activités
revue électronique

L'organisation temporelle des engagements visuels dans des situations de multi-activité équipée en milieu urbain

Christian Licoppe

LTCI - Telecom ParisTech. 46 rue Barrault, 75 013 Paris
christian.licoppe@telecom-paristech.fr

Julien Figeac

LTCI - Telecom ParisTech. 46 rue Barrault, 75 013 Paris
julien.figeac@telecom-paristech.fr

ABSTRACT

The temporal organization of visual involvements in multi-activity urban settings. To understand how the temporal management of activities which characterize multi-activity situations is organized, we report on the development of a method for recording the different uses of mobile phones in situations of mobility. The method is based on the synchronization of captures of video screen activity and videos of user activity made using camera glasses. This will allow us to analyze how users appropriate their smartphones during daily commutes. These data show the temporal organization of gaze switching (towards and away from the mobile screen) and provide crucial empirical information to understand how users actually manage mobile communications and other activities in everyday multi-activity settings. To speak of multi-activity is to reflect a particular stance that actors take in respect of the potential sequential implications which emerge from the temporal intrication of events relating to the use of mobile devices and to the mobility environment; these events create occasions to visually switch from one field of activity to another. We discuss how the 'pragmatic texture' of interfaces should be structured to offer such sequential opportunities so as to project a temporal organization of uses which is better suited to multi-activity settings.

KEYWORDS

mobility, mobile phone, multi-activity, camera glasses, gaze switch

1.- Introduction

Bien qu'omniprésente dans nos modes de vie contemporains et de plus en plus objet d'une préoccupation sociétale, la notion de multi-activité est remarquablement difficile à saisir empiriquement. La recherche sur les temporalités distingue tant bien que mal plusieurs formes du phénomène, d'un côté le multitâche consécutif, lorsqu'il s'agit d'entreprendre quelque chose puis autre chose et de reprendre l'activité initiale ; d'un autre côté le multitâche simultané, plus en phase avec l'idée de sens commun de faire plusieurs choses à la fois (Kenyon, 2010). Ces configurations ont en commun l'idée d'une concurrence entre des cours d'activité reconnaissables du point de vue de l'allocation des ressources temporelles, cognitives ou matérielles, dans une situation donnée (Circella, Mokhtarian, P., & Poff, K., 2012 ; Kaufman-Scarborough, 2003).

Le multitâche consécutif est plus proche de la fragmentation temporelle que de l'idée de sens commun de faire deux choses à la fois. C'est seulement à un niveau extrême de fragmentation que les frontières temporelles entre les tâches ou entre les activités s'estompent ; le *zapping* incessant entre différents cours d'action peut alors être éprouvé comme de la multi-activité (Licoppe, 2008). Cette version de la multi-activité comme point limite d'une fragmentation extrême des tâches a toutefois le mérite de suggérer au chercheur une analyse empirique des façons dont peut s'accomplir le passage d'une activité à l'autre, notamment lors d'interruptions. Fragmentation et interruptions sont des objets d'études privilégiés dans le domaine du travail, en particulier des activités managériales et informationnelles (Mintzberg, 1973). La recherche contemporaine sur les interruptions au travail souligne leur caractère ambivalent, ainsi le fait qu'elles puissent être traitées aussi bien comme des perturbations que comme des ressources positives dans la réalisation d'activités (Hudson, Christensen, J., Kellogg, & Erockson, 2002 ; O'Connail & Frohlich, 1995). D'autres travaux ont par ailleurs récemment établi le lien entre l'occurrence d'événements potentiellement perturbateurs et l'usage des technologies de communication (Gonzalez & Mark, 2005 ; Wajcman & Rose 2011).

La recherche en management échoue toutefois à rendre compte de comment la multi-activité peut être non seulement une manière de catégoriser des situations, mais également vécue comme une préoccupation [1] (Datchary & Licoppe, 2007), des membres, endogène à l'organisation de leur action. Le courant de recherche qui a récemment émergé autour de l'action située prise dans ses dimensions incorporée et multimodale (Goodwin, 2002 ; Stivers & Sidnell, 2005 ; Streeck, Goodwin, & LeBaron, 2011) a montré dans le détail la façon dont les membres accomplissent collaborativement de multiples cours d'activité dont chacun produit sa propre organisation temporelle, en alternant engagement et désengagement à travers des performances incorporées, par exemple pour parler tout en dînant (Goodwin, 1984) ou tout en conduisant (Laurier, 2000 ; Mondada, 2012 ; Nevile, 2012). Il ne faudrait toutefois pas confondre multi-activité et multimodalité, dans la mesure où si toute action humaine ou presque peut être décrite comme multimodale, toute situation n'est pas nécessairement caractérisable comme multi-activité (sinon le concept n'a plus de sens et on perd de vue le phénomène). Comme le remarque Laurier dans son étude sur les travailleurs mobiles en voiture (Laurier, 2000), l'activité de conduite exige à elle seule l'articulation d'une multitude de domaines sémiotiques, et diffère peu de ce point de vue de l'activité de conduire tout en travaillant en voiture.

Comment peut-on préciser davantage la notion de multi-activité ? Un premier appui se trouve dans la distinction entre action et opération proposée par la théorie de l'activité (Leont'ev, 1974) qui montre que le fait de passer une vitesse constitue pour un conducteur débutant une action en soi, un projet distinct des autres préoccupations simultanément pertinentes associées à la conduite (comme garder un œil sur la route ou tourner le volant), alors qu'un conducteur plus averti fera l'expérience d'un cours d'activité unifié, sans heurt et accompli en toute fluidité. Le conducteur débutant, parce qu'il n'a pas encore appris à intégrer ces différentes composantes de l'action, peut les éprouver comme des projets distincts à accomplir simultanément. La conduite devient pour lui un lieu de multi-activité, à la différence du conducteur averti. La multi-activité implique donc une écologie fragmentée, composée de multiples domaines de préoccupation, et peut ainsi être définie de façon plus restrictive comme un type de situations dans lesquelles les acteurs s'orientent vers différents cours d'action en tant que projets séparés, reconnaissables et simultanément pertinents (et donc dans une certaine mesure en concurrence).

Cet article va développer l'idée qu'une des manières dont la multi-activité se rend observable pour un observateur extérieur, et sur la base de compétences ordinaires qui sont

[1] La notion de préoccupation n'est pas entendue ici, et tout au long de l'article, au sens psychologique mais bien ethnométhodologique, comme orientation démontrable des membres.

aussi celles des membres, c'est dans l'imbrication temporelle des cours d'activité, dans la manière et le timing selon lequel l'acteur s'engage dans une activité donnée puis dans une autre. Chaque activité pratique en tant qu'elle est reconnaissable projette une certaine organisation temporelle et séquentielle, configurant certains instants comme propices à l'accomplissement de certaines actions. Par rapport à ces attentes, les acteurs peuvent se percevoir et être perçus par des observateurs ordinaires comme « en retard » ou en « avance » dans l'allocation de leurs engagements. Comme nous le verrons plus bas, le conducteur arrêté derrière une file de voitures et qui démarre de telle sorte qu'un espace s'est ouvert devant lui et éventuellement qu'une voiture est venue se glisser dans cet espace apparaît « en retard » par rapport à l'organisation visible des activités collectives de trafic, et la manière un peu précipitée dont il se désengage de ses activités en cours pour redémarrer témoigne de la manière dont il est comptable ("*accountable*") de ce retard, et de l'organisation morale sous-jacente à cette situation de multi-activité. C'est par rapport à cet ordre visible, et aux inférences ordinaires qu'il permet sans avoir à y réfléchir, ou "*witnessable order*" (Livingston, 2008) que les acteurs sont potentiellement comptables (au sens du mot anglais "accountability") de leurs engagements. En ce sens nous ne nous intéresserons pas ici aux logiques d'action « cachées » des utilisateurs [2].

