

Impact de la médiatisation technologique sur la mobilisation des modalités non-verbales dans les références démonstratives.

Defays¹ Aurore, Jeunejean² Aurélie, Giboin³ Alain

¹ *LECIT, Université de Liège, Liège, Belgique*

² *LUCID, Université de Liège, Liège, Belgique*

³ *Wimmics, Inria, I3S, Université de Nice, Sophia Antipolis, France*

L'objectif de ce papier est de comprendre comment la médiatisation technologique et les contraintes d'utilisation d'un outil vont modifier la manière dont les agents vont mobiliser les modalités de communications pour collaborer efficacement. Dans ce travail, nous nous sommes centrés sur les références démonstratives, c'est-à-dire les références verbales qui nécessitent d'être accompagnées de gestes pour que leur interprétation soit complète. Nous avons mené cette recherche dans le domaine de la conception architecturale et avons examiné deux situations de travail collaboratif (une en coprésence et une outillée) dans le but de pouvoir les comparer. Pour chaque situation, nous décrivons comment les architectes s'adaptent aux modalités potentiellement disponibles pour désigner l'objet à quoi ils font référence dans leur discours. Nos résultats montrent que lorsque la modalité gestuelle ne peut être utilisée de manière effective, comme c'est le cas dans la situation outillée, les architectes vont développer des *stratégies de compensation*. En effet, la modalité graphique va pallier la transmission partielle des gestes en prenant le relais de la fonction de désignation.

Introduction

Dans leur globalité, nos recherches s'intéressent à la manière dont les personnes arrivent à se comprendre et à communiquer efficacement. Avant toute communication ou toute collaboration, les personnes partagent une

certaine quantité de connaissances communes. Cet ensemble de connaissances communes est appelé *common ground* ou terrain commun, et désigne l'ensemble des connaissances, croyances ou suppositions pertinentes qui peuvent servir de base commune pour communiquer et agir efficacement (Clark & Schaefer, 1989).

Nos recherches concernent l'analyse du processus de compréhension mutuelle dans des situations de travail collaboratif. En particulier, nous nous intéressons à la façon dont les modalités de communication non-verbales (visuelle, gestuelle et graphique) vont supporter la modalité verbale dans le processus de compréhension mutuelle (appelé processus de *grounding* par Clark & Schaefer, 1989). Il est aujourd'hui bien connu que la plupart des messages ou énoncés échangés entre deux personnes sont de nature multimodale (Stein & Meredith, 1993). Les avantages de cette multimodalité sont également bien connus (de Gelder & Bertelson, 2003 ; Dumas et al., 2009). En effet, la multimodalité garantit certaines substitutions ou compensations sensorielles qui permettent de maintenir une signification à ce qui est perçu, même dans des situations non optimales ou dégradées, notamment grâce à la redondance et/ou la complémentarité des informations (Campanella et al., 2007 ; Stein & Meredith, 1997).

Plusieurs études expérimentales ont montré l'avantage de la multimodalité, mais souvent dans des tâches simples et contrôlées, plutôt artificielles et de type laboratoire (Darses et al., 2001 ; Dumazeau, 2005 ; Tajariol, 2006). Rares sont les études qui étudient le rôle des modalités de communication sur les mécanismes collaboratifs dans des situations écologiques, c'est-à-dire impliquant des activités réelles de travail. Rares sont également les études qui étudient le rôle de la multimodalité dans le processus de *grounding* (Chellali, 2009 ; Dumazeau, 2005 ; Nomura & Hutchins, 2007).

Nous savons que le processus d'élaboration et de maintien du *common ground* est intimement lié au processus de *référenciation* (Clark et al., 1983), c'est-à-dire au fait de faire référence à des objets (présents ou non dans l'environnement). Lorsque le locuteur produit un énoncé en rapport avec un référent donné, l'auditeur doit correctement identifier ce à quoi le locuteur fait référence pour qu'ils se comprennent correctement. Le processus de compréhension mutuelle va donc passer par la bonne identification et la bonne compréhension des référents utilisés par l'un et l'autre dans la discussion. Cette identification et cette compréhension des référents sont d'autant plus rapides et efficaces que le *common ground* partagé entre les partenaires est grand (Clark et al., 1983).

