
Situation Awareness en conception système

Xavier Chalandon

Dassault Aviation (DGT / DTS / IHS)
78, quai Marcel Dassault, 92552 St Cloud
mail : xavier.chalandon@dassault-aviation.fr

RESUME

Cette communication présente deux études internes de la société DASSAULT AVIATION centrées sur le concept de Situation Awareness (SA) et son utilisation possible en conception de systèmes d'avions d'armes. Après avoir rappelé les différentes approches de la littérature, la SA est inscrite dans un processus de prise de conscience en soulignant deux de ses caractéristiques essentielles : sa finalisation et sa fonction anticipatrice (« être devant l'avion »). Un passage vers la conception système est proposé en faisant l'hypothèse que certaines contraintes de l'environnement sont en correspondance avec les invariants opératoires de schèmes de gestion du risque. Il convient de valider cette hypothèse puis de formaliser et d'explicitier ces contraintes au niveau système et interface afin d'améliorer le processus d'adaptation des opérateurs. Dans le cadre de notre application, le « temps de contact à la cible » est identifié comme candidat potentiel pour le développement d'une aide à la gestion du risque tactique.

MOTS-CLES

environnement dynamique; situation awareness; risque; adaptation; temps de collision

1 INTRODUCTION

Depuis plus de vingt ans, la communauté de la psychologie aéronautique présente le concept de *Situation Awareness*¹ (SA) comme un élément clé des processus cognitifs en environnement dynamique. La littérature est assez homogène sur les fonctions principales associées à la SA : donner cohérence aux événements externes, créer des attentes et orienter la prise d'information, servir d'ancrage aux décisions et actions ultérieures, permettre l'anticipation des évolutions de la situation et des effets d'actions (Endsley, 1995 ; Klein, 1995 ; Grau et al., 1996). De plus les différentes approches de la SA sont unanimes sur sa modalité consciente ou du moins conscientisable (Endsley, 1995 ; Sarter & Woods, 1991 ; Smith & Hancock, 1995). Ces traits généraux sont cohérents de deux caractéristiques de la SA soulignées par les définitions empiriques : sa finalisation (« savoir ce qui se passe pour savoir quoi faire ») et sa fonction anticipatrice (« être devant l'avion ») qui sont révélatrices de la hantise des équipages de ne pas disposer du temps nécessaire pour anticiper sur les événements et donc d'échouer dans l'ajustement de leur activité au dynamisme de la situation (Valot, 1996). Par contre des divergences sensibles existent dans la littérature quant à la nature de la SA et aux processus cognitifs impliqués. La SA apparaît comme un concept « mou », tantôt centré sur la perception immédiate de la situation, tantôt censé embrasser tout le champ de la cognition en environnement dynamique (Chalandon, 1998).

Dans un premier temps, nous nous proposons ici d'établir une taxonomie des différentes approches de la SA puis de tenter une unification au travers des mécanismes de prise de conscience². Une étude expérimentale brièvement décrite viendra appuyer notre propos.

Dans un deuxième temps, nous aborderons la problématique de la conception système, i.e. du passage d'un modèle cadre de la SA à un modèle formel inscriptible dans les logiques système. L'hypothèse avancée est que la notion de « temps de contact à la cible » est un candidat potentiel pour le développement d'une aide à la gestion du risque tactique.

¹ La terminologie anglo-saxonne est maintenue car la traduction en français de l'expression *situation awareness* n'est pas neutre vis-à-vis de la définition du concept signifié.

² Étude limitée au champ de la cognition individuelle. Aborder la problématique de la SA collective nécessiterait d'élargir notre cadre au domaine de l'interaction et de la coopération Homme-Homme.

2 DE LA PERCEPTION A LA CONSCIENCE DE LA SITUATION

Les principales approches de la SA peuvent être regroupées en quatre familles de modèles schématiquement décrits ci-après.

2.1 L'approche « linéaire »

Cette approche définit la SA comme le produit de trois activités hiérarchiquement et temporellement organisées : « *la perception des éléments de l'environnement dans un volume de temps et d'espace, la compréhension de leur signification et la projection de leur état dans le futur proche* » (Endsley, 1995). Cette approche envisage une SA « contemplative » intervenant de façon indépendante et sérielle en amont de la prise de décision et n'intégrant pas les possibilités d'action. Cette position théorique est également la base d'une approche d'évaluation normative où la SA, en tant qu'état instantané de connaissances sur le monde, doit être la plus précise et complète possible, au pire par rapport à un référentiel absolu, au mieux par rapport à un référentiel constitué de l'agrégat d'expertises.

Cette approche est assez paradoxale car elle est centrée sur le sujet quant à la description de mécanismes cognitifs sous-tendant la SA mais elle nie l'autonomie du sujet en spécifiant la production de ces mécanismes et donc les modes opératoires. De plus, pour reprendre l'expression d'Amalberti (1996), le modèle sériel perception-compréhension-anticipation-décision-action, qui suppose une dépendance à sens unique de l'amont vers l'aval, est d'une « *désarmante naïveté* » et il existe un parallélisme et des interdépendances qui sont amplifiés par l'expertise (Chalandon, 1998).

2.2 L'approche « décisionniste » ou « réactive »

En réaction à l'approche linéaire (Prince & Salas, 1997), l'approche décisionniste considère la reconnaissance implicite de patterns significatifs pour l'action comme une solution essentielle au problème de la synchronisation activité-environnement dans les situations à forte dynamique (Klein, 1995). La situation est donc imposée à l'opérateur qui la « découvre » et dont l'expertise se traduit par une reconnaissance finalisée de cette situation. Ce « coup d'œil » de l'expert qui l'oriente vers un cours d'action consiste en un appariement plus ou moins explicite entre les événements situationnels et les connaissances fonctionnelles possédées en mémoire. Ces auteurs sont assez prudents sur les notions de représentation et de conscience et insistent sur le rôle de la perception dans la cognition, se rapprochant ainsi de la notion d'affordance³ développée par Gibson (1979). Dans ce cadre, l'évaluation consciente (la compréhension) de la situation n'intervient qu'en cas de résistance du réel et est centrée sur un « problème » (Orasanu & Fisher, 1997).

Devant la multiplicité des significations possibles et des affordances offertes par la situation, un des problèmes rencontrés est celui de « *la prise en compte de la signification dans la saillance perceptive* » (Dubois et al., 1993), phénomène qui ne semble s'expliquer en dehors d'une fonction anticipatrice assurée par des construits internes tels que les schémas ou les schèmes (Neisser, 1976 ; Denecker, 1999).

2.3 L'approche « centrée sur la régulation »

Pour cette approche, issue de l'ergonomie centrée sur l'activité, le processus de compréhension consiste en l'ajustement permanent d'une représentation fonctionnelle selon l'historique de la situation et des buts à atteindre (Hoc & Amalberti, 1994 ; Sarter & Woods, 1991 ; Prince & Salas, 1997). Cette activité de révision de la représentation dite « *occurrente* » fonctionne en temps partagé avec des rendez-vous de synchronisation par rapport à la dynamique de la situation : la conduite synchronique est guidée par la partie de la représentation jugée cohérente et les ajustements ou recompositions s'effectuent en parallèle à des niveaux plus élevés d'abstraction (Amalberti, 1996). Un des ressorts essentiels de ce modèle est l'acceptation par l'opérateur de « *ne pas comprendre* » (au sens d'une compréhension optimale) car les ressources qu'il devrait affecter à cette activité sont incompatibles avec la dynamique de la tâche. Cette compréhension minimale (ou du moins sub-optimale) est le

³ Le concept d'affordance peut se traduire comme une propriété de l'environnement révélant au sujet les potentialités d'interaction avec cet environnement. L'affordance permet une adaptation immédiate de l'action basée sur la perception, le sujet ayant appris, par développement ou apprentissage, à « saisir » ces informations sans les traiter.

produit d'un compromis cognitif contrôlé par les méta-connaissances du sujet (Amalberti, 2001). Une double dualité est soulevée par ce modèle :

- La situation est un système dual pilote-tâche dont les éléments se co-déterminent.
- Le produit et le processus de l'activité de compréhension se co-déterminent et la SA ne peut s'envisager en dehors des processus qui la sous-tendent.

2.4 L'approche « centrée sur les contraintes »

La problématique de la SA est ici interprétée comme imposant un nécessaire changement de paradigme de la cognition centrée sur le sujet vers l'étude du couplage dynamique entre l'homme et l'environnement. Cette approche souligne la nécessité pour l'opérateur d'avoir une conscience explicite des contraintes et ressources de l'environnement et définit la SA comme «*une conscience dirigée vers l'externe*» (Smith & Hancock, 1995). Refusant d'identifier la SA à un produit ou à une étape des processus cognitifs, la SA est identifiée à un invariant structurant la connaissance et le comportement permettant de répondre aux exigences normatives de l'environnement socio-technique (Flach, 1995).

Si cette approche est explicitement normative, elle est très différente de la prescription de l'approche linéaire en ce sens qu'elle ne cherche pas à nier l'autonomie de l'opérateur en spécifiant la totalité de la SA-produit mais à « l'éclairer » en explicitant les contraintes de l'environnement nécessaires à l'action efficace (Flach & Rasmussen, 2000). À ce titre, Vincente (1999) parle d'approche «*formatrice* ».

2.5 Tentative d'unification

Il nous semble possible d'unifier ces différentes approches en inscrivant la SA dans un processus de prise de conscience (Piaget, 1974 ; Vermersch, 1994). La psychologie piagétienne envisage la connaissance comme le résultat d'un processus d'adaptation qui tend à l'équilibre cognitif entre un mécanisme d'assimilation (intégration de l'environnement aux structures existantes de l'activité) et un mécanisme d'accommodation (modification de ces structures pour assimilation). Les abstractions successives de la prise de conscience modifient le rapport du sujet au monde en accroissant le champ spatio-temporel de l'adaptation. La SA s'inscrit pleinement dans ce processus d'abstraction qui permet l'assimilation de réalités de plus en plus éloignées de l'action propre et une accommodation de cette action au possible considéré. Produit de l'engagement de l'opérateur dans la situation passée et des activités d'anticipation, la SA permet la réduction du champ structurellement possible (situation externe aux combinaisons quasi-infinies) au champ matériellement possible (situation intériorisée que l'opérateur estime pouvoir contrôler), elle donne sens aux états présents du réel par pré-construction de cadres d'assimilation et elle est le support de l'engagement futur. De cette approche découlent certaines caractéristiques de la SA déjà identifiées par les approches que nous avons appelées « réactive » et « centrée sur la régulation » :

- L'accommodation et l'élévation du niveau d'abstraction sont déclenchés par une résistance du réel à l'assimilation et par un échec de l'accommodation à un certain niveau d'abstraction. La SA est associée aux obstacles rencontrés ou anticipés par le sujet dans la réalisation de ses objectifs.
- L'équilibration est un processus dynamique dans lequel coexistent plusieurs niveaux d'abstraction, ce qui rend délicat la traduction de l'*awareness* par la « conscience » au sens strict de pensée réflexive thématisée. La SA serait une représentation dynamique composée de multiples niveaux d'abstraction, dont les éléments sont conscientisables mais non effectivement conscients à tout instant.
- Les modes de fonctionnement les plus abstraits étant les plus volontaires et les plus complexes, ils sont les plus coûteux à mettre en jeu. De plus la prise de conscience est toujours en décalage par rapport au vécu : il y a une dissymétrie temporelle entre le comprendre et le réussir. La (re)construction de la SA est donc à envisager sur un mode diachronique.
- Toute action, même la plus abstraite, la plus conceptualisée du fait des connaissances et des objectifs dont elle suppose la maîtrise, contient des connaissances pré-réfléchies ou non conscientisées. Il y a une part de l'activité irréductible à la SA.

L'approche centrée sur les contraintes a elle-même toute sa place dans ce modèle cadre. En effet, Vergnaud (1985) souligne qu'une des composantes d'un schème (les invariants opératoires) doit être en adéquation avec le réel afin d'assurer l'efficacité des autres composantes (inférences, règles d'action, prédiction). Placés dans ce cadre, Flach & Rasmussen (2000) et Vincente (1999) cherchent à réduire le champ des structurellement possibles en énonçant une condition **nécessaire** à l'efficacité de l'action en termes de performance et sécurité : certains invariants de l'environnement **doivent** être en correspondance avec des invariants opératoires.

3 ETUDE EXPERIMENTALE

L'instanciation de ce cadre théorique par rapport aux modalités fonctionnelles de sujets concrets a été l'objet d'une étude expérimentale menée consécutivement avec 4 pilotes d'essais d'avion d'armes (Chalandon, 1998).

Centré sur les stratégies adoptées par les pilotes pour la régulation d'une situation tactique, le protocole expérimental a consisté à faire préparer une mission aux opérateurs (1h30, avec le support de la technique de l'information à la demande), à faire réaliser la mission en simulation (40 minutes) puis à restituer la mission en utilisant la technique de l'auto-confrontation (2h00). S'inspirant de l'approche centrée sur la régulation, ce protocole a permis de formaliser des stratégies de gestion du risque planifiées avant la mission et de mettre en évidence des indices contextuels amenant les opérateurs à réviser, voire à abandonner, les stratégies prévues.

Pendant les phases de préparation, les pilotes ont été unanimes sur le fait que la réussite d'une telle mission repose avant tout sur la discrétion vis-à-vis des moyens adverses. Cette discrétion est pour partie construite à l'aide d'un plan de vol très basse altitude, ce qui diminue l'intervisibilité entre l'avion et les acteurs ennemis. Les pilotes n'envisagent un abandon de cette stratégie qu'en cas de diagnostic sûr d'un danger immédiat (information de présomption de tir donnée par les capteurs de bord), ce qui est symptomatique de la compréhension optimale revendiquée par l'approche linéaire de la SA.

Pendant la mission, les pilotes ont parfois abandonné cette stratégie de discrétion sur la base de diagnostics qualifiés d'incertains en préparation de mission (i.e. absence de présomption de tir), ce qui est symptomatique de l'approche centrée sur la régulation (compréhension sub-optimale). Les cas rencontrés correspondaient à des actions de diagnostic (compréhension par l'action) et/ou à des actions visant à éviter un rapport de force anticipé comme défavorable dans le cas d'un engagement lui-même anticipé (compréhension par simulation de l'action). L'analyse d'activité a ainsi mis en évidence que les pilotes régulent la situation tactique en fonction de la situation relative anticipée, soulignant l'importance de la représentation de l'action comme préalable à la compréhension.

Se rapprochant ensuite de l'approche centrée sur les contraintes, la SA peut alors être considérée comme un espace d'actions possibles dont les frontières sont déterminées par des facteurs géométriques et cinématiques (position et trajectoires des avions et leurs capacités d'évolution), des facteurs système (capacités de détection des différents acteurs tactiques, leurs capacités offensives et défensives), des facteurs procéduraux (règles d'engagement, heures de passage à respecter, etc.), des facteurs organisationnels (structure des dispositifs aériens), etc. A de très rares exceptions près, ces facteurs acquièrent une signification non pas sur un mode absolu mais sur un mode relatif (différence d'altitude ou d'armement avec les ennemis par exemple), ce qui permet de réduire la complexité externe par sélection et interprétation. Cet espace intègre le temps car non seulement il est lui-même dynamique (évolution continue au cours de la mission) mais ses « déformées » possibles sont également anticipées sur des emplans temporels variables.

4 VERS LA CONCEPTION SYSTEME : QUEL TYPE D'ASSISTANCE ?

À partir des sorties de cette première étude, l'ambition industrielle est de spécifier sous 2 ans un système d'assistance à la gestion du risque tactique. Deux types d'assistance sont a priori envisageables :

- Des aides dites « intelligentes », reposant le plus souvent sur des systèmes experts et fonctionnant en mode anticipatif (proposition de solution) ou réactif (proposition d'alternative) (Amalberti, 1996). Ce type d'assistance doit posséder non seulement un savoir-faire (une capacité d'expertise sur la

situation) mais également un « savoir-coopérer » (Hoc, 1998), nécessitant un modèle de l'opérateur à assister.

- Des assistances non coopératives (i.e. ne gérant pas d'interférence avec l'opérateur (Hoc, 1998) qui peuvent soutenir la mise en œuvre de processus cognitifs (fonctionnalité aide-mémoire par exemple) ou rendre compte d'états ou de contraintes de l'environnement (Vincente, 1999).

Les contraintes industrielles de notre projet (budget, délais) interdisent le recours aux solutions type aide intelligente. Plus fondamentalement, ce type de solution pose le problème de la robustesse d'un savoir-faire système face aux variabilités non prévisibles en conception. Moins ambitieuses, les assistances non coopératives, en particulier celles basées sur les contraintes, semblent présenter l'avantage de ne pas fermer la conception sur des procédures, quand bien même proviendraient-elles d'analyses d'activité réalisées sur le terrain, mais d'envisager le système comme une ressource facilitant la gestion des variabilités par les opérateurs. Le pari est donc fait d'une plus grande stabilité de certaines contraintes de l'environnement que des règles d'action.

5 LE TEMPS DE COLLISION : SOURCE D'UN INVARIANT OPERATOIRE ?

Une des ambitions que nous poursuivons actuellement est la formalisation des facteurs de « frontière » de l'espace des actions matériellement possibles (cf. § 3) afin de les expliciter au niveau système et interface pour faciliter le processus d'adaptations des opérateurs. Ces facteurs sont assez hétérogènes et variables et les hypothèses que nous formulons sont que certains de ces facteurs (en particulier les facteurs géométriques, cinématiques et certains facteurs système des capacités offensives et défensives) sont intégrables dans une variable unique qui elle-même correspond à un invariant opératoire des schèmes de gestion du risque tactique. Cette variable « candidate » est dérivée de la notion de temps de collision.

5.1 Temps de collision : (brève) revue de la littérature

5.1.1 La vision écologique et le facteur TAU

Un des éléments de l'approche écologique développée par James Gibson (1979) est la thèse de la « perception directe » (ou vision écologique) qui soutient que le flux optique est structuré de manière corrélative à l'environnement et que cette structuration géométrique dynamique spécifie directement les propriétés de l'environnement, y compris les affordances.

Dans le cadre de cette approche, Lee (1980) a montré que l'estimation du temps de contact avec une cible ne nécessite pas le calcul « cognitif » ($\text{Distance} / \text{Vitesse}$) mais est spécifiée par une variable baptisée TAU, directement disponible dans le flux optique (l'inverse de la variation relative de la taille angulaire de la cible). Lee et ses collègues ont par la suite montré l'implication du facteur TAU dans la régulation de l'action pour un certain nombre d'activités d'évitement d'obstacle ou d'interception de mobile (applications à la psychologie animale, à la psychologie du sport, à la psychologie de la conduite automobile, etc. (Bruce & Green, 1993). Le TAU a maintenant des applications en ingénierie, par exemple le développement d'aide au pilotage (Padfield et al., 2001) et la robotique mobile (Duchon et al., 1998).

Bien que des travaux remettent en cause le TAU comme unique indice visuel du temps de contact (Cavallo, 1989 ; Tresilian, 1999), l'importance du temps à la cible comme facteur de régulation de l'action semble être confirmée pour de multiples activités impliquant une vision non médiatisée.

5.1.2 Temps de collision en aéronautique civile

Parallèlement au mouvement écologique, la notion de temps de collision a été développée en aéronautique civile (Morrel, 1958) et sert à l'élaboration des alarmes des systèmes embarqués d'anticollision avec le sol ou avec le reste du trafic aérien.⁴

⁴ Systèmes GPWS (Ground Proximity Warning System) et TCAS (Traffic alert and Collision Avoidance System). Le système TCAS détecte des conflits potentiels et « propose » des évolutions de trajectoire pour les résoudre. Ces systèmes prennent également en compte des valeurs minimales de distance pour la génération d'alarmes.

Au cours d'une étude sur la détection de conflit entre avions réalisée avec 10 pilotes de lignes sur simulateur de Boeing 757, Smith et al. (1997) ont mis en évidence les points suivants :

- Au delà de la distance réglementaire minimum de séparation entre avions, le temps de collision est un invariant structurant le choix de stratégie des pilotes entre ne « rien faire » et « modifier la trajectoire pour gérer en anticipation un conflit potentiel ». Il est à noter que ce temps de collision n'était pas indiqué dans l'interface utilisée au cours de cette étude et donc que les pilotes « extraient » cette information à partir des informations de trafic qui leur sont présentées.
- La valeur seuil de déclenchement de ces stratégies d'évitement était d'environ 3 minutes, ce qui est bien supérieur aux seuils d'alarme des systèmes embarqués gérant l'anticollision air (de 15 à 50 secondes selon l'altitude et le type d'alarme). Cette constatation est symptomatique de la volonté des pilotes « d'être devant l'avion », de gérer les conflits par anticipation et non pas sur un mode réactif.

Dans le cadre de cette application, Smith & Hancock (1995) définissent opérationnellement la SA par un « *espace de risque* » constitué par la distance entre avions et leur vitesse de rapprochement : la position d'un avion dans cet espace indique le temps de collision et contraint la stratégie d'interaction à adopter.

5.2 Application en aéronautique militaire : une hypothèse à tester

L'analyse des données recueillies pendant la première étude ainsi que la (brève) revue de la littérature ci-dessus (qui devrait être complétée par d'autres domaines d'application comme le contrôle aérien par exemple, (Morineau, 2000) nous amène à faire l'hypothèse qu'un invariant opératoire de la gestion du risque tactique est associé à la notion de temps de collision. L'approche de l'aéronautique civile n'est cependant pas applicable directement à l'aéronautique militaire pour les raisons suivantes :

- L'important n'est pas le temps de collision entre mobiles mais le « temps d'interception », i.e. le temps de rentrée dans le domaine de tir des autres acteurs. Il convient donc de prendre en compte dans les calculs cette enveloppe dont les 3 dimensions dépendent des altitudes, des vitesses, etc.
- Le temps « civil » est estimé pour une situation courante. Or les avions de combat se distinguent par leurs capacités de manœuvrabilité et les tactiques aériennes induisent des fortes évolutions des avions dans des temps très courts. Il convient donc de prendre des hypothèses un peu plus « paranoïaques » que la situation instantanée et de parler de « Temps d'Interception Potentielle » (TIP).

Des développements sont actuellement en cours pour prendre en compte ces spécificités. Dans un deuxième temps, l'hypothèse du TIP en tant que déterminant de l'activité de régulation tactique sera testée expérimentalement. Si l'hypothèse est confirmée, l'utilisation du TIP sera expérimentée aussi bien pour la hiérarchisation instantanée des acteurs tactiques que pour la hiérarchisation du risque externe associé aux routes possibles de l'avion (fonctionnalité de type « what if »). Ce dernier point est particulièrement important car il doit permettre à l'équipage d'anticiper le niveau de risque pour la suite de la mission (les plans de vol des missions étant rarement rectilignes) ou de l'aider à trouver une solution en cas de nécessaire déroutement.

6 DISCUSSION

En évacuant en partie le contexte et le sujet, notre approche présente certaines limites tant du point de vue des contraintes externes qu'internes :

- Un grand nombre de données externes, tout aussi déterminantes pour la gestion du risque, ne sont pas prises en compte par le TIP : aussi bien des données instantanées (quantité de carburant restant, états des équipiers, données contre-mesures, etc.) que le contexte de fond (finalité de la mission, règles d'engagement, structure des dispositifs amis et ennemis, règles de coopération au sein de la patrouille, etc.). En particulier, l'étude expérimentale avait suggéré que la gestion du risque était distribuée dans le dispositif et que les pilotes font certaines hypothèses concernant la distribution de la SA au sein du dispositif adverse (par exemple, interconnexion des acteurs ennemis) et donc que la SA d'un pilote intègre une confrontation des SA collectives du dispositif ami et du dispositif ennemi.
- Bien que l'élaboration du TIP soit l'objet d'un travail de conception en équipe intégrée Pilote-Ergonome-Ingénieur, le sujet est absent du TIP, aussi bien au niveau de ses objectifs, de sa

planification et de son histoire d'engagement qu'au niveau de ses compétences. En particulier, contraindre le champ des actions structurellement possibles ne renseigne en rien sur les compétences nécessaires aux trajectoires d'action efficace restant possibles (Grant, 1995).

La suite de l'étude s'attachera aux conséquences théoriques de ces limites et à déterminer s'il est possible d'intégrer certains de ces paramètres dans l'assistance envisagée. Sans reprendre l'argumentaire du § 4, nous espérons que la structuration de données que nous essayons de réaliser soit déjà un facteur d'amélioration de l'efficacité du couple Homme-Système et puisse être considérée comme une brique de base d'un système d'aide plus ambitieux à venir.

