Étude de l'impact du couplage geste et parole sur un robot

Céline Jost, Brigitte Le Pévédic, Dominique Duhaut

Université de Bretagne Sud
Lab-STICC
dominique.duhaut@univ-ubs.fr

On ne peut pas ne pas communiquer (Watzlawick, 1967). Si cet acte paraît anodin, sa compréhension est plus complexe. Dans un contexte d'interaction homme-robot, la communication entre les deux interlocuteurs révèle une inconnue : la communication du robot qui est une boîte vide *a priori*. Concernant la communication entre les humains, Grandgeorge et al (2011) indique que « si l'identité proposée par l'individu A (*e.g.* statut, but, …dans l'interaction) est conforme à l'identité représentée de A, et réciproquement pour ces deux identités de l'individu B, nous pouvons alors parler d'intercom-préhension. » Mais est-ce qu'un humain attend vraiment d'un robot une communication similaire à la sienne ? La réponse à cette question est un pas fondamental vers l'intercompréhension puisqu'elle ne peut exister que si l'être humain perçoit correctement les signaux du robot.

Pour comprendre comment le robot est perçu par l'homme, deux études ont été réalisées. Ces deux expérimentations avaient pour objectif de déterminer la place du geste dans la communication du robot d'une part, et l'acceptabilité du robot

en tant que partenaire social d'autre part. Les deux études proposaient aux participants d'observer un robot. Dans la première, celui-ci prononçait plusieurs fois la même phrase en effectuant des mouvements différents. Il a été montré **que les humains trouvaient le robot crédible et sincère lorsque la gestuelle était cohérente**. Dans la deuxième, le robot et un expérimentateur dialoguaient ensembles. Quelques fois le robot effectuait une gestuelle non adaptée. Il a été montré que **le robot est apprécié et son comportement accepté** même si, de temps en temps, sa gestuelle n'est pas cohérente.

La communication homme-robot aujourd'hui

La littérature montre plusieurs types de robots : ceux qui ressemblent aux humains, aux animaux et ceux qui représentent des créatures inventées (Shibata, 2004). La première catégorie est intéressante car l'anthropomorphisme des machines augmente le jugement positif des humains et la qualité de l'interaction (Kim, 2011). De plus, lorsque le robot imite le comportement humain, l'interaction est fluide et naturelle (Kirby, 2010). Ainsi, la communication du robot est verbale et non verbale. C'est une erreur de privilégier l'une ou l'autre car aucun canal n'est prédominant sur un autre dans la communication homme-homme (Ekman, 1980) et il est plus facile de décoder les expressions humaines lorsqu'on analyse plusieurs canaux (Castellano, 2008). La plupart des études s'inscrivent dans le domaine des émotions (Scherer, 1984 ; James, 1884 ; Darwin, 1872). Le processus émotionnel n'est pas bien connu et révèle des inconnues alors la majorité des chercheurs travaille sur les six émotions de base qui sont la joie, la peur, la tristesse, la surprise, le dégoût et la colère (Ekman, 2003). La communication verbale du robot fait appel à la synthèse vocale. Elle ne permet pas actuellement

d'exprimer des émotions car il est difficile de trouver le ton, la vitesse et la prosodie associés (Breazeal, 2001). La communication non verbale est importance pour préciser le discours ou le rendre plus compréhensible (Tojo, 2000). Des expressions sur les robots humanoïdes (Breazeal, 2003) aux expressions sur les ours en peluche (Saint-Aimé, 2009), la recherche sur la communication homme-robot est très active comme l'explique en détail Breazeal dans son ouvrage (Breazeal, 2005). Mais à quel point le non verbal du robot affecte-t-il l'être humain qui interagit ?

Étude 1 : Importance de l'adéquation entre le discours et les gestes

Cette expérimentation utilisait le même protocole expérimental que celui de Rivière *et al.* (2011) dans leur étude sur les agents virtuels. Le robot effectuait trois séries de gestes différents pour une même phrase. Il était mis en présence d'un observateur qui donnait ensuite son ressenti.

Participants

Soixante personnes (17 femmes et 43 hommes) âgées de 10 à 57 ans (moyenne de 18,1 ans) ont participé à cette évaluation.

