

LA COMPREHENSION DE LA MACHINE A TRAVERS L'EXPRESSION NON-VERBALE

Magalie Ochs, Elisabetta Bevacqua, Ken Prepin, Quoc Anh Le, Yu Ding, Jing Huang,
Radoslaw Nieviadomski, Catherine Pelachaud

*Greta Team, CNRS LTCI, Télécom ParisTech, Paris, France
greta@telecom-paristech.fr*

Résumé: Pour permettre une compréhension mutuelle dans une interaction, le comportement non-verbal joue un rôle prépondérant. Dans cet article, nous explorons comment une machine, et plus particulièrement des agents conversationnels animés ou des robots humanoïdes, peut utiliser le comportement non-verbal pour *se faire comprendre* par l'utilisateur et montrer leur *compréhension* de ce dernier. Nous présentons les problématiques liées au développement de tels artefacts et des méthodes pour y répondre.

Mots clés: Agent Conversationnel Animé (ACA), comportement non-verbal, émotion, synchronie.

INTRODUCTION

Plusieurs recherches ont montré que les utilisateurs tendent à interagir avec les ordinateurs comme s'ils étaient de véritables personnes, lorsque ces machines possèdent des capacités communicatives similaires à celles des humains (Reeves *et al.*, 1996). En conséquence, les interfaces homme-machine sont implémentées de plus en plus avec le but de simuler les interactions humaines. Plus ces interfaces empruntent les caractéristiques de la communication humaine, plus les utiliser devient simple et accessible (Ball *et al.*, 2001). Ce rapprochement de la communication humaine dans les interfaces virtuelles peut être atteint en utilisant des artefacts humanoïdes capables de simuler la richesse de la conversation humaine. Les récents progrès dans les technologies informatiques ont rendu possible la création de ces artefacts, appelés *Agents Conversationnels Animés* (ACAs). Un *agent conversationnel animé* est un personnage virtuel créé par l'ordinateur qui peut converser avec les utilisateurs d'une façon naturelle, similaire à celle appliquée par les humains (Cassel, 2000). Les ACAs ont trouvé place dans plusieurs applications : par exemple, comme assistants pour aider les utilisateurs à exécuter une tâche ou encore comme présentateur (André *et al.*, 2001).

Pour permettre une interaction efficace entre les individus et ces agents jouant des rôles typiquement incarnés par des humains, il est essentiel qu'humains et agents se *comprennent*. Les utilisateurs doivent comprendre le message que veut transmettre l'agent virtuel et réciproquement l'agent virtuel doit être capable de comprendre le message explicitement et implicitement exprimé par l'utilisateur. La compréhension d'un message ne passe pas uniquement par la compréhension de la parole. Les *comportements non-verbaux* du locuteur et de l'interlocuteur jouent en effet un rôle essentiel en fournissant des informations qui peuvent être aussi bien complémentaires que redondantes par rapport au discours (Kendon, 1994). Les interactions humain-machine impliquent de plus une dimension intrinsèquement

émotionnelle. Les utilisateurs ressentent en effet de nombreuses émotions, telles que la joie ou la frustration, durant leur interaction avec une machine (Picard, 1997). Les agents virtuels doivent être alors capables à la fois de comprendre ces émotions pour y répondre, et à la fois d'exprimer eux même des émotions pour permettre une communication sur un plan émotionnel.

Dans cet article, nous abordons différentes problématiques liées aux développements d'agents conversationnels animés capables de *se faire comprendre* (Section 1) et capables de montrer de la *compréhension* (Section 2) à travers leurs comportements non-verbaux. Pour chaque problématique présente, une méthode pour y répondre est introduite.

1. VERS UNE MACHINE CAPABLE DE SE FAIRE COMPRENDRE

Pour *se faire comprendre* par l'utilisateur, à l'image des individus dans une interaction interpersonnelle, un agent conversationnel animé doit être capable d'utiliser son corps pour accompagner son discours et exprimer des émotions. Dans cette section, nous présentons les différentes problématiques liées au développement d'ACAs capables d'une communication émotionnelle, gestuelle et faciale pour un enrichissement de l'interaction humain-machine.