Dans ce présent papier, nous nous centrons sur un type en particulier de références : les *références démonstratives*. Il s'agit de références verbales qui nécessitent d'être accompagnées de gestes pour que leur compréhension et interprétation soient complètes et correctes (Clark et al., 1983).

Elles sont donc particulièrement intéressantes à étudier puisqu'elles sont multimodales par essence. Les gestes qui accompagnent ce type de références sont appelés *déictiques* (Falzon, 1994).

Nous avons choisi comme terrain d'étude l'architecture puisque les références démonstratives verbales et les gestes déictiques sont décrits comme particulièrement fondamentaux dans le cadre de la conception architecturale. En effet, les gestes déictiques permettent la co-référenciation aux éléments du plan ou aux objets architecturaux (Falzon, 1994).

Objectifs et hypothèses

Dans un premier temps, nous avons voulu connaître comment les architectes utilisaient les références démonstratives lorsqu'ils travaillent sur un projet réel, dans leur environnement naturel, avec leurs outils habituels, etc. Comment les architectes mobilisent-ils habituellement les multiples modalités de communication qui sont à leur disposition pour montrer à leur interlocuteur ce à quoi ils font référence verbalement? Comment les modalités non-verbales sont-elles mobilisées pour venir soutenir le processus de monstration ?

Dans un second temps, nous avons observé les mêmes architectes dans une autre situation de travail collaboratif, en leur permettant de réaliser leur projet via un outil technologique multimodal, le Studio Digital Collaboratif. Dans cette situation, les architectes placés à distance ont également travaillé sur un projet réel, mais l'ensemble de leurs échanges multimodaux était retransmis via une webcam (pour les modalités auditives, visuelles et gestuelles) et via une table partagée (pour la modalité graphique).

L'intérêt majeur de cette étude est qu'elle permet de comparer les stratégies mises en place par les architectes pour désigner des objets de l'environnement dans les deux situations envisagées. Cette situation d'étude est particulièrement intéressante puisqu'elle permet d'étudier le rôle de la modalité graphique dans le processus de monstration, jusqu'alors très peu envisagée.

Les gestes *déictiques* (ou de pointage) n'étant que partiellement retransmis par la webcam, nous postulons qu'en situation de travail distant, les architectes vont éprouver des difficultés à faire référence aux objets dans l'environnement numérique au moyen de leurs gestes de pointage habituels. Nos observations préalables suggéraient des stratégies de compensation de l'inefficacité de la modalité gestuelle par l'usage de la modalité graphique. Cette micro-analyse des deux situations a pour objectif de confirmer cette hypothèse.

Contexte de l'étude

Cette étude a été réalisée dans le domaine de la conception collective d'un projet architectural. Pour répondre à nos hypothèses, nous avons filmé et comparé des architectes travaillant dans deux situations différentes :

- une situation de **travail en coprésence** où les architectes collaborent en un même lieu, face à face, et où ils disposent des outils traditionnels : plans, calques, crayons, photos, etc. Nous avons observé et suivi tout le processus de conception du projet architectural, à savoir les dix réunions de travail organisées pour permettre aux architectes de travailler ensemble sur le projet. Sur les dix réunions observées, trois réunions ont été sélectionnées : une au début du processus de travail, une au milieu, et une à la fin du travail de conception. Au sein des trois réunions, 10 minutes ont été analysées en détails. Notre choix dans la sélection des extraits n'a pas été fait au hasard. Nous avons choisi pour chacune des trois réunions sélectionnées un extrait de 10 minutes, situé à la moitié de leur durée respective afin de garantir que les architectes soient bien engagés dans leur tâche de conception. L'ensemble des trois réunions représente donc un total de 30 minutes. La tâche était la réalisation d'un avant-projet concernant la réorganisation de l'espace architectural d'une entreprise et la construction d'une extension. Deux acteurs étaient présents lors des réunions : l'architecte responsable et un de ses collaborateurs internes depuis plusieurs années.
- une situation de **travail distant** où les acteurs collaborent à partir de lieux différents grâce à un outil technologique, le Studio Digital Collaboratif (SDC). Il s'agit d'un outil multimodal offrant aux utilisateurs la possibilité d'utiliser les mêmes modalités de communication que celles qu'ils peuvent mobiliser en situation de coprésence (verbale, visuelle, gestuelle et graphique). Ces informations multimodales sont échangées en temps réel, et à distance, via un système de vidéoconférence couplé à une table partagée qui permet la