Cette perspective d'inspiration ethnométhodologique permet donc des analyses empiriques concrètes de ce phénomène difficile à saisir que constitue la multi-activité, pourvu que l'on puisse disposer d'enregistrements vidéo des activités assez fins pour rendre visible l'allocation des engagements, leur rythme ou leur ordonnancement temporel, et les manifestations incorporées par lesquels les participants rendent perceptibles des « écarts » par rapport à un ordre temporel projeté. Pour le démontrer plus précisément, et saisir toutes les implications de cette approche pour l'analyse des activités et même la conception des environnements artefactuels, nous allons analyser un cas concret, celui d'une conductrice automobile qui se partage entre conduite automobile en milieu urbain et consultation des sites de réseaux sociaux sur son smartphone. Le caractère potentiellement dangereux (et même illégal) de ce type d'activité qu'elle conduit spontanément rend particulièrement tendue la gestion temporelle des engagements dans deux domaines d'activité bien séparés et reconnaissables, la conduite et l'usage du smartphone, tant du point de vue pratique que du point de vue moral.

Dans cet article, nous montrerons plus précisément comment les données contextuelles fournies par un dispositif original d'enregistrement des activités en mobilité (captures d'écran et lunettes-caméra), permet de documenter et d'analyser les situations de multi-activité durant lesquelles l'utilisateur est engagé dans une activité communicationnelle nomade et dans l'activité de déplacement. Il va s'agir de décrire l'organisation temporelle de ces activités en rendant compte de la manière dont les utilisateurs distribuent leur attention entre la mobilité et l'écran de leur smartphone. Au fil de nos analyses, nous mobiliserons ces données en rendant compte des préoccupations locales et situées des utilisateurs dans le but d'apporter quatre contributions distinctes et imbriquées qui vont de l'observation ethnographique à la conception du design, en passant par la construction de phénomènes pouvant être analysés en appui sur une réelle démarche probatoire :

— Cet article décrit une méthode qui permet d'enregistrer les activités réalisées depuis les écrans des smartphones, en situations d'usage naturelles, grâce à un système de capture d'écran qui est synchronisé avec l'enregistrement de la prise de vue contextuelle recueillie par des lunettes caméra.

— Cette méthode est appliquée à un échantillon d'utilisateurs de smartphones, sélectionnés en raison de leur utilisation assidue des applications dédiées aux réseaux sociaux

[2] En ce sens nous ne chercherons pas ici à identifier des logiques d'action « cachées » des utilisateurs individuels et qui seraient susceptibles de donner lieu à des formes particulières de réflexivité et de vigilance.

numériques. L'analyse des données ainsi recueillies fait émerger un phénomène central : l'organisation temporelle des orientations visuelles des usagers et les basculements de leurs regards entre l'écran du smartphone et des éléments saillants de l'environnement proximal. Nous documentons ces basculements visuels en décrivant la manière dont ils sont accomplis en lien avec différents événements liés à l'interface du téléphone mobile ou au contexte d'utilisation. Ces données permettent donc de décrire l'accomplissement pratique du flux des activités ainsi que leur imbrication, notamment sur un mode séquentiel.

— Ce papier montre comment l'analyse de la structuration séquentielle des engagements visuels des utilisateurs renouvelle notre compréhension de la mobilité et de la manière dont les smartphones l'équipent. Il fait apparaître le phénomène de la multi-activité, soit le fait que les engagements visuels des utilisateurs sont organisés sur un mode séquentiel. Il existe une orientation vers la multi-activité dans la mesure où certains événements, qui émergent de l'activité liée à l'écran ou à la conduite du déplacement, sont identifiés et traités comme des moments opportuns pour réorienter (temporairement) son attention en (ré) investissant un autre domaine d'activité. En permettant de décrire empiriquement l'organisation séquentielle du flux des activités, ce papier présente une méthode qui ouvre des perspectives de recherche nouvelles sur la multi-activité et l'engagement conjoint dans plusieurs cours d'action.

— Enfin, ce papier soulève des pistes de préconisations pour la conception des services mobiles en invitant à considérer la manière dont les interfaces s'inscrivent dans l'organisation séquentielle des flux d'activités. Chaque service projette, selon son interface et la manière dont son contenu est hiérarchisé, plus ou moins de moments pertinents pour désinvestir temporairement l'écran des yeux afin de regarder autour de soi, de gérer l'engagement conjoint dans une autre activité telle que le déplacement. Nous évoquerons alors l'idée d'une « texture pragmatique » des interfaces pour entrevoir une conception des services moins absorbante et davantage favorable à la gestion visuelle de la multi-activité.

2.- Méthodologie

La production de smartphones de plus en plus performants a favorisé l'essor des usages de l'Internet mobile, notamment la participation aux réseaux sociaux numériques en situation de mobilité (Lee, Kim, & Kim, 2005). L'étude de ces usages d'Internet soulève cependant de réelles difficultés méthodologiques liées aux contraintes du nomadisme. L'appropriation des applications exploitant les ressources de l'Internet mobile a été étudiée sur de longues périodes grâce au recueil et au traitement statistique de journaux de connexion (Tossell, Kortum, Rahmati, Shepard, & Lin, 2012). Malgré leur intérêt propre, ces méthodes ne donnent pas la possibilité de documenter convenablement les situations d'utilisation réelles, si ce n'est par des inférences indirectes. D'autres approches plus qualitatives (Isomorsu, Marika, Soili, & Kuuti, 2007), basées sur des entretiens ou sur la méthode des carnets d'utilisation, rendent compte de manière plus fine de l'expérience des utilisateurs et de leur sensibilité aux contextes d'usage. Cependant, malgré leur granularité d'analyse relativement fine, ces approches ne permettent pas d'articuler avec précision le déroulement des activités technologiques avec les situations de mobilité (Tamminen, Oulasvirta, Toiskallo, & Kandaine, 2004), afin notamment de proposer des interfaces dont l'usage s'ajuste plus efficacement aux déambulations des téléphonistes dans les lieux publics (Kane, Wobbrock, & Smith, 2008). C'est l'objectif que nous avons essayé d'atteindre en développant un dispositif d'observation basé sur des enregistrements vidéo.

Le recours à des enregistreurs portatifs offre en effet la possibilité de développer des méthodes prometteuses pour accéder au plus près de l'expérience des utilisateurs en enregistrant leurs usages, en récoltant des données sur la façon dont ils gèrent les

applications mobiles de leur smartphone d'une manière qui est sensible aux contextes et aux autres activités qu'ils réalisent en parallèle de leurs utilisations. Ces caméras portatives ont déjà été utilisées par l'anthropologie visuelle pour étudier la mobilité humaine (Pink, 2007). L'analyste filme l'engagement et l'itinéraire de la personne tout en enregistrant ses propres commentaires pour produire une visite visuelle égocentrée de différents environnements urbains (Oulasvirta, Tamminen, Roto, & Kuolahti, 2005). Cette prise de vue audiovisuelle, réalisée à la première personne, a également été développée dans le cadre d'approches ethnographiques des usages nomades des téléphones mobiles (Zouinar, Relieu, Salembier, & Calvet, 2004). Dans ce cas-là, les caméras portatives sont surtout utilisées pour documenter la multimodalité des interactions sociales en ciblant les contextes de mobilité où il est facile de « suivre » les utilisateurs, comme par exemple en équipant les automobiles de caméras (Laurier, Lorimer, Brown, Jones, Juhlin, Noble, et al., 2008 ; Mondada, 2012). Certains de ces travaux montrent ainsi comment les passagers gèrent collaborativement, en situation de conduite, les événements et les perturbations liés aux téléphones mobiles (Esbjornsson , Juhlin, & Weilenmann, 2007 ; Haddington & Rauniomaa, 2011) ou au guidage GPS (Brown & Laurier, 2012). Cependant, les protocoles vidéos déployés dans les habitacles des véhicules ne sont pas adaptés aux autres situations de mobilité, comme par exemple les transports en commun. C'est dans ces circonstances denses et complexes qu'il est alors pertinent d'utiliser des dispositifs d'enregistrement vidéo portatifs qui puissent être « portés » par les utilisateurs.