1.1. La communication émotionnelle

De nombreux travaux de recherche montrent que l'expression d'émotions des ACAs permet d'améliorer l'interaction humain-machine (Russel et al., 2009). Cependant, pour garantir cet effet positif, les expressions des émotions doivent être appropriées aux événements et au contexte social de l'interaction. En effet, lorsque les expressions d'émotions de l'agent sont inappropriées, celui-ci est perçu plus négativement : moins crédible, moins agréable, etc. (Ochs, et al., 2010). C'est en effet l'incompréhension par l'utilisateur des expressions des émotions de l'agent qui détériore l'interaction. Pour une communication émotionnelle compréhensible, il est donc important que les émotions de l'agent soient à la fois correctement exprimées et exprimées au moment opportun durant l'interaction. Pour ce faire, un ACA doit tout d'abord être doté d'un riche répertoire d'expressions. Les personnages virtuels existant utilisent souvent des expressions stéréotypées, très intenses, limitées au visage, leur conférant un aspect très caricatural et limitant leur capacité de communication émotionnelle à quelques émotions typiques (joie, colère, surprise, tristesse, dégoût et peur). Or, une reconnaissance des émotions est une condition nécessaire à une bonne compréhension de l'agent. Comme le montrent (Niewiadomski et al., 2011), certaines émotions comme l'embarras ou la fierté sont difficilement perceptibles à travers uniquement l'expression faciale. Pour pallier à ces limites, et pour doter les ACAs de la capacité de communiquer des émotions plus subtiles tels que la soulagement ou l'embarras, un algorithme permettant la génération des séquences de signaux sur différentes modalités (gestes, visage, torse, etc.) pour l'expression d'émotions a été développé. Ce modèle a été construit à partir d'annotations manuelles de vidéos d'individus, dans lesquelles les relations entre les différentes modalités lors de l'expression d'émotion ont été extraites (Niewiadomski et al., 2011).

1.2. La communication gestuelle

De nombreux travaux ont montré l'importance du comportement non-verbal dans la communication. Plus de la moitié des informations dans une interaction interpersonnelle sont en effet exprimées à travers les comportements non-verbaux (Mehrabian, 1969). Dans une

interaction humain-machine, il est donc essentiel que les agents virtuels soient capables de produire des gestes expressifs communicatifs. Pour spécifier le comportement non-verbal d'un agent virtuel accompagnant son discours, des langages de représentation, tel que BML (Behavior Markup Language) ont été développés (Vilhjálmsson *et al.*, 2007). Cependant, les agents n'ont pas toujours les mêmes capacités d'expressivité. Par exemple, l'ACA *Greta* (Pelachaud, 2009) et le robot humanoïde *Nao* (Gouaillier *et al.*, 2009) n'ont pas les mêmes degrés de liberté dans les gestes. Il faut donc que le robot et l'agent virtuel communiquent des informations similaires mais pas forcément en utilisant des gestes identiques. Pour permettre la sélection et la planification des gestes quel que soit l'agent, la notion de familles gestuelles introduit par Geneviève Calbris (Calbris, 1983) peut être utilisée. Les gestes d'une famille transmettent un même sens, mais ils sont différents en forme. Par exemple, le robot n'a que deux configurations de la main, ouverte et fermée; il ne peut pas étendre un doigt seulement. Par exemple, pour faire un geste d'ictique, le robot peut pointer l'ensemble de son bras vers la cible plutôt que seulement son index comme cela est fait par l'agent virtuel.

De plus, le comportement non-verbal produit doit être synchronisé avec la parole. De même que pour le point précédent, le robot et l'agent virtuel ont des propriétés physiques très différentes. L'expressivité gestuelle, caractérisée par différentes dimensions, telles que l'ampleur des gestes, leur vitesse, leur puissance d'exécution, leur fluidité et leur rapidité, doit être adaptée aux agents. Le robot est en effet une entité physique avec une masse corporelle, des articulations physiques avec une limite de vitesse et de déplacement. Ce n'est pas le cas de l'agent virtuel. Un mécanisme de synchronisation pour le comportement du robot est donc développé en tenant compte de ses caractéristiques physiques (Anh Le *et al.*, 2011).