réalisation de croquis collaboratifs (pour une description complète du système, voir Safin et Leclercq, 2009). Nous avons analysé une réunion de travail distant réalisé avec l'outil ; nous avons filmé et codé l'ensemble de la réunion, couvrant une durée totale de 60 minutes. La tâche était de mettre au point une esquisse sommaire pour l'aménagement d'une parcelle en zone mixte, comprenant notamment une crèche, des résidences-services, des commerces de grande distribution et des logements. Trois acteurs étaient présents : un architecte responsable d'atelier, le même que celui observé dans la première situation, un autre des collaborateurs internes que celui observé dans la situation de coprésence, et enfin un des ses collaborateurs externes. Ces trois architectes sont également habitués à travailler ensemble depuis plusieurs années.

Dans le cadre d'une expérimentation exploratoire, la situation de coprésence constitue la *situation de référence* à laquelle nous comparons les échanges multimodaux en situation outillée. En effet, nous considérons cette situation comme « ligne de base » à partir de laquelle nous déduirons des stratégies d'adaptation et de compensation de la part des architectes.

Méthodologie

Notre méthodologie d'analyse a consisté dans un premier temps à repérer et à caractériser les modalités mobilisées dans les communications entre agents. Nous avons distingué les modalités verbale, graphique, visuelle et gestuelle.

Nous avons considéré les indicateurs suivants pour caractériser chaque modalité :

- Indicateurs de la modalité verbale : les *communications verbales* entre acteurs ont été enregistrées et retranscrites. Les protocoles verbaux ont ensuite été analysés et décomposés en « tours de parole ». Un « tour de parole » débute et se termine avec la prise de parole d'un locuteur ;
- Indicateurs de la modalité graphique : nous avons recensé toutes les *productions graphiques* effectuées par les acteurs au cours de l'activité de collaboration. Nous les avons classées en deux catégories, selon leur fonction : (a) annoter ou compléter et (b) pointer ou localiser ;
- Indicateurs de la modalité visuelle : nous avons enregistré *l'activité des acteurs* en utilisant une vidéo et analysé à postériori les *regards* des différents acteurs (en distinguant ceux dirigés vers le partenaire distant de ceux dirigés vers les plans) ;

- Indicateurs de la modalité gestuelle : les *gestes de pointage* ou *gestes déictiques* (Mc Neil, 1992 ; Falzon, 1994) ont été enregistrés et analysés en détails.

Pour une description complète de la méthodologie utilisée, voir Defays et al. 2013.

Supports logiciels

Le codage et l'analyse des modalités de communication ont nécessité l'utilisation d'outils informatiques.

Dans un premier temps, nous avons utilisé le logiciel *ELAN*¹, spécifiquement destiné au codage des données multimodales (Rohlfing *et al.*, 2006).

Nous avons ensuite développé une application, appelée « COMMON Tool² », permettant d'importer les données codées dans ELAN dans une base de données. Cette base de données est ensuite interrogée via des requêtes SQL pour réaliser des analyses fines sur les modalités et déduire des articulations multimodales (mobilisation simultanée de deux ou plusieurs modalités).

Résultats

Comparaison du nombre de mobilisations

Dans un premier temps, nous avons comparé le nombre de fois où les modalités gestuelle et graphique ont été mobilisées pour les deux situations de travail collaboratif (en coprésence / outillée). Les durées des deux situations étudiées étant différentes (coprésence: 30min / outillée : 60 min), et le nombre d'architectes étant différent (2 en coprésence, et 3 en outillée), nous avons ramené l'ensemble de nos résultats par heure et par acteur.

Comme le montre la Figure 2, nous observons une faible augmentation du nombre de traces graphiques, et une très forte diminution des gestes de pointage en situation outillée. Les différences observées entre les deux situations (coprésence vs outillée) étant significatives ($p < .001$).

