

# Améliorer le contexte mutuellement partagé lors de communications distantes avec un outil de désignation.

*Cécile Dumazeau*

GRIC-IRIT  
118 rte de Narbonne  
31062 Toulouse  
[dumazeau@irit.fr](mailto:dumazeau@irit.fr)

CENA  
7 av. Edouard. Belin  
BP 4005  
31055 Toulouse

*Laurent Karsenty*

Intuilab  
Prologue 1, La Pyrénéenne  
31672 Labège Cedex  
[karsenty@intuilab.fr](mailto:karsenty@intuilab.fr)

## RESUME

De nombreux outils d'aide au travail collaboratif et de communication médiatisée aident à créer un contexte partagé pour faciliter la communication à distance. Mais des études montrent qu'il est aussi important d'établir la preuve de partager mutuellement une information que de partager l'information elle-même. Notre expérience vise à évaluer si un outil de pointage dans un espace partagé suffit pour avoir la certitude de partager une information. Le but final est de concevoir un outil aidant les contrôleurs aériens à se coordonner à distance. Toutefois, pour contourner certaines contraintes opérationnelles, une tâche artificielle reproduisant la plupart des contraintes et des exigences de ce domaine a été créée. Dans cette expérimentation, les performances liées à la tâche et à la communication sont comparées, selon que les sujets soient co-présents ou distants avec un outil de pointage. Nous discutons des éclaircissements théoriques et des recommandations de conception qui découlent des résultats.

**MOTS CLES :** Communication médiatisée par ordinateur, contexte partagé, contrôle aérien, ergonomie cognitive.

## ABSTRACT

Many CSCW tools try to make a shared context available in order to ease remote communication and collaboration. However, studies have shown that the evidence that the context is shared is as important as the shared information itself. Our study tries to assess how confident remote collaborators are regarding shared information that is established with a telepointer. The target domain of application is air traffic control (ATC), where remote coordination is a core activity. However, to cope with operational constraints, we have set up an artificial task with almost the same communication requirements and constraints as ATC. In this experiment we have compared task and communication performance with one group of subjects working side-by-side and the other remotely with context sharing tools. Finally we raise theoretical issues and provide some design recommendations.

**KEYWORDS :** Computer-mediated Communication, Shared Context, Air Traffic Control, Cognitive ergonomics.

## INTRODUCTION

Les outils d'aide au travail collaboratif (CSCW en anglais) et les outils de communication médiatisée par ordinateur (CMC en anglais) ont ouvert de nouvelles perspectives pour le travail collectif. Ces disciplines, au départ très portées sur les aspects technologique et innovant, cherchent maintenant à faire le lien avec les sciences humaines qui se sont attachées à modéliser le travail collectif. Notre étude vise à mieux comprendre la communication humaine au travail dans le but de concevoir des outils d'aide aux communications synchrones et distantes. Un axe de recherche important dans ce domaine est la notion de *conscience de la situation* du collègue distant. De nombreux travaux de psychologie linguistique [3, 14], d'ethnométhodologie [10] et d'ergonomie cognitive [11] montrent que le fait d'être face à face et de pouvoir suivre en permanence les activités de l'autre favorise la création d'un contexte partagé. Ces travaux montrent que ce contexte partagé rend l'interprétation des intentions de l'autre plus aisée et globalement facilite la communication et la collaboration. Mais à distance ce contexte partagé n'existe pas. De nombreuses solutions technologiques permettent la transmission d'informations sur l'activité d'une personne distante [5, 9]. Mais est-ce suffisant de transmettre du contexte sur la situation de l'autre pour construire un contexte partagé ? Ne faut-il pas en plus des indices assurant au locuteur que son partenaire a vu ces informations et les partage réellement avec lui, comme cela se fait naturellement en face à face ?

Cette question est particulièrement importante pour les contrôleurs aériens car ils ont souvent à communiquer sous forte pression temporelle avec l'impératif d'être immédiatement bien compris. Les contrôleurs aériens, distribués dans divers lieux, se partagent le travail en contrôlant chacun un volume aérien. Ils doivent donc coordonner leurs actions pour éviter des collisions aux frontières de ces volumes. En France ces coordinations se font par téléphone. La possibilité d'avoir mutuellement conscience de l'environnement de travail de l'autre devrait permettre aux contrôleurs aériens de communi-

quer plus facilement tout en diminuant les risques d'incompréhension.

Pour explorer cette question, nous précisons dans un premier temps les modèles théoriques de construction d'un contexte partagé dans la communication humaine naturelle. Nous présentons ensuite une expérimentation est visant à évaluer dans quelle mesure le partage d'une information visuelle suffit pour que cette information soit mutuellement partagée par des personnes distantes. Pour cela nous avons comparé les performances de communication et de réalisation d'une tâche collective de personnes présentes dans un même lieu et de personnes distantes. Pour contourner certaines contraintes opérationnelles, la tâche choisie pour cette expérimentation est une tâche artificielle reproduisant la plupart des propriétés des tâches de coordination dans le contrôle aérien.

## CONTEXTE PARTAGE ET COMMUNICATION

### Notion de contexte cognitif

Cette recherche se place dans le domaine de l'ergonomie cognitive, ainsi par contexte, nous désignons le contexte cognitif. Le contexte est souvent défini comme l'environnement physique (bruit, luminosité etc.) dans lequel se trouvent les individus, les artefacts qui les entourent (outils, interfaces, contenu informationnel de ces interfaces etc.) et les activités et caractéristiques de ces individus (localisation, statut, état de stress etc.) [4]. Il est objectivement descriptible par un observateur extérieur.