7 BIBLIOGRAPHIE

- Amalberti, R. (1996) - *La conduite des systèmes à risques*. Paris : Presses Universitaires de France.
- Amalberti, R. (2001) – La maîtrise des situations dynamiques. *Psychologie Française*, 46(2), 107-118.
- Bruce, V. & Green, P.R. (1993) - *La perception visuelle - Physiologie, psychologie et écologie*. Grenoble : Presses Universitaires de Grenoble.
- Cavallo, V. (1989) - Le rôle des informations visuelles dans l'anticipation d'une collision. *Psychologie Française*, 34(1), 33-40.
- Chalandon, X. (1998) - *Situation Awareness : de la perception à la conscience de la situation - Etude de la régulation d'une situation tactique par des pilotes de combat*". Mémoire de DEA d'Ergonomie. EPHE.
- Denecker, P. (1999) - Les composantes symboliques et subsymboliques de l'anticipation dans la gestion des situations dynamiques. *Le Travail Humain*, 62(4), 363-385.
- Dubois, D., Fleury, D. & Mazet, C. (1993) - Représentations catégorielles : perception et/ou action ?. In A. Weill-Fassina, P. Rabardel & D. Dubois (Eds.) *Représentations pour l'action*. Toulouse : Octarès.
- Duchon, A.P., Warren, W.H. & Kaelbling, L.P. (2001) – Ecological robotics. *Adaptive Behavior*, 6(3/4), 473-507.
- Endsley, M.R. (1995) - Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37 (1), 32-64.
- Flach, J.M. (1995) - Situation Awareness : proceed with caution. *Human Factors*, 37 (1), 149-157.
- Flach, J.M. & Rasmussen, J. (2000) - Cognitive engineering : designing for situation awareness. In N.B. Sarter & R. Amalbert (Eds.) *Cognitive engineering in the aviation domain*. New Jersey : Lawrence Erlbaum Associates.
- Gibson, J.J. (1979) - *The ecological approach to visual perception*. Boston : Houghton-Mifflin.
- Grant, S. (1995) - Safety systems and cognitive models. In *Proc. 5th International Conference on Human-Machine Interaction and Artificial Intelligence in Aerospace*, Toulouse : EURISCO.
- Grau, J.-Y., Menu, J.-P. & Amalberti, R. (1996) - La conscience de la situation en aéronautique de combat. In AGARD conference proceedings 575 *La conscience de la situation : les limitations et l'amélioration en environnement aéronautique*. Neuilly-sur-Seine : AGARD.
- Hoc, J.-M. (1998) – L'ergonomie cognitive : un compromis nécessaire entre des approches centrées sur la machine et des approches centrées sur l'homme. In Actes du colloque *Recherche et Ergonomie*. Toulouse : SELF.
- Hoc, J.-M. & Amalberti, R. (1994) - Diagnostic et prise de décision dans les situations dynamiques. *Psychologie Française*, 39 (2), 177-192.
- Klein, G. A. (1995) - Studying situation awareness in the context of decision-making incidents. In D.J. Garland & M.R. Endsley (Eds.) *Experimental analysis and measurement of situation awareness*. Florida : Embry-Riddle Aeronautical University Press.
- Lee, D.N. (1980) - The optic flow held : The foundation of vision. *Philosophical Transactions of the Royal Society of London, Series B*, 290, 169-179.
- Morineau, T. (2000) - Time-to-collision and action sequencing on aircraft conflicts in air traffic control. In *Proc. 10th European Conference on Cognitive Ergonomics*, Linköping : Sweden.
- Morrel, J.S. (1958) - The mathematics of collision avoidance in the air. *Journal of Navigation*, 11, 18-28
- Neisser, U. (1976) – *Cognition and Reality*. San Francisco : Freeman.

-
- Orasanu, J. & Fisher, U. (1997) - Finding decisions in natural environments : the view from the cockpit. In C.E. Zsombok & G.A. Klein (Eds.) *Naturalistic Decision Making*. New Jersey : Lawrence Erlbaum Associates.
- Padfield, G.D., Lee, D.N. & Bradley, R. (2001) - How do helicopter pilots know when to stop, turn or pull up ?. *American Helicopter 57th annual forum*. Washington DC.
- Piaget, J. (1974) - *La prise de conscience*. Paris : Presses Universitaires de France.
- Prince, C. & Salas, E. (1997) - The role of situation assessment in the conduct of flight and in decision making. In D. Harris (Ed.) *Engineering psychology and cognitive ergonomics, volume one*. Aldershot : Ashgate.
- Sarter, N.B. & Woods, D.D. (1991) - Situation awareness : a critical but ill-defined phenomenon. *The International Journal of Aviation Psychology*, 1(1), 45-57.
- Smith, K. & Hancock, P.A. (1995) - Situation awareness is adaptative, externally directed consciousness. *Human Factors*, 37 (1), 137-148.
- Smith, K., Lewin, J. & Hancock, P.A. (1997) - The invariant that drives conflict detection. In D. Harris (Ed.) *Engineering psychology and cognitive ergonomics, volume one*. Aldershot : Ashgate.
- Tresilian, J.R. (1999) - Visually timed action : time-out for 'tau' ?. *Trends in Cognitive Sciences*, 3(8), 301-310.
- Valot, C. (1996) - Gestion du temps. Gestion du risque (à travers quelques situations aéronautiques). In J.-M. Cellier, V. De Keyser & C. Valot (Eds.) *La gestion du temps dans les environnements dynamiques*. Paris : Presses Universitaires de France.
- Vergnaud, G. (1985) – Concepts et schèmes dans une théorie opératoire de la représentation. *Psychologie Française*, 30(3/4), 245-252.
- Vincente, K. (1999) - *Cognitive Work Analysis*. New Jersey : Lawrence Erlbaum Associates.
- Vermersch, P. (1994) - *L'entretien d'explicitation en formation initiale et en formation continue*. Paris : Editions Sociales Françaises.

Télétravail : contraintes temporelles et ordonnancement de l'action quotidienne

Maunier Laurent, Carreras Ophélie, Valax Marie-Françoise

Laboratoire Travail et Cognition

Université Toulouse-Le-Mirail

Maison de la Recherche, UMR 5551

05, allées Antonio Machado, 31058 Toulouse cedex 9

maunier@univ-tlse2.fr

RESUME

Le télétravail - fondé sur l'usage des NTIC – se caractérise par un mode d'organisation du travail plus flexible. Dans cette étude, des concepts issus des modèles automatiques d'ordonnancement dynamique ont été utilisés pour préciser la notion de flexibilité temporelle. Une technique d'ordonnancement par contraintes temporelles a ainsi été appliquée à des télétravailleurs pendulaires et nomades. Les résultats indiquent que la flexibilité temporelle est plus sensible à la nature des tâches qu'au type de télétravail et que cette flexibilité temporelle évolue sous l'action de différents facteurs. Enfin, nos résultats semblent aller dans le sens d'une utilisation par l'homme de méthodes d'ordonnancement dynamique avec prises de décisions hors-ligne (avant exécution) et surtout en-ligne (pendant l'exécution).

MOTS-CLES

Télétravail, NTIC, ordonnancement, planification, flexibilité.

1 NTIC ET NOUVELLES FORMES DE TRAVAIL : LE TELETRAVAIL

1.1 Télétravail : définition

Le développement des Nouvelles Technologies de l'Information et de la Communication (NTIC) est à l'origine de nouvelles formes d'organisation du travail qui transforment le rapport du travailleur au temps. Par rapport à l'organisation conventionnelle, centralisée, le télétravail ouvre la voie à une organisation virtuelle, décentralisée et fragmentée, qui offre une plus grande latitude vis-à-vis des horaires de travail. Le télétravail (Breton, 1994 ; eWork, 2002) n'est pas un nouveau métier, mais une nouvelle forme de réalisation du travail, qui se caractérise par : (1) la distance du travailleur avec l'endroit où le résultat de son travail est attendu ; (2) l'impossibilité physique pour le donneur d'ordre de surveiller de façon permanente l'exécution de la prestation ; (3) l'usage des technologies de l'information et de la communication.

Cette définition générale recouvre différentes modalités d'exercice du télétravail que l'on peut décrire sur deux dimensions, l'une géographique, l'autre organisationnelle.

D'un point de vue géographique, il s'exerce : (1) au *domicile*, c'est-à-dire dans le lieu de vie familiale ; (2) dans *l'entreprise-mère*, c'est-à-dire dans le lieu de travail traditionnel ; (3) en *centre de proximité*, c'est-à-dire dans un local professionnel qui, situé à proximité du domicile, en zone rurale ou périurbaine, est géré soit par l'entreprise-mère soit par un opérateur public ou privé qui met ce local à disposition d'entreprises ou de travailleurs indépendants.

D'un point de vue organisationnel, le télétravail est le plus souvent : (1) *pendulaire*, lorsque les personnes répartissent leur temps de travail entre l'entreprise-mère, un centre de proximité et/ou le domicile ; (2) *nomade*, lorsque les personnes se déplacent constamment pour offrir leurs services aux clients (commerciaux ou consultants par exemple) ; (3) *spontané*, lorsque les personnes prolongent leur journée ou semaine de travail à la maison.

Le télétravail est souvent associé aux personnes travaillant exclusivement à leur domicile, toutefois, ce mode de télétravail n'est pourtant pas le plus répandu (Bates, & Huws, 2002).

1.2 Télétravail et flexibilité temporelle

Dans une étude réalisée par Stewart (2000), 90% des télétravailleurs déclarent bénéficier de plus de flexibilité temporelle qu'auparavant. Cependant, plusieurs études notent les limites de cette flexibilité. D'une part, les horaires de travail étant parfois précisés dans le contrat de travail, la flexibilité dépend des relations contractuelles entre le télétravailleur et l'entreprise ou le client (Konradt, Schmook et Mälecke, 2000). D'autre part, l'environnement familial peut limiter la flexibilité temporelle lorsque le télétravailleur aménage son temps professionnel en fonction des horaires de travail de l'époux, de l'école, ... (Gottlieb, Kelloway, & Bartham, 1998). Par ailleurs, flexibilité ne signifie pas réactivité totale. Dans la plupart des cas, notamment dans les cas de réussite du télétravail, la flexibilité temporelle tient plus de la modification des contraintes temporelles que de leur annulation. Ceci apparaît notamment dans la catégorisation des modes d'organisation temporelle des télétravailleurs proposée par Stewart (2000). Cet auteur définit trois modes d'organisation : (1) *Conservation du cadre temporel antérieur*. Les sujets travaillent de 9h à 17h, en phase avec les horaires de leurs clients ou de leur hiérarchie, et toute déviation est difficilement tolérée. (2) *Organisation du télétravail "déphasé"*. Les sujets concernés n'arrivent pas à trouver un équilibre entre les temps de la sphère domestique et ceux de la sphère professionnelle. Ils peuvent travailler pendant des journées entières sans tenir compte des repas, des pauses, du temps passé en famille. Ce mode d'organisation se traduit souvent par un échec. (3) *Construction d'un nouveau cadre temporel*. Les sujets, souvent travailleurs indépendants, parviennent à construire de nouveaux cadres temporels qui intègrent loisirs, vie domestique et vie sociale.

Toutefois, les travaux sur le télétravail manquent de rigueur méthodologique (Konradt & al., 2000) ; si les outils, tels qu'entretiens et questionnaires, permettent d'appréhender le vécu du télétravail, ils ne permettent pas une description objective de ses exigences (McCloskey & Igarria, 1998). De plus, la diversité des situations de télétravail, des tâches et des modes d'exercice est négligée.

2 PROBLEMES DE PLANIFICATION-ORDONNANCEMENT

Différentes disciplines font écho aux travaux sur la planification alors que les travaux sur l'ordonnancement, qui leur sont pourtant complémentaires, restent souvent confinés au domaine de la productique et du génie industriel. En automatique, l'organisation des tâches est abordée en termes de résolution de problèmes. Deux types de problèmes sont distingués : la planification et l'ordonnancement. La planification consiste à choisir dans le temps des tâches qui vont permettre à l'agent d'atteindre un but donné. Cependant, lorsque des ressources sont impliquées dans l'exécution, le plan tel que défini ci-dessus n'est pas exécutable directement : l'étape de planification est suivie d'une étape d'ordonnancement. Plus proche de l'exécution, l'ordonnancement précise le plan de tâches.

2.1 Résolution des problèmes de planification-ordonnancement

La résolution de problème de planification-ordonnancement doit faire face à la complexité et à l'incertitude. L'approche prédictive (Morris, Muscettola, Vidal, 2001 ; Huguet, Lopez, Vidal, 2002) construit des plans optimaux, mais applicables uniquement dans des environnements statiques. Plus récemment, la prise en compte de l'incertitude a conduit au développement d'approches dynamiques.

L'approche dynamique de la planification pose comme préalable l'incertitude du futur : les données ne sont pas toutes disponibles en début de raisonnement et les données se modifient constamment pendant la planification et aussi pendant l'exécution. Pour pallier les difficultés liées à l'incertitude, la solution consiste à combiner deux types de décisions, qui se distinguent par le moment où elles sont prises : les décisions "hors-ligne", c'est-à-dire avant l'exécution ; et les décisions "en-ligne", c'est-à-dire pendant l'exécution. Selon les modèles, les décisions à prendre hors-ligne ou en-ligne varient : (1) *Modèles réactifs* (Szelke & Kerr, 1994) : plan et ordonnancement sont construits hors-ligne et, en cas d'aléas, sont corrigés en-ligne. (2) *Modèles à horizon glissant* (Erschler, Huguet, & de Terssac, 1997) : le plan est construit hors-ligne et l'ordonnancement en-ligne, mais en tâche de fond. Le principe est de construire l'ordonnancement progressivement, sur un intervalle de temps restreint proche du présent. (3) *Modèles proactifs* (Morris et al, 2001 ; Huguet et al. 2002) : les décisions d'ordonnancement sont prises hors-ligne et en-ligne. Les décisions en ligne sont préparées dans l'ordonnancement hors-ligne qui intègre les solutions associées aux aléas probables et possibles.

Cependant, la difficulté de l'approche dynamique, et notamment des modèles proactifs, réside dans la robustesse des décisions hors-ligne ; décisions qui doivent, quoi qu'il arrive, autoriser des décisions en-ligne. Deux propriétés sont associées à celle de robustesse : la contrôlabilité dynamique, qui concerne le positionnement des tâches dans le temps, et la séquentiabilité dynamique qui concerne le séquençage des tâches en conflit sur une ressource. La contrôlabilité dynamique intéresse des problèmes de satisfaction de contraintes temporelles avec incertitude (Vidal, & Fargier, 1999) alors que la séquentiabilité dynamique garantit que les tâches pourront être exécutées grâce aux ressources nécessaires quelles que soient les conditions d'exécution.

En référence à des travaux sur la propagation de contraintes en ordonnancement sous incertitude (Huguet, & Lopez, 2000), le problème consiste à déduire des interdictions de précédence entre tâches en considérant les dates limites et les durées de ces tâches ainsi que l'ensemble des tâches nécessitant cette ressource.

2.2 Application des modèles d'ordonnancement automatique à l'homme

Seuls quelques travaux sur le travail humain ont posé la réalisation de l'action humaine en termes d'ordonnancement (e.g., Sanderson, 1991). Pourtant en considérant que l'homme dispose de ressources pour réaliser ses tâches et qu'il est lui-même une ressource polyvalente à capacités énergétique et temporelle variables, la réalisation de l'action quotidienne est non seulement un problème de séquençage de tâches visant à un but, c'est-à-dire un problème de planification, mais aussi et surtout un problème de positionnement des tâches dans le temps compte tenu des ressources disponibles, c'est-à-dire un problème d'ordonnancement. De plus, à la fois organisateur et exécuteur de l'action dans un environnement dynamique, l'homme peut être assimilé à un système autonome qui prend des décisions hors-ligne et surtout des décisions en-ligne.

En référence aux problèmes d'ordonnancement avec incertitude, nous poserons que l'homme est assimilable à un système de planification-ordonnancement dynamique. Dans cette perspective, l'homme assure la cohérence de ses décisions d'exécution grâce à un plan de tâches flexible assimilable à un système de contraintes. Notre intérêt se porte sur les contraintes temporelles sur les tâches et sur les ressources. Pour ces dernières, considérons l'homme lui-même en tant que ressource à capacité temporelle et énergétique variable. A ce titre divers types de contraintes peuvent être envisagées, de nature physiologiques, psychologiques et/ou sociales organisées par des cadres temporels réguliers, journaliers et hebdomadaires. Cette présentation portera principalement sur les contraintes temporelles sur les tâches.

3 METHODOLOGIE

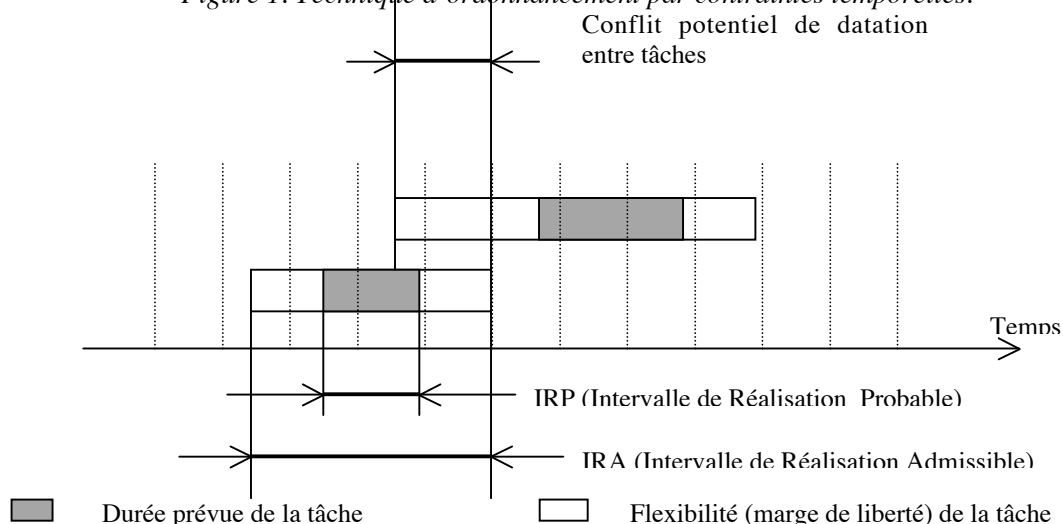
L'étude présentée porte sur deux types de télétravailleurs. (1) des *nomades* (1 femme et 6 hommes, de 26 à 48 ans) : consultants, ils partagent leur temps professionnel entre un (ou plusieurs) télécentre(s) périurbain(s), les clients, l'entreprise-mère et leur domicile. (2) Des *pendulaires* (2 femmes et 4 hommes, de 29 à 57 ans) : administratifs (gestion, comptabilité,...), ils travaillent en alternance entre l'entreprise-mère et un télécentre périurbain.

Ce travail porte sur l'observation systématique de l'ordonnancement des tâches quotidiennes. Les données sont recueillies grâce à la technique d'ordonnancement par contraintes temporelles (Valax, 1998). Inspirée du concept d'intervalle utilisé en automatique, cette technique (figure 1) consiste à représenter spatialement une tâche donnée sur une grille horaire par deux intervalles de temps emboîtés : (1) un *Intervalle de Réalisation Admissible* (IRA), emboîtant, qui symbolise l'horaire de début au plus tôt et l'horaire de fin au plus tard de la tâche, en d'autres termes les positions possibles de la tâche dans le temps. (2) un *Intervalle de Réalisation Probable* (IRP), emboîté, qui symbolise l'horaire de début et l'horaire de fin de la tâche, en d'autres termes la position la plus probable de la tâche (et en conséquence sa durée).

Préalablement à la tâche d'ordonnancement, des entretiens ont été menés et des grilles de budget-temps prévisionnels ont été remplies par les sujets (sur le jour et la semaine suivante) de façon à permettre de définir les différentes tâches réalisées par cette population.

Pour l'ordonnancement, les sujets doivent prévoir leur emploi du temps sur les trois jours suivants. Le résultat de l'ordonnancement effectué est un plan de tâches dont nous allons présenter les caractéristiques.

Figure 1. Technique d'ordonnement par contraintes temporelles.



4 RESULTATS

4.1 Nature des tâches

Sur l'ensemble des budgets-temps prévisionnels nous avons obtenu 36 intitulés différents regroupés en 11 tâches distinctes en fonction des caractéristiques fonctionnelles précisées par les sujets. Nous pouvons distinguer : (1) *rédaction*, (2) *prestation*, (3) *rendez-vous client*, (4) *information et communication médiatisée*, (5) *réunions internes*, (6) *gestion*, (7) *veille technologique*, (8) *formation*, (9) *trajet*, (10) *déjeuner*, (11) *non professionnelle*.

Parmi les tâches professionnelles, les données de budgets-temps et d'entretiens font apparaître deux classes de tâches centrales pour atteindre l'objectif de travail : (1) la *production* qui se rapporte à des tâches génératrices de facturation et de chiffre d'affaire (*rédaction*, *prestation* et *rendez-vous client*). (2) l'*entretien* qui se rapporte aux tâches qui ne génèrent pas directement de chiffre d'affaire, mais qui lui sont nécessaires (*gestion*, *réunions internes*, *information et communication médiatisée*, *veille technologique* et *formation*). Et une classe de tâches que nous nommerons *périphériques* qui concerne les *trajets* et les *déjeuners*.

4.2 Durée des tâches

La durée des tâches est calculée à partir des IRPs relevés sur les ordonnancements par contraintes temporelles. Les durées analysées sont des volumes horaires correspondant à la somme des durées de chaque IRP relevant d'une tâche donnée, sur l'ensemble des jours de recueil.

En moyenne journalière, l'intervalle de temps couvert par l'ordonnement est égal à 11h36. Les tâches de *production* et d'*entretien* couvrent 79,7% de cet intervalle et représentent une durée moyenne égale à 9h15. Ces valeurs varient en fonction des groupes de télétravailleurs, elles sont inférieures pour les *nomades* (10h05 / 75,7% / 7h36) par rapport aux *pendulaires* (12h47 / 82,8% / 10h31).

Le temps affecté aux tâches est différent entre les *pendulaires* et les *nomades* avec moins de tâches de *production* pour les *pendulaires* mais plus de temps alloué aux tâches d'*entretien*. Le temps affecté aux tâches *périphériques* par les *nomades* est supérieur à celui affecté aux tâches d'*entretien* alors que pour les *pendulaires* il est nettement inférieur à celui des autres classes de tâches.

Enfin, pour les deux populations, on peut constater le dépassement non négligeable de la durée légale du travail dans l'ordonnement réalisé, et ceci, même en excluant les tâches *périphériques*.

4.3 Flexibilité de la datation des tâches

Dans la technique de recueil par contraintes temporelles, la flexibilité de datation des tâches peut être saisie grâce au rapport entre les positions possibles de la tâche et sa position probable dans le temps. Les positions possibles de la tâche sont représentées par l'IRA, la durée probable est représentée par l'IRP qui représente non seulement la date de la tâche, mais aussi sa durée probable. La technique d'ordonnement par contraintes temporelles permet d'estimer la flexibilité de la datation

des tâches grâce au rapport entre la taille de l'IRP et la taille de l'IRA. L'indice de flexibilité des tâches (F_x) a été calculé comme suit : $F_x = 1 - [IRP/IRA]$. Quand F_x égale 0, IRA égale IRP, la flexibilité est minimale, quand F_x égale 1, IRP égale 0, la flexibilité est maximale.

L'indice de flexibilité moyen est relativement faible, égal à 0,34. Si la valeur de l'indice ne varie pas en fonction du type de télétravail, il varie en fonction des tâches [$F(9, 971)=3,738$; $p<.00$]; globalement, les tâches de *production* (0,21) obtiennent un indice de flexibilité plus faible que les tâches d'*entretien* (0,32), elles-mêmes obtenant une valeur plus faible que les tâches *périphériques* (0,41). Notons que parmi les tâches d'*entretien*, les tâches *informations et communication médiatisée* et *veille technologique* sont associées à des valeurs équivalentes aux tâches *périphériques*.

Ainsi, certaines tâches étant plus flexibles que d'autres, elles peuvent permettre de conserver des marges de manœuvre.

4.4 Conflits de datation entre tâches

Le conflit de datation entre tâches a été étudié sur les données d'ordonnement par contraintes temporelles à partir de la fréquence et de l'ampleur des recouvrements entre IRA.

Pour appréhender la fréquence des conflits potentiels, nous avons mesuré l'écart entre le nombre de recouvrements réels et théoriques¹ pour chaque paire possible de tâches. Les recouvrements *intra-classes* sont peu nombreux, inférieurs aux valeurs théoriques alors que les recouvrements *inter-classes* concernent principalement la paire *production* / tâches *périphériques*.

Lorsque deux IRA se recouvrent en tout ou partie, nous avons mesuré l'ampleur du recouvrement grâce à l'indice R_c suivant : $R_c = [2 \times (T1 \cap T2)] / (T1 + T2)$. $T1$ et $T2$ représentent les tâches constitutives de la paire considérée. $T1 \cap T2$ représente l'ampleur du recouvrement entre $T1$ et $T2$. l'indice R_c varie entre 0 et 1. Le recouvrement est minimum lorsque R_c tend vers 0. Le recouvrement est maximum lorsque R_c égale 1, dans ce cas $T1$ est égal à $T2$.

L'indice de recouvrement moyen des IRAs est faible (0,31) et reste constant pour les deux groupes de sujets [$F(1,80)=0,14$; NS]. En considérant les valeurs d'indices supérieures à la moyenne, on peut dire que 3 tâches montrent des recouvrements importants : (1) *information et communication médiatisée* et *rédaction* qui obtiennent chacune 4 valeurs supérieures à la moyenne pour 6 conflits potentiels, et qui appariées sont associées à l'indice de recouvrement le plus important (0,71) ; (2) *veille technologique* qui obtient 3 valeurs supérieures à la moyenne pour 5 conflits potentiels. Enfin, notons que si *trajets* produit de nombreux recouvrements, leur ampleur est relativement faible.