1.3. La communication faciale

Le comportement non-verbal du visage accompagnant la parole, comme, par exemple, un léger hochement de tête ou un rapide soulèvement des sourcils, permet d'augmenter la compréhension de la parole (Munhall *et al.*, 2004), et surtout de refléter l'émotion du locuteur. Même si des recherches ont confirmé que le comportement non-verbal est lié fortement à la parole synchronisé (Munhall *et al.*, 2004), la production des comportements réels pour transmettre des messages subtils est encore un défi de recherche en interface homme-machine. Une méthode qui a été explorée, consiste à utiliser un modèle d'apprentissage pour synthétiser le comportement non-verbal (surtout les expressions du visage et les mouvements de tête) par rapport aux indices acoustiques de la parole. Dans un premier temps, un corpus est recueilli. Des acteurs équipés par des capteurs dispersés sur tout le visage récitent des phrases avec différentes émotions. Les phrases sont phonétiquement balancées, telles que celles utilisées par (Bänziger and Scherer, 2011). Il est demandé aux acteurs de jouer les émotions correspondant à un scénario décrivant diverses situations. Un modèle d'apprentissage est alors utilisé pour estimer la relation entre les caractéristiques de la voix (telles que l'intonation et l'intensité) et les déplacements des capteurs sur le visage.

En dehors de la synchronisation entre la parole et l'animation du visage, les détails d'un visage jouent un rôle important pour transmettre une émotion (Courgeon *et al.*, 2009). Par exemple, dans le domaine de la bande dessinée, le contraste des couleurs et les ombres sont souvent employés pour définir l'ambiance d'une scène. Les ombres noires apportent en effet un aspect effrayant et mystérieux aux personnages et accentuent l'expression d'une émotion négative. Ces effets sont par exemple utilisés pour le personnage de *Batman*, connu aussi sous le nom de «chevalier noir». Dans le domaine des ACAs, un modèle permettant de considérer ces

effets d'ombres noires pour modifier la perception qu'a l'utilisateur de l'ambiance d'une scène a été développé (Huang *et al.*, 2011).

Pour augmenter le réalisme des animations, les détails des rides lors de l'expression d'une émotion sont à considérer. En effet, comme le soulignent (Melo *et al.*, 2009), dans le domaine des personnages virtuels, les rides causées par une déformation des couches de la peau suite à une contraction musculaire jouent un rôle important dans la reconnaissance des émotions par les utilisateurs, et donc dans la compréhension des émotions exprimées. Différentes techniques peuvent être utilisées pour modifier les effets des rides (Mikkelsen, 2010, Melo *et al.*, 2009). Une première méthode repose sur la déformation géométrique du visage. Elle donne de très bons résultats mais reste très coûteuse en termes de temps de calcul. Une deuxième méthode consiste à appliquer une texture sur le visage donnant l'illusion de rides en utilisant la lumière; une méthode qui ne modifie pas la géométrie du visage et reste donc peu coûteuse en terme de temps de calcul.

2. VERS UNE MACHINE CAPABLE DE COMPRENDRE

Dans une interaction humain-machine, les ACAs doivent être compris par l'utilisateur mais l'utilisateur doit aussi se sentir compris par l'agent ; une compréhension à la fois de son discours et de ses émotions.