A l'inverse, le contexte cognitif de quelqu'un est ce que cette personne perçoit subjectivement de cet environnement. La perception de l'environnement est déterminée par des aspects externes sur lesquels on peut agir (accessibilité des informations, audibilité). Mais cette perception va aussi fortement dépendre de l'état interne de l'individu : but immédiat, interprétations précédentes, attentes sur la situation ou encore état de stress. Selon ce contexte interne, certaines connaissances à long terme seront plus facilement accessibles que d'autres. Dusire S. [7] montre ainsi que l'individu se construit une représentation fonctionnelle de l'environnement en sélectionnant seulement une partie des informations disponibles et en créant de nouvelles informations par des inférences reposant sur ses attentes et sa connaissance du domaine. Ainsi, deux personnes dans un même environnement objectif ne seront pas nécessairement dans le même contexte cognitif.

### Différents types de partage de contexte

La majorité des travaux sur le partage de contexte à distance comme [5, 9] se base sur la définition objective du contexte et sur la notion de *conscience de la situation* qui en découle (*Situation awareness*). Il s'agit de maintenir à jour la connaissance de l'environnement de tra-

vail des autres, de sa configuration dans le temps et l'espace, et des gens qui y interagissent. Chacun peut ainsi vérifier que sa contribution individuelle est cohérente avec l'activité du groupe. Cela peut être suffisant quand les tâches des uns et des autres sont faiblement couplées. Mais, pour communiquer, il est indispensable de pouvoir faire des hypothèses sur les informations contextuelles dont l'auditeur dispose afin d'adapter le contenu verbal de son message aux capacités inférentielles de l'auditeur. Pour autant, inférer de manière unilatérale le contexte de l'autre n'est pas toujours suffisant pour une communication implicite et réussie. Il faut que locuteur et auditeur fassent du mieux possible les mêmes hypothèses sur le contexte qu'ils partagent. Pour cela, il faut qu'ils aient des raisons suffisantes de croire qu'ils *partagent mutuellement* un même contexte. De nombreuses incompréhensions sont en effet dues à des mauvaises représentations de ce qui est partagé avec l'autre. Sperber et Wilson [13] introduisent la notion d'*environnement cognitif mutuel* dans lequel toute information contextuelle manifeste est *mutuellement manifeste* aux deux personnes, c'est-à-dire à la fois manifeste pour chacun d'entre eux et manifeste que l'autre a accès à cette information. Clark et Marshall [3] utilisent les termes de *savoir mutuel*. Un savoir est dit mutuel lorsqu'on peut, a posteriori, lui appliquer un raisonnement équivalent au suivant : « J'ai lu le livre X, je sais que mon partenaire a lu ce livre et réciproquement, mon partenaire sait que j'ai lu le livre X, mais je sais aussi que mon partenaire sait que je sais qu'il a lu le livre », et ainsi de suite à l'infini.

Il n'est pas toujours utile d'avoir un niveau de certitude élevé sur le contexte effectivement partagé, mais nous estimons que c'est une exigence forte dans le cas de coordinations verbales distantes demandant une efficacité maximale.

### Un modèle cognitif du contexte mutuellement partagé

Clark et Marshall [3] estiment que trois heuristiques permettent de faire l'hypothèse qu'une information contextuelle est mutuellement partagée :

- la *co-présence physique* : les deux personnes sont ou ont été simultanément en présence physique avec l'information contextuelle,
- la *co-présence linguistique* : tout fait évoqué verbalement au cours d'une conversation verbale,
- l'*appartenance à une même communauté* : à chaque communauté (culturelle, scientifique, familiale etc.) correspond un ensemble de connaissances partagées (linguistiques, épisodiques etc.).

Les heuristiques de co-présence reposent elles-mêmes sur trois hypothèses :

- *Hypothèse de simultanéité* : à un moment donné les deux personnes et l'objet sont ouvertement co-présents, dans un même espace, au même moment,

- *Hypothèse d'attention* : les interlocuteurs ont, ont eu ou peuvent facilement avoir à un moment donné tous les deux leur attention tournée sur cet objet,
- *Hypothèse de rationalité* : chacun des interlocuteurs suppose que l'autre a les capacités perceptives et cognitives de faire les hypothèses précédentes.