5 DISCUSSION

Il semble important de constater le dépassement prévisionnel de la durée légale de travail. Ce dépassement est de 49% pour les pendulaires et de 9% pour les nomades (sans considérer les trajets et les déjeuners). La différence entre les deux types de télétravail pourrait être expliquée en partie par des stratégies différentes de gestion du temps. En effet, si la somme des durées des tâches de production et d'entretien est plus importante chez les pendulaires que chez les nomades, inversement l'indice de flexibilité est plus important chez les nomades. Ainsi, on peut penser que la stratégie des pendulaires consisterait plutôt à surestimer la durée des tâches, alors que celle des nomades serait de se laisser une marge de liberté plus importante.

La flexibilité du plan de tâches, relative à la datation des tâches, se répartit sur chacune d'entre elles en fonction de leurs exigences intrinsèques et de leur contexte de réalisation. La flexibilité de datation des tâches est un des critères essentiels pour la définition du plan de tâches. Dans les entretiens concomitants au remplissage des budgets-temps prévisionnels, il apparaît qu'une des règles qui régit la construction du plan de tâches est «*placer en premier les tâches les moins flexibles*». Cette règle se manifeste par une procédure de construction qui consiste (1) à rechercher et fixer les tâches associées à un indice de flexibilité faible, comme *prestation*, *réunions internes* et *rendez-vous* ; (2) à dater les tâches restantes dans les intervalles de temps laissés libres. Notons, par ailleurs que les tâches peu flexibles, qui ancrent le plan de tâches, sont principalement des tâches collectives.

¹ Le nombre de recouvrement théorique est donné par un modèle dans lequel toute paire de tâches se recouvre. Sachant que nous avons 55 paires de tâches possibles (en excluant les paires de type de tâche A / tâche A) et 81 recouvrement effectifs, chaque paire doit obtenir un nombre de recouvrement théorique égal à 1,5.

La flexibilité du plan de tâches induit des conflits potentiels de datation entre tâches qui devront être réglés lors de l'exécution. Nos résultats montrent que l'éventualité d'un conflit entre tâches est fonction de la nature de l'appariement entre tâches. Ce résultat peut s'expliquer par l'usage d'un cadre temporel journalier, préexistant à l'ordonnancement dynamique, dans lequel chaque créneau horaire serait associé à une classe de tâches ou dans lesquels les tâches d'une même classe s'excluraient les unes les autres (Valax, 1986). Par ailleurs, certaines tâches – les moins collectives - génèrent plus de conflits que d'autres, notamment les tâches *rédaction, information et communication médiatisée, veille technologique et trajet*.

L'étude réalisée nous permet de dire que les notions développées dans les modèles automatiques de planification-ordonnancement peuvent s'appliquer à la gestion temporelle des tâches quotidiennes chez les télétravailleurs. Cette correspondance est avantageuse car elle permet (1) d'envisager le développement de techniques d'observation standardisables, (2) la définition de critères univoques et stables pour définir les tâches humaines et (3) le développement d'outils d'aides à la gestion du temps plus adaptés au travail humain.

6 BIBLIOGRAPHIE

- Bates, P. & Huws, U. (2002) *Modelling eWork in Europe, Estimates, models and forecasts from the Emergence project*. The Institute for Employment Studies, report n°388. Brighton, Great Britain.
- Breton, T. (1994). *Le télétravail en France, situation actuelle, perspectives de développement et aspects juridiques*. Paris : La documentation.
- Erschler, J., Huguet, M.-J., & de Terssac, G. (1997). Décision distribuée en gestion de production. In J.-C. Hennet (Ed), *Concepts et outils pour les systèmes de production* (pp.105-132). Toulouse : Cépadués.
- eWork (2002). *Les nouvelles méthodes de travail dans la société de l'information en 2002*. Commission Européenne, rapport juillet 2002. Luxembourg, Belgique.
- Gottlieb, B.-H., Kelloway, E.-K., & Bartham, E. (1998). *Flexible Work Arrangements : Managing the Work-family Boundary*. Chichester : John Wiley.
- Huguet, M.-J., & Lopez, P. (2000). Approche par contraintes de problèmes d'ordonnancement de tâches et d'affectation de ressources : *CIFA 2000*.
- Huguet, M.-J., Lopez, P., Vidal, T. (2002). *Dynamic task sequencing in temporal problems with uncertainty*, AIPS workshop on on-line planning and scheduling, 2002.
- Konradt, U., Schmook, R., & Mälecke, M. (2000). Impacts of telework on individuals, organizations and families – a critical review. *International Review of Industrial and Organizational Psychology*, 15, 63-99.
- McCloskey, D.-W., & Igarria, M. (1998). A review of the empirical research on telecommuting and directions for future research. In M. Igarria, M. Tan (Eds.), *The virtual Workplace*, 338-358. Hershey, PA : IdeaGroup Publishing.
- Morris, P., Muscettola, N., Vidal, T. (2001). *Dynamic control of plans with temporal uncertainty*, IJCAI 2001.
- Sanderson, P. (1991). Towards the Model Human Scheduler. *International Journal of Human Factors in Manufacturing*, 1(3), 195-219.
- Steward, B. (2000). Changing times – the meaning, measurement and use of time in teleworking. *Time & Society*, 9(1), 57-74.
- Szelke, E., & Kerr, R.-M. (1994). Knowledge-Based reactive scheduling. *Production Planning and Control. Production Planning & Control*, 5(2), 124-145.
- Valax, M.-F. (1998). Technique de recueil de données pour l'étude de la position des tâches quotidiennes. *Le Travail Humain*, 61(2), 187-199.
- Vidal, T., Fargier, (1999). *Handling contingency in temporal constraint network : form consistency to controlabilities*, Journal of Experimental and Theoretical Artificial Intelligence, vol 45, pp11-23.

Session 3

Effets de l'interaction

Scénarios d'accidents impliquant des deux-roues à moteur : une question d'interaction

Pierre Van Elslande

INRETS-MA - Chemin de la Croix Blanche - 13300 Salon de Provence

pierre.van.elslande@inrets.fr

RESUME

L'objet de cette communication est de montrer en quoi les accidents impliquant des deux-roues à moteur présentent une spécificité du point de vue des mécanismes qui les sous-tendent, par rapport aux autres types d'accidents de la route. L'analyse s'appuie sur des données détaillées recueillies sur ces accidents. Plus précisément, elle porte sur l'activité des conducteurs en situation critique, les défaillances qui y correspondent et les facteurs qui leur sont associées. Elle vise la mise en évidence d'un certain nombre de différences entre les processus de dysfonctionnement caractérisant les conducteurs de deux-roues et ceux des automobilistes. Elle vise également à dégager les difficultés particulières des automobilistes impliqués dans des accidents avec deux-roues à moteur, par rapport à celles des automobilistes impliqués dans d'autres types d'accidents. L'échantillon étudié comporte 80 accidents de deux-roues à moteur. L'analyse de ces 80 cas, comparée aux résultats issus de travaux préalables sur des accidents d'automobiles (392 cas), permet de dégager des paramètres typiques aux deux-roues et aux automobilistes qui y sont confrontés, ainsi qu'aux scénarios d'interaction qu'ils mettent en oeuvre conjointement.

MOTS-CLES

Deux-roues à moteur - Erreur humaine - Conduite automobile - Accident - Interaction.

1 INTRODUCTION

Depuis une vingtaine d'années, le nombre d'accidents et de morts sur la route diminue progressivement, en relation avec les améliorations apportées aux organes de sécurité des véhicules, les aménagements de l'infrastructure routière, la formation dispensée aux conducteurs, ainsi que certaines mesures légales (e.g. limitations de vitesse). Pourtant, le deux-roues à moteur reste aujourd'hui un moyen de transport particulièrement dangereux : en 2001, le nombre de conducteurs de deux-roues à moteur victimes d'accidents représentait 17 % du total des tués et 22 % du total des blessés alors même que ces véhicules ne représentaient que 6,5 % du parc circulant et qu'ils parcouraient en moyenne près de quatre fois moins de Km par an que les automobiles (ONISR, 2002). Les mesures prises ces dernières années à leur égard ne sont pas parvenues à faire baisser significativement ces taux. Ce constat atteste de l'intérêt de mener des études approfondies qui, au-delà de l'exploitation des fichiers statistiques nationaux, visent à comprendre et à expliquer la genèse des problèmes rencontrés par les conducteurs de deux-roues à moteur. C'est dans cet objectif qu'a été réalisée la présente recherche qui investit les données correspondant à un échantillon de 80 Etudes Détaillées d'Accidents (EDA) sous un éclairage cognitif.

Une première section fait une synthèse des travaux portant sur les particularités de la conduite des deux-roues motorisés et de leur accidentalité. Une seconde partie est l'occasion d'un rapide exposé du cadre d'analyse utilisé pour notre étude. La présentation des résultats, objet de la troisième section, vise à montrer en quoi, par rapport à d'autres types d'accidents, les accidents de deux-roues à moteur présentent des spécificités du point de vue des processus fonctionnels engagés. Elle pointe également la nature des difficultés rencontrées par les automobilistes impliqués dans ces mêmes accidents. La discussion finale insiste sur le caractère interactif des scénarios d'erreurs confrontant ces deux types d'usagers de la route.

2 PROBLEMATIQUE : CONDUITE ET ACCIDENTOLOGIE DES DEUX-ROUES MOTORISES

Le véhicule deux-roues à moteur présente de nombreuses spécificités, du point de vue de ses caractéristiques, du mode de conduite induit et de l'accidentalité qui en résulte.

2.1 Un comportement spécifique

Les particularités dynamiques de ces véhicules, comme l'ont souligné plusieurs études (Têtard, 1994 ; Obenski, 1994 ; Chinn, 1990), les rendent particulièrement difficiles à conduire, en liaison notamment avec la nécessité du contrôle de l'équilibre. Du fait de ces paramètres, l'expérience spécifique du véhicule est une variable beaucoup plus importante dans la conduite d'un deux-roues à moteur que dans la conduite automobile (Obenski, 1994 ; Têtard, 1994). Deux véhicules de même cylindrée peuvent ainsi avoir des comportements dynamiques très différents selon leur conception et selon le type d'utilisation prévu (routière, « trail », sportive), et le temps d'apprentissage de la conduite est plus long quel que soit le deux-roues (Têtard, 1994). En comparaison de la conduite d'une automobile, un deux-roues à moteur est moins maniable, notamment en situation d'urgence (Têtard, 1994), les capacités d'évitement sont plus faibles, le freinage d'urgence est moins commode, voire plus dangereux (Obenski, 1994).

Par ailleurs, les caractéristiques mêmes de ces engins permettent des comportements souvent imprévisibles pour les automobilistes : le faible espace occupé sur la chaussée leur permet de s'intercaler entre les voitures et d'effectuer des changements fréquents de trajectoire. Les performances très différentes entre voitures et deux-roues à moteur, ainsi que l'espace restreint occupé par ces derniers amènent une hétérogénéité importante dans les vitesses pratiquées par ces deux catégories dans certaines situations, particulièrement en agglomération. De plus, les capacités d'accélération de ces véhicules sont généralement bien au-dessus de celles de la plupart des automobiles (0 à 100 Km/h en 40 mètres pour les plus rapides...), capacités souvent méconnues des automobilistes. Une étude de Brooks & Guppy (1990) a montré que les représentations qu'ont les conducteurs d'automobiles des capacités dynamiques des deux-roues motorisés sont le plus souvent erronées. Ils considèrent que ces véhicules sont plus maniables, ont une tenue de route et un freinage meilleurs que les automobiles. Ces conceptions orientent leurs attitudes de conduite, et amènent fréquemment des évaluations fausses des situations rencontrées et, en conséquence, l'exécution de manœuvres inappropriées.

2.2 Une place à part au sein du trafic

Les particularités de ces véhicules et de leur mode de conduite posent ainsi un problème d'intégration au sein du trafic global. Le différentiel qu'ils introduisent est considéré dans la littérature amener des difficultés d'ordres perceptif, évaluatif et représentationnel dans leurs relations aux autres. Leur gabarit plus faible est souvent évoqué pour poser un problème de détectabilité. Obenski (1994) souligne qu'un obstacle, même très petit peut les masquer entièrement dans certaines conditions, ce qui peut être paradoxalement le fait de certains éléments d'aménagement. Ce même auteur rend compte par ailleurs de comportements de conducteurs de deux-roues à moteur, plus spécialement les inexpérimentés, qui ne tiennent pas compte de leur faible détectabilité et adoptent des modes de conduite qui ne facilitent pas leur détection par les automobilistes. C'est dans la perspective de réduire ce problème que l'allumage des feux de croisement de jour a été promulgué pour les motocyclistes. Mais l'efficacité de cette mesure est contestée par certains auteurs (dont Têtard, 1994), notamment parce qu'elle ne permet pas de résoudre cet autre problème-clé qui réside dans les erreurs d'estimation des vitesses d'arrivée que ces véhicules suscitent, notamment en situation de dépassement et d'intersection. Howells *et al* (1980) ont ainsi montré que les automobilistes n'avaient pas de meilleures performances d'évaluation lorsque les motocycles ont les feux de croisement allumés. Caird & Hancock (1995) ont vérifié que la taille du véhicule était fortement utilisée comme un indice de jugement et qu'elle serait donc un facteur important d'erreurs d'estimation de la distance et ainsi, du temps d'arrivée. Mais ces problèmes font également intervenir les connaissances et représentations qui sont développées à l'égard des deux-roues. Têtard (1994) note ainsi que les automobilistes tendraient à moins les prendre en considération dans la mesure ou ils se sentent moins vulnérables qu'eux. Par ailleurs, la rencontre de deux-roues à moteur dans la circulation constitue un événement relativement rare et les automobilistes auraient tendance à leur attribuer des vitesses du même ordre de grandeur

que celles des voitures, ce qui aura des conséquences à la fois sur leur détection (du fait des stratégies de recherche d'information mises en place) et sur l'évaluation de leur vitesse d'approche. Howells *et al* (1980) ont montré expérimentalement que le fait de pratiquer la conduite d'un motorcycle avait une incidence sur les évaluations réalisées à leur rencontre en situation d'intersection simulée. Des sujets qui la pratiquent sont très peu impliqués dans des situations conflictuelles avec les motocyclistes et perçoivent mieux leurs variations d'allure. Des sujets qui ne conduisent pas de moto sous-estiment les vitesses d'approche et décident de passer alors qu'ils n'en ont pas le temps.

En bref, le problème pour les conducteurs de deux-roues à moteur est triple. Ils sont moins visibles en approche du fait de leur faible gabarit. Lorsqu'ils sont détectés, les évaluations de leurs vitesses et distance sont souvent erronées¹. Un autre facteur, plus basique, fait référence à la notion de « loi du plus fort » (Favero *et al*, 1980 ; Têtard, 1994) : les automobilistes tendraient à moins prendre en considération les deux-roues à moteur du fait d'un sentiment de moins grande vulnérabilité.

2.3 Une accidentalité particulière

Ces différents problèmes caractérisant les deux-roues dans leur relation au trafic ont une incidence directe sur leur accidentalité. En effet, environ les deux tiers des accidents de deux-roues motorisés impliquent au moins un autre véhicule, le plus souvent un véhicule léger (VL). Les deux situations les plus accidentelles sont l'arrivée en intersection et le dépassement. Les autres renvoient principalement aux ouvertures de portière, traversées de piétons hors des passages cloutés, franchissements de feux rouges, mouvements dangereux des autres usagers, mais dans des proportions moindres (Pibault & Bilman, 1997).

Les recherches accidentologiques sur les deux-roues à moteur ont évolué en trois phases. Dans les années 70, les études portaient principalement sur les types de choc. Dans les années 80, on s'est intéressé au processus de conduite, à la perception du risque, à l'expérience, le conducteur étant regardé comme le « composant central ». Depuis le début des années 90, on s'intéresse au conducteur comme « agent actif », et aux processus qui sous-tendent son traitement des informations. On étudie également les attitudes, les croyances sur la sécurité, ainsi que les effets de l'expérience et de la formation sur les habiletés (Chesham *et al*, 1993). Les travaux les plus répandus sur la conduite des deux-roues à moteur sont basés sur des analyses de procès-verbaux d'accidents. L'intérêt de ces études est indéniable sur de nombreux points : détecter les populations les plus exposées (Têtard, 1994), les types de véhicules les plus impliqués (Filou, 1998), décrire l'accident-type de motocyclette (Preusser *et al*, 1995 ; Têtard, 1994), les principaux facteurs d'accidents de deux-roues à moteur (Hurt *et al*, 1981 ; Rutter & Quine, 1995) et proposer en conséquence des mesures de prévention et de formation. Cependant, on peut émettre plusieurs critiques envers la majorité de ces études. Une première est de s'en tenir à la recherche des causes de l'accident d'une part, et des causes chez le conducteur de deux-roues d'autre part. Quelques études seulement tiennent compte véritablement de l'interaction entre véhicules dans la recherche d'explications. Par ailleurs, le procès-verbal est un document qui n'est pas établi dans un objectif de recherche (si ce n'est de responsabilité) et comporte à cet égard de nombreux biais. Il n'est pas rare que plusieurs variables soient mal renseignées (Têtard, 1994). A l'arrivée, certains éléments utiles à la compréhension des accidents ne peuvent être exploités.

Les principaux facteurs d'accidents des deux-roues à moteur, dégagés par les différentes études sont l'âge, l'expérience de la conduite, celle du véhicule, la vitesse excessive (notamment en approche d'intersection ou en virage), la cylindrée du véhicule et les attitudes vis à vis de la prise de risque. Ces différents facteurs permettent d'expliquer les accidents sur le versant du conducteur de deux-roues. On soutiendra dans la section suivante qu'il est trop simple de réduire tous les accidents à ces seuls éléments. Comme on l'a vu, dans plus des deux tiers des accidents de deux-roues à moteur, un VL est impliqué. La compréhension réelle de l'émergence des erreurs et des facteurs d'apparition des défaillances doit donc passer par l'étude des interactions entre les différents protagonistes et de leur participation réciproque à la dégradation des situations. Un point marquant des différentes études d'accidents évoquées ci-dessus est le peu de référence faite aux aspects cognitifs de la conduite et à leurs implications dans l'accidentalité des deux-roues à moteur.

¹ Les variations dans les performances des différents types de deux-roues à moteur ajoutent encore à cette difficulté d'évaluation pour les automobilistes.

3 CADRE D'ANALYSE : LES ERREURS ET LEURS FACTEURS ²

Dans l'analyse qui suit, on considérera qu'au-delà des particularités évoquées dans la section précédente, la conduite d'un deux-roues à moteur fait globalement appel à des processus cognitifs similaires à ceux qui sont mis en oeuvre dans la conduite d'une automobile, notamment en termes de mécanismes et de fonctions de « traitement de l'information » impliqués dans la réalisation de l'activité, et qu'ils peuvent par conséquent bénéficier des mêmes méthodes d'analyse. Des différences essentielles existent bien évidemment -que la présente étude se propose d'éclairer sous l'angle des dysfonctionnements- en ce qui concerne les degrés de sollicitation des différents processus psychomoteurs et perceptivo-cognitifs engagés, en liaison notamment avec les particularités du pilotage de ce type de véhicules et leurs répercussions dans le trafic.

3.1 L'Étude Détaillées des Accidents (EDA)

Si elles sont loin de pouvoir tout expliquer, et en dépit des difficultés de leur mise en oeuvre (Grayson & Hakkert, 1987), les analyses d'accidents réalisées dans un objectif de recherche offrent une contribution indéniable à la compréhension des problèmes d'insécurité. Elles permettent d'appréhender des phénomènes complexes, tels « l'erreur » et ses processus de production, à la fois sous l'angle des situations dans lesquelles elles émergent et des facteurs qui y contribuent. C'est dans cet objectif qu'est réalisée l'EDA au département Mécanismes d'Accidents de l'INRETS. Lorsqu'un accident de la route se produit dans le secteur d'intervention, une équipe pluridisciplinaire se rend sur les lieux en même temps que les services de secours et les forces de l'ordre, dans l'objectif de recueillir les données les plus éphémères sur le déroulement de l'événement. Un psychologue réalise des entretiens « à chaud » auprès des personnes impliquées³ et des témoins ; un technicien recueille des informations sur les véhicules et l'infrastructure. Les informations ainsi obtenues sont affinées par un recueil complémentaire le lendemain ou surlendemain. L'étude de l'ensemble de ces éléments -traces, indices, mesures, interviews- permet une analyse fine du déroulement de l'accident. Différentes disciplines scientifiques sont requises pour ce travail : psychologie, dynamique du véhicule et mécanique, génie civil, médecine.

L'analyse de ces données recueillies s'appuie sur *i*) une approche système qui met l'accent sur les interactions entre l'impliqué, son véhicule et l'environnement ; *ii*) un modèle cinématique permettant la reconstruction et le paramétrage, dans le temps et l'espace, de la dynamique du phénomène ; *iii*) un modèle du fonctionnement humain qui exploite les formalismes de la psychologie cognitive ; *iv*) un modèle de découpage en phases – de conduite, de rupture, d'urgence, de choc – du déroulement de l'accident (Ferrandez, 1995). Le résultat de cette démarche est une collection de cas d'accidents analysés cliniquement, intégrant un diagnostic sur les mécanismes et les facteurs concourant au déclenchement, au déroulement, et aux conséquences de ces accidents. Un recoupement de ces monographies sur la base de différents critères de sélection permet la mise en place d'études thématiques comme celle qui est présentée ici, dont l'objet est la mise en évidence des scénarios d'erreurs dans les accidents impliquant des deux-roues à moteur.

3.2 Défaillances fonctionnelles

Dans la lignée des développements récents des travaux sur le thème de « l'erreur » (e.g. Reason, 1995 ; Leplat, 1999 ; Amalberti, 2001)⁴, celle-ci ne sera pas analysée comme la cause première de la dégradation des situations mais comme le symptôme et le vecteur des failles d'un système (Van Elslande, 2003). Nous en rendrons compte par la suite sous le label de « défaillance fonctionnelle », et ce pour trois raisons : d'abord, pour se démarquer des ambiguïtés de la notion d'erreur dans son acception courante (souvent synonyme de « faute »), ensuite pour cibler l'analyse sur les erreurs non récupérées (qui sont par définition celles que l'on étudie à partir de données d'accidents...), enfin pour inclure dans cette analyse des phénomènes plus larges liés aux capacités, notamment physiologiques, des individus. Autrement dit, cette notion rend compte de la défaillance de l'une (au moins) des

² Cette section fait l'objet d'un résumé nécessairement succinct. Pour plus d'information, le lecteur intéressé pourra se référer à Van Elslande 2000 et 2003.

³ Avec leur accord, bien entendu, et sous réserve d'une autorisation préalable des services de secours.

⁴ Pour n'en citer que quelques uns.

fonctions cognitives, sensori-motrices ou psychophysiologiques qui permettent, en temps normal, à l'opérateur de s'adapter aux difficultés qu'il rencontre dans l'accomplissement de sa tâche.

Des travaux préalables ont été réalisés dans cette perspective, visant à montrer comment les dysfonctionnements dans les interactions du système Homme-Véhicule-Environnement s'expriment au plan des fonctions engagées par celui qui en a le contrôle direct : le conducteur. Ils ont amené l'élaboration d'une grille de classification des défaillances fonctionnelles sur la base d'une particularisation des modèles « classiques » d'analyse des erreurs (Rasmussen, 1986 ; Reason, 1993) aux spécificités des processus à l'œuvre dans les accidents de la route (Van Elslande *et al*, 1997 ; Van Elslande 2000, 2003). Cette grille rend compte d'une vingtaine de défaillances spécifiques impliquées aux différentes étapes de la réalisation de l'activité. Ces défaillances fonctionnelles sont réparties en 6 catégories : les 5 premières réfèrent à l'une ou l'autre des différentes étapes de « traitement » mises en œuvre dans la logique d'un modèle séquentiel (perception, diagnostic, pronostic, décision, exécution de l'action⁵) ; la dernière correspond à une altération de l'ensemble de cette chaîne fonctionnelle, identifiée dans certains types d'accidents. Mais s'il est important de connaître la nature précise des difficultés rencontrées par l'opérateur, il ne faut pas perdre de vue que les défaillances fonctionnelles dont il peut être l'objet s'inscrivent dans un contexte et s'expliquent par un certain nombre de facteurs qu'il est essentiel de mettre en évidence si l'on cherche à comprendre l'origine et le processus qui conditionnent une telle issue. Une fois identifiée, chaque défaillance est ainsi réintroduite au sein du scénario occurrence qui en a autorisé l'émergence et qui rend notamment compte des combinaisons d'éléments qui l'expliquent, liés au conducteur et à l'environnement de sa tâche. On signalera à ce propos que si ces défaillances sont analysées sous un angle cognitif, la recherche de leurs causes humaines fait également intervenir des paramètres relatifs aux motivations (e.g. rattraper un retard, trajet ludique), aux émotions (impatience, peur), à l'état psychologique (degré d'attention) et physiologique (niveau de vigilance), etc.