¹ <http://www.lat-mpi.eu/tools/elan/>

² Pour plus d'informations sur cette application: <http://www.commontool.be>

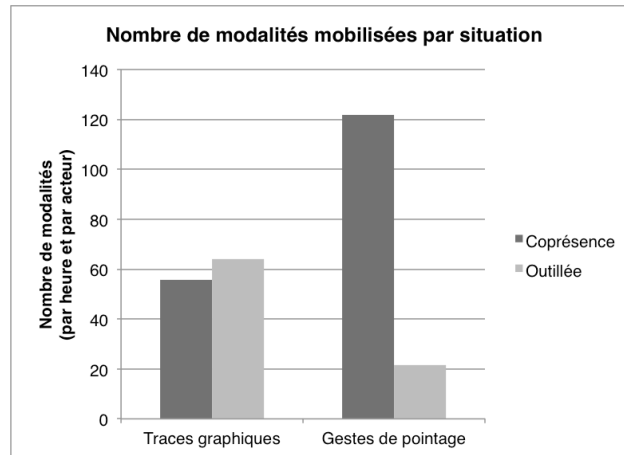


Fig. 2 Mobilisation des modalités graphique et gestuelle pour chaque situation (coprésence ou outillée)

Nous expliquons ce résultat par le fait que les gestes de pointage n'étant que partiellement retransmis par la webcam, la mobilisation de cette modalité devient inefficace, et donc nettement moins utilisée en situation outillée. Nous postulons que les architectes vont compenser cette incapacité à pouvoir mobiliser la modalité gestuelle par l'usage de la modalité graphique, même si le nombre de traces graphiques n'augmente que légèrement.

Pour vérifier cette hypothèse, nous avons examiné plus en détail la nature des traces graphiques mobilisées dans ces deux situations. Nos résultats montrent que :

- Pour la situation de coprésence, 57% des traces graphiques ont pour rôle de compléter et d'annoter les plans, et 43% ont pour fonction le pointage ou la localisation d'une zone d'intérêt ;
- Pour la situation outillée, seulement 41% des traces graphiques ont pour rôle l'annotation des plans, et 59% ont pour fonction la localisation d'une zone d'intérêt.

Les différences observées dans la répartition de la nature des traces graphiques entre les deux situations (coprésence vs outillée) étant également significatives ($p < .05$).

Ces résultats suggèrent qu'en situation outillée, la modalité graphique va être davantage mobilisée pour pointer ou localiser le référent de la discussion que pour dessiner à proprement parler. Ces résultats suggèrent également qu'en situation outillée, les architectes compensent l'inefficacité de la modalité gestuelle par l'usage de la modalité graphique, qui va dans plus de 50% des cas, avoir un rôle de monstration. Pour pallier la

dégradation de la modalité gestuelle en situation outillée, les architectes semblent donc devoir produire davantage de traces graphiques *déictiques*.

Comparaison de la durée de mobilisation

Si on considère la durée de mobilisation moyenne de chaque modalité de communication pour les deux situations, nos résultats montrent que les quatre modalités de communication sont mobilisées sur une durée plus longue en situation outillée par rapport à la situation de coprésence (voir Defays et al. 2013).

Cette différence est d'autant plus importante pour la modalité gestuelle puisqu'il faut en moyenne 2 secondes pour pointer une zone du plan ou un élément précis en coprésence, et près de 5 secondes en outillée. Conformément à nos observations et hypothèses, ce résultat montre que les architectes vont laisser plus longtemps leur main ou leur doigt sur les plans numériques que sur les plans papier afin de garantir que leurs gestes de pointage soient bien perçus par l'autre via la webcam.

Comme cela a déjà été montré, les acteurs ont également tendance à prendre plus de temps pour produire leurs traces graphiques avec un dispositif numérique que sans (Safin, 2011).

Ces résultats suggèrent que l'utilisation de l'outil transforme l'activité de conception collective puisqu'il semble amener les architectes observés à mobiliser les modalités de manière différente, et pendant une durée moyenne plus longue.

Analyse de la multimodalité des références démonstratives

Dans ce présent papier, nous nous centrons sur le processus de production des *références démonstratives*. Nous cherchons à comprendre comment les architectes vont s'aider des modalités non-verbales qui sont à leur disposition pour désigner le référent de leurs discours. En d'autres mots, nous analysons la manière dont les architectes mobilisent simultanément les modalités non-verbales pour accompagner leurs références verbales démonstratives. En outre, nous voulons déterminer si la médiatisation technologique va modifier la manière dont ces modalités non-verbales sont mobilisées.

