine, non seulement en ce qui concerne les capacités de dialogue de l'agent, mais aussi l'adéquation de l'environnement à la tâche. La création de corpus riches est une condition préalable à cet objectif. Nous avons présenté dans cet article une plateforme semi-automatique répondant à ces besoins : offrir les possibilités d'une interaction humain-machine écologique dans un environnement de réalité virtuelle pour ensuite développer une architecture finale du système en implémentant les fonctionnalités de tous les modules.

## 7. ACKNOWLEDGEMENTS

Ce travail est financé par l'Agence Nationale de la Recherche (ANR) pour le projet ACORFORMED (ANR-14-CE24-0034-02) et soutenu par les subventions ANR-16-CONV-0002 (ILCB), ANR-11-LABX-0036 (BLRI) et ANR-11-IDEX-0001-02 (A\*MIDEX).

## 8. REFERENCES

- [1] A. Andrade, A. Bagri, K. Zaw, B. Roos, and J. Ruiz. Avatar-mediated training in the delivery of bad news in a virtual world. *Journal of palliative medicine*, 13(12) :1415–1419, 2010.
- [2] M. P. Aylett and C. J. Pidcock. The cerevoice characterful speech synthesiser sdk. In *IVA*, pages 413–414, 2007.
- [3] W. Baile, R. Buckman, R. Lenzi, G. Glober, E. Beale, and A. Kudelka. Spikes-a six-step protocol for delivering bad news : application to the patient with cancer. *Oncologist*, 5(4) :302–311, 2000.
- [4] J. N. Bailenson, K. Swinich, C. Hoyt, S. Persky, A. Dimov, and J. Blascovich. The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence : Teleoperators and Virtual Environments*, 14(4) :379–393, 2005.
- [5] B. Bigi. Sppas : a tool for the phonetic segmentations of speech. In *The eighth international conference on Language Resources and Evaluation*, pages 1748–1755, 2012.
- [6] P. Boersma and D. Weenik. Praat : a system for doing phonetics by computer. report of the institute of phonetic sciences of the university of amsterdam. *Amsterdam : University of Amsterdam*, 1996.
- [7] J. Cassell. *Embodied conversational agents*. MIT press, 2000.
- [8] S. Garrod and M. J. Pickering. Why is conversation so easy? *Trends in cognitive sciences*, 8(1) :8–11, 2004.
- [9] M. Gerhard, D. J. Moore, and D. J. Hobbs. Continuous presence in collaborative virtual environments : Towards a hybrid avatar-agent model for user representation. In *International Workshop on*

*Intelligent Virtual Agents*, pages 137–155. Springer, 2001.

- [10] P. Lison and C. Kennington. OpenDial : A Toolkit for Developing Spoken Dialogue Systems with Probabilistic Rules. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Demonstrations)*, pages 67–72, Berlin, Germany, 2016. Association for Computational Linguistics.
- [11] K. Monden, L. Gentry, and T. Cox. Delivering bad news to patients. *Proceedings (Baylor University Medical Center)*, 29(1), 2016.
- [12] M. Ochs, G. Montcheuil, J.-M. Pergandi, J. Saubesty, B. Donval, C. Pelachaud, D. Mestre, and P. Blache. An architecture of virtual patient simulation platform to train doctor to break bad news. In *International Conference on Computer Animation and Social Agents (CASA)*, 2017.
- [13] C. Pelachaud. Studies on gesture expressivity for a virtual agent. *Speech Communication*, 51(7) :630–639, 2009.
- [14] C. Porhet, M. Ochs, J. Saubesty, G. Montcheuil, and R. Bertrand. Mining a multimodal corpus of doctor’s training for virtual patient’s feedbacks. In *Proceedings of 19th ACM International Conference on Multimodal Interaction (ICMI)*, Glasgow, UK, 2017.
- [15] M. Rosenbaum, K. Ferguson, and J. Lobas. Teaching medical students and residents skills for delivering bad news : A review of strategies. *Acad Med*, 79, 2004.
- [16] T. W. Schubert. The sense of presence in virtual environments : A three-component scale measuring spatial presence, involvement, and realness. *Zeitschrift für Medienpsychologie*, 15(2) :69–71, 2003.
- [17] H. Sloetjes and P. Wittenburg. Annotation by category : Elan and iso dcr. In *LREC*, 2008.
- [18] R. Stalnaker. Common ground. *Linguistics and Philosophy*, 25(5) :701–721, 2002.