3.3 Scénarios-types

Le concept de scénario-type d'accident de la circulation a pour visée de synthétiser et de généraliser les connaissances qui sont extraites des études monographiques de cas d'accidents. C'est une construction issue de la synthèse de plusieurs cas considérés comme semblables. Le scénario-type se définit donc comme un déroulement récurrent d'accidents : ceux-ci présentent de telles similitudes dans l'enchaînement des faits et des relations de causalité qu'on peut considérer leur ensemble comme une forme de modèle. Si l'intérêt immédiat du concept de scénario-type est d'identifier des groupes d'accidents homogènes, l'objectif opérationnel est de se fonder sur cette homogénéité pour ouvrir de nouvelles perspectives de prévention (Fleury & Brenac, 2001).

Par rapport à cette conception globale, la spécificité de la présente analyse consistera à mettre l'accent sur la composante humaine en définissant des scénarios types de défaillances fonctionnelles, s'exprimant sous la forme d'enchaînements typiques qui relient une tâche à réaliser, des éléments explicatifs (« facteurs »), une défaillance fonctionnelle, l'activité qui en découle et la situation d'accident résultante (Van Elslande, 2003). Le déplacement de l'analyse du système dans son ensemble à celle du fonctionnement du conducteur en particulier ne revient pas à nier le caractère pluricausal de l'accident. Il s'agit simplement de décomposer cette multicausalité pour l'éclairer sous un angle particulier, celui des traitements cognitifs qui n'ont pu empêcher les conditions situationnelles de dégénérer.

La section suivante présente une application de ce cadre d'analyse pour la mise en évidence des spécificités des mécanismes d'accidents impliquant des deux-roues à moteur, par rapport aux autres types d'accidents.

⁵ Il faut préciser que le caractère séquentiel de cette grille de classification des défaillances ne sous-entend en aucun cas un fonctionnement séquentiel du conducteur (Van Elslande, 2003). Il s'agit bien d'une grille d'analyse des dysfonctionnements et non d'un modèle de fonctionnement de l'opérateur.

4 RESULTATS : SPECIFICITE DES MECANISMES EN JEU DANS LES ACCIDENTS IMPLIQUANT DES DEUX ROUES A MOTEUR

Un échantillon de 80 dossiers d'accidents impliquant un conducteur de deux-roues à moteur a donc été extrait de la base EDA de l'INRETS-MA. Il se compose de 64 accidents impliquant une automobile et de 16 pertes de contrôle véhicule seul. Les données correspondantes ont été analysées en termes de conditions d'occurrence, de défaillances fonctionnelles et d'éléments explicatifs, selon la logique définie plus haut. Il faut préciser que chacun de ces accidents a été étudié par (au moins) deux experts en accidentologie clinique, et que tous l'ont été dans une logique d'analyse globale -c'est-à-dire sans thématique spécifique définie a priori- ce qui garantit une absence de biais quant à la façon d'appréhender différenciellement les mécanismes d'accidents avec deux-roues. La présente section rend compte des résultats de cette analyse, en confrontation avec les données caractérisant un échantillon de 392 conducteurs impliqués dans un accident ne faisant pas intervenir de deux-roues à moteur. Ce second corpus, qui a fait l'objet d'une étude antérieure (Van Elslande *et al*, 1997), sera utilisé ici à titre d'échantillon « témoin » pour appréhender la spécificité des difficultés rencontrées non seulement par les conducteurs de deux-roues, mais également par les automobilistes qui leur sont confrontés.

4.1 Les conducteurs de deux roues

4.1.1 Spécificité des défaillances fonctionnelles

La comparaison des catégories de défaillances produites par les uns et les autres (figure 1) montre que les 80 conducteurs de deux-roues sont plus sujets aux erreurs de pronostic sur l'évolution des situations ainsi qu'aux défaillances dans l'exécution des actions que les 392 conducteurs de l'échantillon témoin ($\chi^2=12.90$, $p>0.024$). En conséquence, ils produisent proportionnellement moins d'erreurs aux étapes de perception et de diagnostic. Ils présentent également un taux plus faible de « défaillances généralisées », qui correspondent à une perturbation globale des capacités de conduite (qu'elle soit due à l'absorption de produits psychotropes, à un endormissement, un malaise, l'âge, etc.).

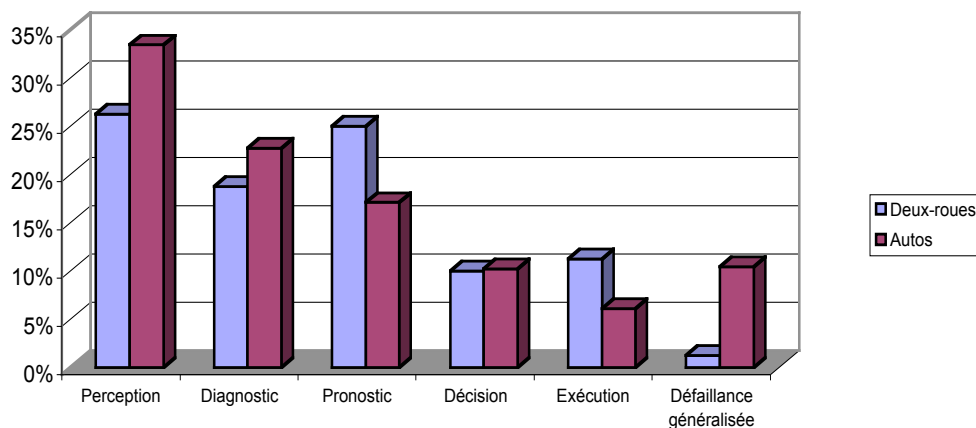


Figure 1. Comparaison des catégories de défaillances caractérisant les conducteurs de deux-roues à moteur (n=80) et un échantillon témoin d'automobilistes (n=392)

Ces résultats globaux peuvent être particularisés de la façon suivante. Tout d'abord, la surreprésentation des défaillances de pronostic provient principalement de la part que prennent deux types d'erreurs spécifiques. La première correspond à une *anticipation erronée d'une régulation par l'autre* usager avec lesquels ils sont en interaction conflictuelle, notamment en situation d'intersection (12 % des défaillances). La seconde consiste en une *prévision erronée d'absence d'obstacle potentiel* dans une situation à faible visibilité (11%). Ensuite, le fait que les conducteurs de deux roues rencontrent plus de problèmes à l'étape d'exécution des actions s'explique aisément par les contraintes dynamiques liées à la manœuvrabilité de ces véhicules, telles qu'en rend compte la littérature (e.g. Obensky, 1994). La question reste toutefois posée de savoir si l'importance que prend ce type de défaillance dans l'échantillon étudié provient d'une occurrence plus fréquente d'erreurs de la part des conducteurs de deux-roues, ou bien d'une récupérabilité plus difficile de telles erreurs dans la conduite

de ces véhicules. Par ailleurs, si les conducteurs de deux-roues sont globalement moins sujets aux défaillances de type perceptif que les autres dans la genèse des accidents dont ils sont victimes, ils sont néanmoins surreprésentés pour une défaillance perceptif spécifique, qui consiste en une *négligence des exigences de recherche d'information* (9%) lors de situations à faibles contraintes (conduite en file, avec un faible trafic, sur un trajet connu, etc.). Dans la même lignée, mais au niveau du diagnostic, cette population montre une tendance à être plus sujette que les autres à la *mauvaise évaluation d'une difficulté routière ponctuelle* (10%), telle qu'un virage ou une zone de travaux. Enfin, au niveau décisionnel, ils tendent à commettre plus souvent une *violation d'une règle de sécurité* (10%), notamment en situation de dépassement.

4.1.2 Spécificité des éléments explicatifs

L'analyse des éléments qui contribuent à l'émergence de ces différentes défaillances (figure 2) montre que pour les conducteurs de deux-roues on trouve plus souvent l'origine de leurs défaillances dans les conditions à la fois internes et externes qui déterminent leur trajet, que les autres conducteurs ($\chi^2=25.28$, $p>0.0001$). Parmi les éléments qui ressortent le plus de cette première catégorie, figurent un *attachement trop rigide au statut prioritaire*, ainsi que la tendance à adopter un *comportement à risque* (vitesse, recherche de sensation). Mais ce sont les facteurs liés au contexte externe (« conditions de circulation ») qui spécifient le plus cette population. Ces facteurs sont liés principalement aux interactions entre véhicules, qu'ils soient relatifs à *l'ambiguïté des indices émis ou des manœuvres réalisées par autrui*. Intervient également, plus indirectement, les conditions de visibilité qui rendent les deux-roues moins facilement détectables, ainsi que les contraintes spatio-temporelles liées à l'intensité du trafic.

On pourra donc noter que les défaillances des conducteurs de deux-roues ne s'expliquent pas seulement par elles-mêmes, mais qu'elles font aussi intervenir les autres conducteurs avec lesquels ils sont en interaction. Les erreurs de pronostic découlent ainsi souvent de situations critiques générées par les automobilistes dont le comportement met en défaut leur anticipation de l'évolution de la situation dans laquelle ils sont engagés.

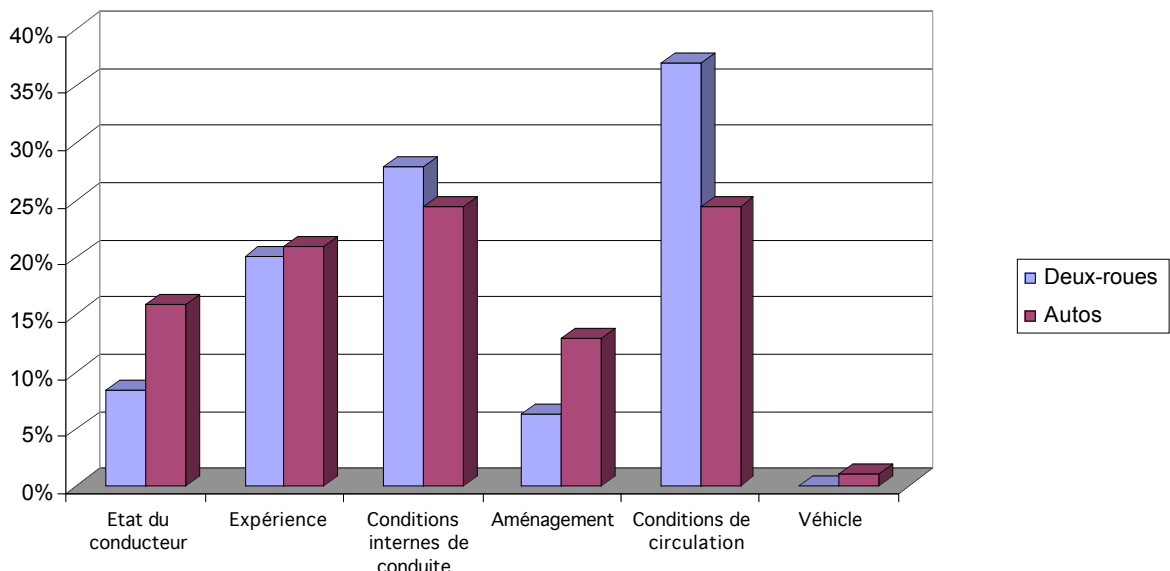


Figure 2. Comparaison des catégories d'éléments explicatifs des défaillances des conducteurs de deux-roues (n=202), et d'un échantillon témoin d'automobilistes (n=1126)⁶

⁶ Une défaillance s'expliquant le plus souvent par une combinaison de facteurs, l'effectif des éléments explicatifs est donc supérieur à celui des défaillances étudiées.

4.2 Les automobilistes accidentés avec les deux roues

Il s'agit par conséquent de ne pas omettre que les deux-roues à moteur interagissent sur la route avec des automobiles et qu'une part importante de leurs accidents implique au moins un Véhicule Léger (VL). L'étude des interactions qu'ils mettent en oeuvre dans les accidents paraît indispensable pour mieux cerner les difficultés respectives rencontrées par les différents conducteurs selon les configurations accidentelles auxquelles ils ont été confrontés. Une première question qui se pose consiste à savoir si les automobilistes accidentés avec un deux-roues font montre de problèmes spécifiques par rapport aux autres automobilistes accidentés. La présente section rend compte d'une comparaison des défaillances et des éléments explicatifs caractérisant les conducteurs de voiture intervenant dans les accidents de deux-roues, et ceux qui caractérisent le même échantillon « témoin » que précédemment (ne faisant pas intervenir de deux-roues).

4.2.1 Spécificité des défaillances fonctionnelles

Comme on peut le noter à la lecture de la figure 3, une nette spécificité se dégage de la comparaison des catégories de défaillances fonctionnelles produites par les automobilistes confrontés à une situation accidentelle avec un véhicule deux-roues, à celles des automobilistes de l'échantillon témoin ($\chi^2=19.96$, $p>0.001$). Ils sont en effet confrontés à près de 20 % de plus de défaillances de type perceptif, avec une répercussion négative sur les autres catégories d'erreurs. Le problème semble donc clair : dans un cas sur deux, les automobilistes ne voient pas le deux-roues. Cela confirme les conclusions des études basées sur des analyses de procès-verbaux, à savoir que les automobilistes confrontés à des deux-roues ont plus souvent des problèmes d'ordre perceptif. Toutefois, la catégorie « perception » se décline, comme les autres, en un ensemble de défaillances plus spécifiques. Ces défaillances ne renvoient pas seulement aux problèmes de visibilité, mais intègrent également les questions de stratégies de recherche d'information qui sont mises en oeuvre par les différents acteurs. Une décomposition de cette catégorie d'erreurs « perceptives » fait ainsi ressortir deux principaux types de défaillances fonctionnelles pour cet échantillon, la première correspondant à une *recherche hâtive d'information* opérée par les automobilistes (20% de l'ensemble des erreurs), l'autre à une *focalisation de cette recherche sur un autre élément* de la scène visuelle (25%). On identifie par contre moins de défaillances du type *non détection par absence de visibilité* que dans l'échantillon témoin. Ces données sont intéressantes en ce qu'elles mettent en question l'hypothèse selon laquelle le problème se situe essentiellement sur le plan de la détectabilité des deux-roues par les automobilistes.

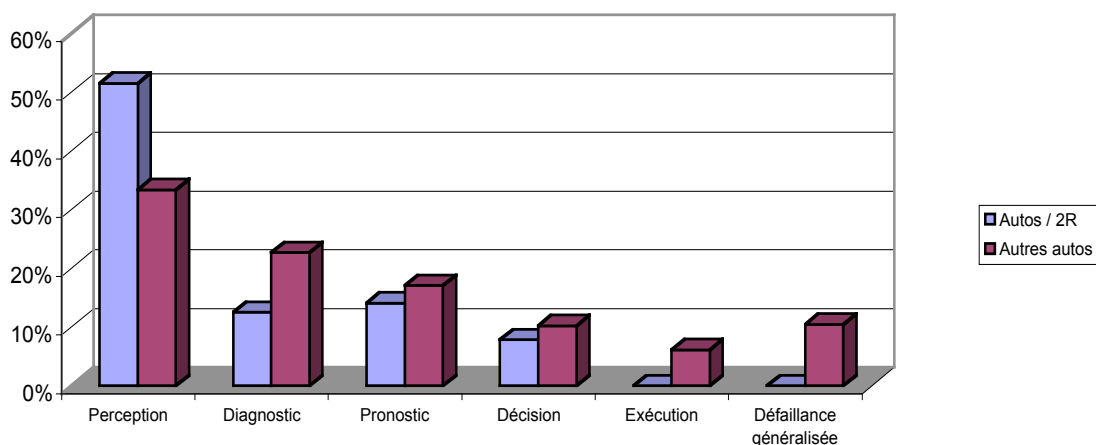


Figure 3. Comparaison des catégories de défaillances caractérisant les automobilistes impliqués dans un accident de deux-roues à moteur (n=64) et un échantillon témoin (n=392)

4.2.2 Spécificité des éléments explicatifs

A l'instar de ce que l'on a pu constater pour les conducteurs de deux-roues, on peut noter sur la figure 4 que les défaillances des automobilistes qui sont impliqués dans les mêmes accidents tendent à trouver -plus souvent que dans les autres types d'accidents- l'origine de leurs défaillances dans le contexte à la fois interne et externe de réalisation de l'activité, et moins dans les autres catégories de

facteurs : variables caractérisant l'état du conducteur, variables d'expérience, éléments liés aux aménagements. Mais, à la différence des conducteurs de deux-roues (cf. 4.1.2), les automobilistes qui leur sont confrontés semblent plus influencés par leur contexte interne de conduite, avec des éléments tels que *l'attention portée sur un autre élément de la scène visuelle*, une *contrainte de temps subjective affectée à leur tâche*, la *réalisation d'une manœuvre sur un mode plus ou moins automatique*, etc., autant d'éléments qui ne leur permettent pas d'opérer une recherche d'information appropriée à la détection d'un deux-roues. Les paramètres caractérisant le contexte externe (tels que le *caractère atypique d'une manœuvre d'autrui*, ou une *gêne à la visibilité*) auraient moins d'influence sur eux que sur les conducteurs de deux roues ($\chi^2=4.08$, $p>0.043$). Un point intéressant qui peut être déduit de ces données est que les défaillances perceptives des automobilistes impliqués dans des accidents avec deux-roues reposent plus sur un arrière-plan « cognitif » que sur un strict problème de visibilité. Le fait que les deux-roues soient moins saillants et peut être moins « attendus » pourrait ainsi expliquer en grande partie leur non détection par les conducteurs de voiture.

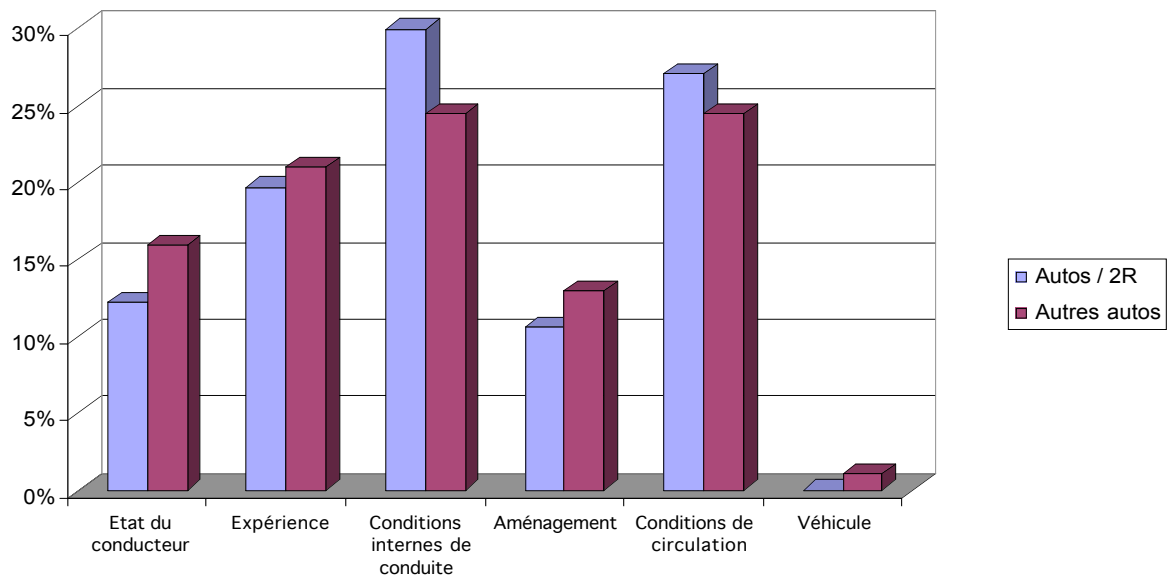


Figure 4. Comparaison des catégories d'éléments explicatifs des défaillances des automobilistes confrontés à un deux-roues (n=195), et d'un échantillon témoin d'automobilistes (n=1126)

Toutefois, la comparaison statistique des variables explicatives des erreurs de ces automobilistes avec celles des automobilistes impliqués dans d'autres types d'accidents ne donne pas de résultats statistiquement significatifs. Ceci tendrait à soutenir l'hypothèse que si les défaillances sont différentes selon que l'on est confronté ou non à un deux-roues, les éléments qui expliquent ces différentes défaillances ont une base commune pour l'ensemble de la population des automobilistes, et se répercuteraient donc plus souvent en un problème de détection dans le cas de la rencontre d'un deux-roues. Une telle hypothèse demande à être explorée, notamment sur un échantillon plus vaste.

4.3 Les accidents de deux-roues véhicule seul

L'échantillon étudié de 80 accidents impliquant un deux-roues comporte 16 cas ne mettant en jeu qu'un seul véhicule. Le faible effectif de ces accidents qualifiés de « pertes de contrôle » ne permet pas d'en faire une analyse représentative, mais révèle cependant un certain nombre d'indications intéressantes. Ces accidents surviennent le plus souvent lors de la rencontre d'un virage, que celui-ci soit ou non connu du conducteur. Les défaillances fonctionnelles qui y correspondent sont globalement homogènes. Elles font principalement intervenir, d'une part *la mauvaise évaluation d'une difficulté de l'infrastructure*, d'autre part une défaillance sensorimotrice de *contrôle dynamique du véhicule* en situation de négociation de trajectoire. Les éléments qui les expliquent sont, eux aussi, assez peu variés et corroborent les facteurs classiquement évoqués à propos des accidents de deux-roues, notamment sur le plan endogène, à savoir : *la faible expérience de la conduite et du véhicule*, le choix d'une *vitesse trop élevée pour la situation*, l'adoption d'une *conduite « à risque »* (recherche de performance). Les

éléments exogènes qui y contribuent tiennent uniquement à des gênes à la visibilité. Un dénominateur commun à cet ensemble d'accidents est de mettre en scène des conducteurs qui roulent aux limites de leurs capacités et se font surprendre par la rencontre d'une difficulté imprévue ou sous-estimée.

5 DISCUSSION

Les conducteurs de deux-roues font montre de défaillances spécifiques, de même que les automobilistes qui se confrontent à eux lors d'accidents de la route. La question qui se pose est celle de l'articulation de ces deux spécificités.

5.1 Des scénarios d'erreurs complémentaires

Les analyses menées en parallèle aux sections 4.1 et 4.2 permettent de vérifier qu'il existe un lien entre les défaillances des automobilistes et celles des conducteurs de deux-roues à moteur. Un travail sur les scénarios-types les plus fréquents pour les deux catégories de véhicules met en lumière ces interactions. Les deux scénarios typiques associés sur la figure 5 montrent bien les correspondances entre les contextes, les comportements, les défaillances associées et leurs conséquences. Ils confirment que l'attachement au statut prioritaire des conducteurs de deux-roues et les niveaux de vitesse qu'ils pratiquent croisent des défaillances de perception des automobilistes. Ils montrent aussi que ces défaillances perceptives des automobilistes s'articulent généralement à des erreurs de pronostic des conducteurs de deux-roues à moteur, surpris par leurs manœuvres.