Là encore, le choix du dispositif portatif va être fonction du type de données empiriques recherché et va orienter le déroulement de l'analyse. Les caméras fixées au niveau du torse de l'utilisateur vont par exemple recueillir une prise de vue contextuelle qui délivre des indications sur les postures et les orientations corporelles de l'utilisateur. Mais la prise de vue de ces enregistrements vidéo ne donne pas réellement accès à l'expérience de l'utilisateur, en permettant par exemple de percevoir ce qu'il fait avec le dispositif technologique. En revanche, lorsque l'enregistrement vidéo est réalisé au niveau de la tête de l'utilisateur, dans le but de reproduire et de donner accès à ce que l'utilisateur voit, il est possible de recueillir des données égocentrées qui fournissent des prises analytiques plus solides pour analyser l'expérience de l'utilisateur et son utilisation d'une technologie nomade. Cette captation vidéo de l'expérience de l'utilisateur via son engagement visuel dans l'action, ou « en action », est aujourd'hui viable en raison de la miniaturisation des caméras portatives qui sont dorénavant relativement discrètes et ne transforment plus le participant en une sorte de « cyborg », comme c'était le cas dans les premières expérimentations de Steve Mann. De même, le perfectionnement des oculomètres permet aujourd'hui à la psychologie cognitive d'affiner l'analyse égocentrée de l'expérience des utilisateurs telle qu'elle peut être appréhendée sous le prisme des orientations oculaires (Yilmaz, Javed, & Shah, 2006). Cette démarche a été appliquée à l'étude de l'utilisation des téléphones mobiles (Drewes, De Luca, & Schmidt, 2007), mais les oculomètres et leurs protocoles expérimentaux restent difficilement transposables aux contextes réels de la mobilité humaine (Roto et al., 2004). Ces contraintes ont justifié le développement d'approches plus naturalistes basées sur le fait de demander à l'utilisateur de porter une caméra frontale (Omodei, McLenna,, & Whitford, 1998) ou des lunettes spécialement équipées d'une micro caméra et d'un microphone ; elles sont désignées de différentes manières au sein de la littérature scientifique : « lunettes vidéo », « subcams » ou « lunettes caméra » (Lahlou, 2006).

Ces lunettes ont été développées pour capturer le point de vue subjectif d'un sujet engagé dans une activité sociale, notamment pour l'amener à visionner les enregistrements afin qu'ils puissent rendre compte, de manière rétrospective, du vécu de son expérience (Lahlou, 2011). Ces lunettes ont ensuite été utilisées pour étudier les usages des téléphones mobiles en situation naturelle (Figeac, 2009 ; Mark, Christensen, & Shafae, 2001 ; Relieu, 2002). Ces enregistrements ont alors été traités de manière autonome ou ils ont été croisés avec l'analyse quantitative des traces d'usage (Oh & Lee, 2005 ; Zouinar et al., 2004). Cependant,

le cadrage et la résolution de ces vidéos ne sont pas suffisamment corrects pour permettre à l'analyste de percevoir le détail de ce qui est affiché sur les écrans des téléphones. Il est donc utile d'essayer de compléter cette prise de vue contextuelle, réalisée avec les lunettes, par l'enregistrement de la capture de l'écran pour avoir un accès précis et indépendant aux activités que les utilisateurs réalisent à l'aide de leur mobile (Relieu, 2009). Nous avons dès lors développé un dispositif qui nous permet de synchroniser ces deux prises de vue, sur le contexte d'usage et sur l'activité d'écran, afin de recueillir des données particulièrement fines sur la façon dont les utilisateurs gèrent leur engagement conjoint dans leur usage du smartphone et dans la conduite de leur déplacement dans un environnement urbain.

Pour analyser les usages des smartphones dans les situations ordinaires, nous proposons des données empiriques basées sur la synchronisation de deux prises de vue : l'enregistrement de la capture d'écran des smartphones et l'enregistrement vidéo des contextes d'usage réalisé à l'aide de lunettes caméra portées par l'utilisateur. Pour capturer l'activité d'écran, nous avons connecté un enregistreur audio et vidéo au smartphone (Figure 1), en reproduisant une méthode qui a été initialement développée pour étudier la visiophonie mobile (Morel & Licoppe, 2011).



Figure 1 : Un dispositif portatif permettant d'enregistrer les usages des smartphones Androïd

Figure 1: A simple portable set-up to record the audio-video flux on Android-based smartphones

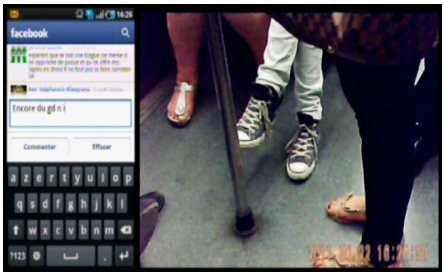
Ce système enregistre l'écran et l'ensemble des actions réalisées par l'utilisateur à l'aide de l'écran tactile du smartphone. Ce dispositif fournit ainsi des données très précieuses et particulièrement détaillées sur les usages « naturels » des services mobiles. Cela dit, hormis la date et l'heure, cette capture des écrans fournit peu d'informations par rapport au contexte d'utilisation. En parallèle de cette capture des écrans, nous avons alors demandé aux utilisateurs de porter des lunettes équipées d'une micro caméra pour obtenir des données sur les situations d'usage. Ce type de lunettes caméra a été utilisé par différents chercheurs pour bénéficier d'une « vue subjective » sur l'engagement des acteurs dans diverses activités sociales (Lahlou, 2011). Nous nous sommes appropriés ce dispositif d'enregistrement portatif pour obtenir des données sur l'utilisation des smartphones en situation de mobilité.

Pour l'analyste, la tentation est grande d'assimiler ce que les lunettes caméra donnent à voir à la perception visuelle réelle du participant. Pourtant, le cadrage de la prise de vue des lunettes caméra ne peut pas être confondu avec la vision humaine. En effet, le champ de la prise de vue de cette micro caméra est d'environ 40 degrés, ce qui est de très loin inférieur au champ de la vision humaine. Ces données vidéo ne procurent alors qu'un accès très relatif aux dimensions périphériques des orientations du regard. Le fait que l'angle de la caméra soit bien inférieur à la vision humaine a une autre conséquence fâcheuse lorsqu'on se sert de ce dispositif pour étudier les usages des smartphones. Quand les utilisateurs manipulent leur téléphone dans des environnements publics, ils ont tendance à le tenir légèrement incliné vers le bas, bien souvent à la périphérie de l'espace perceptif partagé (Figeac, 2010). Il apparaît alors que ce positionnement des téléphones est problématique puisque l'écran n'est

filmé qu'en partie (Figure 2a) ou alors il est situé hors du champ des lunettes caméra (Figure 2b).



(a)



(b)

Figure 2 : Dans les figures de cet article, la capture de l'activité d'écran apparaît sur la gauche et l'enregistrement vidéo réalisé avec les lunettes caméra est positionné sur la droite. Dans ces deux figures, l'utilisateur est assis et il oriente son smartphone vers le bas. Dès lors, au fil des basculements visuels de l'utilisateur, l'écran du smartphone est en partie dans le champ des lunettes caméra (a) ou pas du tout (b).

Figure 2: In all images the screen capture appears on the left, and the camera glasses recording on the right (they have been synchronized). In both images the user is sitting and using the smartphone on his lap, the phone being partly (a) or not at all (b) visible.

Cependant, ces contraintes peuvent être contournées afin de traiter ces enregistrements avec prudence. Tout d'abord, l'analyste voit sa tâche simplifiée par la petite taille de l'écran du smartphone. En effet, pour regarder le terminal, l'utilisateur doit orienter sa tête et son regard dans une direction assez précise, en général vers le sol (Figures 2a et 2b), qui indique assez clairement qu'il regarde en direction de l'écran. Cette indication peut alors être confirmée par la capture d'écran qui visibilise les opérations qu'ils réalisent depuis l'écran tactile du téléphone et prouve qu'il est effectivement impliqué visuellement dans cette activité. Toutefois, lorsque l'utilisateur se désengage visuellement de l'écran pour regarder autour de lui, il devient bien plus difficile de déduire ce qu'il regarde ou ce à quoi il accorde son attention. Par conséquent, le phénomène qui émerge de ces données vidéo recueillies à l'aide des lunettes caméra, telles que nous les synchronisons avec la captation de l'activité d'écran, renvoie à l'organisation des basculements visuels entre les phases où les utilisateurs investissent et désinvestissent visuellement le cours de l'activité réalisée avec le smartphone. Ces basculements des orientations visuelles peuvent être interprétés comme des phases de transitions entre les diverses activités et les divers engagements des participants. Dans les circonstances qui nous intéressent ici, ces basculements sont avant tout relatifs à deux types de pré-occupations, celles qui sont alimentées par les activités communicationnelles liées au smartphone et celles qui émergent de l'environnement urbain. Ces données fournissent ainsi des informations précieuses par rapport à la façon dont les utilisateurs de smartphones peuvent gérer leurs diverses implications au sein des situations de mobilité.