Vers une compréhension des émotions

Une machine doit être capable de comprendre, non seulement le message explicitement exprimé par l'utilisateur, mais aussi celui plus implicite des émotions. De nombreuses recherches ont en effet montré que les agents virtuels capables d'exprimer des émotions d'empathie envers l'utilisateur permettent d'améliorer l'interaction humain-machine (Russel *et al.*, 2009). L'empathie se définit généralement comme «la capacité de se mettre mentalement à la place d'autrui afin de comprendre ce qu'il éprouve » (Pacherie, 2004). Être empathique suppose donc d'être capable d'évaluer la charge émotionnelle d'une situation pour un autre individu. Une machine doit, pour ce faire, connaître dans quelles circonstances quelles émotions peuvent être ressenties. Doter une machine d'empathie signifie donc lui apporter une représentation des conditions de déclenchement des émotions de l'utilisateur. A partir de ces informations, la machine peut en effet déduire d'une situation d'interaction les émotions potentiellement ressenties par l'utilisateur. Pour ce faire, une méthode consiste à analyser des situations réelles humain-machine ayant amené l'utilisateur à exprimer des émotions. Pour comprendre les causes de ces émotions, les théories de l'évaluation cognitive des émotions en psychologie cognitive peuvent être utilisées (Scherer, 1988 ; Ortony *et al.*, 1988). A partir de ces informations théoriques et empiriques, un modèle formel des émotions d'un agent virtuel a par exemple été construit (Ochs *et al.*, 2010). Les émotions y sont définies par leurs conditions de déclenchement représentées par des états mentaux particuliers, *i.e.* des combinaisons de croyances, d'incertitudes et d'intentions. Cette formalisation permet à un agent virtuel de déterminer dans une situation donnée l'émotion d'empathie qu'il devrait exprimer. Une évaluation auprès d'un panel de personnes a permis de mettre en évidence l'amélioration de la perception qu'a l'utilisateur de l'agent lorsque celui-ci exprime des émotions d'empathie à partir de ce modèle (Ochs *et al.*, 2010).

Vers une compréhension de l'orateur

La communication a réellement lieu lorsqu'une réponse est obtenue. L'interlocuteur ne reste pas immobile pendant une interaction en assimilant passivement le discours du locuteur; activement il prend part à la conversation en montrant, à travers des signaux verbaux et non verbaux, qu'il écoute, qu'il comprend, qu'il est d'accord et ainsi de suite. Ces informations permettent au locuteur de savoir si l'interaction est bien amenée jusqu'à la fin et si son message a été correctement transmis. Plusieurs recherches ont montré que les signaux émis par la personne qui écoute permettent de faire avancer l'interaction et que l'absence de ce type de signaux peut faire écrouler la conversation (Nespolous *et al.*, 1986). Nous nous référons aux signaux émis par l'interlocuteur avec le terme *backchannel* proposé par Yngve (Yngve, 1970). Un *backchannel* se définit comme un signal verbal ou non verbal émis par l'interlocuteur, sans interrompre le locuteur, avec le but de transmettre des informations sur ce qu'il pense du discours (Allwood *et al.*, 1993).

Étant donné l'importance du comportement de celui qui écoute durant une interaction, un système informatique pour la génération de ce type de comportement pour un agent conversationnel animé a été développé. La première étape consistait à identifier les différents signaux de backchannel non verbaux (Heylen *et al.*, 2007). A travers des tests perceptifs durant lesquels les utilisateurs observaient des signaux joués par un ACA, les associations entre un ensemble de signaux non verbaux et les significations qu'ils peuvent avoir quand ils sont montrés par un ACA qui écoute, ont été identifiés. Par exemple, le hochement de tête et le sourire utilisés simultanément reflètent la compréhension ; alors que le froncement des sourcils et l'inclinaison de tête reflètent plutôt l'incompréhension. A partir de ces résultats, un modèle computationnel de l'interlocuteur a été développé. Fondé sur la littérature (Maatman *et al.*, 2005; Ward *et al.*, 2000), un ensemble de règles stochastiques a été défini dans ce modèle pour déterminer quand un signal de backchannel devrait être émis. Le système analyse les comportements verbaux et non verbaux du locuteur pour identifier ceux qui pourraient déclencher une action par l'interlocuteur. Par exemple, un mouvement de la tête ou une variation dans le ton de la voix de l'utilisateur provoquent un backchannel avec certaine probabilité. Une évaluation du modèle a permis de montrer qu'un tel comportement donnait l'impression aux utilisateurs que l'ACA les écoutait. De plus, ils ont trouvé l'interaction avec l'ACA plus facile, plus agréable et on perçoit l'ACA plus positif, chaleureux, intéressé et sincère (Bevacqua *et al.*, 2010).