Méthode

L'expérimentation s'est déroulée pendant un concours de robotique, Robofesta, dans un stand isolé. Deux types de participants ont pu être mis en évidence, ils étaient soit inscrits au concours, soit spectateurs du concours. Ceci nous a laissé penser qu'ils avaient une bonne connaissance de la robotique ou tout au moins un intérêt pour l'objet de l'expérimentation.

Individuellement, le participant devait observer le robot et donner son impression en remplissant un questionnaire.

Équipement

Le robot (Figure 1) était positionné sur une table, face au participant. L'expérimentateur était assis sur le côté, face à l'ordinateur de contrôle.

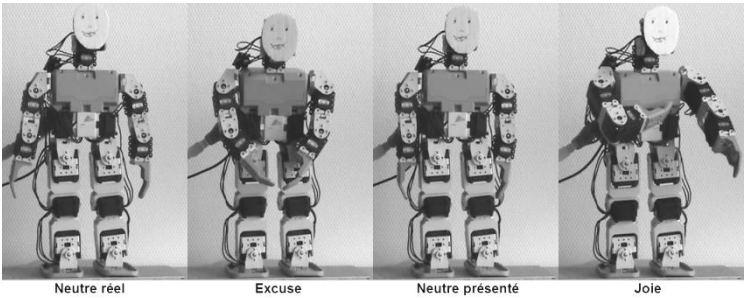


Figure 1: Les positions clefs du robot

Déroulement de l'expérimentation

Tout d'abord, l'expérimentateur demandait au participant d'imaginer le contexte suivant : le robot leur avait conseillé d'aller voir un film. Au retour du cinéma, deux scénarios se mettaient en place. Scénario A : il n'avait pas aimé le film et faisait des reproches au robot. Scénario B : il avait aimé le film et remerciait le robot. Pour chacun des deux scénarios, le participant devait observer trois réactions et indiquer si le robot était crédible et sincère dans sa réponse. Pour le scénario A, le robot disait «Je suis vraiment désolé. Ce film avait de bonnes critiques. Je pensais qu'il te plairait. ». Pour le scénario B, le robot disait «Alors tu as aimé le film. Je suis content de t'avoir si bien conseillé. ». Les mouvements proposés (Figure 1) pour les deux scénarios étaient identiques. Le mouvement 1 (*Excuse*) représentait le comportement que nous proposions pour symboliser l'excuse. Le mouvement 2 (*Neutre présenté*) était une expression neutre très légèrement négative (tête qui se baisse légèrement). Le mouvement 3 (*Joie*) représentait le

comportement que nous proposions pour symboliser une joie intense. L'expression *Neutre réel* montre la position de base du robot, considéré comme sa position neutre. La situation congruente pour le scénario A était le mouvement 1 et pour le scénario B le mouvement 3.

Collecte de données et analyses

Après chaque mouvement, les observateurs devaient répondre à un questionnaire en répondant aux deux questions suivantes : Le robot semble-t-il exprimer ce qu'il pense ? (Sincérité) Son expression vous paraît-elle plausible ? (Crédibilité) Les réponses disponibles étaient : « pas du tout », « plutôt pas », « plutôt » ou « tout à fait ».

Les analyses statistiques ont été réalisées avec le logiciel Minitab 15©. Le test du Chi-deux a permis de vérifier les réponses significatives. Le seuil de significativité p a été fixé à 0,05.

Résultats

Scénario A : le participant n'a pas aimé le film. Dans ce scénario, les participants n'ont pas répondu au hasard (X^2=189,1, dl=33, p<0.001). Concernant le mouvement 1 (*Excuse*), les participants ont très significativement répondu qu'il était sincère (60% plutôt et 36,7% tout à fait, X^2=132.7, dl=3, p<0.001; Figure 2) et crédible (48,3% plutôt et 35% tout à fait, X^2=63.0, dl=3, p<0.001). Concernant le mouvement 2 (*Neutre présenté*), près de la moitié des participants l'ont trouvé plutôt sincère (40%, X^2=20.1, dl=3, p<0.001; Figure 2). Néanmoins, il faut noter des réponses contrastées puisque si 18.3% l'ont estimé pas du tout sincère, 15% l'ont tout à fait estimé sincère. A l'inverse, 43% des participants ne l'ont pas