Pour les deux situations envisagées (coprésence / outillée), nous avons repéré dans le discours ce que nous avons appelé des indices de monstration. Ces indices sont :

- les adjectifs démonstratifs suivants : « ce, cet, cette, ces » ;
- les pronoms démonstratifs suivants : « celui, celui-ci, celui-là, ceux, ceux-ci, ceux-là, celle, celle-ci, celle-là, celles, celles-ci, celles-là, ceci, cela, ça ».

L'application COMMON Tool permet de filtrer l'ensemble des verbalisations en fonction de leurs contenus, mais également de les lier à la mobilisation simultanée des autres modalités non-verbales. Nous avons utilisé cette application pour sélectionner les verbalisations contenant au moins un indice de monstration, et pour savoir si chacune d'entre elles était accompagnée ou non de l'une ou l'autre modalité non-verbale.

Nos résultats montrent que les verbalisations contiennent relativement le même pourcentage d'indices de monstration dans les deux situations analysées (coprésence: 28,92% / outillée: 30,03%). Le fait de montrer ou de désigner un endroit précis ou une zone semble occuper une très grande part de l'activité de conception architecturale.

En examinant les verbalisations contenant au moins un indice de monstration, nos résultats montrent que près de 90% d'entre elles sont accompagnées de modalités non-verbales, et ce, pour les deux situations envisagées (coprésence: 87,5% / outillée: 91,21%). Ce résultat, sans grande surprise, confirme que les modalités non-verbales sont indispensables dans le processus de monstration. Le corollaire de ce résultat est que, en l'absence de ces modalités non-verbales, les processus de production et de compréhension des références démonstratives seront perturbés.

Globalement, ces deux résultats montrent, que pour les deux situations envisagées, la proportion de verbalisations contenant des indices de monstration et la proportion de celles-ci qui sont accompagnées de modalités non-verbales restent comparables.

Nous avons alors jugé intéressant d'analyser si les modalités non-verbales qui accompagnaient les verbalisations contenant des indices de monstration restaient les mêmes d'une situation à l'autre.

Nos résultats montrent que les modalités non-verbales qui les accompagnent ne sont pas représentées dans les mêmes proportions lorsque les architectes travaillent en coprésence et avec un outil technologique (voir Figure 3). Les différences observées entre ces deux situations (coprésence vs outillée) étant significatives ($p < .001$).

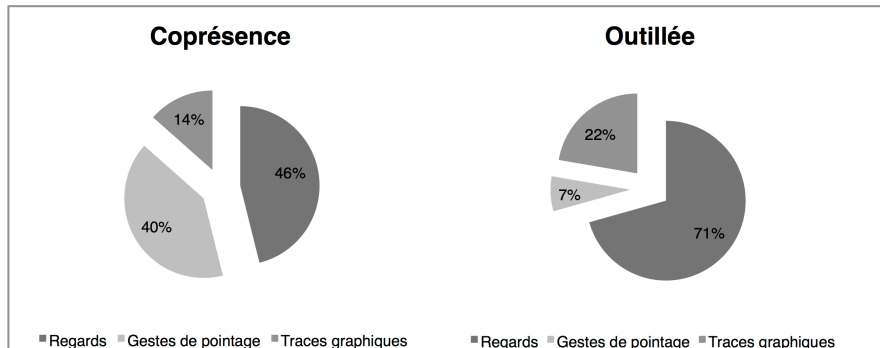


Fig. 3 Modalités non-verbales accompagnant les verbalisations contenant des indices de monstration

En coprésence, nos résultats montrent que :

- lorsque les architectes observés utilisent des références verbales démonstratives, ils vont moins souvent chercher le regard d'autrui pour vérifier la compréhension de leur partenaire qu'en situation outillée. Leurs regards, étant moins dirigés vers leur partenaire, sont davantage dirigés vers les plans ;
- c'est la modalité gestuelle qui va être principalement mobilisée pour désigner le référent dont il est question verbalement ;
- la modalité graphique est peu utilisée (14%) pour supporter la fonction de monstration.