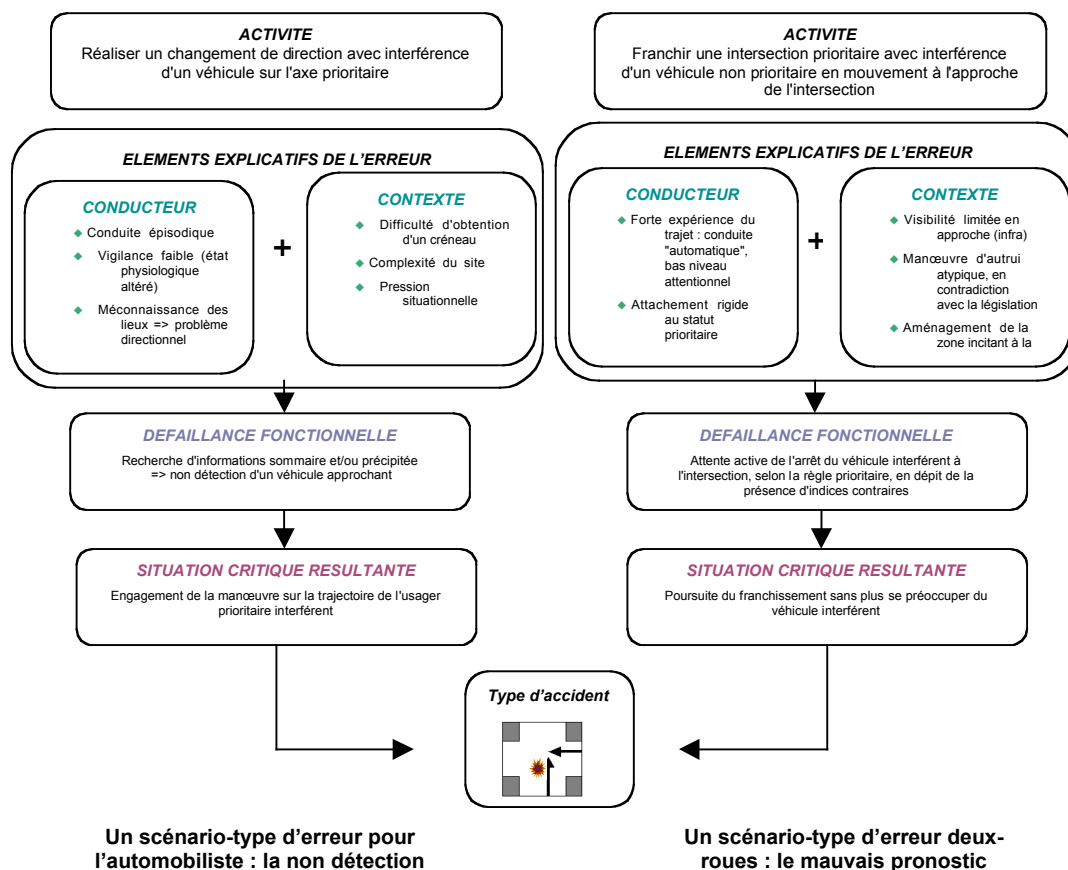


Figure 5. Scénarios-types d'erreurs les plus représentés pour les conducteurs de deux-roues et les automobilistes avec lesquels ils sont impliqués

5.2 Une question d'interactions

Les données qui caractérisent les accidents impliquant des deux-roues à moteur suscitent les réflexions suivantes :

- Ils mettent globalement en jeu des défaillances de pronostic, d'évaluation, et de contrôle de la part des conducteurs deux-roues. Un point commun qui semble sous-tendre ces différentes défaillances est le problème de la confiance que ces conducteurs ont tendance à porter à leurs attentes ainsi qu'à leurs capacités de régulation.
- Pour les accidents impliquant un autre véhicule, les erreurs d'anticipation des conducteurs de deux-roues interviennent souvent en étroite interaction avec une défaillance perceptive de la part des automobilistes. Au-delà de la question de l'inadaptation des stratégies de recherche d'information des automobilistes, ces défaillances de détection soulèvent le problème de la visibilité physique et « sociale » du deux-roues. Visibilité physique, au sens où leur plus faible gabarit peut les rendre moins détectables ; mais aussi, pour certains accidents, au sens où des vitesses d'approche supérieures à la normale peuvent mettre en défaut les prises d'information des automobilistes. « Visibilité sociale », au sens où la probabilité d'interagir avec un deux-roues étant plus faible, les automobilistes s'attendent moins à les rencontrer, les recherchent moins, et par conséquent les détectent moins (cécité inattentionnelle). On pourrait également discuter de l'illusion de visibilité de la part des conducteurs de deux-roues, au sens où ils voient l'autre et pensent être vus en retour de la même manière.
- Ce qui ressort par ailleurs de l'analyse détaillée des accidents, c'est qu'en majorité les deux-roues à moteur sont confrontés à une perturbation dont ils ne sont pas à l'origine, si ce n'est indirectement par le mode de conduite adopté. C'est un événement extérieur, qu'il soit lié à une difficulté de l'infrastructure ou à une interaction conflictuelle avec un autre usager, qui vient surimposer une contrainte supplémentaire aux conditions de réalisation de leur activité. Le problème étant que les caractéristiques du pilotage que ces conducteurs ont tendance à développer les rend inaptes à compenser cette perturbation. Il s'agit en effet souvent d'une conduite à la limite des performances du système, c'est-à-dire à la fois des capacités du pilote, de sa machine, de l'infrastructure, des interactions entre ces composants et des interactions avec les autres usagers de la route. Il en résulte qu'ils ne disposent d'aucune latitude de régulation pour faire face à une sollicitation accrue, d'aucune marge de tolérance pour récupérer un conflit. Et ceci aboutit en fin de compte à un dépassement des compétences lors de la rencontre d'une difficulté imprévue par le conducteur, qu'elle soit liée à l'infrastructure ou au comportement d'autrui, mais qui correspond malgré tout à des événements qui sont susceptibles d'être rencontrés sur la route.

C'est donc une question d'interactions. Bien que communément définie comme un moyen de transport individuel, la conduite automobile place néanmoins souvent les opérateurs en situation, même ponctuelle, de « co-action » (Rogalsky, 1994), de « coopération » (Hoc, 1996), parfois sous forme de compétition. Ces interactions sont sous la dépendance des représentations (sociales et opératives) des conducteurs, qui vont constituer autant de systèmes d'attentes vis-à-vis des situations rencontrées (Mundutégui, 2001). Un accident de la route constitue, non seulement l'issue négative, mais surtout le révélateur d'une mauvaise gestion d'un conflit d'interaction avec les autres usagers et/ou l'infrastructure. La composante interactive de l'activité de conduite et les dysfonctionnements qui peuvent en résulter sont tout spécialement mis en relief dans la confrontation de ces deux groupes que constituent les deux-roues et les automobiles, régis par des caractéristiques, des statuts, des normes, des règles, des représentations, des motivations et des pratiques spécifiques, parfois antinomiques.

6 CONCLUSION

Le recours aux données détaillées d'accidents, qui sont spécifiquement recueillies dans un objectif de recherche, permet de dégager un certain nombre d'indications qui viennent en complément des résultats issus des analyses quantitatives exploitant des fichiers descriptifs issus des procédures de police. Ces dernières analyses mettent en évidence l'importance des risques auxquels sont confrontés les deux-roues à l'échelon national, ainsi qu'un certain nombre de facteurs qui y correspondent. Les données issues des EDA autorisent, quant à elles, d'investiguer des phénomènes plus complexes qui sous-tendent les mécanismes de ces accidents, tels que les « erreurs » et les combinaisons d'éléments qui les expliquent. L'analyse comparative présentée ici, a ainsi permis de montrer que les conducteurs de deux-roues à moteur sont plus enclins à certains types de défaillances fonctionnelles spécifiques,

qui se situent notamment sur le plan de la fonction de pronostic sur l'évolution potentielle des situations. L'analyse des interactions dans la production des accidents permet par ailleurs d'établir que de telles erreurs trouvent souvent leur origine dans un comportement inattendu des automobilistes, lui-même lié au différentiel que les deux-roues introduisent dans le trafic. Le problème des deux-roues est donc à la fois et pas seulement le problème des deux-roues. Il provient d'une part d'un mode de conduite suscité par les caractéristiques de ces véhicules ainsi que par les motivations qui orientent vers leur usage. Il correspond d'autre part aux variations qu'ils engendrent au sein du système routier et qui génèrent des perturbations dans les interactions avec les autres usagers du même espace. Tout accident étant par définition la résultante du dysfonctionnement d'un système, il est essentiel d'éclairer les difficultés d'interaction qui contribuent à sa genèse, d'une part dans l'objectif de mieux comprendre les mécanismes qui le sous-tendent, d'autre part pour tenter de définir des mesures qui tiennent compte de l'ensemble des paramètres en jeu. Les actions visant à réduire l'accidentalité des deux-roues devraient ainsi s'adresser non seulement aux conducteurs de ces véhicules, de façon à susciter l'adoption de plus grandes marges de tolérance dans leur mode de conduite ; elles devraient également concerner les automobilistes dans le sens d'une meilleure intégration de la dimension « deux-roues », notamment dans leurs stratégies de prise d'information.

Au-delà des perspectives d'application, de tels résultats témoignent une fois de plus du caractère primordial de l'interactivité dans la plupart des conduites humaines situées et suscitent un élargissement de l'investigation sous l'angle des modèles psychologiques d'interaction pour l'éclairage qu'ils apportent à l'analyse de l'activité sur cet espace partagé qu'est la route (Saad *et al*, 1999).

7 BIBLIOGRAPHIE

- Amalberti, R. (2001). La maîtrise des situations dynamiques. *Psychologie Française*, 46(2), 107-118.
- Brooks, P., Guppy, A. (1990). Driver awareness and motorcycle accidents. *International motorcycle safety conference proceedings*. Motorcycle safety foundation.
- Caird, J.K., & Hancock, P.A. (1995). The perception of arrival time for different oncoming vehicles at an intersection. *Ecological Psychology*, 6(2), 83-109.
- Chesham, D.J., Rutter, D.R., & Quine, L. (1993). Motorcycling safety research: a review of the social and behavioural literature. *Société de Sciences Médicales*, 37(3), 419-429.
- Chinn, B.P. (1990). *Motorcycle safety*. Transport and Road Research Laboratory. Safety report 91.
- Favero, J.-L., Ferrandez, F., & Fleury, D. (1980). Les accidents de deux-roues : diagnostic et perspectives. *TEC*, 38, 32-38.
- Ferrandez, F. (Ed.) (1995). *L'étude détaillée d'accidents orientée vers la sécurité primaire: méthodologie de recueil et d'analyse*. Paris : Presses de l'Ecole Nationale des Ponts et Chaussées.
- Filou, C. (1998). Constitution d'un fichier des principales caractéristiques des motocycles et mise en relation avec les données d'accidents. (Rapport DERA MACIF/INRETS n°9806). Arcueil, F : INRETS.
- Fleury D., & Brenac T. (2001). Accident prototypical scenarios, a tool for road safety research and diagnostic studies. *Accident Analysis and Prevention*, 33(2), 267-276.
- Grayson, G.B., & Hakkert, A.S. (1987). Accident analysis and conflict behaviour. In T. Rothengatter & R. de Bruin (Eds.), *Road users and traffic safety*. Assen, Netherlands: Van Gorcum.
- Hoc, J.M. (1996) *Supervision et contrôle de processus : la cognition en situation dynamique*. Grenoble : Presses Universitaires de Grenoble.
- Howells, R.A., Buck, J.R., Weiss, S.M., & McMillan, A.G. (1980). *Decision making in intersection entry accidents*. International Motorcycle Safety Conference Proceedings. Volume 4, 1585-1611. Linthicum, Maryland.
- Hurt, H.H., Ouellet, J.V., & Thom, D.R. (1981). *Motorcycle accident cause factors and identification of countermeasures*. Technical report. Traffic Safety Centre, University of Southern California.
- Leplat, J. (1999). Analyse cognitive de l'erreur. *Revue Européenne de Psychologie Appliquée*, 49(1), 31-41.
- Mundutégui, C. (2001). *Reconnaissance d'intention et prédiction d'action pour la gestion des interactions en environnement dynamique*. Thèse, Conservatoire National des Arts et Métiers, Paris.

-
- Obenski, K.S. (1994). *Motorcycle accident reconstruction: understanding motorcycles*. New York: Lawyers and Judges Publishing Co.
- ONISR (2002). *La sécurité routière en France : bilan de l'année 2001*. Paris : La Documentation Française.
- Pibault, C., & Bilman, G. (1997). *Les accidents corporels des deux-roues motorisés à Paris*. Paris: CERET.
- Preusser, D.F., Williams, A.F., & Ulmer, R.G. (1995). Analysis of fatal motorcycle crashes: crash typing. *Accident Analysis and Prevention*, 27 (6), 845-852.
- Rasmussen, J. (1986). A framework for cognitive task analysis in systems design. In E.Hollnagel, G. Mancini, & D.D. Woods (Eds.), *Intelligent decision support in process environments*. Berlin: Springer-Verlag.
- Reason, J. (1993). *L'erreur humaine* (J.M. Hoc, Trad.). Paris: Presses Universitaires de France. (Edition originale, 1990).
- Reason, J. (1995). A model of organizational accident causation. *Ergonomics*, 38(8), 1708-1721.
- Rogalsky, J. (1994). Formation aux activités collectives. *Le Travail Humain*, 57(4), 367-386.
- Rutter, D.R., & Quine, L. (1995). Age and experience in motorcycle safety. *Accident Analysis and Prevention*, 28(1), 15-21.
- Têtard, C. (1994). *Etude approfondie d'accidents impliquant des deux-roues : le cas des motocyclistes*. (Rapport final sur convention). Arcueil, F : INRETS..
- Saad, F., Mundutéguy, C., & Darses F. (1999). Managing interactions between car drivers: an essential dimension of reliable driving. *Proceedings of the Seventh European Conference on Cognitive Science Approaches to Process Control*. Villeneuve d'Asq, F : Presses Universitaires de Valenciennes.
- Van Elslande, P. (2000). L'erreur humaine dans les scénarios d'accident : cause ou conséquence ? *Recherche Transports Sécurité*, 66, 7-33. (Numéro spécial « Psychologie de la conduite »).
- Van Elslande, P. (2003). Erreurs de conduite et besoins d'aide : une approche accidentologique en ergonomie. *Le Travail Humain*, 66(3), 197-224.
- Van Elslande, P., Alberton, L., Nachtergaële, C., & Blancher, G. (1997). *Scénarios-types de production de « l'erreur humaine » dans l'accident de la route*. (Rapport de recherche No.218). Arcueil, F : INRETS.

Un système d'observation et d'analyse en direct de séances d'enseignement

Éric Allègre

GPE-CREFI, 5 allées A.-Machado, Université Toulouse-II-Le-Mirail
31058 Toulouse CEDEX 1
EAllegre@aol.com

Philippe Dessus

LSE, 1251 av. Centrale, BP 47, Université Pierre-Mendès-France
38040 Grenoble CEDEX 9
Philippe.Dessus@upmf-grenoble.fr

RESUME

Nous présentons les fonctionnalités et les premiers tests d'un système d'observation et d'analyse de séances d'enseignement, en direct. Nous montrons que ce système prend mieux en compte certains paramètres importants des situations scolaires que les méthodes d'observation classiques, comme leur temporalité complexe et le fait que les actions de l'enseignant et des élèves ne sont pas seulement dépendantes de l'action qui précède.

MOTS-CLES

Situation dynamique — temporalité — analyse quantitative — analyse de la sémantique latente

1 INTRODUCTION

L'objet de cet article est de présenter un système d'observation et d'analyse d'une activité complexe, l'enseignement, en utilisant le cadre théorique de la psychologie ergonomique (Rogalski, 2000). Nous nous référons, pour définir ce qu'est une activité complexe, à Funke (2001). Il décrit deux caractéristiques typiques de ce genre de problèmes : — la connectivité entre variables, — la dynamique du processus supervisé. Ces caractéristiques concernent également un environnement scolaire. Doyle (1986), rend ainsi compte de caractéristiques de situations d'enseignement : — *multidimensionnalité*, de très nombreux événements surviennent et de très nombreuses tâches sont à effectuer ; — *simultanéité*, de nombreux événements peuvent survenir simultanément ; — *immédiateté*, les événements surviennent de plus à un rythme rapide ; — *imprévisibilité*, les événements peuvent avoir une issue non prévue par l'enseignant, les interruptions et diversions sont fréquentes. Les deux premiers points renvoient à la connectivité, les deux derniers à la dynamique de la situation.

Le fait que ces caractéristiques aient été exprimées n'a pas conduit les chercheurs en éducation à se doter de méthodes d'observation et d'analyse de situations scolaires qui leur soient compatibles. La quasi-totalité de ces méthodes sont en directe ligne de la grille de Flanders (1976), qui part du principe que les interactions entre enseignant et élèves s'*enchaînent*, c'est-à-dire qu'elles ont une forte dépendance deux à deux. Même si des méthodes plus récentes ont été élaborées (*e.g.*, la grille d'organisation et de gestion pédagogique, Altet *et al.*, 1996), elles restent très inspirées de la grille de Flanders et sont loin de prendre en compte les caractéristiques énoncées ci-dessus. Il existe d'autres dispositifs qui, en s'inspirant de démarches ergonomiques issues de la cognition située (*e.g.*, Clot *et al.*, 2001) mettent plus l'accent sur l'action en situation : la séquence est filmée et on effectue une analyse qualitative de quelques segments significatifs, choisis par le chercheur. Ici, l'analyse de la séance est faite en profondeur, et très rarement dans son ensemble, à cause de la quantité de l'information récupérée.

Ces différentes méthodes ont deux inconvénients : elles ne considèrent pas que les différents événements observés se déroulent dans des temporalités multiples et, de plus, elles ne supposent pas que les actions de l'enseignant non contiguës sont interreliées, ce qui empêche l'analyse d'éventuelles régularités de l'action de l'enseignant, à un niveau plus large de description (*e.g.*, l'épisode, le plus petit ensemble d'événements concourant à remplir un objectif d'enseignement). Notre approche, tout

en essayant de pallier ces deux inconvénients, se situe à mi-chemin des deux précédentes : nous ne récupérons pas un grand nombre d'informations difficilement analysables (vidéo), préférant passer par une phase de tri des événements sur-le-champ, par l'observateur. En revanche, nous essayons de nous assurer que les informations récupérées sont les plus objectives possibles et qu'elles peuvent faire l'objet d'un traitement quantitatif à un niveau suffisamment proche des événements observés. L'objet de cet article est donc de proposer une méthode d'observation, de recueil et d'analyse automatisée des événements d'enseignement et d'apprentissage qui tienne compte de ces caractéristiques.

2 FONCTIONNEMENT GENERAL DU SYSTEME

Une manière de tenir compte des spécificités évoquées *supra* sera de proposer une méthode, assistée par ordinateur, qui ait les caractéristiques suivantes : — prise en compte des événements de classe à un niveau de grain suffisamment précis, les épisodes ; — saisie facilitée des événements observés, *via* des raccourcis-clavier correspondant à des épisodes prédéfinis ; — gestion de la temporalité multiple des variables observées, *via* des chronomètres multiples, au niveau de chaque élève ; — possibilité de considérer les relations entre chaque événement et entre chaque épisode. Cette dernière fonctionnalité, non présente dans la première version du système, est réalisée par le biais d'une méthode d'analyse factorielle que nous décrirons plus bas.

Voici dans quel contexte l'observateur travaille. Posté au fond de la classe devant un ordinateur, il utilise un logiciel dans lequel il saisit des codes prédéfinis correspondant à des événements de classe, segmente les différents événements en différents épisodes. Afin d'augmenter la validité écologique de l'observation, il n'intervient en aucune manière dans le déroulement de l'activité de l'enseignant ou des élèves, ni en ce qui concerne le choix des contenus. Pour les mêmes raisons, l'enseignant n'est pas informé de l'objet sur lequel porte l'observation : l'enseignant observé ne voit qu'un chercheur prenant des notes en direct à partir d'un logiciel de traitement de textes courant. Le fait de disposer de certains codes prédéfinis va lui faciliter la prise de notes en temps réel, tout en continuant d'observer ce qu'il se passe, et permettre de manipuler un grand nombre de chronomètres, afin de noter les plages de recouvrement des différents événements. Ainsi, le niveau de grain des observations est assez fin pour permettre la saisie d'événements rapides, comme les échanges maître-élève. À titre de comparaison, le grain des événements observés dans d'autres études vont de 4 à 5 min pour une séance d'une heure à 5-10 s pour une séance de 17 min (*e.g.*, Berdot *et al.*, 2003). Les différentes données recueillies sont enregistrées dans un fichier, pour analyses ultérieures.

Si l'observateur est assisté dans la gestion des différents chronomètres, c'est tout de même lui qui va réaliser deux tâches importantes, requérant des prises de décision : le découpage en épisodes et la dénomination d'un événement. *Le découpage en épisodes* nécessite de prendre une décision quant à la continuité des événements (liés à l'enseignant ou aux élèves). En effet, si l'on imagine une séance de mathématiques, dans laquelle les élèves vont passer à un épisode de recherche (dans *Look-Cum 1*, appui sur la touche *R*), c'est bien à l'observateur de décider à quel moment précis cet épisode va débiter et finir. L'événement déclencheur de cet appui de touche requiert bien une réponse à la question suivante : « quand commencent-ils à chercher ? » Et les réponses peuvent être multiples : Quand ils ont fini de lire ; quand ils ont découpé et collé l'énoncé ; quand l'enseignant a interrogé tel ou tel élève ; quand un silence s'installe ; quand l'enseignant commence à passer dans les rangs, etc. *La dénomination d'un événement*, à notre avis, est facilitée si on la compare à celle des grilles d'observation classiques, où des indices très subjectifs sont évalués (« accepte les sentiments », « louanges ou encouragements », etc.). Toutefois, certaines distinctions entre événements sont à faire par l'observateur : par exemple, lorsque l'enseignant énonce une consigne d'exercice ou bien lorsqu'il pose une question à toute la classe. Passons à la description de notre système.

3 DESCRIPTION DE LOOK-CUM 1

3.1 Architecture et fonctionnalités de *Look-Cum 1*

Look-Cum 1 (Maurice & Allègre, 2002) est une macro-commande du logiciel *WordPerfect* en *Visual Basic*. Elle permet très simplement d'horodater, à la seconde près, toute séquence de touches saisie par l'observateur, et de l'écrire dans le « fichier espion ». De plus, si la séquence correspond à un événement prédéfini, c'est ce dernier qui va s'écrire dans le fichier. Pour faciliter la lecture, les

indications « début » et « fin » sont automatiquement ajoutées à chaque paire identique de séquences. De plus, diverses informations élémentaires sur la séance observée peuvent être recueillies par le système : la date et l'heure de l'observation, le nom de l'enseignant, l'effectif des élèves présents et absents, la matière ainsi que le contenu de la séance et toute autre observation nécessaire dans la mesure où tout peut être noté et horodaté. Ces différentes données sont rassemblées dans une base de données, pour un accès ultérieur. Certains événements sont déjà répertoriés dans *Look-Cum 1*. L'observateur peut les noter en un appui de touche (*i.e.*, *M* code « l'enseignant regarde sa montre », *T* code « l'enseignant écrit au tableau »). Par ailleurs, chaque élève est repéré par un numéro et chaque événement (*e.g.*, une question posée, un déplacement de l'enseignant vers un élève, un envoi d'élève au tableau, ou qu'il est le protagoniste d'un quelconque événement, cela est noté par l'observateur et horodaté par le système (*voir Tableau I*). Le fait d'avoir recours à une macro-commande requiert un délai incompressible de l'ordre d'une seconde, ce qui n'est pas gênant du fait de la temporalité des événements scolaires, au-dessus de cette limite.

À la fin de l'observation, tous les événements connus sont énumérés par ordre chronologique, avec les remarques éventuelles de l'observateur ainsi que les durées locales et cumulées calculées par le système (*voir Tableau II*). À la fin du document figurent également quelques statistiques : pourcentages des différents épisodes (*i.e.*, énoncé, recherche, correction), des durées de parole, par rapport à la durée totale de la séance ou de la séquence. Ce logiciel peut permettre de rendre compte d'un certain nombre de phénomènes scolaires : d'une part la répartition des durées de parole et d'action de l'enseignant et des élèves (intra- ou interépisodes) ; d'autre part la détermination d'un « élève-repéré » (*i.e.*, le plus supervisé, en fréquence ou en durée), en sommant les durées de visite par l'enseignant de chaque élève.

Tableau I — Trace du système Look-Cum 1 à propos d'une séance de mathématiques (extrait)

Trace	Commentaires
Synthèse début = 40788	Temps de début de la synthèse du problème
Remarque : on met au tableau toutes les procédures qui ont permis de répondre = 40806 ;	L'enseignant affiche au tableau les processus de résolution utilisés par quelques élèves repérés pendant qu'il passait dans les rangs durant la recherche
tdi = 40810 ; tdt = 40812 ;	tdi/tfi = temps de début/fin d'intervention orale de l'enseignant
Remarque : se pose alors le problème de 104 soulevé par le dernier élève = 40893 ;	
tft = 40961 (loc = 149) ;	tdt/tft = temps de début/fin d'intervention écrite de l'enseignant au tableau
tfi = 40964 (loc = 154 ; cum = 599) ;	loc/cum = durée locale/cumulée de visite de l'enseignant
Remarque : on arrive aux fractions = 40994 ;	
tdi = 41007 ; tfi = 41009 (loc = 2 ; cum = 601) ;	
Synthèse fin = 41012 (loc = 224 ; cum = 224)	Temps de fin de la synthèse

Tableau II — Données calculées (extrait) par Look-Cum 1 à partir du traitement d'une séance observée

Données calculées	Commentaires
Global	Durée totale de la séance
Durée globale = 2475	
Énoncé	Données concernant l'épisode « énoncé »
Durée d'énoncé = 197 (%glo = 7,96)	%glo = % du temps global
Durée de consigne propre à l'énoncé = 189 (%glo = 7,64) (%eno = 95,94) [...]	%eno = % du temps d'énoncé
Durée de parole enseignant = 189 (%glo = 7,64) (%eno = 95,94)	
Durée de parole élève = 0 (%glo = 0,0) (%eno = 0,0)	
Recherche	Données concernant l'épisode « recherche »
Durée de recherche = 1094 (%glo = 44,20)	%rch = % du temps de recherche
Durée d'aide collective = 0 (%glo = 0,0) (%rch = 0,0)	
Durée d'aide individuelle = 933 (%glo = 37,70) (%rch = 85,28)	
Durée de parole maître = 466 (%glo = 18,82) (%rch = 42,56)	
Durée de parole élève = 466 (%glo = 18,82) (%rch = 42,60)	

3.2 Test de *Look-Cum 1*

Le test de la première version du système a été réalisé par l'observation de 260 séances dans sept classes de cycle 3 du primaire, durant deux années. Ce premier test permet de préciser de quelle manière les enseignants prennent en charge la temporalité de l'environnement qu'ils supervisent. Un même enseignant laisse toujours la même proportion de temps à ses élèves pour chercher une solution à une situation-problème, quelle que soit la discipline (mathématiques, français ou histoire-géographie sciences) et quelle que soit la durée du problème (de 4 min 30 à 64 min). En revanche, cette proportion est différente selon les enseignants (Maurice & Allègre, 2002). De plus, une étude plus précise révèle que les enseignants regardent plus fréquemment leur montre lorsque la situation qu'ils supervisent présente une nouveauté pour eux. Présentons maintenant la deuxième version de notre système, en cours de développement, qui ajoute une fonctionnalité d'analyse statistique de données.