3.- Terrain

Nous avons recruté 10 participants, 5 femmes et 5 hommes âgés de 18 à 35 ans, qui consultent et participent fréquemment à Facebook à l'aide de leur smartphone. Ils ont accepté de porter des lunettes caméra pour enregistrer les usages qu'ils réalisent durant leur temps de transport, essentiellement entre leur domicile et leur lieu de travail, sur une période allant d'une semaine à dix jours. Après avoir recueilli les enregistrements, nous avons réalisé des « entretiens d'autoconfrontation » auprès des participants (Theureau, 2006) en leur faisant visionner les vidéos pour les amener à décrire et analyser le déroulement de leurs activités. Dans le cadre de cet article, nous n'exploiterons pas ces données recueillies *a posteriori* grâce à ces entretiens, car l'objectif ici est de développer une méthode d'analyse qui soit autonome, comme nous nous proposons de le faire en traitant de l'organisation

temporelle des basculements visuels. Précisons toutefois le fait que ces entretiens d'autoconfrontation représentent un moment de dialogue avec les participants où ils sont informés de l'usage que nous ferons de leurs enregistrements vidéo et où nous leur donnons la possibilité de remettre en question leur participation à cette recherche s'ils considèrent notre démarche comme trop intrusive par rapport à leur vie privée.

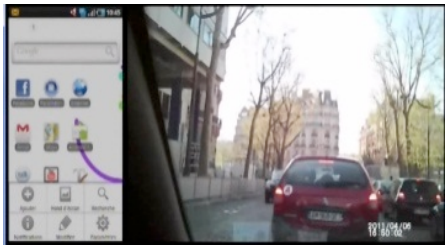
Notre corpus de données vidéo est composé de 25 heures d'enregistrements. Il regroupe essentiellement des utilisations des smartphones réalisées dans les transports en commun de la région Île-de-France. La consigne donnée aux utilisateurs était de ne rien changer à leurs pratiques habituelles. Toutefois, deux membres de notre échantillon ont également enregistré leur usage de Facebook pendant qu'ils étaient en voiture (usage illégal, mais qui leur est habituel), et accepté de nous laisser utiliser ces données de manière anonymisée. Même s'il s'agit là d'un cas plus particulier, nous avons décidé de le situer au cœur de cet article. L'intérêt d'analyser les utilisations faites en voiture tient au fait qu'elles contiennent moins d'informations que les usages réalisés dans les transports en commun. En effet, l'utilisatrice à laquelle nous allons nous intéresser est seule dans sa voiture. Dès lors, elle n'est pas en situation de coprésence directe avec autrui et elle n'a pas à échanger des regards avec un quelconque passager. Le fait qu'elle soit seule facilite notre projet d'analyser les basculements visuels dans la mesure où, dans les transports en commun, les usagers déploient divers types de regards autour de l'écran de leur smartphone pour surveiller le déroulement de leur trajet, pour se coordonner avec les autres passagers, etc. (Figeac, 2010). Lorsque l'utilisateur se trouve en situation de conduite automobile, les choses sont plus simples puisque ces désengagements visuels de l'activité d'écran peuvent être facilement imputables au déroulement de l'activité de conduite. C'est du moins ce que nous allons montrer. Sur cette base, il nous sera possible d'analyser l'organisation temporelle des basculements de regards dans cette forme de multi-activité développée au volant d'une voiture. Par la suite, dans le cadre d'autres articles, nous pourrons alors prendre appui sur cette première démonstration pour montrer comment ce phénomène des basculements visuels s'opère durant les activités de déplacement plus complexes, du moins d'un point de vue sociologique, lorsqu'elles sont réalisées au moyen des transports publics.

4.- L'observation de la gestion des activités conjointes : la communication mobile et le transport

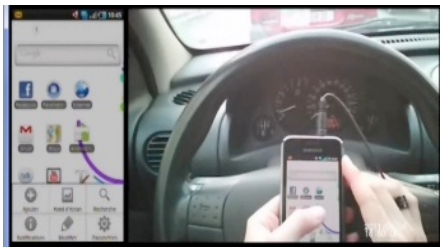
L'utilisation des smartphones en situation de mobilité suppose de traiter à la fois les contingences liées aux activités d'écran et celles liées à l'environnement plus global. En utilisant nos données empiriques, nous allons décrire le placement des basculements de regards, dédiés à la gestion de chacune de ces activités, pour mettre en lumière les aspects les plus importants de l'organisation temporelle de ces formes d'engagements multiples et enrichir le concept de multi-activité.

4.1. - Les basculements visuels et l'organisation temporelle des engagements multiples

Dans l'exemple que nous allons développer ici, l'utilisateur du smartphone est une femme qui se connecte à Facebook pendant qu'elle conduit sa voiture. Ce qui va nous intéresser ici, c'est la manière dont elle gère sur un plan pratique son engagement dans ces deux cours d'action simultanés liés à la conduite et à l'application Facebook. Pour montrer ce que révèle notre méthode d'observation quant aux phénomènes de multi-activité, nous allons analyser une séquence où la participante arrive à un feu rouge et immobilise sa voiture à la suite de la file des voitures qui la précèdent (Figure 3a). Il apparaît alors qu'une fois à l'arrêt, elle baisse légèrement les yeux et maintient son smartphone en appui au centre du volant de son véhicule (Figure 3b) avant d'ouvrir l'application Facebook et de consulter les dernières publications de ses amis.



(a)



(b)

Figure 3 : a) Immobilisation du véhicule à un feu rouge ; b) Saisir cette opportunité pour regarder vers le bas, placer le téléphone en appui au centre du volant et de lancer l'application Facebook.

Figure 3: a) Arriving at a red light behind the slowing traffic; b) Using this as an opportunity to look down, place the phone in the middle of the steering wheel and launch the Facebook application.

Un feu rouge est un dispositif de signalisation régulant le trafic automobile qui est faussement simple et familier. Il produit une sorte d'instruction aux conducteurs en leur ordonnant de s'arrêter pendant une durée qu'ils ont appris, par expérience, à évaluer. Dès lors, ce dispositif n'agit pas par lui-même. Comme le « gendarme couché », étudié par la sociologie des sciences et des techniques (Latour, 1999), le feu rouge forme la face visible d'un réseau vaste et hétérogène, d'un assemblage de personnes et d'objets à qui a été « délégué » une part de l'agentivité en jeu dans la coordination entre automobilistes. C'est cette infrastructure en réseau qui agit en arrière-plan pour que le feu rouge soit décodé par les conducteurs comme l'instruction d'immobiliser leur véhicule. Tel un indice visuel et sémiotique inscrit dans la « boîte noire » de l'infrastructure du réseau routier, le feu rouge projette l'injonction de marquer une pause. Les conducteurs partagent alors la compétence pratique de pouvoir évaluer la durée de cette pause. Pour l'utilisateur de smartphone au volant, ce feu rouge forme un moment opportun, un point de transition pertinent qui lui permet de se désengager temporairement de l'activité de conduite pour prendre en main son smartphone.