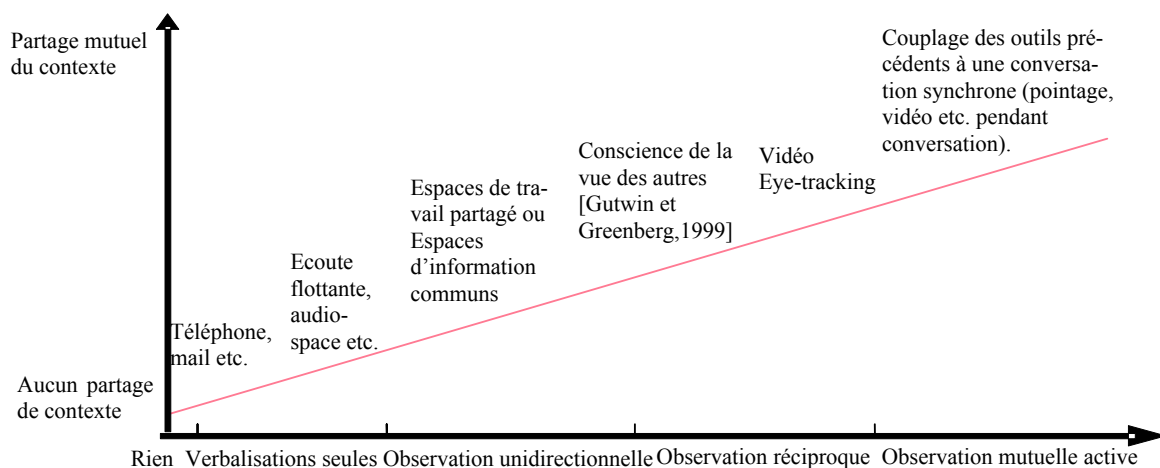
A distance, il manque l'heuristique la plus forte : la co-présence physique. Les actes déictiques, c'est à dire les gestes de pointage des objets, permettent à la fois de rendre ces objets ouvertement co-présents et d'attirer l'attention de l'autre dessus : ils assurent une co-présence physique. Dans une expérimentation, Wilke-Gibbs et Clark [14] ont isolé d'autres indices qui permettent d'établir l'assurance de partager un même contexte. Ainsi, si une personne distante peut voir l'objet en même temps qu'elle écoute les verbalisations d'autrui sur cet objet, alors l'assurance de partager avec eux la connaissance de l'objet est plus forte que si cette personne ne peut entendre que leurs verbalisations. Le fait d'avoir en plus un contact visuel mutuel permet d'être encore plus sûr de partager un même contexte et le fait de participer activement à la conversation est ce qui permet d'avoir le sentiment le plus fort de partager un même contexte.

## OUTILS DE PARTAGE DE CONTEXTE

Pour favoriser un contexte mutuellement partagé à distance, il ne suffit donc pas de rendre les informations contextuelles accessibles, il faut aussi fournir des indices permettant aux interlocuteurs de s'assurer que ces informations sont mutuellement perçues par l'autre. Un premier axe de recherche consiste à donner la possibilité

du visage aide à gérer les tours paroles dans le cas de conversations à plusieurs et améliore l'efficacité des dialogues dans lesquels l'émotion de l'autre a une importance, la vidéo n'a généralement pas d'effets significatifs sur les conversations de travail entre deux personnes [8]. Une étude en cours [1] étudie aussi cette question en testant l'utilité d'un système de capture du regard pour suivre où se pose le regard de l'autre à distance. Un deuxième axe de travail vise à coupler la transmission d'informations contextuelles à la possibilité d'y agir manifestement dessus, au moment de la conversation. Chaty, Girard et Sire [2] préconisent de fusionner les fonctions de communication et de production de la tâche collective, là où la plupart des collecticiels séparent ces fonctions. De nombreux outils visent en effet à recréer artificiellement une unité de lieu et de temps malgré la distance. Les propositions technologiques sont généralement basées sur la reproduction la plus exhaustive et fidèle possible de la situation de face à face, par exemple, par des systèmes de réalité virtuelle ou de superposition de l'image des personnes sur une représentation électronique de l'espace de la tâche. Mais ces outils n'ont pas forcément fait l'objet d'une évaluation ergonomique. La figure 1 résume et met en relation les indices disponibles, le niveau de certitude de partage de contexte qu'ils permettraient selon Wilkes-Gibbs et Clark [14] et les solutions technologiques existantes pour transmettre ces indices à des personnes distantes.

Dans l'expérimentation présentée dans la suite, le but était d'évaluer dans quelle mesure la simple désignation d'un objet dans un espace virtuel partagé suffit à créer la



**Figure 1** : Niveau de certitude de partage mutuel d'un contexte en fonction du type d'informations partagées à distance.

aux opérateurs distants de valider l'hypothèse d'attention. La transmission vidéo du visage ou de ce que voit l'autre est la solution la plus souvent envisagée dans les recherches actuelles (ex., [12]). Mais si la vidéo

certitude de partager mutuellement la connaissance de cet objet.

## METHODE

Dans cette expérimentation, nous comparons la performance de communication et de résolution d'une tâche collective de sujets distants disposant d'un espace partagé et d'un outil de pointage avec des sujets placés dans un même lieu. Cette expérimentation, en plus de fournir des connaissances sur le contexte partagé, devait nous aider à concevoir un outil pour les contrôleurs aériens.

### Tâche collaborative

Nous avons conçu un jeu reproduisant partiellement les propriétés du contrôle aérien. L'espace de ce jeu est un rectangle coupé en deux zones identiques par un « mur » vertical percé de trois portes numérotées de haut en bas (figure 2). Chacun des deux joueurs a la responsabilité d'une moitié de l'espace de jeu, dans laquelle des billes se déplacent horizontalement et à vitesse constante. Sur la bille est écrit le numéro de la porte par laquelle il faut la faire passer et l'heure à laquelle elle arrivera au mur.