4 DESCRIPTION DE *LOOK-CUM 2*

La version 2 reprendra les fonctionnalités de *Look-Cum 1*, et comprendra de plus un moteur d'inférences fondé sur une méthode d'analyse statistique de grands corpus textuels, l'analyse de la sémantique latente. L'idée principale est de permettre un repérage des régularités des événements observés en classe. Il faut noter que cette analyse se fait avec les données saisies qui ne représentent pas des durées. Dans les sections suivantes, nous décrivons cette méthode, puis ses applications dans le domaine qui nous concerne.

4.1 Présentation générale de la méthode

L'analyse de la sémantique latente (LSA, *Latent Semantic Analysis*) est une méthode simulant l'acquisition de connaissances et de la compréhension humaines (Landauer & Dumais, 1997). Son principe est de définir statistiquement le sens d'un mot à partir de l'ensemble des contextes (*i.e.*, paragraphes, phrases) dans lesquels ce mot apparaît. Par exemple, le mot *avion* va apparaître souvent conjointement à des mots comme *décoller*, *aéroport* et très peu fréquemment à des mots comme *dinosaure*. Cependant, cette information statistique sur le contexte d'un mot n'est pas suffisante pour en définir le sens, puisqu'elle ne dit rien quant aux liens sémantiques avec tous les autres mots n'apparaissant jamais conjointement à ce mot : les informations de co-occurrence propres à chaque mot doivent être croisées. En d'autres termes, LSA repose sur la définition suivante : deux mots sont similaires s'ils apparaissent dans des contextes similaires. Deux contextes sont similaires s'ils comportent des mots similaires. Les vecteurs des différents mots du corpus sont projetés dans un espace d'environ 300 dimensions (axes), constitué à partir d'une réduction de la matrice mots/contextes initiale par une décomposition aux valeurs singulières. Dans le domaine qui nous concerne ici, les mots seront des événements de classe et les paragraphes les épisodes dans lesquels ces mots apparaissent. Les comparaisons interépisodes et interévénements sont simplement réalisées en calculant le cosinus des vecteurs représentant deux épisodes ou deux événements : des épisodes ou événements similaires auront une valeur de proximité égale à 1, des épisodes indépendants une proximité de 0, des épisodes opposés une valeur théorique de -1. Le principe de cette méthode est celui à partir duquel sont établies les différentes méthodes d'analyse factorielle — analyse en composantes principales, analyse des correspondances (Lebart, Morineau, & Piron, 1997), à ceci près que le grand nombre d'axes conservés rend leur interprétation impossible. Son avantage est qu'elle permet de traiter des matrices mot/contexte très creuses (comprenant de nombreux 0), ce qui est le cas ici.

4.2 Applications de la méthode à l'analyse d'environnements dynamiques

L'analyse de la sémantique latente a récemment été utilisée pour rendre compte de l'activité de supervision d'environnements dynamiques. Quesada et ses collègues ont mené une série d'études (Quesada, Kintsch, & Gomez, 2001, 2002) validant la méthode de LSA à ce champ de travail. Ils ont ré-analysé avec LSA les données provenant de l'activité de sujets dans le simulateur d'environnement dynamique *Firechief*. Les sujets ont pour tâche d'éteindre le plus rapidement possible un feu de forêt, en agissant de trois manières différentes : — jeter de l'eau sur un segment de terrain donné ; — contrôler le feu sur un segment de terrain ; — effectuer un déplacement avec un appareil. Certains paramètres, en influant sur la vitesse et la direction de propagation du feu, créent le caractère

dynamique de ce phénomène. Leur travail a consisté à ré-analyser, avec LSA, les « fichiers espions » de 3 400 sessions comportant au total 360 000 actions environ. Cette méthode est capable, d'une part, de relever que des actions proches, bien que formellement différentes, ont le même but puisque mises en œuvre dans des contextes similaires. À l'inverse, des actions formellement proches ont été mesurées par LSA comme éloignées, car mises en œuvre dans des contextes, et pour remplir des buts, différents. D'autre part, une comparaison entre des jugements humains et les évaluations de LSA montrent une adéquation de 57 %. Notons que LSA traite les données liées à un épisode « paquets » : aucune prise en compte de l'ordre des actions à l'intérieur d'un épisode n'est effectuée, ce qui peut être, *a priori*, un inconvénient majeur. Toutefois, les nombreux tests effectués sur différentes données montrent que cet inconvénient est atténué, pour peu que le corpus traité soit important. De plus, comme dans notre système, LSA effectue un traitement conjointement à une analyse temporelle des données, cet inconvénient est encore diminué. Les tests réalisés à ce jour par Quesada et ses collègues sont réalisés dans des micro-mondes, présentant des simulations d'environnements dynamiques. Nous essayons ici de répliquer leurs résultats en nous centrant sur un environnement scolaire réel.

4.3 Architecture et fonctionnalités de Look-Cum 2

Comme déjà signalé, l'interface de *Look-Cum 2* sera identique à celle de la version 1. La seule différence est que les différentes données saisies par l'observateur seront analysées en temps réel par LSA. Ce dernier introduit régulièrement les événements saisis dans l'espace multidimensionnel, comme autant de vecteurs placés dans l'espace, sans pour autant le recalculer en entier. Cela permet ainsi leur analyse en temps réel et *a posteriori*. L'analyse en temps réel permet d'afficher la liste des épisodes atypiques d'une séance, c'est-à-dire les épisodes qui n'entretiennent pas une grande proximité avec les autres épisodes de même propos. L'analyse *a posteriori* permet de comparer les séances entre elles, et de détecter d'éventuelles similitudes, notamment entre séances menées par des novices et des experts. Ces deux types d'analyse peuvent avoir un intérêt dans un contexte de formation d'enseignants.

4.4 Test du moteur de Look-Cum 2

Nous avons recodé (voir *Tableau III*) et réanalysé avec LSA 41 séances de mathématiques réalisées par le même enseignant (soit un corpus d'environ 2 300 événements dans 123 épisodes, énoncé, recherche ou correction). Détaillons deux types d'analyse, tout d'abord au niveau des épisodes, puis des événements. Nous avons comparé chaque épisode avec tous les autres et conservé les valeurs supérieures à un écart type de la moyenne des proximités interépisodes. Dans la quasi-totalité des cas (trois exceptions sur près de 3 400 comparaisons), les épisodes sont bien de même catégorie, ce qui signifie que LSA est capable de discriminer, par la seule analyse des événements les composant, les trois types d'épisodes. Enfin, en utilisant le même critère de seuil, nous avons comparé chaque événement avec chacun des 18 autres événements observés, ce qui nous permet de les organiser en un réseau de proximité. Deux réseaux principaux ont été mis au jour, reliés par l'événement le plus proche de tous les autres, *consulte_montre*. Ce qui montre combien la gestion du temps est centrale dans l'activité d'enseigner. Un réseau comprend les événements liés à la gestion du contenu, un autre a trait à la gestion de la classe (voir *tableau IV*).

5 DISCUSSION

Look-Cum permet de recueillir, en direct, des informations concernant la dynamique temporelle des événements d'enseignement et d'apprentissage. Les premiers tests de ce système montrent qu'il est capable de rendre compte, d'une part, de la gestion du temps en classe, à un niveau de grain tel que des régularités temporelles indépendantes du contenu enseigné ont pu être montrées. D'autre part, l'ajout à ce système d'une méthode d'analyse statistique analysant les co-occurrences d'événements au sein d'épisodes a dévoilé la manière dont ils surviennent selon les protagonistes de la situation, et donc leurs finalités. Ainsi, un tel système permet une analyse prenant en compte la complexité des multiples événements d'un environnement scolaire.

Tableau III — Liste partielle d'événements recodés à partir des données de Look-Cum 1

Événement	Signification de l'action
demande_calme	L'enseignant demande du calme
visite_eleve[n]	L'enseignant se rend auprès de l'élève <i>n</i>
aide_collective	L'enseignant donne une indication collective
leve_doigt[n]	L'élève <i>n</i> lève le doigt
reponse_d_eleve[n]	L'élève <i>n</i> répond à l'enseignant
synthese_locale	L'enseignant résume les informations à propos d'un exercice à toute la classe
consulte_montre	L'enseignant consulte sa montre
son	Un élève non déterminé fait du bruit

Tableau IV — Réseau de proximités entre événements, tous reliés à *consulte_montre*. Les valeurs sont entre parenthèses, toutes supérieures au seuil fixé

Gestion de la classe	Événement de référence	Gestion du contenu
demande_calme (0,49)		
leve_doigt[i] (0,50)		synthese_locale (0,45)
aide_collective (0,51)	consulte_montre	reponse_d_eleve (0,47)
son (0,55)		
visite_eleve[i] (0,59)		

Remerciements

Nous remercions Jean-Jacques Maurice pour sa participation aux différents tests de *Look-Cum*, ainsi que Jacques Baillé et Erica de Vries pour leur commentaires d'une version précédente de cet article.

6 BIBLIOGRAPHIE

- Altet, M., Bressoux, P., Bru, M., & Lambert, C. (1996). Étude exploratoire des pratiques d'enseignement en classe de CE2. *Les dossiers d'éducation et formations*, 70.
- Berdot, P., Blanchard-Laville, C., & Chaussecourte, P. (2003). Analyse clinique. In C. Blanchard-Laville (Ed.), *Une séance de cours ordinaire* (pp. 159-198). Paris : L'Harmattan.
- Clot, Y., Faïta, D., Fernandez, G., & Scheller, L. (2001). Entretiens en autoconfrontation croisée : une méthode en clinique de l'activité. *Education Permanente*, 146, 17-25.
- Doyle, W. (1986). Classroom organization and management. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 392-431). New York : McMillan.
- Flanders, N. A. (1976). Analyse de l'interaction et formation. In A. Morrison & D. McIntyre (Eds.), *Psychologie sociale de l'enseignement* (Tome 1, pp. 57-69). Paris : Dunod.
- Funke, J. (2001). Dynamic systems as tools for analysing human judgement. *Thinking and Reasoning*, 7(1), 69-89.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem : the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge. *Psychological Review*, 104, 211-240.
- Lebart, L., Morineau, A., Piron, M. (1997). *Statistique exploratoire multidimensionnelle*. Paris : Dunod.
- Maurice, J.-J., & Allègre, E. (2002). Invariance temporelle des pratiques enseignantes : le temps donné aux élèves pour chercher. *Revue Française de Pédagogie*, 138, 115-124.
- Quesada, J., Kintsch, W., & Gomez, E. (2001). A computational theory of complex problem solving using the vector space model (part II). *Proc. Conf. Cognitive Res. with Microworlds*. Grenade.
- Quesada, J., Kintsch, W., & Gomez, E. (2002). A computational theory of complex problem solving using Latent Semantic Analysis. In W. D. Gray & C. D. Schunn (Eds.), *Proc. 24th Ann. Conf. of the Cognitive Science Society* (pp. 750-755). Mahwah : Erlbaum.
- Rogalski, J. (2000). Approche de psychologie ergonomique de l'activité de l'enseignant. *Actes du XXVI^e Colloque COPIRELEM* (pp. 45-66). Limoges : IREM de Limoges.

Effet de la modalité d'interaction avec un système de dialogue en langue naturelle

Ludovic Le Bigot

France Télécom Recherche & Développement DIH/D2I

ludovic.lebigot@francetelecom.com

Eric Jamet

CRPCC, Université de Rennes II

eric.jamet@uhb.fr

Jean-François Rouet

LaCo, CNRS, Université de Poitiers

jean-francois.rouet@univ-poitiers.fr

RESUME

Ce papier propose d'étudier l'effet de modalité lors d'une interaction avec un système de dialogue en langue naturelle dans une tâche de recherche d'informations. Vingt-quatre participants adultes ont effectué douze tâches de recherche d'informations avec un système informatisé de dialogue en langage naturel, en communiquant avec lui soit à l'écrit soit à l'oral. L'analyse des résultats a montré un effet de la modalité sur les performances et sur l'utilisation d'indicateurs discursifs. Le nombre de tours de parole (TP), la durée de l'interaction et la charge de travail mental perçue étaient plus faibles à l'écrit qu'à l'oral. De plus, le nombre moyen de mots par TP et la fréquence du pronom-sujet "Je" par dialogue étaient plus importants à l'oral lors des premières tâches. Les implications de l'effet de modalité pour le développement de ce genre de système sont brièvement discutées.

MOTS-CLES

Modalité, dialogue, langue naturelle, IHM

1 INTRODUCTION

1.1 Le dialogue en langue naturelle

Aujourd'hui, il est possible d'être assisté dans des tâches de recherche d'informations en dialoguant avec des systèmes coopératifs en langage naturel. Un tel système à base de technologie Artimis a été développé à France Télécom R&D (Sadek, Bretier & Panaget, 1997). Ce type de dialogue permet à l'utilisateur de formuler des requêtes en langue naturelle sous n'importe quelle forme, et dans la mesure du possible le système doit se comporter comme un opérateur humain (pour un exemple, cf. encart 1). La difficulté pour cerner les mécanismes cognitifs que sous-tendent le dialogue et la recherche d'informations en langage naturel réside principalement dans leur multiplicité et leur entrecroisement lors de l'interaction. Dans une étude précédente, Le Bigot, Jamet & Rouet (soumis) ont montré l'existence d'un effet d'apprentissage et de modalité dans ce genre d'interaction pour des individus plutôt familiarisés avec l'informatique. Dans une tâche de recherche d'informations avec un système de dialogue en langue naturelle, la performance (en termes de nombre de TP, temps d'interaction totale, charge mentale...) s'améliorait au fur et à mesure des échanges avec le système. De plus, l'interaction à l'oral était plus consistante qu'à l'écrit (meilleure performance à l'écrit, et structure discursive différente entre les deux modalités). L'étude proposée ici tente d'élargir et de compléter les résultats sur l'effet de modalité.

Depuis Turing, l'intérêt porté à la communication avec un système, et plus particulièrement en langue naturelle, n'a cessé de croître. Les communications médiatisées *avec (ou par)* un système ont été depuis plusieurs décennies étudiées en utilisant le paradigme du magicien d'Oz (e.g., Bubb-Lewis & Scerbo, 2002 ; Ferrara, Brunner & Whittemore, 1991 ; Oviatt, 1995 ; Whittaker, Walker, & Moore, 2002 ; Zoltan-Ford, 1991...). Néanmoins, en plus des contraintes conversationnelles intrinsèques à la

communication humaine (Whittaker, 2003), il existe de fortes contraintes liées au mode de communication (e.g., Brennan & Schober, 1998). Par exemple, Brennan (1991) a montré que le simple fait de converser avec une machine transformait la structure des énoncés produits par les utilisateurs. Elle a constaté que les utilisateurs s'adaptent à leur partenaire interactif en concevant des requêtes se rapprochant du type de réponses du système.

Bubb-Lewis et Scerbo (2002) ont étudié l'usage d'un système interactif adaptatif dans des tâches de planification de trajets. Ils ont observé des différences de performance maximales entre les groupes qui pouvaient communiquer librement et ceux où la communication était restreinte ou nulle. À mesure que la restriction augmentait, la performance et la verbosité des participants diminuaient. Ils ont conclu que lorsque des utilisateurs échangeaient avec un système, ils avaient tendance à se focaliser sur la tâche. Les différents modes d'interaction (de contrôlé à libre) ont affecté la disposition de l'humain à recevoir le feedback du système, et par conséquent la structure du discours et la performance. Bien qu'un groupe contrôle de communication homme / homme n'ait pas été inclus dans leur étude, Bubb-Lewis et Scerbo ont suggéré que certains des processus, qui rendent la communication humaine si efficace, sont absents dans la communication homme / système sous sa forme actuelle.

Encart 5 : Exemple de dialogue oral avec l'application PlanResto (les critères peuvent être donnés en une fois)

<p>Système : Bienvenue sur PlanResto [...] Que désirez-vous ?</p> <p><i>Utilisateur</i> : Je désire un restaurant dans le onzième arrondissement à Paris</p> <p>Système : Vous recherchez un restaurant dans le 11e arrondissement. J'en ai trouvé plus de 10. Vous pouvez donner plus de précisions [...]</p> <p><i>Utilisateur</i> : alors un restaurant indien</p> <p>Système : Vous recherchez un restaurant indien dans le 11e arrondissement. J'en ai trouvé 7. Vous pouvez donner plus de précisions, par exemple le lieu, consulter les solutions ou effectuer une nouvelle recherche.</p> <p><i>Utilisateur</i> : un restaurant pour vingt-cinq euros</p> <p>Système : Vous recherchez un restaurant indien pour environ vingt-cinq euros dans le 11e arrondissement. J'en ai trouvé un qui s'appelle "Le Village de Gannath". Il propose des repas à environ vingt-quatre euros le midi et des repas à environ vingt-quatre euros le soir. Désirez-vous plus d'informations ou effectuer une nouvelle recherche?</p> <p><i>Utilisateur</i> : Est-ce que vous pouvez répéter le nom du restaurant</p> <p>Système : ...</p>
--

1.2 La distinction entre écrit et oral

La modalité (oral ou écrit) d'un dialogue induit des contraintes spécifiques pour l'individu. Par exemple, les phrases sont plus longues et élaborées à l'écrit, mais le nombre total de mots est plus important à l'oral (e.g., DeVito, 1966, 1967 ; Drieman, 1962 ; Fraisse & Breyton, 1959 ; Poole & Field, 1976,...). Oviatt (1995) a re-analysé trois expériences de simulation d'un système de service de transaction avec un input oral / écrit (verbal, alpha-numérique, tablette graphique) en s'intéressant aux discontinuités du discours oral et écrit. Elle a conclu que ces discontinuités étaient très sensibles aux sollicitations cognitives et à la génération d'énoncés longs (dans le sens du coût de la macroplanification, Levelt, 1989). De plus, elle a interprété les différences entre écrit et oral par le fait que les processus cognitifs liés le sont de manières différentes. Des auteurs (e.g., Bereiter, 1980) ont soutenu que le langage écrit était plus complexe que le langage oral, ou bien que le langage oral était le langage ordinaire de la conversation (Olson, 1977, cité dans Hidi & Hildyard, 1983), que le texte était destiné avant tout à l'argumentation ou à la prose expositive. Le langage conversationnel serait spécialisé pour les situations sociales et serait sensible au contexte basé sur la compréhension réciproque des interlocuteurs (Brennan & Clark, 1996 ; Brennan & Schober, 1998 ; Clark, 1996, 1997). Le langage oral resterait sensible aux demandes de la mémoire et à l'élaboration de stratégies de mémorisation (Sachs, 1974), tandis que le langage écrit requerrait d'autres spécialisations mémorielles. Par exemple, la présence continue des mots à l'écrit réduirait la charge de mémoire (Hidi & Hildyard, 1983). L'établissement de relations entre les concepts serait ainsi facilité.

Aux différences entre oral et écrit viennent s'ajouter les contraintes technologiques de l'interaction avec un système. Par exemple, Ferrara, Whittemore & Bruner (1991) ont noté une

caractéristique des difficultés imposées par la technologie, lors d'interactions communicatives écrites synchrones à travers un système interactif, par l'omission des copulatifs, des pronoms-sujets et des articles (voir aussi McCarthy & Monk, 1994 pour le traitement des pronoms-sujets). Dans le même sens, Falzon (1989) a fait le constat que la majorité des travaux sur le dialogue en langue naturelle mettaient en évidence l'apparition de restrictions de la langue. Néanmoins, les résultats devaient être relativisés car ils portaient pour la plupart sur l'interaction écrite (taper), la nature de cette dernière favorisant l'apparition de ces comportements. Dans le cas d'interactions orales, des résultats différents seraient certainement observés (Falzon, 1989, p.37). Brennan (1991) a montré que les pronoms à la 3^{ème} personne (he, she, him, her, his, it), dont la fonction est de connecter les TP et les références, étaient utilisés dans les mêmes proportions en dialogue homme / homme et en dialogue homme / machine (DHM). Au contraire, l'occurrence des pronoms à la 1^{ère} et à la 2^e personne (I, you, me) diminuait pour le DHM du fait de leur fonction souvent méta-cognitive (apparition dans des questions indirectes, typiquement plus polies et correspondant au contexte social). Cependant, Brennan a aussi montré que même si les individus traitaient les partenaires machines différemment des partenaires humains, les attentes concernant l'interaction conversationnelle continuaient à jouer un rôle significatif dans le DHM en langage naturel. Dans une autre expérience, Brennan & Ohaeri (1999) ont constaté que les individus se sentaient mieux dans une communication médiatisée par ordinateur lorsqu'ils parlaient que lorsqu'ils tapaient. En définitive, l'écrit pourrait-il être vu comme un support de dialogue analogue à l'oral lors d'une interaction avec un système en langage naturel ?

1.3 Objectif

Les difficultés liées aux dialogues avec un système ont été étudiées aussi bien du point de vue de la production verbale (e.g., restriction de la langue, planification) que de la compréhension (e.g., compréhension des feedbacks), et des processus mêmes du dialogue (e.g., grounding dans les IHM) mais rarement en manipulant la modalité d'interaction. L'objectif de cette étude visait à déterminer les processus sous-tendus par l'utilisation des différentes modalités d'interaction avec le système lors d'une recherche d'informations. En raison des contraintes cognitives plus fortes (exigences attentionnelles, temps de planification réduit), le dialogue utilisateur-système devrait être plus difficile à l'oral qu'à l'écrit. Ces difficultés devraient se manifester aussi bien lors du déroulement de la tâche (TP, temps d'exécution de la tâche) qu'en terme de charge mentale perçue par l'utilisateur. Par ailleurs, l'hypothèse a été faite que l'interaction orale maintiendrait davantage les indicateurs classiques du dialogue que l'interaction écrite.

2 METHODE

2.1 Sujets

Les participants étaient au nombre de 24 (16 hommes, 8 femmes). L'âge moyen était de 26.5 ans ($E.T.=3.90$). Deux tiers des participants étaient plutôt familiarisés avec la technologie, le dernier tiers était plutôt peu ou pas familiarisé dans ce domaine. Le niveau d'étude était assez élevé (en moyenne 5.2 années d'étude après le bac, $E.T.=1.59$). Il faut noter que les participants se sont définis eux-même avec un niveau en informatique moyen ($M=3.54$, $E.T.=0.88$, sur une échelle en 5 points de très faible à fort), et qu'ils utilisaient tous assez régulièrement l'outil informatique (courrier électronique, Internet et ordinateur). Les participants ont été répartis de manière équivalente dans les deux groupes expérimentaux. Une autre caractéristique était qu'aucun des participants n'employait régulièrement des services interactifs en ligne ou par téléphone.

2.2 Matériel

Le système *PlanResto* a été utilisé dans cette expérience. *PlanResto* est un prototype de service Grand Public, à base de synthèse et de reconnaissance vocale, et de dialogue intelligent en langage naturel pour la recherche de restaurants sur Paris, utilisant la technologie Artimis. L'utilisateur pouvait effectuer une recherche avec 3 critères : emplacement du restaurant, prix, spécialité. Le système lui proposait en retour les solutions correspondant ou s'approchant le mieux de sa requête (cf. encart 1). Ce service fonctionnait à l'oral par téléphone (synthèse et reconnaissance vocale) mais également à

l'écrit avec une interface de type Web avec des applets (saisie du texte dans une fenêtre et validation par la touche entrée – historique affiché sous la fenêtre – pas d'utilisation de la souris).

Scénarios de dialogue. Afin de mettre l'utilisateur en situation de test, des tâches de recherche d'informations sous forme de scénarios lui ont été fournies. Deux ou trois critères de recherche étaient donnés à l'utilisateur (lieu, spécialité, prix) pour effectuer sa recherche (cf. encart 2). En manipulant le nombre et l'ordre des critères, 12 scénarios simples ont été créés. Il existait un scénario par ordre possible et l'ordre a été contrebalancé. Les scénarios ont été regroupés par six en deux sessions (S1 & S2).

Encart 2 : Exemple de critères pour un scénario

Lieu :	Louvre
Spécialité :	traditionnelle
Prix :	vingt euros

Questionnaire NASA-TLX d'évaluation de la charge de travail mentale. La charge mentale est un concept général établi depuis plusieurs décennies à partir des théories des ressources cognitives et de l'attention afin d'étudier les systèmes homme-machine surtout dans le domaine des transports (Parasuraman & Hancock, 2001). Dans les recherches, plusieurs techniques ont été utilisées pour estimer le niveau de charge relatif à une tâche : la double tâche ou la triple tâche (en terme d'effort cognitif, cf. Piolat & al., 1999), les mesures physiologiques (Jørgensen, Garde, Laursen, & Jensen, 1999) ou les questionnaires subjectifs (Scerbo, 2001). Cette dernière technique présente un bon niveau de validité car elle dépend directement de la tâche. De plus, elle a l'avantage d'être facile à faire passer et est très utilisée pour l'évaluation de la charge cognitive. Le questionnaire *NASA-TLX* (Task-Load Index, Hart & Staveland, 1988) est une échelle subjective bipolaire pondérée ayant pour but d'évaluer le niveau de la charge de travail mental imposée pour une tâche particulière. Elle fournit une estimation globale à partir de six dimensions, trois se rapportant à la tâche et trois à l'investissement de l'opérateur dans la tâche, que l'on combine entre elles. Les dimensions sont (1) l'activité mentale (effort mental), (2) la pression temporelle (enchaînement des sous-tâches), (3) la fatigue (travail), (4) la performance (efficacité de réalisation de la tâche), (5) le stress, (6) l'activité physique (demande physique de réalisation de la tâche). Pour cette étude, chaque dimension a été opérationnalisée par une question à l'exception de la dimension performance pour laquelle deux items étaient proposés (e.g., pour une question de la dimension performance : "Comment jugez-vous votre niveau d'efficacité dans le dialogue ?"). Les sujets répondaient avec des échelles de type Likert comprenant 9 points (le niveau 1 représentant en général le niveau de charge de travail mental le moins élevé et le niveau 9 représentant le plus haut niveau). L'indice de charge mentale a été obtenu en calculant la moyenne des estimations du sujets sur chacune des six dimensions (homogénéité du test : alpha de Cronbach = .85 à .90). La passation originale du *NASA-TLX* étant altérée, le modèle de la moyenne pondérée, calculée à partir de 21 mesures, i.e. 6 mesures prises sur chacune des dimensions et 15 mesures de comparaisons prises deux à deux, n'a pas été retenu (pour un exemple d'utilisation du *NASA-TLX*, cf. Raufaste & al., 2001).