jugé crédible (X^2=39.7, dl=3, p<0.001). Dans la dernière condition, où le robot exprime le mouvement 3 (*Joie*), aucune différence statistique n'est notée sur la sincérité du robot (X^2=5.9, dl=3, p=0.117; Figure 2). Il est à noter une différence selon le genre : les femmes ont significativement plus fréquemment estimé le robot "pas du tout sincère" par rapport aux hommes (41,2% des femmes et 23,3% des hommes; X^2=10.5, dl=3, p=0.015). Concernant la crédibilité, les participants donnent des réponses significativement contrastées (X^2=12.4, dl=3, p=0.006). En effet, 30% l'ont jugés pas du tout crédibles et 31,67% tout à fait crédibles. Les situations intermédiaires sont moins représentées (25% et 13,3%).

Scénario B : le participant a aimé le film. Dans ce scénario, les participants n'ont pas répondu au hasard (X^2=88.4, dl=15, p<0.001). Concernant le mouvement 1, 38,33% participants ont répondu que le robot était plutôt pas sincère et 33,33% plutôt sincère (X^2=26, dl=3, p<0.001; Figure 2). La réponse à la question sur la crédibilité est plus marquée : seuls 3,3% l'ont trouvé tout à fait crédible alors que la majorité ne l'ont plutôt pas ou pas du tout trouvé crédible (38,3% et 38,3% respectivement, X^2=46.7, dl=3, p<0.001; Figure 2). Concernant le mouvement 2, 38.3% participants ont répondu qu'il était plutôt pas sincère et 33.3% plutôt sincère (X^2=29.9, dl=3, p<0.001; Figure 2). La réponse à la question sur la crédibilité est là aussi plus marquée : seuls 8,3% l'ont trouvé tout à fait crédible alors que la majorité ne l'ont plutôt pas ou pas du tout crédible (43,3% et 33,3% respectivement, X^2=42.7, dl=3, p<0.001; Figure 2). Lorsque le robot a exprimé le mouvement 3, les participants ont très significativement répondu qu'il était sincère (38,3% plutôt et 41,7% tout à fait, X^2=56.5, dl=3,

p<0.001; Figure 2) et crédible (36,7% plutôt et 35% tout à fait, X^2=26.2, dl=3, p<0.001; Figure 2).

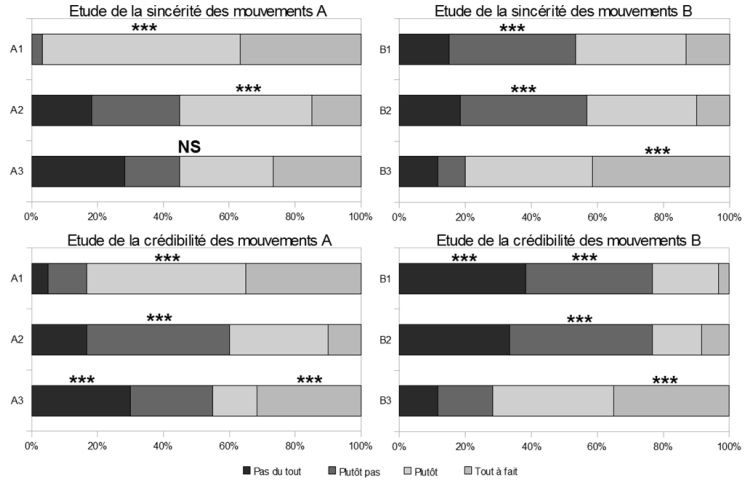


Figure 2: Résultats sur la crédibilité et la sincérité en fonction des scénarios A et B et des mouvements du robot.
*Niveau de significativité : * p<0,05, ** p<0,01, *** p<0,001*
NS : non significatif (Test du Chi-deux).

Discussion

Concernant le scénario A, le robot s'excusait d'avoir mal conseillé le participant. Le robot a été clairement jugé crédible et sincère lorsqu'il a effectué le mouvement d'excuse. Cette situation est congruente et comprise par un panel large de personne. Pour les deux autres mouvements, les réponses sont contrastées. Les personnes ne semblent pas réussir à se positionner. On peut imaginer que les réponses mitigées reflètent un niveau d'hésitation due à une incompréhension de la situation. Par exemple, pour la crédibilité du mouvement 3, le choix semble être fait au hasard (répartition équivalente des

réponses dans les 4 choix). Ainsi, les mouvements 2 et 3 semblent difficiles à évaluer par les participants. L'intention du robot semble peu claire.