Figure 2 : Surface de jeu d'un sujet

A chaque fois qu'une bille apparaît, une cible lui correspondant apparaît dans l'espace de jeu du partenaire. Le but du jeu est de mettre le plus grand nombre possible de billes dans la cible qui lui correspond : chaque bille qui rejoint correctement sa cible rapporte des points. Il s'agit d'un jeu coopératif dans la mesure où les contributions de chacun rapportent des points aux deux joueurs : le score est unique. Les sujets doivent détourner les billes de leur trajectoire pour les transférer à leur partenaire par la porte indiquée sur la bille ou vers la cible adéquate : ils ont donc deux flux de billes à gérer. Mais, parfois deux billes sont prévues pour arriver à la même heure sur une même porte qui ne permet de faire passer qu'une bille à la fois. Les joueurs ont pour consigne de détecter ce type de problème et de se concerter pour éviter la collision et perdre le moins de points possible.

### Dispositifs comparés

*Côte à côte* : Les deux joueurs sont assis devant leurs écrans placés l'un à côté de l'autre dans une même pièce. Ils peuvent communiquer quand ils le veulent, se voient en permanence, et peuvent à tout moment regarder l'écran de l'autre et pointer du doigt dessus si besoin.

*A distance* : Les joueurs sont dans des pièces différentes. Ils disposent d'un micro-casque pour communiquer dès que nécessaire. Ils peuvent désigner sur l'écran de leur partenaire une ou plusieurs billes de leur choix à volonté, en pointant avec la souris la bille ou la cible correspondante et en appuyant sur le bouton droit ; la bille est alors entourée d'un cercle jaune clignotant sur leurs deux écrans distants (figure 2).

### Participants

36 sujets ont participé : des hommes et des femmes de 18 à 56 ans (âge d'une population active), se servant d'ordinateur tous les jours ou presque pour 89% d'entre eux et ayant pour 97% d'entre eux au moins un BAC+2. Ils sont répartis en deux groupes indépendants: 9 couples en côte à côte et 9 en situation distante.

### Procédure

4 scénarii de 12 minutes ont été conçus : la première phase est lente, et lors des 5 dernières minutes, la cadence d'arrivée des billes double, augmentant ainsi le nombre de billes à gérer en même temps (charge de travail élevée). Ils contiennent tous 11 conflits au niveau des portes répartis équitablement au cours des deux phases et se produisant dans des contextes semblables d'un scénario à l'autre, mais dans un ordre différent. Chaque couple a reçu strictement les mêmes instructions grâce à une consigne écrite. Dans une *phase de familiarisation*, l'expérimentatrice leur montrait les fonctionnalités puis ils jouaient le scénario de familiarisation durant lequel il était possible de poser des questions. Ensuite, chaque couple a joué successivement les trois scénarii dans un ordre différent. Après la dernière partie, une auto confrontation était menée avec les deux joueurs ensemble. Ils visionnaient leur dernière partie et écoutaient leurs échanges. A chaque coordination l'expérimentatrice cherchait par des questions ouvertes à faire expliciter les représentations qu'ils avaient au cours de la partie. Des questions plus générales étaient aussi posées sur les stratégies, les problèmes rencontrés ou sur l'intérêt ressenti pour l'outil. Enfin, les sujets devaient remplir un questionnaire portant sur leurs caractéristiques individuelles et leur sentiment de facilité de collaboration.

### Recueil des données et mesures

Les coordinations verbales ont été enregistrées et transcrites. Les actions réalisées sur l'interface ont été aussi enregistrées de manière à transcrire et analyser les verbalisations en contexte. Les sujets en côte à côte étaient filmés afin d'observer leurs gestes et regards. Par ailleurs, d'après l'observation des contrôleurs aériens [6] et les pré-tests nous nous attendions à ce que les coordinations soient divisées en une phase d'*identification des billes*, une phase d'*établissement du diagnostic* pendant laquelle les sujets se mettent d'accord sur l'existence d'un conflit et enfin, une phase de *résolution du problème*. Nous n'étudions que la phase d'identification car

initiant la première prise de contact elle est très dépendante du fait d'avoir ou pas un contexte mutuellement partagé et pose des difficultés chez les contrôleurs.

Nous présentons ci-dessous les variables que nous avons étudiées.

### **Mesure de l'efficacité du dialogue de désignation**

- *Nombre d'interventions.*

Une intervention est une prise de parole d'un même locuteur abordant un thème et un seul.

- *Nombre de descriptifs.*

Ces descriptifs sont la couleur, la porte par laquelle la bille doit passer (exemple : « J'en ai une qui passe en 3. »), l'heure de passage de la porte, et l'ordre d'arrivée sur la porte, par rapport à d'autres billes (exemple : « La première »).

- *Nombre de billes verbalisées.*

Il peut être suffisant de ne décrire qu'une bille du conflit et de laisser le partenaire inférer les autres ou de ne verbaliser aucune bille (exemple : « Ca va pas passer, là. »).

- *Nombre d'indices de localisation.*

Il s'agit d'adverbes comme « en bas », du numéro de la porte sur laquelle le conflit aura lieu ou encore d'un possessif qui indique dans laquelle des deux moitiés de l'espace de jeu la bille se situe (« J'ai une 03:44 et tu as une 03:48 »).

- *Difficulté de compréhension.*

Elles sont révélées par des demandes de désambiguïsation (exemple : « Tu parles de la bille 02 : 12 ou de la 02 : 16 ? »), des demandes ouvertes comme « Tu parles de quelle bille ? » ou encore par des silences et/ou mots montrant une hésitation. Sont comptées aussi les incompréhensions récupérées par le locuteur (par exemple : « C'est la bleue, pas celle-là. »).