En situation outillée, nos résultats montrent que :

- les partenaires se regardent très souvent pour vérifier leur compréhension mutuelle du référent désigné. En effet, les regards vers le partenaire accompagnent dans 71% des cas (au lieu de 46% en coprésence) les références démonstratives des architectes observés. Leurs regards, étant plus dirigés vers leur partenaire, sont moins dirigés vers les plans ;
- l'usage de la modalité gestuelle dans la fonction de désignation est diminué plus d'un quart (passant de 40% en situation de coprésence à 7% en situation outillée) ;
- la modalité graphique va davantage supporter la fonction de monstration puisqu'elle accompagne les références démonstratives dans 22% des cas (au lieu de 14% en coprésence) ;

Ce résultat confirme l'effet de compensation de l'inutilisabilité de la modalité gestuelle par l'usage de la modalité graphique. Cela confirme que dans la situation outillée observée, l'usage de la modalité graphique a davantage pour fonction la monstration et la mise en évidence d'une zone

d'intérêt, que de supporter l'activité de conception architecturale à proprement parler.

La distance et les contraintes de l'outil, rendant certaines modalités de communication difficilement mobilisables, vont donc changer la manière dont les architectes vont mobiliser ces modalités pour faire référence à des éléments du plan et construire le sens dans leurs interactions.

Discussion et conclusion

Cette étude a pour objectif de décrire comment les architectes vont mobiliser et combiner les modalités visuelle, gestuelle et graphique pour désigner ce à quoi ils font référence dans leurs discours. En particulier, nous voulons investiguer si la médiatisation technologique va modifier la manière dont ces modalités non-verbales sont mobilisées, et si des stratégies de compensation entre modalités de communication sont observées.

Nos résultats montrent que la fonction de monstration occupe une place importante dans les activités de conception envisagées, que ces activités prennent place en coprésence ou avec l'outil. Ensuite, nous montrons que les références démonstratives verbales sont presque systématiquement accompagnées de modalités non-verbales, rendant ces dernières indispensables dans le processus de production et de compréhension de ce type de références. Ce résultat suggère également que lorsque ces modalités non-verbales ne sont pas (ou sont difficilement) mobilisables, les processus de monstration et de compréhension des références monstratives vont être perturbés, ce qui a nécessité une adaptation de la part des opérateurs et donc des changements par le développement de compensations. En effet, en cas d'absence de l'une ou l'autre modalité non-verbale, la monstration ne se réalisera pas correctement, et donc l'identification et la compréhension du référent désigné sera également perturbée.

Nos résultats montrent qu'en cas d'inefficacité de la modalité gestuelle (transmission partielle des gestes de pointage), nous retrouvons un effet de compensation par l'utilisation de la modalité graphique. Ces résultats ont corroboré nos observations in situ qui postulaient que lorsque certaines modalités ne pouvaient être mobilisées, des stratégies de compensation par d'autres modalités allaient être observées.

D'un point de vue cognitif, ces stratégies d'adaptation sont parfois « coûteuses » pour les architectes. Par exemple, ceux-ci doivent souvent effacer les traces graphiques qu'ils ont produites dans le but de désigner un

référent. Durant l'activité analysée, les architectes ont également vérifié plusieurs fois que leur partenaire distant avait bien accès à leurs gestes de pointage, et comprenait bien le référent dont il était question. En outre, nous avons observé que les utilisateurs se tenaient debout pour pouvoir avoir une meilleure vue sur les gestes du partenaire distant.

Dans d'autres recherches, nous montrerons que ces gestes de pointage sont indispensables en coprésence, surtout lors de vérification verbale de la compréhension du partenaire, ou encore lors de clarifications verbales. Le *common ground* est également essentiel dans ce processus de monstration puisqu'il va permettre aux partenaires de pouvoir faire les bonnes inférences sur le référent désigné par l'une ou l'autre modalité (Clark et al. 1983).