Une telle organisation temporelle des activités, en phases successives d'action et d'inaction, est structurée par la médiation d'artefacts techniques. Elle est révélée ici grâce à l'observation du rôle du feu tricolore et du déroulement de l'activité de conduite. Et elle émerge également de l'enregistrement des activités réalisées depuis les écrans tactiles des smartphones. En effet, lors de la navigation dans une application mobile, comme Facebook par exemple, il est très courant de voir afficher à l'écran un cercle de progression (Figure 4) après avoir « cliqué » sur l'interface pour initier une opération. Cette représentation graphique indique un temps de chargement et manifeste à l'utilisateur une phase de pause dans le déroulement de l'activité médiatique. Cet indicateur l'invite à attendre, voire elle lui préconise de ne plus cliquer sur l'interface pour ne pas initier une nouvelle opération. Ce cercle de progression ne délivre pas une instruction qui soit du même ordre que celle véhiculée par le feu rouge, dans la mesure où ce dernier s'inscrit à l'interface d'une plus large chaîne de médiations humaines et matérielles. Même si l'information délivrée n'est pas de même nature que celle délivrée par un feu tricolore, sur un plan juridique par exemple, il n'en demeure pas moins que cette représentation graphique projette également, sur un plan pratique, une phase de pause dans le déroulement de l'activité médiatique. Cet indicateur fonctionne alors comme un « indexical prospectif » (Goodwin, 1996)[3], qui indique ici à l'utilisateur un temps de chargement et l'invite à attendre la fin de l'opération avant d'en

[3] L'indexical prospectif est un dispositif qui pointe vers un moment futur.

initier une autre. Autrement dit, ces cercles de progression forment des médiations qui structurent le déroulement de l'activité médiatique tels des « points » de transition entre les phases d'action et d'inaction.



Figure 4 : L'affichage visuel typique de la barre de progression circulaire.

Figure 4: A typical visual display for the circular progress bar.

Revenons à notre exemple. Trois minutes après ce premier arrêt à un feu rouge, notre participante immobilise à nouveau son véhicule à un feu tricolore et elle baisse à nouveau son regard en direction de son smartphone. Durant l'intégralité de ce coup d'œil jeté aux dernières actualités de son compte Facebook, elle maintient l'écran en appui au centre de son volant. En procédant ainsi, elle manifeste son orientation vers cette multi-activité et elle adopte un positionnement ajusté à cette forme d'engagement dans la mesure où il facilite les basculements des orientations visuelles entre l'écran du smartphone et la route. Lors de cet arrêt au feu rouge, elle rédige le commentaire « joyeux anniversaire » avant de le publier sur le mur du compte Facebook d'un ami. Après avoir « cliqué » sur la fonctionnalité « Envoyer » de l'interface, la barre de progression circulaire apparaît à l'écran associée à l'information « Envoi du message » (voir la Figure 4 et la Figure 5a ci-dessous).

Durant l'affichage du cercle de progression, elle déplace sa main droite vers le bas pour enclencher une vitesse. Elle saisit alors cette pause dans l'activité réalisée avec le smartphone comme une opportunité pour passer une vitesse, soit une opération qui prouve qu'elle se réoriente vers la reprise en main de la conduite. Il convient ici d'insister sur l'importance de la prédictibilité de ces pauses, tant du point de vue de leur occurrence que de leur durée, car elles sont traitées comme des moments opportuns pour réinvestir la conduite malgré le fait que la reprise en main de cette activité ne soit pas encore nécessaire. Après avoir enclenché une vitesse, elle regarde en direction de la route (Figure 5c) et voit que le feu tricolore est toujours rouge. Le placement de ce coup d'œil révèle bien que son engagement dans la conduite n'était pas lié à un événement extérieur émergeant de l'environnement routier. Son placement séquentiel est bien à référer à la pause projetée par l'activité communicationnelle sous la forme d'un cercle de progression. De même, on peut considérer que le fait d'avoir immobilisé son véhicule à un feu rouge peut être saisi comme un moment opportun pour baisser à nouveau son regard en direction du smartphone (Figure 5d), plutôt que de continuer à surveiller l'état du trafic.

Ces observations révèlent notre compétence à exploiter l'organisation temporelle des activités, où elles apparaissent ponctuées de phases d'inactivité récurrentes et prédictibles, comme une ressource qui permet de gérer l'engagement simultané dans deux activités. Plus généralement, elles éclairent l'expérience de la multi-activité avec une réelle originalité empirique. On peut ainsi mieux cerner les contours de cette notion floue qui est bien souvent interprétée à partir de son sens commun, soit le fait de faire deux choses en même temps (ce qui n'est presque jamais le cas si on procède à une analyse détaillée de l'encastrement des cours d'action) ou à partir du cas extrême où la fragmentation des tâches est telle qu'il n'est plus possible d'identifier les basculements des engagements entre les différentes activités (Kenyon, 2010). Avec nos données, nous pouvons dire que l'engagement des participants dans des formes de multi-activité devient manifeste lorsque les phases d'inaction (voire éventuellement d'autres événements) que l'organisation temporelle d'une activité rend prévisible sont systématiquement traitées comme des moments opportuns pour basculer d'une activité à l'autre. Si l'on définit ainsi les propriétés reconnaissables de la multi-activité, il

apparaît que les utilisations de l'application Facebook durant la conduite des déplacements, en automobile ou en transport en commun, relèvent d'un engagement dans une forme de multi-activité.



Figure 5 :
a) Affichage de la barre de progression circulaire suite à l'envoi du message ;
b) la main droite est orientée vers la droite pour enclencher une vitesse ;
c) elle lève son regard en direction de la route et « découvre » que le feu est toujours rouge ;
d) elle regarde à nouveau en direction du smartphone.

Figure 5:
a) Sending a message and getting the circular progress bar;
b) the right hand moves to the right to get into gear;
c) She looks up at the road and "discovers" that the light is still red;
d) she immediately looks back down at the smartphone.

Cette manière d'organiser les basculements visuels autour des pauses projetées par un cours d'action forme une tendance récurrente au sein de notre corpus. On la retrouve notamment dans les séquences où les participants se trouvent dans un métro, même si ce type de mobilité ne requiert pas le même degré d'attention que la conduite automobile. Il apparaît en effet que les désengagements visuels de l'écran, opérés au profit de la gestion du déplacement, sont fréquemment placés lorsque les rames entrent en gare. Un tel placement des basculements visuels traduit l'orientation des utilisateurs à l'égard de la progression de leur trajet, mais aussi des autres passagers qui entrent et sortent. En parallèle de cette propension, il apparaît que les utilisateurs déploient également leurs basculements visuels lors des pauses projetées par l'activité d'écran, notamment lors de l'affichage de la barre de progression circulaire. Ils traitent là aussi ces phases de pause comme des moments opportuns pour lever les yeux et balayer visuellement l'environnement ambiant, et ce quelles que soient les circonstances, c'est-à-dire même quand la rame n'entre pas en gare et qu'il n'y a pas de mouvements de passagers.

Notons enfin que la littérature appréhende souvent la multi-activité à travers le prisme des arrangements situés entre des postures, des engagements corporels et l'environnement. Notre

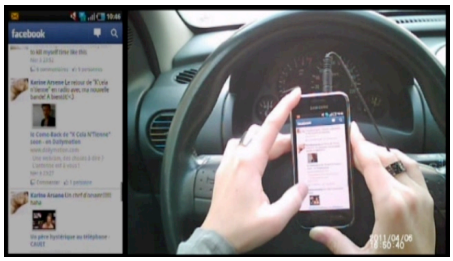
approche prolonge cette perspective en montrant que le positionnement des smartphones en appui sur le volant de la voiture ou, dans le métro, vers le bas en appui sur les cuisses forment un arrangement situé ajusté aux basculements visuels entre l'écran et l'environnement proximal.

4.2.- Les discordances dans l'organisation temporelle des différents cours d'action

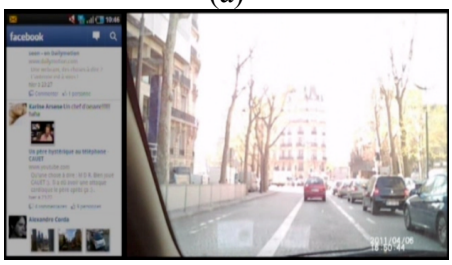
Le fait d'aborder ici une forme de multi-activité en lien avec la mobilité humaine et la conduite d'un déplacement renvoie inéluctablement au caractère désordonné des environnements urbains et à la manière dont ils peuvent contraindre les usages des technologies nomades. Ce « désordre » a été décrit pour rendre compte de l'hétérogénéité des infrastructures sociotechniques qui équipent les environnements urbains ainsi que de la diversité des arrangements interactionnels que des inconnus anonymes doivent opérer pour cohabiter (Bell & Dourish, 2007). On peut alors se demander comment les individus font l'expérience de ce désordre et composent avec les multiples sollicitations qui ponctuent leur mobilité.