- *Niveau d'explicité de l'accord de l'auditeur.*

En cas d'incompréhension, seul l'accord après discussion sur le désaccord est codé. Trois niveaux sont définis :

- (1) *Aucune marque d'accord.* L'auditeur ne dit rien et ses actions ne permettent pas de dire s'il y a accord.
- (2) *Accord implicite.* L'auditeur montre son accord indirectement par ses actions (il déplace la bille dont on lui a parlé comme convenu) ou par son discours (exemple : L1 : « Il va y avoir un problème avec ma 2 et ta rouge. » L2 : « Je descends la mienne en porte 3. »).
- (3) *Accord explicite.* L'interlocuteur verbalise son accord par des mots comme « OK » ou en répétant le nom des billes (« Ah oui, ta noire 5:08 pourra pas passer. »), accompagné ou pas d'une interrogation, invitant l'autre à confirmer.

### **Mesures de la performance de la tâche**

- *Score :* Total des points obtenus à un scénario.

- *Réussite de la coordination :* Il y a réussite si le conflit objet de la coordination est évité, échec si les billes du conflit entrent en collision.

- *Collision en parallèle :* C'est le cas si d'autres billes entrent en collision pendant la conversation (un autre conflit ou une cible est oublié).

- *Optimisation de la solution :* La solution est optimale si un minimum de billes passe par la mauvaise porte et sous optimale dès qu'une bille est détournée alors qu'elle pourrait passer.

- *Diagnostic :* Si les sujets concluent à un problème là où il n'y en a pas ou vice versa, il s'agit d'un mauvais diagnostic. Si après un bon diagnostic la mise en oeuvre rate (oubli, raté etc.), le diagnostic est considéré comme bon.

### **Hypothèses**

#### *Hypothèse 1 : Performance de la tâche*

Les sujets distants partagent beaucoup d'informations sur l'état de la tâche et en partie sur les actions de l'autre. Mais ils n'ont pas autant d'informations que dans une situation de co-présence physique. Les sujets placés en côte à côte devraient donc avoir des facilités pour collaborer et donc mieux réussir la tâche que les sujets distants.

#### *Hypothèse 2 : Efficacité du dialogue de désignation*

Les sujets distants ont des informations sur la tâche et sur certaines actions du partenaire, mais ils ne disposent d'aucune information visuelle sur le visage et le comportement d'autrui. Les sujets distants ne devraient donc pas avoir une assurance aussi forte de partager un même contexte que ceux placés en côte à côte, ce qui devrait affecter l'efficacité du dialogue. A distance le locuteur devrait faire plus d'effort pour désigner les billes (plus de descriptifs, de billes verbalisées et d'indices de localisation). L'auditeur devrait montrer plus explicitement sa compréhension et avoir plus de difficultés de compréhension, engendrant plus d'interventions verbales.

#### *Hypothèse 3 : Efficacité des actes déictiques*

Conformément au cadre théorique sur lequel nous nous appuyons, la possibilité de désigner les billes devrait tout de même renforcer l'assurance de partager un même contexte. Dans les cas où l'outil de désignation des billes est utilisé par les locuteurs distants, le processus de référence aux billes devrait être plus efficace que dans les autres cas et devrait plus se rapprocher de la situation en côte à côte (sans être aussi efficace).

### **RESULTATS**

#### **Effet sur la forme des coordinations**

Certaines coordinations ont des structures différentes de nos attentes. Parfois le locuteur se contente d'informer qu'il a détourné une des billes sans désigner le conflit à son partenaire. Dans d'autres cas, il s'agit d'un raisonnement à voix haute, dans lequel le locuteur évoque de manière floue un ensemble de billes dans lequel il pense qu'il peut y avoir un conflit mais c'est en parlant avec l'autre qu'il le détermine. Enfin, certains sujets préfèrent

une stratégie de recherche systématique, ils énoncent verbalement toutes les billes au fur et à mesure qu'elles apparaissent et détectent ainsi des conflits très en amont, avant que les billes ne soient dans la zone partagée. Cependant, les coordinations conformes à nos attentes, incluant une phase de désignation du conflit, sont les plus nombreuses (63,6% de l'effectif) et offrent 457 coordinations, nombre suffisant pour poursuivre une étude quantitative sur ce type de coordination. Notons qu'à distance les coordinations correspondent plus à nos attentes, alors qu'en côte à côte les sujets choisissent plus souvent des stratégies de recherche systématique. Dans ces cas, ils gardent un contact verbal presque permanent, et se racontent ce qu'ils voient et font.

### Effets sur la performance de la tâche

Les scores moyens sont à peu près identiques dans les deux conditions : le score moyen des sujets en côte à côte est de 2197 points (sd=84) et celui des sujets distants de 2189 points (sd=93) sur un score maximal de 2280 points. Sur l'ensemble des autres indicateurs de performance, nous n'observons aucune différence significative entre les deux conditions expérimentales (figure 3) ni selon la charge de travail.