Vers l'émergence d'une compréhension mutuelle

Pour donner à l'humain une sensation d'interaction naturelle et de compréhension mutuelle, un agent doit être capable de se synchroniser avec lui aux moments opportuns. Depuis la définition de l'hétéro-synchronie (Condon et Ogston, 1966), la synchronie inter-personnes et son lien avec la qualité de l'interaction et l'inter-compréhension ont été étudiés tant dans des études comportementales (Murray and Trevarthen, 1985; Nadel and Tremblay-Leveau, 1999; Poggi and Pelachaud, 2000; Prepin and Gaussier, 2010) que dans des études de neuro-imagerie (Chammat *et al.*, 2010; Dubal *et al.*, 2010; Dumas *et al.*, 2011). Pour modéliser ce lien entre inter-compréhension et synchronisation entre partenaires, une approche très répandue dans le *design* des agents conversationnels consiste à considérer la *synchronisation* comme un *acte communicatif* (Poggi et Pelachaud, 2000) : l'ACA cherche à se synchroniser avec l'utilisateur lorsqu'il veut communiquer sa satisfaction, son engagement et sa bonne compréhension. Contrastant avec cette approche, la relation entre synchronie et inter-compréhension a récemment été modélisée comme intrinsèquement liée aux relations entre

dynamiques non-verbales des agents et sémantique de leur échange verbal (Prepin et Pelachaud, 2011) : la synchronie qui apparaît lors d'un dialogue y est modélisée comme un effet de bord de la compréhension mutuelle; la compréhension mutuelle entre les interlocuteurs permet leur couplage, et une conséquence de ce couplage est l'émergence de synchronie; la synchronie n'est alors pas un *acte communicatif* fait ou non par l'agent, c'est une variable dyadique qui apparaît nécessairement lorsque les interlocuteurs se comprennent mutuellement et se couplent; si les partenaires d'une dyade se comprennent, la synchronie émerge, et inversement, si ils ne se comprennent pas, la synchronie est rompue.

En plus de retrouver le fait que l'émergence de synchronie au sein d'une dyade dépend du degré de compréhension partagé, ce modèle et son étude mettent en avant deux propriétés essentielles du lien entre compréhension et synchronie : P1, tout comme la présence de synchronie signifie qu'il y a compréhension mutuelle, l'absence de synchronie montre qu'il y a incompréhension; P2, synchronisation et désynchronisation émergeant de la compréhension mutuelle sont deux phénomènes très rapides. Ces deux propriétés rendent le lien entre synchronie et inter-compréhension fondamental pour le *design* d'agents interactifs : la présence de synchronie donne une information forte sur la compréhension mutuelle (P1, il y a équivalence entre présence de synchronie et inter-compréhension); si les agents sont capables de détecter la synchronie, ils ont alors une information *temps-réel* sur le fait qu'ils se comprennent mutuellement ou non (P2, synchronisation et désynchronisation sont deux phénomènes très rapides).

CONCLUSION

En conclusion, l'un des enjeux dans le domaine de l'interaction humain-machine est de créer des agents virtuels et physiques compréhensibles par l'utilisateur, *i.e.* capable de transférer une information de manière intelligible. A ceci s'ajoute la problématique de créer une interaction engageante, agréable et satisfaisante pour l'utilisateur. Il est donc essentiel de concevoir des modèles permettant aux agents d'adopter un comportement non-verbal pour accompagner le discours et exprimer des émotions et des états cognitifs tels que la compréhension ou l'incompréhension. Pour définir *comment* se comporter «physiquement», une méthode consiste à observer le comportement des individus et définir des algorithmes permettant de faire jouer ces comportements par les agents. Les modèles permettant de déterminer *quand* adopter *quel* comportement peuvent être construits à partir d'analyse d'interaction humain-machine ou en se fondant sur des phénomènes tels que la synchronie.