En des termes plus pragmatiques, on peut se demander comment l'utilisateur d'une technologie nomade procède pour négocier ce système d'activités et d'engagements multiples ? S'il exploite, comme nous l'avons montré, les pauses projetées par les activités dans lesquelles il est engagé, il va de soi que ces pauses émergent dans le « désordre ». Autrement dit, l'émergence des points de transition d'un cours d'action n'est pas corrélée au déroulement des autres cours d'action qui composent la multi-activité. Les différentes activités gérées simultanément engagent des attentes temporelles différentes (le prochain arrêt en station, le temps de chargement d'un fil de discussion, etc.) qui peuvent converger ou qui peuvent au contraire entrer en concurrence en générant des opportunités séquentielles discordantes. Au cours de la précédente partie, nous avons détaillé un exemple de ce type de discordances. Lorsqu'elle est immobilisée au feu rouge, la conductrice profite du temps de chargement nécessaire à l'envoi de son message (Figure 5b) pour lever la tête et percevoir que le feu est toujours rouge (Figure 5c) avant de réinvestir temporairement son usage de Facebook (Figure 5d). Le fait d'arriver à gérer conjointement les « attentes » de l'activité communicationnelle et de la conduite révèlent la singularité de cette compétence pratique. Elle consiste à réussir à réguler les discordances temporelles entre les différents cours d'action pour accomplir les différentes activités, tout en respectant les exigences de la situation telles celles instituées par le code de la route.

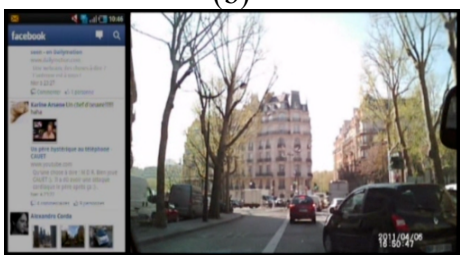
De tels phénomènes émergent de l'observation minutieuse de l'accomplissement situé des activités sociales. Le recours aux enregistrements vidéo répond à cette ambition de chercher à observer et à documenter ce type de phénomènes en affinant la granularité des analyses qui peuvent être menées à l'aide des techniques d'observation. Revenons par exemple à notre utilisatrice au moment où elle traite l'arrêt à un feu rouge comme une opportunité pour consulter son compte Facebook (Figure 3). Lorsque le feu passe au vert, elle ne lève pas immédiatement les yeux. Elle reste le regard orienté en direction du smartphone pendant un certain laps de temps malgré le fait que les autres voitures, visibles depuis l'enregistrement des lunettes caméra (Figure 6a), ont commencé à avancer. On peut donc déduire de son comportement que son absorption dans son usage est telle ici qu'elle s'est temporairement désolidarisée de la circulation automobile au point de reprendre en main sa conduite en retard, en décalage par rapport aux autres automobilistes (Figure 6b). Ce constat n'est pas une vue de l'esprit de l'analyste, il est fondé sur les données vidéo qui montrent qu'un automobiliste à exploiter le délai de réponse de l'utilisatrice pour changer de voie de circulation et prendre sa place dans l'organisation de la file d'attente au feu rouge (Figures 6b et 6c).



(a)



(b)



(c)

Figure 6 : a) La participante survole sa liste de notifications Facebook. Le mouvement des voitures environnantes est très nettement détectable dans les données, notamment à travers l'angle inférieur gauche du pare-brise ; b) une fois qu'elle a fini de consulter les notifications Facebook récentes, et que les notifications reçues « Hier à x heures » sont affichées à l'écran, elle réinvestit visuellement la route. Une grande distance sépare dorénavant sa voiture de celle qui la précède, révélant ainsi le caractère différé de sa réponse ; c) le conducteur de la voiture noire coupe sa trajectoire par la droite pour se positionner dans l'espace de la file d'attente qu'elle a laissée inoccupée.

Figure 6: a) The driver is running down her list of Facebook notifications. Whilst not visible in the picture, the movement of surrounding cars is clearly visible from the data, particularly through the side window; b) When she has finished reading new notifications and when notifications received "Yesterday at x o'clock" are displayed on the screen, she looks back at the road. There is now a large gap between her car and the car in front, demonstrating the delayed nature of her reaction; c) the black car on the right 'jumps' into the gap she has left, before she finally starts to move forward again.

La matérialisation du retard de sa réponse à cette sollicitation, qui émerge de l'activité de conduite, montre comment l'utilisation d'un smartphone peut venir interférer sur la conduite, même si cette interférence est mineure. Pour aller plus loin dans la compréhension du caractère différé de cette réponse, il est nécessaire de disposer de l'enregistrement vidéo des activités réalisées depuis le smartphone. En effet, au moment où le feu passe au vert, il apparait que l'utilisatrice était en train de consulter la page des notifications de son compte Facebook où sont listés divers messages : ceux qui lui ont été adressés à titre privé, l'annonce de la date de l'anniversaire de ses amis, les invitations à des événements, etc. Pour découvrir les notifications récentes, elle a fait défiler la page avec son index en effectuant un mouvement du haut (où les dernières notifications sont affichées) vers le bas (où les anciens messages sont mémorisés). Cette trace de son usage doit être analysée avec prudence, car cette capture de l'écran ne permet pas de dire si cette utilisatrice recherche ou attend une notification précise (c'est-à-dire s'il s'agit d'une activité polarisée vers un but) ou si elle vérifie simplement les dernières notifications avec l'espoir de voir sa curiosité captée par un message pertinent (c'est-à-dire s'il s'agit d'une activité pilotée selon un mode d'exploration opportuniste). Quoi qu'il en soit, ce qu'il y a de significatif dans cette page des notifications Facebook, c'est que leur structuration visuelle et pragmatique sous la forme d'une liste ne propose pas de fin. Cette liste de notifications Facebook, de par sa continuité répétitive, ne laisse pas envisager un point de transition à venir, qui soit instantanément identifiable et dont l'apparition serait suffisamment récurrente pour pouvoir être prédite par habitude, et qui constituerait pour les utilisateurs nomades des moments opportuns pour se désengager temporairement de l'activité communicationnelle et réorganiser le cours de leurs divers engagements. L'utilisateur doit définir ces points de transition par lui-même en faisant défiler la liste vers le bas jusqu'au moment où il va atteindre les notifications qu'il jugera non pertinentes (tout en considérant que les notifications découvertes lors d'un usage précédent peuvent toujours être pertinentes pour lui, par exemple s'il a différé sa réponse à un message qui lui avait ainsi été notifié). Cette définition des moments transitoires constitue

une opération profondément absorbante dans la mesure où les notifications inscrivent l'utilisateur au carrefour entre les sollicitations extérieures les plus récentes, qui viennent animer son compte, celles qui sont en attentes de réponse ainsi que l'historique des actualités passées. Si l'on croise maintenant cette activité d'écran singulière avec la prise de vue contextuelle, il apparait alors que l'utilisatrice s'est désolidarisée de la conduite au point de ne pas traiter un événement aussi important que le passage du feu au vert. On peut donc considérer que malgré l'interdiction d'utiliser son smartphone au volant et malgré le fait que les voitures avancent suite au passage du feu au vert, l'inertie liée à l'exécution de cette opération et à l'identification des points de transition pertinents est venue renforcer l'absorption de l'utilisatrice au point de l'amener à se désolidariser du trafic routier.