Performance	Côte à côte	Distance
Echec	3,49%	4,82%
Solution sous-optimale	3,13%	4,00%
Collisions en parallèle	4,37%	3,56%
Mauvais diagnostics	5,56%	5,97%

Figure 3 : Réussite de la tâche collective

Cette performance est atteinte dès la première ou deuxième partie et ne s'améliore pas par la suite. Globalement les sujets réalisent une bonne performance et cette performance n'est dégradée ni par la distance ni par une charge de travail élevée. Lors du questionnaire l'ensemble des sujets ont répondu que la collaboration était facile (sur une échelle en 5 points, de très facile à très difficile 86,67% des sujets ont choisi les deux premiers points et les autres ont choisi le point neutre : moyennement facile). Contrairement à notre première hypothèse, ces outils de partage de contexte semblent donc suffire pour collaborer aussi efficacement qu'en côte à côte.

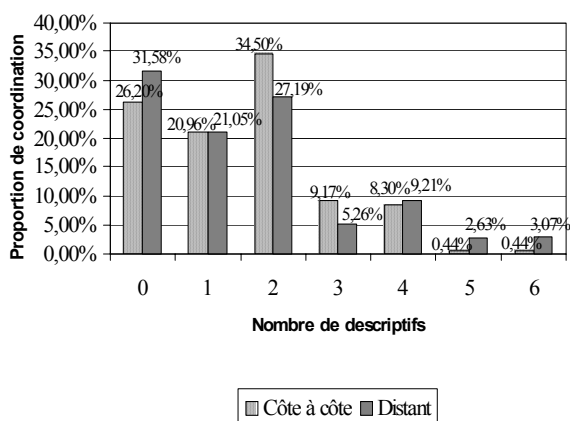


Figure 4 : Nombre de descriptifs

### Effets sur le dialogue

La figure 4 montre que les cas de coordinations avec seulement 2 ou 3 descriptifs verbaux sont plus nombreux en côte à côte (43,67%) qu'à distance (32,46%;  $\chi^2=5,6251$ ;  $p<0,01$ ). Inversement les cas de coordinations nécessitant 4, 5 ou 6 descriptifs verbaux pour décrire les billes tendent à être plus nombreux à distance (14,9%) qu'en côte à côte (9,2 %;  $\chi^2=3,0364$ ;  $p<0,05$ ). Conformément à l'hypothèse 2, à distance les sujets semblent éprouver un peu plus souvent le besoin de donner un grand nombre de caractéristiques des billes. Contre toute attente, les coordinations avec absolument aucun descriptif sont nombreuses à distance comme en côte à côte ( $\chi^2=1,3576$ ;  $p>0,1$ ).

Les cas dans lesquels les sujets décrivent verbalement plus de 2 billes sont aussi peu nombreux à distance (8,1%) qu'en côte à côte (4,4%,  $p>0,05$ ). Les sujets donnent souvent des indices de localisation et cela dans des proportions semblables que ce soit en côte à côte (59,83%) ou à distance (60,53%). Dans une très grande majorité des cas (82,06%), au maximum trois interventions verbales suffisent pour arriver à une compréhension commune. L'intercompréhension semble donc assez aisée. Cependant, les coordinations avec seulement 3 interventions sont significativement plus nombreuses en côte à côte (17,90%) qu'à distance (10,90%,  $\chi^2=3,9082$ ,  $p=0,02$ ), alors qu'inversement, les cas de coordinations longues, avec plus de 4 interventions verbales sont un peu plus nombreux à distance (10,96%) qu'en côte à côte (6,11%,  $\chi^2=2,8514$ ,  $p<0,05$ ). Ainsi, les quelques cas à problème demandant des interventions verbales supplémentaires sont légèrement plus nombreux à distance.

Les difficultés de compréhension sont assez fréquentes (19,5% des coordinations), mais il n'y a pas de différence significative entre les deux conditions ( $p=0,7$ ). Par contre le processus d'accord entre les deux interlocuteurs diffère légèrement (figure 5). Les cas d'absence totale de signe d'accord sont significativement plus fréquents en côte à côte (8,77%) qu'à distance (1,77%;  $\chi^2=9,7594$ ;  $p=0,001$ ). Alors qu'inversement, l'auditeur donne son accord de manière explicite plus souvent à distance (65,04%) qu'en côte à côte (52,19%;  $\chi^2=7,2053$ ,  $p<0,005$ ). Les auditeurs distants semblent ainsi éprouver plus souvent le besoin de montrer leur compréhension de manière explicite que les sujets en présence physique, qui se permettent un peu plus souvent de montrer leur accord de manière non verbale.

### Effet de la désignation des billes

Les sujets distants désignent assez souvent les billes (29,4% des coordinations) alors qu'en côte à côte ils ne pointent presque jamais les billes du doigt (1,3% des coordinations,  $p<0,001$ ). En moyenne, les sujets distants verbalisent significativement moins de descriptifs quand ils désignent au moins une bille (1,03) que quand ils n'en désignent aucune (1,83;  $F=8,5529$ ;  $p<0,001$ ). Plus