Concernant le scénario B, le robot se « félicitait d'avoir si bien conseillé le film ». Les participants ont jugé le mouvement 3 crédible et sincère. C'est la situation congruente qui a donc été choisie par les participants. Les deux autres mouvements ont été jugés fortement non crédibles alors que la sincérité est plus mitigée. Il semble donc que la sincérité soit plus difficile à évaluer que la crédibilité dans des contextes plus mitigés.

De façon générale, les participants ont plus facilement pensé que le robot semblait exprimer ce qu'il pensait même si son expression ne leur paraissait pas plausible. Dans les deux scénarios, les situations congruentes ont été choisies sans équivoque. Concernant les mouvements non congruents, il ne semble pas y avoir d'hésitation dans le cas d'une émotion positive. Au contraire, dans le cas d'une émotion négative, le robot est tout de même jugé plutôt sincère, même si sa crédibilité est diminuée et même si sa gestuelle est positive. De plus, le mouvement 2, qui était un neutre négatif n'a pas été retenu. Pourtant, le robot baissait la tête et effectuait un léger mouvement des mains contre ses jambes. Le mouvement 3 était exagéré. Ainsi puisque c'est le mouvement exagéré qui a été choisi, et puisque la sincérité est la même pour un mouvement neutre ou pour un mouvement positif, il semble important d'accentuer la gestuelle, dans le cas d'une émotion négative, pour rendre le robot d'avantage crédible et sincère. Rien ne peut être conclu sur la gestuelle d'une émotion positive.

En comparant les résultats de la première expérimentation avec ceux de Rivière et al (2011) et leur agent virtuel, on

remarque que les résultats du robot sont au moins aussi bons que ceux de l'agent virtuel (Tableau 1). En moyenne, le robot semble plus crédible et sincère qu'un agent virtuel. Cela amène à la question suivante : est-ce qu'un robot est un meilleur partenaire social qu'un agent virtuel ? Pour répondre à cette question, une étude future sera réalisée pour comparer l'impact d'un robot et d'un agent virtuel sur l'interaction homme-machine.

		Agent virtuel (GRETA)	ROBOT
S'EXCUSE	**SINCERITE**	65,00%	96,67%
		65,00%	96,67%
	CREDIBILITE	70,00%	83,33%
		70,00%	83,33%
SE REJOUIT	**SINCERITE**	74,00%	80,00%
		74,00%	80,00%
	CREDIBILITE	78,00%	71,67%
		78,00%	71,67%

Tableau 1: Comparaison de la sincérité et de la crédibilité d'un agent virtuel et du robot

Étude 2 : Acceptabilité d'un robot malgré quelques incohérences dans la communication.

Cette étude a permis de déterminer le degré d'acceptabilité d'un robot lors d'une interaction avec un humain. Nous avons utilisé le protocole d'évaluation de Rivière et al (2011). Chaque

participant devait observer un homme (l'expérimentateur) et un robot dialoguant ensemble, dans 4 contextes différents. Nous avons placé de façon aléatoire une gestuelle non adaptée du robot. Notre objectif était d'étudier si ces incohérences avaient un impact sur la perception de l'humain.

Participants

Vingt-six personnes (9 femmes et 17 hommes) âgées de 18 à 28 ans (moyenne de 22,5 ans) ont participé à cette évaluation. Les participants n'avaient pas de connaissance spécifique en robotique. Les participants étaient des étudiants d'informatique, de mathématiques et de statistiques et étaient de niveau bac+2 et bac+4. Le choix de ces différences entre les cursus des étudiants avait pour objectif d'éviter d'obtenir des réponses biaisées par un environnement commun.

Méthode

Chaque participant était isolé dans une pièce en présence de l'expérimentateur et du robot. Pendant toute la phase d'interaction, le participant devait observer et remplir un questionnaire au fur et à mesure.