Nous recommandons dès lors l'implémentation d'un système de reconnaissance des gestes de pointage dans les systèmes visant à supporter les activités de conception architecturale collectives. Dans le dispositif multimodal que les architectes ont utilisé, le Studio Digital Collaboratif, il existe actuellement un système de pointeur représenté par un carré rouge. Ce pointeur indique la position du stylo numérique du partenaire distant, au survol de la table graphique. Nous recommandons un système de pointeur plus visible et plus dynamique que celui implémenté lors de ces analyses³. L'amélioration de la fonctionnalité de pointage faciliterait l'identification du référent, le processus de référenciation, et par là également le maintien d'un *common ground*. Le lien entre le *common ground* et la multimodalité dans le processus de monstration doit cependant être encore approfondi et cette investigation constituera l'objectif de nos travaux ultérieurs.

Remerciements

Cette recherche a été financée dans le cadre des Actions de Recherche Concertées (communauté française de Belgique) et via une bourse de doctorat Non-FRIA de l'Université de Liège. Une partie de l'analyse rapportée ici a été réalisée lors d'un séjour doctoral effectué par le premier auteur dans l'équipe Wimmics du Centre Inria Sophia Antipolis-Méditerranée.

³ Notons que récemment un système de « traînée », appelé COMETE, permettant de visualiser dans un dégradé d'opacité les dernières positions du curseur a été développé. La validité de cette représentation devra faire l'objet de travaux ultérieurs.

Références

1. Campanella, Salvatore, & Belin, Pascal. (2007). Integrating face and voice in person perception. *Trends in Cognitive Sciences*, 11(12), 535-543.
2. Chellali, A. M. (2009). *Étude des interactions homme-homme pour l'élaboration du référentiel commun dans les environnements virtuels collaboratifs* (Doctoral dissertation, Université de Nantes).
3. Clark, H. H., & Schaefer, E. F. (1989). Contributing to discourse. *Cognitive science*, 13(2), 259-294.
4. Clark, H. H., Schreuder, R., & Buttrick, S. (1983). Common ground and the understanding of demonstrative reference. *Journal of Verbal Learning and Verbal Behavior*, 22(2), 245-258.
5. Darses, F., Détienne, F., Falzon, P., & Visser, W. (2001). A method for analysing collective design processes. *Rapport de recherche de l'INRIA-Rocquencourt*.
6. Defays, A., Leclercq, P., & Nyssen, A. S. (2013). Influence des communications multimodales sur le common ground. Proposition d'une méthodologie d'analyse. *Interfaces Numériques*, 2(3), 591-609.
7. de Gelder, B., & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends in Cognitive Sciences*, 7(10), 460-467.
8. Dumas, Bruno, Lalanne, Denis, & Oviatt, Sharon. (2009). Multimodal Interfaces: A Survey of Principles, Models and Frameworks. In D. Lalanne & J. Kohlas (Eds.), *Human Machine Interaction* (pp. 3-26): Springer-Verlag.
9. Dumazeau, C. (2005). *Favoriser l'établissement d'un contexte mutuellement partagé dans les communications distantes*. Cnam, Paris.
10. Falzon, P. (1994). Dialogues fonctionnels et activité collective. *Le Travail humain*, 57(4), 299-312.
11. Nomura, S., & Hutchins, E. (2007). *The Multimodal Production of Common Ground Understandings in Intercultural Flight Training*. Paper presented at the 14th International Symposium on Aviation Psychology.
12. Rohlfing, K., Loehr, D., Duncan, S., Brown, A., Franklin, A., Kimbara, I., . . . Wellinghoff, S. (2006). Comparison of multimodal annotation tools: Workshop report. *Gesprächsforschung*, 7, 99-123.
13. Safin, S. (2011). *Processus d'externalisation graphique dans les activités cognitives complexes : le cas de l'esquisse numérique en conception architecturale individuelle et collective*. Thèse de doctorat, Université de Liège, Belgique.
14. Safin, S. & Leclercq, P. (2009). Studio Digital Collaboratif : un environnement de conception collaborative à distance. *Actes du colloque Interaction Homme-Machine, IHM2009*, Grenoble, France.
15. Stein, B. E., & Meredith, M. A. (1993). *The Merging of the Senses*. Cambridge MA: The MIT Press.
16. Tajariol, F. (2006). *Les effets des indices non-verbaux sur les activités de communication à distance. Deux études expérimentales sur le dialogue tutoriel*. Université Joseph Fourier, Grenoble.