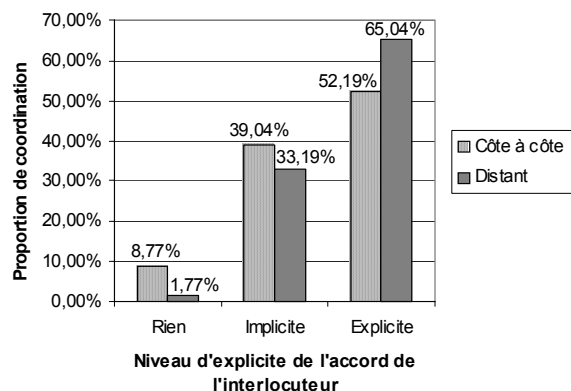


Figure 5 : Niveau d'explicité de l'accord

les locuteurs distants désignent de billes sur l'écran de l'autre, moins ils verbalisent de descriptifs des billes ( $t=11,7168$ ;  $p=0,0271$ , test de corrélation de Pearson). De plus, les locuteurs font un peu moins souvent l'effort de localiser les billes quand ils les désignent avec l'outil (50,00% de localisation dans ces cas) que quand ils ne les désignent pas (65,22% de coordinations avec localisation,  $\chi^2=3,0087$ ;  $p=0,04$ ). Par ailleurs, la désignation de deux billes augmente significativement le taux d'accord implicite (57,89%) par opposition aux cas où aucune (33,33 % d'accord implicite) ou une seule bille (22,92%) est particularisée (respectivement,  $\chi^2=3,4436$ ;  $p<0,05$  et  $\chi^2=6,0486$ ;  $p<0,01$ ). Par contre, la désignation n'a aucun effet significatif sur le nombre d'interventions ni sur la compréhension. L'hypothèse selon laquelle la désignation aide le locuteur distant à indiquer le conflit et l'auditeur à comprendre est ainsi vérifiée pour certaines des variables.

Pourquoi la performance de dialogue en côte à côte est supérieure alors que les sujets ne désignent presque jamais les billes par un geste ? L'analyse des verbalisations autres que les coordinations de quelques binômes de sujets (4 en côte à côte et 3 à distance) révèle qu'à part certaines plaisanteries toutes ces verbalisations peuvent contribuer à construire un contexte partagé utile aux coordinations. Mais à distance il s'agit de monologues hors téléphone donc inutiles. Par ailleurs les locuteurs en côte à côte contrôlent plus souvent le contexte de leur partenaire (11,56 % des coordinations) que ceux à distance (3,51% des coordinations distantes,  $\chi^2=9,4352$ ;  $p=0,002$ ). Ce contrôle se fait par un regard vers l'écran du partenaire (88,46% des cas de contrôle), jamais par un regard vers le visage du partenaire et rarement verbalement en posant des questions du style « Tu vois, la bille bleue, en bas ? (3,85% des contrôles). De plus, il se peut que la quantité de contrôles visuels soit sous-estimée : on peut percevoir beaucoup de chose de son partenaire et de son écran sans bouger la tête. Or la caméra ne permettait pas de voir le regard des sujets. Ce contrôle visuel du contexte de l'autre pourrait compenser l'absence de geste et expliquer la possibilité pour l'auditeur de donner son accord implicitement.

Nous avons recherché les paramètres pouvant influencer l'usage de la désignation des billes. La charge de travail n'a pas d'effet significatif. Par contre, plus les billes en conflit sont éloignées les unes des autres, plus les sujets

utilisent la désignation des billes. Cela peut s'expliquer par le fait qu'en général, moins on a de temps, moins on fait l'effort de s'adapter à l'autre. Au-delà de 22 cases d'écart, le nombre de désignations diminue à nouveau : à cette distance au moins l'une d'entre elles n'est pas encore dans la zone partagée. La confiance en ce qui est partagé intervient donc peut être : on hésite à utiliser la désignation si l'on n'est pas sûr que l'autre voit la bille.

Nous avons observé par ailleurs 11 coordinations sans aucune communication verbale dans lesquelles les sujets se contentent de désigner les billes du conflit. La moitié de ces cas consiste à signaler un problème chez le partenaire. Le fait que 4 de ces 6 problèmes soient finalement résolus par le partenaire laisse penser que la désignation des billes a été vue et comprise. Un des deux échecs est suivi d'un appel téléphonique pour s'excuser d'avoir vu trop tard le problème. Il y aurait donc une véritable communication non verbale à distance, via le partage de vision de l'interface et de l'outil de désignation. Les autres cas sont des conflits sur les portes : généralement le sujet désigne sa propre bille et la détourne aussitôt pour résoudre le conflit. Il n'a pas les moyens de savoir si l'autre a vu et compris. Dans ces cas la résolution du problème semble primer sur l'obtention d'une compréhension mutuelle du problème.

## CONCLUSION ET PERSPECTIVES

### Conclusion

Contrairement à notre première hypothèse, la distance n'a pas affecté la performance de la tâche collective, même en période de forte charge de travail. Par contre, ce maintien de la performance à distance a exigé des efforts de communication supplémentaires (un peu plus de descriptifs, de coordinations longues et d'accords explicites) et cela malgré l'usage plus fréquent d'actes déictiques. L'hypothèse selon laquelle, malgré les outils de partage de contexte, la communication n'est pas aussi optimale qu'en co-présence physique est ainsi vérifiée. Nous notons cependant que quand les billes sont particularisées, les sujets distants utilisent moins de descriptifs et d'indices de localisation et expriment leur accord de manière plus implicite. Ces résultats confortent la troisième hypothèse selon laquelle la désignation des billes améliore les communications distantes. La désignation des billes a même permis des coordinations sans verbalisation. Cet outil a cependant des limites : il demande un effort trop important dans les cas de pression temporelle forte, il est moins utile quand il n'est pas possible de voir réciproquement la même chose et enfin certains sujets ont estimé que la couleur n'était pas assez perceptible.