2.3 Procédure

Les sujets ont été accueillis individuellement et le service PlanResto leur a été présenté. Ils ont reçu une consigne de résolution de problème, i.e. trouver des restaurants à l'aide du service. Après chaque scénario, les sujets ont répondu au questionnaire évaluant la charge mentale. Après chaque session, les participants ont répondu à un questionnaire de satisfaction générale. L'ensemble de l'expérimentation n'excédait pas 40 min.

2.4 Facteurs manipulés

Les facteurs manipulés étaient le facteur ordre de la modalité de départ (oral puis écrit, écrit puis oral) en intergroupe, et le facteur session en intragroupe (S1 & S2), chaque session étant composée de six scénarios (facteur imbriqué). Douze utilisateurs ont commencé par la modalité orale en session 1 et ont poursuivi par la modalité écrite en session 2, les douze autres ont fait l'inverse.

2.5 Mesures

Quatre mesures de performance ont été retenues : le temps total d'interaction (libre) en seconde pour réaliser la recherche, le nombre de TP utilisateur minoré par le nombre de TP supplémentaires dus à une erreur de reconnaissance vocale ou d'orthographe (cet indice a été calculé afin de pallier aux erreurs inhérentes à la reconnaissance vocale), le temps moyen par TP (temps total d'interaction par dialogue divisé par le nombre de TP) et la charge mentale par dialogue. Tous les fichiers de dialogues ont été retranscrits et annotés afin d'analyser la structure du discours.

Deux indicateurs discursifs ont été recueillis : la longueur moyenne (en mots) des énoncés utilisateur (Zoltan-Ford, 1991 ; Oviatt, 1995), la fréquence d'apparition du "Je" en fonction du nombre de mots par dialogue.

3 RESULTATS

Résultats préliminaires : les ANOVAs sur les différentes mesures n'ont pas montré de différences entre les scénarios. Pour l'ensemble des résultats cf. tableau 1.

Tableau 5: Tableau des moyennes pour les différentes mesures

	Groupe 1			Groupe 2		
	S1 oral	S2 écrit	Les deux	S1 écrit	S2 Oral	Les deux
Temps d'interaction (en s)	84.06	52.98	68.52	63.98	67.09	65.53
TP	3.78	2.32	3.05	2.65	2.99	2.82
Temps par TP (en s)	9.53	11.25	10.39	13.18	9.37	11.27
Charge mentale (min.1 / max. 9)	3.04	1.88	2.46	2.82	3.09	2.96
Nombre de mots par TP	5.72	4.84	5.28	3.80	4.19	3.99
"Je" utilisateur par nombre de mots (racine carrée)	0.50	0.13	0.32	0.03	0.24	0.13

3.1 Mesures de performance

Une analyse de variance (ANOVA) pour plan mixte a été utilisée pour les trois premières VD. Une analyse de covariance (ANCOVA) pour la charge mentale a été utilisée avec le temps total d'interaction et le nombre de TP supplémentaires pour chaque scénario en covariants. Des comparaisons planifiées ont été menées afin d'étudier les effets d'interaction.

Le temps moyen d'interaction était significativement plus important en session 1 ($M=74.01$) qu'en session 2 ($M=60.04$), $F(1,22)=4.60$, $P<.05$, mais pas en fonction de la modalité de départ (oral puis écrit $M=68.52$, écrit puis oral $M=65.53$), $F(1,22)=.09$, $P>.1$. L'analyse a révélé un effet d'interaction entre les deux facteurs, $F(1,22)=6.88$, $P<.05$. Dans les deux groupes, le temps moyen d'interaction était plus important à l'oral. De plus, les comparaisons planifiées ont montré que le temps moyen d'interaction pour le groupe ayant commencé à l'oral en session 1 ($M=84.06$) a très fortement diminué à l'écrit en session 2 ($M=52.98$), $F(1,22)=11.37$, $P<.01$, alors que pour le groupe ayant commencé par l'écrit en session 1 ($M=63.98$) le temps moyen d'interaction n'a que très légèrement augmenté à l'oral en session 2 ($M=67.09$), $F(1,22)=.11$, $P>.1$.

Le nombre de TP était significativement plus important en session 1 ($M=3.22$) qu'en session 2 ($M=2.65$), $F(1,22)=4.41$, $P<.05$, mais pas en fonction de la modalité de départ (oral puis écrit $M=3.05$, écrit puis oral $M=2.82$), $F(1,22)=.33$, $P>.1$. L'analyse a révélé un effet d'interaction entre les deux facteurs, $F(1,22)=11.20$, $P<.01$. Dans les deux groupes, le nombre de TP était plus important à l'oral. De plus, les comparaisons planifiées ont montré que le nombre de TP pour le groupe ayant commencé à l'oral en session 1 ($M=3.78$) a très fortement diminué à l'écrit en session 2 ($M=2.32$), $F(1,22)=14.84$, $P<.001$, alors que pour le groupe ayant commencé par l'écrit en session 1 ($M=2.65$) le nombre de TP n'a que très légèrement augmenté à l'oral en session 2 ($M=2.99$), $F(1,22)=.78$, $P>.1$.

Le temps moyen par TP de parole n'a pas été significativement différent entre les sessions 1 ($M=11.35$) et 2 ($M=10.31$), $F(1,22)=.30$, $P>.1$, et en fonction de la modalité de départ (oral puis écrit $M=10.39$, écrit puis oral $M=11.27$), $F(1,22)=.72$, $P>.1$. L'analyse a révélé un effet d'interaction entre les deux facteurs, $F(1,22)=8.00$, $P<.01$. Que ce soit en session 1 ou en session 2, le temps moyen par

TP à l'oral était toujours inférieur à l'écrit (respectivement à l'oral en session 1 puis en à l'écrit session 2, $M=9.53$ et $M=11.25$, vs à l'écrit en session 1 puis à l'oral en session 2, $M=13.18$ et $M=9.37$).

La charge mentale était significativement plus importante en session 1 ($M=2.93$) qu'en session 2 ($M=2.49$), $F(1,20)=11.28$, $P<.01$, mais pas en fonction de la modalité de départ (oral puis écrit $M=2.46$, écrit puis oral $M=2.96$), $F(1,20)=1.95$, $P>.1$. L'analyse a révélé un effet d'interaction entre les deux facteurs, $F(1,20)=9.46$, $P<.01$. Dans les deux groupes, la charge mentale était plus importante à l'oral. De plus, les comparaisons planifiées ont montré que la charge mentale pour le groupe ayant commencé à l'oral en session 1 ($M=3.04$) a très fortement diminué à l'écrit en session 2 ($M=1.88$), $F(1,20)=17.16$, $P<.001$, alors que pour le groupe ayant commencé par l'écrit en session 1 ($M=2.82$) la charge mentale n'a que très légèrement augmenté à l'oral en session 2 ($M=3.09$), $F(1,20)=.14$, $P>.1$.

3.2 Indicateurs discursifs

Une analyse de variance (ANOVA) pour plan mixte a été utilisée pour chacune des deux VD. Des comparaisons planifiées ont été menées afin d'étudier les effets d'interaction.

Le nombre de mots par TP était marginalement plus important pour le groupe ayant commencé à l'oral en session 1 puis continué à l'écrit en session 2 ($M=5.28$) que pour le groupe ayant commencé à l'écrit en session 1 puis continué à l'oral en session 2 ($M=3.99$), $F(1,22)=1.25$, $P<.06$. La différence est expliquée par un nombre de mots par TP en session 1 plus important à l'oral ($M=5.72$) qu'à l'écrit ($M=3.80$), $F(1,22)=8.47$, $P<.01$. Le nombre de mots en session 2 n'était pas différent en fonction des modalités (à l'écrit $M=4.84$, à l'oral $M=4.19$). De manière générale, aucune différence entre sessions 1 ($M=4.76$) et 2 ($M=4.51$) n'a été observée. L'interaction entre les deux facteurs n'était pas significative, $F(1,22)=3.49$, $P>.07$.

Le nombre de "Je" étant très variable en fonction des dialogues et des modalités, et sa fréquence étant souvent égale à 0 ou 1, plutôt qu'une transformation logarithmique, une transformation des données par la racine carrée a été effectuée sur la fréquence d'apparition du "Je" en fonction du nombre de mots par dialogue afin d'homogénéiser la variance et de normaliser la distribution.

Le nombre de "Je" sur le nombre total de mots par dialogue était plus important pour le groupe ayant commencé à l'oral en session 1 puis continué à l'écrit en session 2 ($M=0.32$) que pour le groupe ayant commencé à l'écrit en session 1 puis continué à l'oral en session 2 ($M=0.13$), $F(1,22)=7.72$, $P<.05$, mais pas significativement différent entre la session 1 ($M=0.26$) et la session 2 ($M=0.19$), $F(1,22)=1.64$, $P>.1$. L'analyse a révélé un effet d'interaction entre les deux facteurs, $F(1,20)=24.04$, $P<.0001$. Dans les deux groupes, le nombre de "Je" sur le nombre total de mots était plus important à l'oral. De plus, les comparaisons planifiées ont indiqué que la différence est expliquée par un nombre "Je" en session 1 plus important à l'oral ($M=0.50$) qu'à l'écrit ($M=0.03$), $F(1,22)=45.73$, $P<.0001$. Le nombre "Je" en session 2 n'était pas significativement différent en fonction des modalités (à l'écrit $M=0.13$, à l'oral $M=0.24$), $F(1,22)=1.03$, $P>.1$.

4 DISCUSSION

Les résultats ont révélé une différence pour la performance et l'utilisation d'indicateurs discursifs entre les modalités orale et écrite lors d'une interaction avec un système de dialogue en langue naturelle. Même si le temps par TP était plus important à l'écrit, le temps d'interaction, le nombre de TP ainsi que la charge mentale étaient plus importants à l'oral, surtout lors des premières interactions (session 1). De plus, le nombre de mots par TP était plus important à l'oral, surtout en session 1. Les différences entre écrit et oral s'estompant en session 2. La présence du "Je" comme indicateur d'implication dans le dialogue était plus importante à l'oral qu'à l'écrit.

Les résultats sont conformes à l'idée que l'écrit permettrait la décharge de certains mécanismes en faveur d'un plus grand contrôle de l'activité. Lors d'échanges avec un partenaire interactif machine, les utilisateurs auraient encore plus tendance à se focaliser sur la tâche et son contenu (Bubb-Lewis & Scerbo, 2002). Même s'il est sûr que pour beaucoup de participants taper était plus lent que parler (moins automatisé, en référence à l'analyse des temps moyens par TP plus importants à l'écrit ; cf. aussi Ferrara & al., 1991 ou Hancock & Dunham, 2001), ils ont optimisé très rapidement leurs recherches et leurs échanges à l'écrit, pour se focaliser sur les informations importantes de recherche (suppression de mots inutiles...). Néanmoins, les résultats en faveur d'une faible charge mentale à l'écrit doivent être relativisés par le fort niveau d'étude et de familiarité avec l'informatique d'une

partie des participants. De plus, les échanges à l'oral ont été contraints par la temporalité du discours (par exemple Fraisse & Breyton, 1959 ; Poole & Field, 1976) : à la fois, parler est beaucoup plus rapide qu'écrire ou taper, et lire est beaucoup plus rapide qu'écouter (Chafe, 1982). L'analyse des mesures des performances à l'oral doit aussi être nuancée car les différentes variables sont fortement interdépendantes (temps d'interaction, nombre d'interaction, charge mentale de travail).

Des travaux déjà anciens indiquaient que langage écrit et langage oral diffèrent l'un de l'autre de plusieurs façons plus que par le médium par lequel ils sont exprimés (Chafe, 1982). En ce sens, la structure du discours avec un système de dialogue en langue naturelle a été ici différente entre écrit et oral. Si l'apparition du pronom-sujet "Je" est une caractéristique d'implication dans le dialogue homme / homme (en terme d'indice métaconversationnel, voir Brennan, 1991, McCarthy & Monk, 1994), alors dans la situation expérimentale présentée ici, les participants à l'oral ont un comportement analogue au dialogue naturel, ce comportement apparaissant dans une moindre mesure à l'écrit. Les résultats observés complètent ceux de Ferrara et al. (1991) qui avaient trouvé que les conversants utilisant un système interactif à l'écrit utilisaient de moins en moins de pronoms-sujets. Les résultats de l'étude laisseraient penser que l'oral est plus propice à l'échange en dialogue naturel que l'écrit (Brennan & Ohaeri, 1999). Il existerait des comportements typiques en fonction de la modalité d'interaction dans une tâche de recherche d'informations (Le Bigot et al., soumis), par exemple le comportement de demande de répétition du nom du restaurant ne se produit pas à l'écrit. Comme Oviatt (1995) l'avait noté, les processus cognitifs liés à la production de l'écrit et de l'oral diffèrent de manière catégorielle (i.e., dans le sens où la relation entre la structure profonde et les systèmes de règles est différente en fonction de la modalité lors de la formation des séquences grammaticales).

De manière générale, et à la différence des expériences avec une simulation, l'utilisation d'un système de reconnaissance vocale induit des erreurs nuisibles à l'interaction. Une question peut se poser : est-ce que la seule présence de ces artefacts entraîneraient la majeure partie des difficultés observées à l'oral ? Même si l'ajout de covariants à l'analyse de la charge mentale a semblé être un bon compromis, les résultats doivent être interprétés avec prudence. En revanche, l'effet de modalité semble se dessiner au moins pour l'utilisation des indicateurs discursifs (nombre de mots, fréquence du "Je") et aller dans le sens des observations faites avec des simulations (cf. par exemple, Bubb-Lewis & Scerbo, 2002 ; Oviatt, 1995 ; Zolan-Ford, 1991).

Plusieurs pistes peuvent être envisagées à la suite de cette étude. Tout d'abord, l'étude des interactions avec un système de dialogue en fonction du niveau d'expertise serait nécessaire afin de déterminer si l'effet de modalité se manifeste toujours (et dans quelle mesure) pour des non-informaticiens, et s'il perdure pour des utilisateurs experts du dialogue ou de la tâche avec un système (en référence à la distinction entre connaissances en informatique et du domaine faite par Thomas, 1978, cité dans Falzon, 1989, p.55). Ensuite, l'intégration de recommandations en fonction des modalités devraient améliorer la construction des modèles utilisateurs (Litman & Pan, 2002). Par exemple, l'utilisation d'indices de construction de terrain commun pourrait être interprétée par le système de manière différente en fonction de la modalité, de ce fait la qualité du feedback système pourrait être optimisée.

5 RÉFÉRENCES

- Bereiter, C. (1980). Development in writing. In: Lee W. Gregg/Erwin R. Steinberg, (Eds), *Cognitive Processes in Writing. An Interdisciplinary Approach*. Hillsdale, NJ: Lawrence Erlbaum, 73-96.
- Brennan, S. (1991). Conversation with and through Computers. *User Modeling and User-Adapted Interaction*, 1, 67-86.
- Brennan, S.E. & Clark, H.H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1482-1493.
- Brennan, S. E., & Ohaeri, J. O. (1999). Why do electronic conversations seem less polite? The costs and benefits of hedging. *Proceedings, International Joint Conference on Work Activities, Coordination, and Collaboration (WACC '99)* (pp. 227-235). San Francisco, CA: ACM.
- Brennan, S.E. & Schober, M.F. (1998). The grounding problem in conversation with and through computers. In S. Kreuz (Ed.), *Social and Cognitive Psychological Approaches to Interpersonal Communication* (pp. 201-225). Hillsdale, NJ: Lawrence Erlbaum.

-
- Bubb-Lewis, C. & Scerbo, M.W. (2002). The effects of communication modes on performance and discourse organization with an adaptive interface. *Applied Ergonomics*, 33, 15-26.
- Clark, H. H. (1996). *Using Language*. Cambridge: Cambridge University Press.
- Clark, H. H. (1997). Dogmas of understanding. *Discourse Processes*, 23, 567-598.
- DeVito, J.A. (1966). Psychogrammatical factors in oral and written discourse by skilled communicators. *Speech Monographs* 33, 73-76.
- DeVito, J. A. (1967). Levels of Abstraction in Spoken and Written Language. *Journal of Communication*, 17, 354-361.
- Drieman, G. H. J. (1962). Differences between written and spoken language – Quantitative approach. *Acta Psychologica*, 20, 36-58
- Falzon, P. (1989). *Ergonomie cognitive du dialogue*. Grenoble, Presses Universitaires de Grenoble
- Ferrara, K., Brunner, H. & Whittemore, G. (1991) Interactive written discourse as an emergent register. *Written Communication*, 8(1), 8-34.
- Fraisse, P., & Breyton, M. (1959). Comparaisons entre les langages oral et écrit. *L'année Psychologique*, 1, 61-71.
- Hancock, J. T., Dunham, P.J. (2001). Language Use in Computer-Mediated Communication: The Role of Coordination Device. *Discourse Processes*, 31(1), 91-110.
- Hart, S.G. & Staveland, L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds), *Human Mental Workload*, (pp.139-183). Amsterdam, The Netherlands: Elsevier.
- Hidi, S.E. & Hildyard, A. (1983). The comparison of oral and written productions in two discourse types. *Discourse Processes*, 6, 91-105.
- Jørgensen, A.H, Garde, A.H., Laursen, B. & Jensen, B.R. (1999). *Applying the Concept of Mental Workload to IT-work*. Proc. CybErg 1999, Australia, August 1999.
- Le Bigot, L., Jamet, E., & Rouet, J.-F. (soumis). Learning modalities for dialoguing with a system in natural language. *Applied Ergonomics*.
- Levelt, W.J.M. (1989). Working models of perception; five general issues. In B.A.S. Elsendorn & H. Bouma (Eds), *Working Models of Perception*. London: Academic Press.
- Litman, D.J. & Pan, S. (2002). Designing and evaluating an adaptive spoken dialogue system. *User Modeling and User-Adapted Interaction*, 12, 111-137.
- McCarthy, J.C. & Monk, A.F. (1994). Measuring the quality of computer-mediated communication. *Behavior & Information Technology*, 13, 311-319.
- Olson, D.R. (1977). From utterance to text: The bias of language in speech and writing. *Harvard Educational Review*. (Reprinted in Masahiko Minami & Bruce P. Kennedy (Eds). 1991. *Language Issues in Literacy and Bilingual/Multicultural Education*. Collection of articles and reviews originally published in *HER*, 1963-1991. Harvard Educational Review Reprint Series. Cambridge, MA: Harvard Educational Review.)
- Oviatt, S. (1995). Predicting spoken disfluencies during human-computer interaction. *Computer Speech and Language*, 9, 19-35.
- Parasuraman, R., & Hancock, P. A. (2001). Adaptive control of mental workload. In P. A. Hancock & P. E. Desmond (Eds.) *Stress, Workload, and Fatigue*. (pp. 305-320). Mahwah, NJ: Erlbaum.
- Piolat, A., Olive, T., Roussey, J.Y., Thunin, O., & Ziegler, J.C. (1999). SCRIPTKELL: a computer assisted tool for measuring the distribution of time and cognitive effort in writing and other complex cognitive activities. *Behavior Research Methods, Instruments, & Computers*, 31(1), 113-121.
- Poole, M.E., & Field, T.W. (1976). A comparison of oral and written code elaboration. *Language and Speech*, 19, 305-311.
- Raufaste, E., Terrier, P., Grabisch, M., Lang, J., & Prade, H. (2001). Étude expérimentale de l'applicabilité de modèles d'agrégation flous à l'étude de la charge mentale. In *Journées d'études en Psychologie Ergonomique* (pp. 171-176), *EPIQUE 2001*, Nantes, 29-30 octobre 2001
- Sachs, J. S. (1974). Memory in reading and listening to discourse. *Memory & Cognition*, 2(1A), 95-100.
- Sadek, D., Bretier, P. & Panaget, F. (1997). Artemis: Natural dialogue meets rational agency. *Proceedings of the 15th International Joint Conference on Artificial Intelligence (IJCAI'97)*, Nagoya, Japan, pp. 1030-1035.

-
- Scerbo, M.W. (2001). Stress, Workload, and Boredom in Vigilance: A Problem and an Answer. In P. A. Hancock & P. E. Desmond (Eds.) *Stress, Workload, and Fatigue*. (pp. 267-278). Mahwah, NJ: Erlbaum.
- Whittaker, S. (2003). Theories and Methods in Mediated Communication. In Graesser, A.C., Gernsbacher, M.A., & Goldman, S. (Eds.). *Handbook of discourse processes*. Mahwah, NJ: Erlbaum.
- Whittaker, S., Walker, M., & Moore, J. (2002). *Fish or Fowl: A Wizard of Oz Evaluation of Dialogue Strategies in the Restaurant Domain*. Paper presented at the Language Resources and Evaluation Conference, Spain.
- Zoltan-Ford, E. (1991). How to get people to say and type what computers can understand. *International of Journal Man-Machine Studies*, 34, 527-547.

La pratique du métier de chauffeur de taxi préserve-t-elle les habiletés spatiales du déclin lié à l'âge ?

D. Ansiau & J.-C. Marquié
Laboratoire Travail et Cognition
UMR 5551 du CNRS
Maison de la recherche – Université de Toulouse II
5, allées Antonio Machado
31058 Toulouse Cedex
ansiau@univ-tlse2.fr & marquie@univ-tlse2.fr

RESUME

Des réorganisations structurales de l'hippocampe proportionnelles à l'ancienneté dans le métier ont été récemment mises en évidence chez des chauffeurs de taxi londoniens (Maguire, Gadian, Johnsrude, Good, Ashburner, Frackowiak, & Frith, 2000). L'objectif de cette étude était de voir si des changements fonctionnels en lien avec la pratique du métier de chauffeur de taxi pouvaient aussi être trouvés dans les performances de navigation dans un environnement virtuel. Nous avons comparé les performances d'apprentissage et de rappel différé de chauffeurs de taxi toulousains (N = 39) et de sujets contrôles (N = 42) dont l'âge varie entre 30 et 65 ans, et examiné ces mêmes performances en fonction de l'ancienneté des chauffeurs de taxi. Les analyses n'ont révélé aucune différence significative entre les deux groupes et pas d'effet de l'ancienneté dans le métier. Elles ont confirmé, par contre, le déclin habituellement observé avec l'âge dans les performances de navigation. Les raisons possibles de l'écart entre ces résultats et ceux obtenus par Maguire et al. (2000) sont discutées. En l'état, ces résultats ne permettent donc pas de vérifier l'hypothèse selon laquelle la sollicitation durable d'une fonction cognitive élémentaire permettrait de préserver cette dernière du déclin habituellement observé avec l'âge.

MOTS-CLES

Vieillesse cognitive, pratique professionnelle, préservation différentielle, cognition spatiale, chauffeur de taxi, environnement virtuel, hippocampe.

1 INTRODUCTION

On a des raisons de penser que ce qui se vit dans le travail¹ a des conséquences à long terme plus ou moins importantes et irréversibles sur le système nerveux central et donc sur le fonctionnement cognitif. De ce fait, les conditions de travail ont probablement un effet modérateur non négligeable sur le vieillissement cognitif (Marquié, 1995). Des travaux récents en neurosciences, réalisés au moyen de techniques d'imagerie cérébrale, ont apporté des indices nouveaux en faveur de cette hypothèse (e.g. Cho, 2001 ; Cho, Ennaceur, Cole, & Suh, 2000 ; Maguire, Gadian, Johnsrude, Good, Ashburner, Frackowiak, & Frith, 2000). Une de ces études, celle de Maguire et al. (2000), portait sur les effets d'une longue pratique cognitive spécialisée dans le cadre de la profession de chauffeur de taxi. Ces auteurs ont trouvé une corrélation significative entre la taille de l'hippocampe droit postérieur et antérieur, d'une part, et l'ancienneté dans le métier de chauffeur de taxi, d'autre part ; la corrélation était positive pour l'hippocampe droit postérieur (HDP) et négative pour l'hippocampe droit antérieur (HDA). Le fait que des changements structuraux puissent intervenir dans certaines régions du cerveau humain comme résultat de l'exposition à certaines conditions de travail a des implications importantes pour la recherche dans le domaine du vieillissement et du travail. En particulier, ces données qui révèlent une plasticité du cerveau adulte en lien avec l'exercice de la profession sont particulièrement pertinentes pour tester l'hypothèse de la préservation différentielle (Bosman & Charness, 1996 ;

¹ C'est-à-dire l'ensemble des conditions cognitives et non-cognitives auxquelles le sujet est exposé tout au long de ses 40 ans de vie professionnelle.