Aujourd'hui, plusieurs outils pour la détection des états émotionnels et cognitifs des utilisateurs sont développés. Ces outils peuvent permettre d'affiner le comportement des agents en considérant les réactions des utilisateurs.

BIBLIOGRAPHIE

- Allwood, J., Nivre, J. and Ahlsén E. 1993. *On the Semantics and Pragmatics of Linguistic Feedback*, Journal of Semantics, 9 (1).
- André E., Klesen, P., Gebhard, P., Allen, S. et Rist, T. 2001. *Integrating models of personality and emotions into lifelike characters*. In Paiva, A., éditeur : *Affective interactions: Towards a new generation of computer interfaces*, pages 150–165.

- Ball, G. and Breese, J. 2001. *Emotion and personality in a conversational agent*. In *Embodied conversational agents*. MIT Press, Cambridge, MA, USA 189-219.
- Bänziger, T. and Scherer, K. R. 2010. *Introducing the Geneva Multimodal Emotion Portrayal (GEMEP) Corpus* Blueprint for affective computing: A sourcebook, pp. 271-294, Oxford, England: Oxford University Press.
- Beale, R. and Creed, C. 2009. *Affective interaction: How emotional agents affect users*. *Int. J. Hum.-Comput. Stud.* 67, 9, 755-776.
- Bevacqua, E., Hyniewska, S. and Pelachaud, C. 2010. *Positive influence of smile backchannels in ECAs*, Workshop on Interacting with ECAs as Virtual Characters (AAMAS).
- Calbris, G. 1983. *Contribution à une analyse sémiologique de la mimique faciale et gestuelle française dans ses rapports avec la communication verbale*. Ph.D. Thesis, Paris 3.
- Cassell, J. 2000. *More than Just Another Pretty Face: Embodied Conversational Interface Agents*. *Communications of the ACM*, 43(4): 70-78.
- Chammat, M., Foucher, A., Nadel, J., and Dubal, S. (2010). Reading sadness beyond human faces. *Brain Research*, In Press, Accepted Manuscript.
- Condon, W. S. and Ogston, W. D. (1966). Sound film analysis of normal and pathological behavior patterns. *Journal of Nervous and Mental Disease*, 143:338– 347.
- Courgeon, S., Buisine, S. and Martin, J-C. 2009. *Impact of Expressive Wrinkles on Perception of a Virtual Character's Facial Expressions of Emotions*, Proceedings of the International Conference on Intelligent Virtual Agents.
- Dubal, S., Jouvent, A. F. R., and Nadel, J. (2010). Human brain spots emotion in non humanoid robots. *Social Cognitive and Affective Neuroscience*, in press:–.
- Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., and Garnero, L. (2010). Inter-brain synchronization during social interaction. *PLoS One*, 5(8):e12166.
- Gouaillier, D., Hugel, V., Blazevic, P., Kilner, C., Monceaux, J., Lafourcade, P., Marnier, B., Serre, J. and Maisonnier, B. 2009. *Mechatronic design of NAO humanoid*. Robotics and Automation.
- Heylen, D., Bevacqua, E., Tellier, M. and Pelachaud, C. 2007. *Searching for Prototypical Facial Feedback Signals*, International Conference on Intelligent Virtual Agents (IVA), Paris, France, September 17th - 19th.
- Huang, J., Boubekeur, T., Ritschel, T., Hollander, M. and Eisemann, E. 2011. *Separable Approximation of Ambient Occlusion*. Eurographics.
- Kendon, A. 1994. *Do gestures communicate? A review*. Research on language and social interaction.
- Lê Q.A. and Pelachaud, C. 2011. *Generating co-speech gestures for the humanoid robot NAO through BML*. International Gesture Workshop: Gesture in Embodied Communication and Human-Computer Interaction.
- Maatman, R. M., Gratch, J. and Marsella, S. 2005. *Natural behavior of a listening agent*. In Proceedings of the 5th International Conference on Interactive Virtual Agents. Kos, Greece.
- Mehrabian, A. 1969. *Significance of Posture and Position in the Communication of Attitude and Status Relationships*, Psychological Bulletin, 71, 5.
- Melo, C.M. and Gratch, J. 2009. *Expression of emotions using wrinkles, blushing, sweating and tears*. International Conference on Intelligent Virtual Agents (IVA), pp. 188–200.
- Mikkelsen, M.S. 2010. *Bump mapping unparametrized surfaces on the gpu*. Journal of graphics, gpu, and game tools 15, 1, 49–61.