Nous montrons donc que la possibilité de désigner des objets dans un espace informatique partagé favorise la communication sans atteindre les performances maximales permises en côte à côte. Le manque de mise à jour du contexte partagé par des verbalisations continues et le

manque d'un contrôle visuel mutuel de l'autre, même périphérique, peut expliquer cette baisse de performance. L'impossibilité de se voir peut en particulier expliquer que le processus d'accord soit plus explicite, fait révélateur d'un manque de certitude de partager la même information.

### Conception d'outils de partage de contexte

Pour favoriser l'assurance de partager la connaissance d'un objet montré à distance, plusieurs pistes sont envisageables :

- fournir un outil permettant de délimiter mutuellement la zone partagée et de vérifier que l'objet montré est dans cette zone ;
- obliger le partenaire distant à entrer en communication en sélectionnant et en acquittant l'objet montré ;
- permettre à la personne distante de montrer qu'elle a vu cet objet, par exemple en partageant la vue du curseur de la souris et en lui permettant à son tour de pointer sur cette information ;
- signaler où se pose le regard de l'autre grâce à des outils de capture et suivi du regard ;
- permettre de rester en communication continue tout le long de la tâche par des audio ou video space.

### Application au contrôle aérien : Duophone

Les contrôleurs aériens ont souvent du mal à localiser les avions dont on leur parle et l'absence de partage de la connaissance de sa route gêne la communication [6]. Le prototype Duophone propose à l'appelant de désigner l'avion et d'afficher automatiquement sa trajectoire prévue sur l'image radar distante ; il affiche aussi les limites de ce que le contrôleur appelé voit sur son image radar, pour s'assurer que l'ensemble des éléments du problème est dans la zone partagée. Dans le cas où l'avion désigné se trouve en dehors de l'image radar de l'appelé, ce dernier peut appuyer sur une touche qui va automatiquement régler son écran de manière à visualiser l'avion particularisé sans perdre la vision des autres avions. Un nouvel appui sur cette touche ou le fait de raccrocher le téléphone ramène au réglage initial. Une option envisageable est d'obliger le contrôleur à prendre l'appel téléphonique en cliquant sur l'avion particularisé, donnant ainsi l'assurance à son partenaire qu'il a bien vu l'avion.

### REMERCIEMENTS

Nous remercions Cédric Mariot et l'équipe PII du CENA sans qui les outils de cette expérimentation n'existeraient pas, Karim Mehadhebi et Pascal Lezaud pour leurs conseils sur les statistiques.

### BIBLIOGRAPHIE

1. Brennan, S., Zelinsky, G., & Gerrig, R. *I.T.R.: Contributions of eye Movements and Shared Attention to Collaborate Tasks*. Disponible à l'adresse : <http://www.psychology.sunysb.edu/psychology/personnel/Susan.htm>.
2. Chatty, S., Girard, P., & Sire, S. Vers un support multimédia au collectif synchrone. *Technique et Sciences Informatiques*, 15(9), 1996, pp 1259-1286.
3. Clark, H. & Marshall R. C. Definite reference and mutual knowledge. *Elements of discourse understanding*. Jushi, S. & Weber (eds.), 1981, pp. 10-63.
4. Dey, A. K., Abowd, G. D., & Salber, D. A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Human-Computer Interaction*, 16, 2001, pp. 97-166.
5. Dourish, P., & Bellotti, V. Awareness and Coordination in Shared Workspaces. In *Proceedings of Conf. CSCW'92* (October 31- November 4, 1992, Toronto, Canada), ACM Press, 1992, pp. 107-114.
6. Dumazeau, C. *Analyse empirique des communications distantes dans le cadre du contrôle aérien*. DEA, CNAM, Université Paris 5, Université Paris 8 et Université Toulouse Le Mirail, Paris, 2001.
7. Dusire, S. *Naviguer dans un espace verbal : la construction de la conscience de la situation*. Thèse de doctorat en ergonomie, Conservatoire National des Arts et Métiers, Paris, 2000.
8. Foulon-Molenda, S. A-t-on besoin de se voir pour coopérer ? Contribution théorique issue de la psycholinguistique. *Le Travail Humain*, 63(2), 2000, pp. 97-120.
9. Gutwin, C., M. Roseman, & Greenberg, S. Usability Study of Awareness Widgets in a Shared Workspace Groupware System. In *Proceedings of ACM Conf. on Computer Supported Cooperative Work, CSCW'96*, ACM Press, Cambridge, 1996, pp.258-267.
10. Heath, C., & Luff, P. Activité distribuée et organisation de l'interaction. *Sociologie du travail*, 4, 1994, pp. 523-545.
11. Karsenty L. & Pavard B. Différents niveaux d'analyse du contexte dans l'étude ergonomique du travail collectif. *Réseaux*, 85, 1997, pp. 73-99.
12. Kraut, R. E., Fussell, S. R., & Siegel, J. Visual Information as a Conversational Resource in Collaborative Physical Tasks. *Human-Computer Interaction*, 18, 2003, pp. 13-49.
13. Sperber, D., & Wilson, D. *La pertinence. Communication et cognition*. (A. Gerschenfeld & D. Sperber, Trans.), Les éditions de minuit, 1989.



14. Wilkes-Gibbs, R., & Clark, H. Coordinating beliefs in Conversation. *Journal of Memory and Language*,

31, 1992, pp. 183-19