Dès lors, notre méthode de recherche nous permet d'observer la manière dont les utilisateurs gèrent leurs engagements multiples dans des activités structurellement hétérogènes qui sont susceptibles de générer des discordances temporelles. Nos données nous offrent également la possibilité de reconsidérer le concept de multi-activité, en faisant un focus ici sur la forme singulière de multi-activité qui peut être organisée autour des usages des technologies de communication en mobilité. Comme nous l'avons mis en évidence, les situations de multi-activité se caractérisent par le fait que les utilisateurs se rendent disponibles pour se laisser capter par l'implication séquentielle des événements qui émergent des activités. En exploitant cette implication, ils peuvent répondre aux événements extérieurs et réorganiser, *in situ*, l'accomplissement pratique de la multi-activité. Dans ce type de situations, on voit émerger de l'engagement conjoint dans différentes activités de nombreuses discordances temporelles. Ce désalignement de l'organisation, de la superposition temporelle des cours d'action est clairement identifiable. Il est même prédictible et il peut être, de ce fait, anticipé. Dans notre exemple, cette discordance est matérialisée par le fait que l'utilisatrice a laissé un espace vide relativement conséquent dans la file d'attente des voitures immobilisées au feu rouge. Ce désalignement est indirectement « sanctionné » par un automobiliste qui est venu positionner son véhicule dans l'espace vacant laissé par l'utilisatrice. Cet événement est banal. Le comportement de cet automobiliste peut être considéré comme prévisible dans la mesure où il n'a pas été sanctionné (par un coup de klaxon). Dès lors, en arrière-plan du caractère ordinaire de ce type d'évènement, on peut considérer que l'étude détaillée de l'utilisation d'un smartphone en situation de conduite révèle une forme de multi-activité qui est organisée en lien avec des préoccupations sociales et juridiques.

Notre démonstration n'aboutit pas sur l'idée que ces discordances temporelles entre les engagements multiples viennent nécessairement interférer sur le déroulement de la multi-activité. Nous avons détaillé un exemple où cela n'était pas le cas. Lorsque l'utilisatrice rédige et publie un commentaire Facebook pendant que son véhicule est immobilisé à un feu rouge (Figure 5), elle exploite le temps de chargement dédié à la publication du message pour enclencher une vitesse avant que le feu ne passe au vert. À travers ce « geste préparatoire » (Datchary & Licoppe, 2007), elle révèle son engagement dans une forme de multi-activité. Elle corrobore l'idée selon laquelle les pauses projetées par une activité tendent à être traitées comme des moments opportuns pour opérer les basculements de regards et/ou d'engagements. Le déploiement de ce geste, tel qu'il est ici séquentiellement ordonné, vient fluidifier la transition entre les activités et optimise la reprise en main de la conduite. Dès lors, le sens et l'importance de ces discordances temporelles ne peuvent pas être définis *a priori*. Ils émergent des circonstances pratiques et de la manière dont l'engagement dans des cours d'action conjoints amène à animer l'environnement et à définir la situation de multi-activité. Dans ces circonstances-là, la compétence pratique des utilisateurs peut être singularisée par leur aptitude à gérer de la manière la plus optimale possible les conséquences des désalignements temporels entre leurs divers engagements.

5.- La texture des interfaces des téléphones mobiles

Nous avons vu comment certains événements d'un flux d'activités émergent en suscitant des implications de type séquentiel et sont traités comme des points de transition potentiellement pertinents pour se désengager de l'activité d'écran et regarder autour de soi. Le fait que de tels événements soient traités et déclenchent une réaction est lié à la situation dans son ensemble et à la configuration de l'environnement. Il apparaît également que ces événements émergeant de l'activité d'écran sont extrêmement subtils et peu visibles. Nous avons notamment montré dans les exemples précédents comment l'événement lié au feu tricolore (le passage au feu vert) était traité en fonction de l'activité d'écran et de la structuration de la liste des notifications.

La méthodologie que nous proposons ici se révèle propice pour dévoiler avec une granularité suffisante l'accomplissement d'engagements corporels et visuels dans ces formes de multi-activité développées autour d'un écran. Un tel comportement repose sur des procédures pratiques non-réflexives qui sont finement articulées avec l'interface. C'est ainsi que les caractéristiques du design, comme l'heure de réception des notifications ou la barre de progression, peuvent occasionner des événements significatifs, en situations d'utilisation, qui acquièrent une réelle pertinence séquentielle. Ces caractéristiques, ces informations et ces indicateurs de l'interface peuvent être considérés comme des « affordances » (Gaver, 1991 ; Gibson, 1979 ; Kaptelinin & Nardi, 2012), c'est-à-dire comme une prise perceptive qui est autant inscrite dans l'interface que dans les habitudes de l'utilisateur et qui médiatise une conduite incarnée, non-réflexive. Dans l'optique de la multi-activité, ces affordances prennent une pertinence séquentielle en fonction de leur placement au sein de l'organisation temporelle du flux des activités, notamment ici en fonction de l'activité de conduite.

Nous pouvons appréhender ce type d'affordances comme des caractéristiques qui forment ce que nous proposons d'appeler la « texture pragmatique » des interfaces. En effet, les interfaces peuvent être décrites, sur un plan pragmatique, en fonction de la manière dont elles projettent ces formes de prises ou d'événements d'écran susceptibles d'interférer, par leur simple apparition, sur le déroulement de la multi-activité. Toutes les informations affichées sur l'interface ou tous les événements d'écran ne sont pas traités par les utilisateurs. Ils ne deviennent pas pertinents, sur un plan pratique, par le simple fait de se manifester à l'attention de l'utilisateur engagé dans un flux d'activités multiples. Dès lors, en suivant la perspective d'une théorie de la multi-activité, il s'agit de parler plus précisément de la « texture séquentielle » des interfaces pour rendre compte des prises et des informations, composant la « texture pragmatique » des services mobiles, rendues saillantes et pertinentes sur un plan pratique en fonction de l'engagement dans la multi-activité et de la manière dont cet engagement multiple amène à animer l'environnement proximal et l'environnement technologique.

Parler de la « texture séquentielle » des interfaces permet de rendre compte de la manière dont certaines informations affichées à l'écran ou certains indicateurs, comme les barres de progression, rendent prévisibles des opportunités de transition. C'est ainsi que nous avons montré comment la navigation dans un service comme Facebook projette de nombreux moments de pause, des phases de réactualisation du contenu, qui peuvent être décrits comme ayant une structuration séquentielle « robuste », tant leur apparition à l'écran se manifeste comme pertinente chez celui qui est au volant d'une voiture immobilisée à un feu rouge. Cette qualification de la texture séquentielle des interfaces, en une structuration robuste, ne se borne donc pas au domaine de l'environnement technologique. Elle doit être transposée et référée à l'environnement proximal. Par rapport à l'exemple précédent, il apparaît à quel point l'environnement routier est structuré séquentiellement, en raison du nombre de feux tricolores implantés pour l'organiser et du nombre d'instructions, temporellement synchronisées, qu'ils délivrent aux automobilistes pour coordonner les flux de la circulation. Plus l'environnement de la multi-activité est ainsi structuré

séquentiellement, plus il offre des points de transitions saillants, prédictibles et synchronisés qui garantissent une réelle fluidité dans le déroulement des basculements visuels et attentionnels entre les différents cours d'action gérés simultanément. Inversement, comme nous l'avons montré dans la partie précédente, plus les points de transitions émergeant des différents cours d'action sont difficiles à synchroniser, plus l'organisation temporelle de la multi-activité repose sur des discordances temporelles. Et par conséquent, plus elle repose sur une structuration séquentielle, de l'environnement ambiant et de la texture pragmatique de l'interface, qui est « lâche ».

Notre méthodologie paraît donc particulièrement en phase avec les exigences de l'analyse des usages des smartphones en situation de mobilité, car elle fournit des données empiriques relativement fines qui prouvent, en observant les situations d'usage naturelles, à quel point les utilisateurs « répondent » aux évènements d'écran en tant qu'ils offrent des opportunités séquentielles. Elle révèle ce que nous appelons la « texture pragmatique » de l'interface, c'est-à-dire les caractéristiques enfouies dans la conception du design, pour structurer et implémenter un service, qui peuvent venir générer des événements si et seulement si elles manifestent une information pertinente par rapport au flux des activités dans lequel l'utilisateur est engagé.