Marquié, 1997 ; Salthouse, 1990 ; Salthouse, Babcock, Skovroneck, Mitchell, & Palmon, 1990). Selon cette hypothèse, l'utilisation intensive, régulière et durable de processus cognitifs spécialisés, comme c'est le cas dans beaucoup de professions, pourrait préserver ces processus des déclinés liés à l'âge habituellement observés. Il est bien connu que l'expérience professionnelle constitue une source importante de connaissances déclaratives et procédurales, lesquelles fournissent elles-mêmes des moyens de compensation des difficultés liées à l'âge (e.g., Dixon & Bäckman, 1995). Mais ce qui est spécifique à l'hypothèse de la préservation, comparativement à celle de la compensation, c'est l'idée que la pratique cognitive ne génère pas seulement de nouvelles connaissances (le « soft ») mais entraîne également des changements dans les processus cognitifs de base impliqués (le « hard »). Jusqu'ici, toutefois, des résultats incohérents ont été obtenus en regard de cette hypothèse, essentiellement en raison de l'utilisation de méthodes inappropriées. (pour une revue, voir Marquié, 1997).

L'étude de Maguire et al. (2000) offre un cadre intéressant pour tester l'hypothèse de la préservation différentielle en utilisant une approche cognitive expérimentale. Pour cela, certaines conditions doivent néanmoins être remplies :

(i) Il faut d'abord pouvoir spécifier un minimum le type d'habiletés sous-tendues par l'HDA et l'HDP, et avoir de bonnes raisons de penser que les changements structurels de l'hippocampe des chauffeurs de taxi sont associés à des changements fonctionnels dans les habiletés cognitives. La plupart des études suggèrent que l'hippocampe humain est impliqué dans la mémoire spatiale et la navigation. Plus précisément, et bien que les connaissances sur ce sujet soient encore largement incomplètes, le rappel d'informations topographiques durant la navigation semble être préférentiellement sous-tendu par l'HDP, alors que l'apprentissage de nouvelles informations topographiques durant le déplacement semble sous-tendu par l'HDA (en combinaison avec l'HDP) (Maguire, Frackowiak, & Frith, 1997 ; Maguire et al., 2000 ; Menon, White, Eliez, Glover, & Reiss, 2000).

(ii) Une deuxième condition est que ces habiletés soient reconnues comme déclinant habituellement avec l'avancée en âge. C'est ce qu'ont montré Moffat, Zonderman et Resnick (2001) dans une tâche d'apprentissage et de rappel de déplacements dans un environnement virtuel (EV), ainsi que d'autres auteurs utilisant diverses tâches spatiales (voir Salthouse, 1992).

(iii) Une dernière considération concerne la possibilité de transférer les habiletés cognitives acquises avec l'expérience comme chauffeur de taxi dans d'autres tâches de navigation. Si certains types d'expériences cognitives sont associés à des changements morphologiques, comme l'ont montré Maguire et al. (2000), alors les effets modérateurs de l'expérience des chauffeurs de taxi sur les déclinés liés à l'âge des habiletés cognitives pertinentes doivent être observables dans d'autres tâches, professionnelles ou non, mais pertinentes au sens où elles mettent en jeu les mêmes structures cérébrales. Pour satisfaire à ce critère, nous avons eu recours à une tâche de navigation dans un environnement virtuel déjà employée par Maguire, Burgess, Donnett, Frackowiak, Frith, and O'Keefe (1998). Au moyen de techniques de neuro-imagerie fonctionnelle, ces auteurs ont mis en évidence une activation de l'hippocampe droit fortement corrélée avec un déplacement correct dans cet environnement virtuel.

En se basant sur les observations de Maguire et al. (2000) et sur les considérations envisagées plus haut, le but de la présente étude était de comparer les performances d'apprentissage et de rappel des sujets sélectionnés (les chauffeurs de taxi) et des sujets non-sélectionnés (contrôles) dans un environnement 3D. Compte tenu des liens actuellement admis entre structures et fonctions hippocampiques, notre hypothèse était que les chauffeurs de taxi devraient avoir des performances d'apprentissage plutôt moins bonnes que les contrôles dans la tâche de navigation en EV, et devraient avoir de meilleures performances dans le rappel différé d'informations spatiales précédemment acquises dans ce même environnement. Pour les mêmes raisons, les performances d'apprentissage et de rappel au sein de l'échantillon des chauffeurs de taxi devraient être, respectivement, négativement et positivement corrélées avec le nombre d'années de pratique du métier.

2 METHODE

2.1 Participants

Les participants étaient 81 hommes volontaires répartis dans deux groupes: les chauffeurs de taxi ($n = 39$) et les sujets contrôles ($n = 42$). Dans le premier groupe, tous les participants étaient employés à temps plein en tant que chauffeurs de taxi dans la ville de Toulouse et avaient au minimum 6 mois d'expérience dans la profession (0.5-37 ans). Les sujets contrôles étaient des personnes en activité n'ayant pas d'expérience dans une profession impliquant des activités de navigation. Tous les participants étaient volontaires. Les deux groupes expérimentaux avaient le même âge ($M : 48.4$ ans, $E.T. : 8.3$ pour les chauffeurs de taxi et $M : 47.0$ ans, $E.T. : 6.1$ pour les sujets contrôles ; $p = .40$), mais différaient significativement quant au nombre d'années de scolarisation ($M : 10.7$ ans, $E.T. : 3.2$ et $M : 13.2$ ans, $E.T. : 13.2$, respectivement ; $p < .001$) et à la familiarité avec l'informatique ($M : 3.0$, $E.T. : 1.7$, et $M : 4.8$, $E.T. : 1.5$, respectivement, le nombre plus élevé indiquant une plus grande familiarité ; $p < .001$).

2.2 Matériel

La tâche de navigation était réalisée sur un ordinateur Hewlett Packard Omnibook XE3 733 MHz Pentium III, avec un moniteur 13''. Les déplacements dans l'environnement étaient contrôlés en utilisant le pavé fléché avec les touches avant, arrière, gauche et droite. L'environnement en trois dimension a été réalisé en utilisant le programme BUILD fourni avec le jeu Duke Nukem 3-D (3-D Realms Entertainment, Apogee Software Ltd., Garland, T.X.). Ce logiciel a déjà été utilisé par Maguire et al. (1998) dans le cadre de l'une de leurs recherches en imagerie cérébrale. Cette étude a montré que le déplacement dans l'environnement virtuel sollicite l'activation de l'hippocampe droit. Il semble donc que le fait de se déplacer dans ce type d'environnement virtuel présente une bonne validité écologique. De plus, les travaux de Péruch et Gaunet (1998) ont insisté sur la pertinence de l'utilisation de ces environnements virtuels pour étudier la cognition spatiale. Sept bâtiments (hôtel, motel, bureau de police, bar, bowling, karaoké et cinéma) servaient de points de repère dans l'environnement virtuel et de cible lors des déplacements.

2.3 Tâche et procédure

Un temps standardisé de 10mn était consacré à la familiarisation du participant avec l'EV. Ensuite, un test de dextérité concernant l'utilisation du pavé fléché pour se déplacer dans l'EV était effectué. La tâche de navigation principale comprenait deux phases séparées : l'apprentissage et le rappel différé. La tâche expérimentale d'apprentissage consistait en l'apprentissage successif des sept trajets. Le participant pouvait utiliser autant d'essais que nécessaire pour chaque trajet. Le critère d'un apprentissage réussi d'un trajet était de réaliser ce dernier par le chemin le plus court trois fois de suite.

Le temps séparant la phase d'apprentissage et de rappel était exactement de trente minutes. Pendant ce temps, divers tests étaient administrés en guise de tâche de remplissage. Dans le même EV que pour la phase d'apprentissage, les participants devaient ensuite rappeler les sept trajets mémorisés précédemment, en essayant de trouver le chemin le plus court. Les durées des trajets et les distances parcourues étaient enregistrées, ainsi que le nombre d'essais d'apprentissage. Ce sont ces paramètres qui ont servi de variables dépendantes dans les analyses statistiques.

3 RESULTATS

Aucune différence significative entre les deux groupes n'a été trouvée quant à la vitesse d'utilisation du pavé fléché au test de dextérité ($p = .80$; $M : 91.67$ sec., $E.T. : 25.3$, et $M : 101.67$ sec., $E.T. : 25.45$, pour les chauffeurs de taxi et les sujets contrôles, respectivement).

3.1 Comparaisons entre les chauffeurs de taxi et les sujets contrôles

Une ANOVA a été réalisée avec la variable groupe comme facteur (chauffeurs de taxi vs. sujets contrôles). Les résultats n'ont révélé aucune différence significative entre les deux groupes pour aucun des scores de navigation : pour la durée d'apprentissage; la distance d'apprentissage ; le nombre d'essais d'apprentissage ; pour la durée de rappel ; et pour la distance de rappel (voir tableau 1). Ces

résultats semblent donc aller à l'encontre de nos hypothèses, dans la mesure où ils montrent que les chauffeurs de taxi et les sujets contrôles effectuent de la même manière les tâches de navigation. Nos efforts pour avoir deux groupes expérimentaux qui ne diffèrent pas significativement l'un de l'autre quant à certaines caractéristiques essentielles n'ont été que partiellement fructueux puisque les sujets contrôles présentaient un niveau de scolarisation et des scores de familiarité avec l'informatique supérieurs à celui des chauffeurs de taxi. Des analyses de covariance destinées à contrôler statistiquement l'effet de ces différences entre nos deux échantillons n'ont pas modifié le patron des résultats antérieurement obtenus.

Parce que la distribution des âges dans les deux groupes le permettait, et afin de tester l'hypothèse de la préservation, nous avons réalisé des analyses de régression simple avec l'âge comme prédicteur. Il est possible, en effet, que les chauffeurs de taxi et les sujets contrôles présentent les mêmes performances, en moyenne, mais des différences quant aux effets de l'âge sur leurs performances de navigation. L'effet de l'âge a donc été examiné séparément chez les chauffeurs de taxi et chez les sujets contrôles, et ensuite comparé entre les deux groupes.

L'âge apparaît être un prédicteur significatif du temps d'apprentissage chez les chauffeurs de taxi, $r^2 = .39$, $p \leq .001$, et chez les contrôles, $r^2 = .13$, $p \leq .05$. Les âges les plus élevés prédisent aussi une distance significativement plus importante d'apprentissage pour les chauffeurs de taxi, $r^2 = .11$, $p \leq .05$, mais pas pour les sujets contrôles, $r^2 = .03$, $p = .24$. Finalement, l'âge est un bon prédicteur pour le nombre d'essais en phase d'apprentissage dans les deux groupes, avec un nombre plus important d'essais requis avec l'avancée en âge chez les chauffeurs de taxi, $r^2 = .36$, $p \leq .001$ et chez les sujets contrôles, $r^2 = .15$, $p \leq .05$. Les comparaisons entre les pentes de régression des chauffeurs de taxi et des contrôles n'ont révélé aucune différence significative, indiquant ainsi des patrons de vieillissement identiques pour ces paramètres de performance en phase d'apprentissage (tous les $p > .05$).

Les mêmes analyses de régression ont été réalisées avec l'âge comme prédicteur pour le rappel différé. Il est apparu que des âges plus élevés prédisaient un temps de rappel plus long pour les chauffeurs de taxi, $r^2 = .31$, $p \leq .001$, mais pas pour les contrôles, $r^2 = .03$, $p = .25$. Des résultats similaires ont été observés pour la distance de rappel, avec des âges plus élevés prédisant des distances significativement plus longues chez les chauffeurs de taxi, $r^2 = .11$, $p \leq .04$, mais pas pour les contrôles, $r^2 = .00$, $p = .94$. Les comparaisons des pentes de régression entre les deux groupes ont révélé un effet équivalent de l'âge sur les performances de rappel différé pour les chauffeurs de taxi et les contrôles (tous les $p > .05$).

Tableau 1. Moyennes (et écarts-type) pour les performances de navigation dans chaque groupe et résultats des ANOVAs

Performance de navigation		Chauffeurs de taxi	Sujets contrôles	F(ddl)	<i>p</i>
Apprentissage	Temps (sec.)	28.15 (9.11)	31.04 (8.93)	F(2, 79) = 2.08	.15
	Distance (mm.)	371.35 (52.99)	383.30 (47.82)	F(2, 79) = 1.13	.29
	Nombre d'essais	6.54 (1.51)	6.76 (1.84)	F(2, 79) = .34	.56
Rappel différé	Temps (sec.)	29.61 (11.98)	33.08 (13.84)	F(2, 79) = 1.45	.23
	Distance (mm.)	413.33 (107.73)	408.00 (127.54)	F(2, 79) = .04	.84

3.2 Effets de l'expérience des chauffeurs de taxi

Dans cette partie, nous avons examiné la relation entre le nombre d'années de pratique en tant que chauffeur de taxi et les performances de navigation. Les analyses ont donc été réalisées dans le groupe des chauffeurs de taxi. Des régressions multiples ont été calculées avec le nombre d'années de pratique en tant que chauffeur de taxi et l'âge comme prédicteurs (voir tableau 2). Des deux facteurs entrés dans le modèle, seul l'âge s'est avéré être un prédicteur significatif de la durée d'apprentissage ($p < .001$), et du nombre d'essais d'apprentissage ($p = .001$), les chauffeurs de taxi plus âgés nécessitant plus de temps et plus d'essais pour apprendre les itinéraires désignés. Pour la distance d'apprentissage, le même modèle de régression a fait apparaître un effet exclusif de l'expérience ($p = .04$), les

chauffeurs de taxi plus expérimentés parcourant une distance plus importante lors de l'apprentissage des trajets, quel que soit leur âge.

Les mêmes modèles de régression ont été appliqués pour les performances de rappel différé, excepté que la performance d'apprentissage correspondante (durée ou distance d'apprentissage, selon l'analyse) a été rajoutée au modèle comme troisième prédicteur. Cela permet d'obtenir une évaluation plus pure des effets de l'âge et de l'expérience, c'est-à-dire indépendante des performances d'apprentissage. Seules ces variables, c'est-à-dire la durée et la distance d'apprentissage se sont révélées être des prédicteurs significatifs, respectivement de la durée ($p < .001$) et de la distance de rappel des itinéraires ($p < .001$).

Tableau 2. Régressions multiples pour les chauffeurs de taxi. Le premier rang fournit pour chaque performance d'apprentissage et de rappel le r^2 , le F et le seuil de signification du modèle de régression total.

			β Standardisé	T(2, 36)	p
<i>Apprentissage</i>	<i>Temps</i>	($r^2 = .39$, F(2, 36) = 11.66, $p \leq .001$)			
		Âge	.631	4.22	$\leq .001$
		Années de Pratique	-.009	-.06	.95
	<i>Distance</i>	($r^2 = .21$, F(2, 36) = 4.79, $p = .01$)			
		Âge	.151	.88	.38
		Années de Pratique	.365	2.14	.04
<i>Essais</i>	($r^2 = .37$, F(2, 36) = 10.75, $p \leq .001$)				
	Âge	.542	3.57	$\leq .001$	
	Années de Pratique	.122	.80	.43	
			β Standardisé	T(3, 35)	p
<i>Rappel</i>	<i>Temps</i>	($r^2 = .64$, F(3, 35) = 20.39, $p \leq .001$)			
		Âge	.118	.82	.42
		Années de Pratique	-.054	-.46	.65
		Temps d'apprentissage	.753	5.62	$\leq .001$
	<i>Distance</i>	($r^2 = .40$, F(3, 35) = 7.68, $p \leq .001$)			
		Âge	.139	.91	.38
Années de Pratique		.006	.04	.97	
	Distance d'apprentissage	.567	3.84	$\leq .001$	

4 DISCUSSION

Nos résultats fournissent peu d'indications en faveur de l'hypothèse de la préservation différentielle. Des performances strictement équivalentes entre les chauffeurs de taxi et les sujets contrôles ont été obtenues dans la phase d'apprentissage, même après le contrôle expérimental et statistique de variables potentiellement parasites. Concernant les performances en rappel différé, sur les quatre conditions analysées (durée et distance de rappel, avec ou sans contrôle statistique de variables possibles de confusion) une différence en faveur des chauffeurs de taxi n'a été trouvée que dans une seule de ces conditions. Cette absence de différence claire entre les groupes expérimentaux est cohérente avec l'absence d'effet de l'expérience professionnelle observé dans l'échantillon des chauffeurs de taxi. Par contre, un effet clair de l'âge sur les performances de navigation a été trouvé, comme prévu. Cet effet de l'âge est en accord avec les résultats obtenus par Moffat et al. (2001) dans une tâche similaire, ainsi qu'avec ce que d'autres auteurs ont trouvé dans une grande variété de tâches spatiales (e.g., Salthouse, 1992).

Au moins quatre raisons permettent d'expliquer l'écart entre les changements morphologiques observés par Maguire et al. (2000) chez les chauffeurs de taxi londoniens, et l'absence de changements fonctionnels observés chez les chauffeurs de taxi toulousains. Une première raison qui pourrait être invoquée est le manque de pertinence de la tâche choisie. Pourtant les environnements virtuels sont de plus en plus fréquemment utilisés pour évaluer les habiletés de navigation spatiale chez l'homme (Maguire, Frith, Burgess, Donnett, & O'Keefe, 1998 ; Maguire, Burgess et al., 1998 ; Maguire, Burgess, & O'Keefe, 1999 ; Péruch & Gaunet, 1998). De plus, se déplacer dans de tels environnements semble activer l'hippocampe droit (Maguire, Burgess et al., 1998), comme indiqué plus haut. Enfin, la relative pauvreté de notre environnement virtuel ne semble pas participer au manque d'effet de la

pratique professionnelle, puisque Waller, Knapp et Hunt (2001) ont montré que cette variable n'affecte que très peu la performance des participants.

Deuxièmement, l'absence d'un effet de la pratique des chauffeurs de taxi dans la présente étude peut aussi résulter de la durée relativement courte de l'expérience, ce qui ne permettrait pas aux participants de construire une représentation mentale de l'environnement qui serait suffisamment élaborée pour révéler une différence entre les chauffeurs de taxi et les sujets contrôles, ou des différences parmi des chauffeurs de taxi ayant une quantité différente de pratique professionnelle. Comme l'ont montré Giraud et Péruch (1988), alors que les chauffeurs de taxi expérimentés ne se distinguent pas des novices en ce qui concerne le rappel d'informations relatives au réseau primaire (grands axes routiers), il sont plus précis que les novices dans le rappel d'informations relatives au réseau urbain secondaire d'une ville. Selon ces auteurs, se déplacer dans le réseau urbain primaire requiert seulement une représentation spatiale égocentrique, alors que se déplacer efficacement dans le réseau secondaire requiert de construire une représentation allocentrique de l'environnement. Les chauffeurs de taxi expérimentés sont supposés fonctionner essentiellement sur la base de représentations allocentriques, lesquelles sont associées à l'activation de l'hippocampe droit (Maguire et al., 1999). Il est donc possible que dans la présente étude, le processus d'apprentissage n'ait pas duré suffisamment longtemps pour permettre aux individus de construire une représentation allocentrique de l'environnement virtuel.

Troisièmement, nos hypothèses sont basées sur un modèle très localisationniste de la cognition. Les processus de navigation ne sont pas seulement dépendants de l'hippocampe mais aussi d'autres structures cérébrales, telles que le cortex pariétal médian et inférieur droit, le cortex cingulaire, des parties de la ganglia basale, le cortex préfrontal gauche, et les régions temporales médianes bilatérales (incluant le gyrus parahippocampique) (e.g., Aguirre, Detre, Alsop, & d'Esposito, 1996 ; Maguire, Burgess et al., 1998 ; Maguire et al., 1999 ; Gron, Wunderlich, Spitzer, Tomczac, & Riepe, 2000). Des différences morphologiques dans seulement une seule de ces zones cérébrales ne sont peut-être pas suffisantes pour conduire à des différences fonctionnelles significatives.

Une quatrième raison possible est que les chauffeurs de taxi toulousains ne sont pas tout à fait identiques à ceux de Londres. La ville de Londres est beaucoup plus grande et plus complexe que celle de Toulouse, et le processus de formation qui précède l'obtention de la licence de chauffeur de taxi est sensiblement plus long chez les chauffeurs de taxi londoniens que chez leurs collègues toulousains. Nous ne pouvons écarter la possibilité que de telles différences puissent jouer un rôle en sélectionnant des propriétés qui favoriseraient une plus ou moins grande plasticité cérébrale ou capacité d'apprentissage spatial.

Quoi qu'il en soit, et indépendamment de la référence au modèle biologique (changements morphologiques de l'hippocampe avec l'expérience), ces résultats apportent un éclairage sur une question importante, souvent en débat au carrefour entre la psychologie cognitive, du travail, et du vieillissement. Ils ne confirment pas, en effet, l'hypothèse de la préservation différentielle selon laquelle la pratique intensive, régulière et durable de fonctions cognitives spécialisées, ici la navigation spatiale, préserve ces fonctions du déclin habituellement observé avec l'âge.

5 BIBLIOGRAPHIE

- Aguirre, G.K., Detre, J.A., Alsop, D.C., & D'Esposito, M. (1996). The parahippocampus subserves topographical learning in man. *Cerebellar Cortex*, 6, 823-829.
- Bosman, E.A., & Charness, N. (1996). Age differences in skilled performance and skill acquisition. In T. Hess & F. Blanchard-Fields (Eds.), *Perspectives on Cognitive Change in Adulthood and Aging*. New York: McGraw-Hill.
- Cho, K. (2001). Chronic "jet lag" produces temporal lobe atrophy and spatial cognitive deficits. *Nature Neuroscience*, 4, 567-568.
- Cho, K., Ennaceur, A., Cole, J.C., & Suh, C.K. (2000). Chronic "jet lag" produces cognitive deficits. *Journal of Neuroscience*, 20, 1-5.
- Dixon, R.A., & Bäckman, L. (1995). *Compensating for psychological deficits and decline. Managing losses and promoting gains*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Giraud, M.D., & Péruch, P. (1988). Spatio-temporal aspects of the mental representation of urban space. *Journal of Environmental Psychology*, 8, 9-17.

-
- Gron, G., Wunderlich, A.P., Spitzer, M., Tomczak, R., & Riepe, M.W. (2000). Brain activation during human navigation: gender-different neural networks as substrate of performance. *Nature Neuroscience*, 3, 404-408.
- Maguire, E.A., Burgess, N., & O'Keefe, J. (1999). Human spatial navigation: cognitive maps, sexual dimorphism, and neural substrates. *Current Opinion in Neurobiology*, 9, 171-177.
- Maguire, E.A., Burgess, N., Donnett, J.G., Frackowiak, R.S., Frith, C.D., & O'Keefe, J. (1998). Knowing where and getting there: a human navigation network. *Science*, 280, 921-924.
- Maguire, E.A., Frackowiak, R.S., & Frith, C.D. (1997). Recalling routes around London: activation of the right hippocampus in taxi drivers. *Journal of Neuroscience*, 17, 7103-7110.
- Maguire, E.A., Frith, C.D., Burgess, N., Donnett, J.G., & O'Keefe, J. (1998). Knowing where things are: parahippocampal involvement in encoding object locations in virtual large-scale space. *Journal of Cognitive Neuroscience*, 10, 61-76.
- Maguire, E.A., Gadian, D.G., Johnsrude, I.S., Good, C.D., Ashburner, J., Frackowiak, R.S.J., & Frith, C.D. (2000). Navigation related structural change in the hippocampi of taxi drivers. *Proceedings of the National Academy of Science U.S.A.*, 97, 4398-4403.
- Marquié, J.C. (1997). Vieillesse cognitive et expérience: l'hypothèse de la préservation. *Psychologie Française*, 42, 333-344.
- Menon, V., White, C.D., Eliez, S., Glover, G.H., & Reiss, A.L. (2000). Analysis of a distributed neural system involved in spatial information processing. *Human Brain Mapping*, 11, 117-129.
- Moffat, S.D., Zonderman, A.B., & Resnick, S.M. (2001). Age differences in spatial memory in a virtual environment navigation task. *Neurobiology of Aging*, 22, 787-796.
- Péruch, P., & Gaunet, F. (1998). Virtual environments as a promising tool for investigating human spatial cognition. *Current Psychology and Cognition*, 17, 881-899.
- Salthouse, T.A. (1990). Influence of experience on age differences in cognitive functioning. *Human Factors*, 32, 551-569.
- Salthouse, T.A. (1992). Reasoning and spatial abilities. In F.I.M. Craik & A.T. Salthouse (Eds.), *The handbook of aging and cognition* (pp. 167-211). Hillsdale, N.J.: Lawrence Erlbaum Associates, Publishers.
- Salthouse, T.A., Babcock, R.L., Skovroneck, E., Mitchell, D.R.D., & Palmon, R. (1990). Age and experience effects in spatial visualization. *Developmental Psychology*, 26, 128-136.
- Waller, D., Knapp, D., & Hunt, E. (2001). Spatial representations of virtual mazes: the role of visual fidelity and individual differences. *Human Factors*, 43, 147-158.