Équipement

Le robot (Figure 1) était positionné sur une table, face à l'expérimentateur. Le participant était assis à côté.

Déroulement de l'expérimentation

Tout d'abord, l'expérimentateur présentait le robot et le contexte de l'interaction au participant. Celui-ci avait 5 minutes pour prendre connaissance du questionnaire. Une fois que le questionnaire était lu, l'expérimentateur démarrait l'expérimentation. Les 4 dialogues s'enchaînaient. Le participant devait noter ses réponses pendant l'interaction.

Cela permettait de ne pas sortir du contexte de l'interaction en la stoppant.

Les questions portaient sur chaque phrase prononcée par le robot, puis sur chaque dialogue dans son ensemble puis sur le ressenti de l'interaction en entière.

Pour chacune des phrases, le participant devait indiquer s'il était d'accord avec les affirmations suivantes :

- Le geste s'accorde avec le contenu de la phrase.
- La vitesse du geste est correcte.
- Le geste semble souple et naturel.
- Le robot est crédible quand il parle.

Pour chaque dialogue dans son ensemble, le participant devait indiquer s'il était d'accord avec les affirmations suivantes :

- L'enchaînement des répliques est correct.
- La vitesse de réponse du robot et de l'homme est correcte.
- Les expressions du robot quand il écoute l'homme sont correctes.
- Le robot est attentif.

À la fin de l'interaction, le participant devait donner son point de vue sur l'affirmation suivante : le robot est agréable.

Les réponses possibles étaient : Pas du tout, Un peu, Beaucoup, Sans idée ; sauf pour la deuxième affirmation où l'échelle de Likert 5 points a été utilisée : Trop lent, Lent, Normal, Rapide, Trop rapide.

Les analyses statistiques ont été réalisées avec le logiciel Minitab 15©. Le test du Chi-deux a permis de vérifier les

réponses significatives. Le seuil de significativité p a été fixé à 0,05.

Résultats

Significativement, le geste est considéré comme peu souple et naturel (50.51% un peu, X^2=318.29, dl=3, p<0,001), alors qu'il s'accorde beaucoup avec le contenu de la phrase (62.31% beaucoup, X^2=365.03, dl=3, p<0,001 ; Figure 3). De plus, la vitesse du geste est normale (66.67%, X^2=623.32, dl=4, p<0,001 ; Figure 3). L'enchaînement des répliques est correcte (67.31% beaucoup, X^2=69.57, dl=2, p<0,001 ; Figure 3).

Seulement 4.81% des participants estiment que la vitesse de réponse du robot et de l'homme n'est pas du tout correcte (X^2=50.21, dl=2, p<0,001 ; Figure 3). Et seulement 6.73% des participants estiment que les expressions du robot quand il écoute l'homme ne sont pas correctes (X^2=31.27, dl=2, p<0,001 ; Figure 3).

Concernant le robot lui-même, il est jugé crédible (60% beaucoup, X^2=353.26, dl=3, p<0,001 ; Figure 3) et attentif (67.31% beaucoup, X^2=64.18, dl=2, p<0,001 ; Figure 3). Pour finir, le robot n'a jamais été considéré comme désagréable (X^2=13.04, dl=2, p=0,001 ; Figure 3).

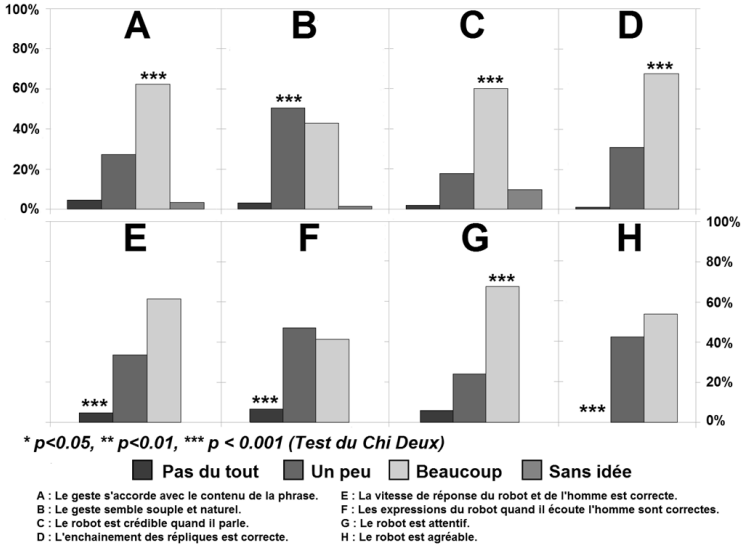


Figure 3: Résultat sur l'acceptabilité du robot au cours d'une interaction moyenne

*Niveau de significativité : * p<0,05, ** p<0,01, *** p<0,001*

NS : non significatif (Test du Chi-deux).