- Munhall, K.G., Jones, J. A., Callan, D.E., Kuratate, T. and Bateson, E.V. 2004. *Visual prosody and speech intelligibility: Head movement improves auditory speech perception*. *Psychological Science*, 15(2):133–137.
- Murray, L. and Trevarthen, C. (1985). Emotional regulation of interactions between two-month-olds and their mothers. *Social perception in infants*, pages 101–125.
- Nadel, J. and Tremblay-Leveau, H. (1999). *Early social cognition*, chapter Early perception of social contingencies and interpersonal intentionality: dyadic and triadic paradigms, pages 189–212. Lawrence Erlbaum Associates.
- Nespolous, J.-L. and Lecours, A. R. 1986. *Gestures: Nature and Function*. The Biological Foundations of Gestures: Motor and Semiotic Aspects Hillsdale, NJ: Lawrence Earlbaum Associates, p. 49-62.
- Niewiadomski, R., Hyniewska, S. and Pelachaud, C. 2011. *Constraint-Based Model for Synthesis of Multimodal Sequential Expressions of Emotions*. *Transactions on Affective Computing*.
- Ochs, M., Sadek, D. and Pelachaud, C. 2010. *A Formal Model of Emotions for an Empathic Rational Dialog Agent*, *Journal of Autonomous Agent and Multi-Agent Systems (JAAMAS)*.
- Pacherie, E. 2004. *L'empathie, L'empathie et ses degrés*, pages 149–181. Odile Jacob.
- Pelachaud, C. 2009. Modelling Multimodal Expression of Emotion in a Virtual Agent, *Philosophical Transactions of Royal Society Biological Science*, 3539-3548.
- Picard, R. 1997. *Affective Computing*. MIT Press.
- Poggi, I. and Pelachaud, C. (2000). Emotional meaning and expression in animated faces. *Lecture Notes in Computer Science*, pages 182–195.
- Prepin, K. and Gaussier, P. (2010). How an agent can detect and use synchrony parameter of its own interaction with a human? In et al., A. E., editor, COST Action2102, Int. Traing School 2009, *Active Listening and Synchrony*. LNCS 5967, pages 50–65. Springer-Verlag, Berlin Heidelberg.
- Prepin K. and Pelachaud C. . Shared understanding and synchrony emergence: Synchrony as an indice of the exchange of meaning between dialog partners. In J. Filipe, editor, *Third International Conference on Agents and Artificial Intelligence, ICAART2011*, pages 1–10. Springer, 2011.
- Reeves, B. and Nass, C. 1996. *The media equation: how people treat computers, television, and new media like real people and places*. Cambridge University Press.
- Vilhj ármsson, H., Cantelmo, N., Cassell, J., Chafai, N.E., Kipp, M., Kopp, S., Mancini, M., Marsella, S., Marshall, A., Pelachaud, C., Ruttkay, Z., Th órisson, K., van Welbergen H. and van der Werf, R. 2007. *The Behavior Markup Language: Recent Developments and Challenges*, *International Conference on Intelligent Virtual Agents*, p. 99-111.
- Yngve, V. 1970. *On getting a word in edgewise*. *Papers from the Sixth Regional Meeting of the Chicago Linguistic Society*, p. 567-577.
- Ward, N. and Tsukahara, W. 2000. Prosodic features which cue back-channel responses in english and japanese. *Journal of Pragmatics*, 23:1177-1207.