Conclusion

Pris indépendamment, les gestes doivent être en adéquation avec le discours pour que le robot semble sincère et crédible. Mais, dans le cas d'une interaction de longue durée, quelques erreurs n'altèrent pas la crédibilité et la sincérité du robot.

Références

Blanchet Philippe, *Linguistique de terrain, méthode et théorie (une approche ethnosociolinguistique)*. Presses Universitaires de Rennes : Rennes, 2000

Breazeal Cynthia, Emotive Qualities in Robot Speech, Intelligent Robots and Systems. *International Conference on IEEE/RSJ*, vol.3, 2001, 1388-1394

Breazeal Cynthia, Emotion and sociable humanoid robots, *International Journal of Human-Computer Studies*, Elsevier, vol 59, n°1, 2003, 119-155

Breazeal Cynthia, *Robot Emotion : A Functional Perspective, Who needs emotions*. Oxford University Press : New York, 2005, 271-310

Castellano Ginevra, Kessous Loic et Caridakis George, Emotion Recognition through Multiple Modalities : Face, Body Gesture, Speech. *Affect and emotion in human-computer interaction*, 2008, 92-103

Darwin Charles, *The expression of the emotions in man and animals*, Oxford University Press, 1872, réédition 2002

Ekman Paul, Friesen Wallace V, O'Sullivan Maureen et Scherer Klaus, Relative Importance of Face, Body, and Speech in Judgments of Personality and Affect, *Journal of Personality and Social Psychology*, vol.38, 1980, 270-277

Ekman Paul et Friesen Wallace V, *Unmasking the face : a guide to recognizing emotions from facial clues*. Ishk, 2003

Gong Li, How social is social response to computers? The function of the degree of anthropomorphism in computer representations, *Computer in Human Behavior*, vol.24, n°4, 2008, 1494-1509

Grandgeorge Marine et MIAC, L'intercompréhension dans les interactions : Modélisation interdisciplinaire autour de l'homme l'animal et la machine, *Congrès III2011,* 2011

James Williams, *What is an emotion? Mind*, vol 9, n°34, 1884, 188-205

Kim Youjeong et Shyam Sundar S., Anthropomorphism of computers: is it mindful or mindless? *Computers in Human Behavior*, vol.28, n°1, 2011, 241–250.

Kirby Rachel., Forlizzi Jodi et Simmons Reid, Affective social robots, *Robotics and Autonomous Systems*, vol 58, n°3, 2010, 322-332

Riviere Jeremy, Adam Carole, Pesty Sylvie, Pelachaud Catherine, Guiraud Nadine, Longin Dominique et Lorini Emiliano, Expressive Multimodal Conversational Acts for SAIBA Agents, *Intelligent Virtual Agents*, 2011, 316-323

Saint-Aimé Sébastien, Le Pevedic Brigitte, et Duhaut, Dominique, iGrace – Emotional Computational Model for EmI Companion Robot, *InTech Education and Publishing*, vol.4, 2009, 51–76

Scherer, Klaus R., Les émotions : fonctions et composantes, *Cahiers de psychologie cognitive*, vol 4, n°1, 1984, 9-39

Shibata Takanori, An overview of human interactive robots for psychological enrichment, *Proceedings IEEE*, vol. 92, n° 11, 2004, 1749–1758

Tojo Tsuyoshi, Matsusaka Yosuke, Ishii Tomotada et Kobayashi Tetsunori. A conversational Robot Utilizing Facial and Body Expressions, *Systems, Man, and Cybernetics, IEEE International*, vol.2, 2000, 858-863

Watzlawick P., Une logique de la communication. Editions Seuil, 1967