Cette mise en évidence de la « texture pragmatique » des interfaces invite dès lors à travailler la conception du design afin de revisiter l'affichage des événements d'écrans en prenant en compte leurs possibles implications séquentielles et leurs effets sur les comportements des utilisateurs. Par exemple, nos observations révèlent que les temps de chargement de l'application Facebook et plus généralement les pauses projetées par le smartphone sont régulièrement traitées comme des moments opportuns pour suspendre temporairement son engagement visuel dans l'activité d'écran et regarder autour de soi. Ces phases d'inaction liées à l'usage du smartphone se présentent donc comme des ressources importantes pour les utilisateurs pour gérer leur engagement dans des activités conjointes. Actuellement, les créateurs d'interfaces représentent ces phases d'inaction à l'aide de barres ou de cercles de progression. Ils cherchent alors à travailler le design de ces indicateurs temporels et de ces représentations graphiques pour atténuer le sentiment d'attente perçu par l'utilisateur (Harrison, Amento, Kuznetsov, & Bell, 2007) et remédier au déplaisir éprouvé durant l'utilisation d'un service dont l'actualisation du contenu est peu réactive. Notre approche propose de compléter ces travaux en suivant la perspective d'une théorie de la multi-activité. Plutôt que de voir ces temps de chargement comme de pures contraintes et de créer des interfaces qui les invisibilisent le plus possible, il peut s'avérer nécessaire de redonner la main à l'utilisateur en tentant d'indiquer le plus clairement possible la durée de ces temps de pauses. Grâce à ces indications, il s'agit de l'aider à optimiser son engagement conjoint dans les différents cours d'action d'une multi-activité en l'aidant à anticiper les pauses projetées par l'application afin d'améliorer la manière dont il va pouvoir organiser ces basculements visuels et attentionnels entre l'environnement technologique et l'environnement proximal.

6.- Conclusion

Pour pallier les problèmes d'observation des usages des technologies de communication en situation de mobilité, nous avons développé un dispositif d'enregistrement vidéo portatif basé sur la synchronisation de deux prises de vue : une prise de vue sur le contexte d'utilisation recueillie à la première personne à l'aide de lunettes caméra et la captation, à l'aide d'un enregistreur audio-vidéo, des activités réalisées avec les smartphones, telles qu'elles apparaissent sur les écrans tactiles en révélant les « clics » et les opérations des utilisateurs. Nous avons montré comment ces enregistrements permettent d'observer deux types de phénomènes : a) les basculements visuels des participants entre l'écran du smartphone et l'environnement proximal ; b) l'agencement des gestes et des postures corporelles dans une forme de multi-activité développée autour des usages des smartphones

et de l'activité de déplacement.

Nous avons testé cette méthode de recherche en demandant à un échantillon d'utilisateurs d'enregistrer leurs utilisations du smartphone durant leurs trajets quotidiens, en voiture ou en transport en commun, entre leur domicile et leur travail. L'observation des basculements visuels se présente alors comme un puissant outil pour comprendre comment les utilisateurs procèdent pour gérer leurs engagements simultanés dans différentes activités, et plus précisément ici dans l'usage d'applications mobiles et la mobilité. En nous basant sur la manière dont les utilisateurs traitent les temps de pause projetés par les smartphones ou l'activité de transport, nous avons été en mesure de proposer une définition originale et centrée sur l'utilisateur de ce que signifie pour une personne de s'engager dans une situation de multi-activité. Il s'agit d'une forme d'engagement dans la gestion simultanée de deux activités qui invite à traiter les pauses et les phases d'inaction comme une occasion de basculer son orientation visuelle et/ou son attention d'une action à l'autre, d'une tâche à une autre, d'une opération à une autre. Par exemple, l'apparition de la barre de progression circulaire, à côté de laquelle est inscrit « Chargement » ou « Envoi du message », forme un indicateur qui projette une pause dans le déroulement de l'activité liée au smartphone. Elle tend à être traitée comme un moment opportun pour vérifier ce qui se passe autour du téléphone, autrement dit pour basculer visuellement vers les zones de l'environnement pertinentes à investir dans ces circonstances pratiques pour gérer la mobilité. Selon cette perspective, l'utilisation des différentes modalités écrites de communication (les publications Facebook, les courriers électroniques, les SMS, les sessions de « Chat », la diffusion de « Tweet », etc.) participent nécessairement, lorsqu'elles se déroulent en situation de mobilité, à cette forme d'engagement caractéristique de la multi-activité.

Ces événements, émergeant des activités, peuvent être caractérisés par le fait qu'ils sont traités comme des créneaux propices au basculement des orientations visuelles ou des engagements. Ils projettent une implication de type séquentielle dans la mesure où leur apparition rend une réaction plus pertinente qu'une autre. Cependant, si le flux d'une activité projette diverses implications séquentielles à des moments différents, il n'est pas forcément pertinent de toutes les traiter et d'y répondre, car elles ne sont pas nécessairement en phase avec les exigences des autres cours d'action qui composent la multi-activité. De telles discordances temporelles interviennent continuellement, nécessitant de réorganiser et de re-synchroniser les implications séquentielles des activités gérées conjointement. Par exemple, nous avons montré comment une participante s'est désolidarisée de l'état du trafic routier, en oubliant de reprendre en main sa conduite au moment où le feu est passé au vert, tant elle était absorbée par la lecture du fil des notifications de son compte Facebook. Nous avons renvoyé son absorption et le fait qu'elle ait réinvesti sa conduite en retard à la texture pragmatique de la liste des notifications qui ne projette pas de fin : elle n'est pas structurée autour de phases d'inaction susceptibles d'être traitées par la conductrice comme des opportunités pour réinvestir sa conduite plus tôt. Le travail d'alignement effectué par les utilisateurs pour traiter ce type de discordances apparait au final comme une compétence pratique fondamentale dans la gestion des situations de multi-activité.

Les interfaces des applications mobiles jouent également un rôle important dans la manière dont la multi-activité est gérée dans la mesure où elles influent différemment, en fonction de leur design et de la structuration de leurs contenus, sur la projection des points de transition, des phases d'action et d'inaction. Nous avons introduit l'idée d'une texture des interfaces pour essayer de rendre compte de cette propension des interfaces de générer des événements d'écran qui interfèrent sur la définition du cadre de la situation d'usage. La « texture pragmatique » d'une interface renvoie à la manière dont une interface visibilise certains événements d'écran comme saillants, tout en les projetant comme pertinents à traiter et ce tout en induisant l'orientation à suivre. Cette texture pragmatique rend dès lors compte de la propension des interfaces à générer des événements d'écran qui projettent une implication de type séquentielle puisque ces événements discriminent les orientations les plus pertinentes à

suivre pour les traiter. Il ne s'agit là que d'une simple implication puisque l'utilisateur peut ne pas la prolonger. Et cette implication ne garantit en rien les conditions de félicité de l'activité dans la mesure où l'utilisateur qui y donne suite peut faire face, comme nous l'avons montré, à un désalignement des cours d'action de la multi-activité dans laquelle il est engagé.

Cette description de l'implication séquentielle qui émane de la texture des interfaces nous permet de documenter plus rigoureusement la problématique sociale liée à l'absorption visuelle dans une activité d'écran. L'absorption « abusive » par les écrans est souvent évoquée pour rendre compte des dangers de certaines pratiques télévisuelles et vidéo-ludiques des enfants et des adolescents. Elle est alors expliquée à l'aide de différentes causes qui vont des manquements parentaux à la forte attractivité du flux des images par rapport à l'écrit. Notre travail permet quant à lui de mieux comprendre les dimensions les plus ordinaires de l'absorption médiatique en contribuant à tenir compte de la manière dont les usagers reconnaissent et gèrent les demandes « sociales » qui émergent autour des écrans, notamment autour des smartphones dans les lieux publics. Nous avons montré ici comment observer et analyser, en tant que processus situé, l'engagement conjoint dans des foyers attentionnels et des cours d'action multiples. Le fait de rester absorbé dans une activité d'écran ou, au contraire, le fait d'arriver à composer avec les diverses exigences des activités semblent être liés à la façon dont les utilisateurs identifient, traitent ou ignorent l'implication séquentielle des événements qui émergent de l'écran ou qui apparaissent dans l'environnement proximal. L'aptitude à se désengager à temps de l'activité faite à l'écran pour gérer convenablement la multi-activité peut être assimilée à une compétence pratique consistant à savoir réparer les décalages temporels générés par les implications séquentielles contradictoires des cours d'action. Il va de soi que le fait d'invoquer cette aptitude ne se situe pas du tout sur le même plan que les politiques éducatives instaurées pour contrecarrer les méfaits d'une consommation médiatique excessive. Seulement, la compréhension de ce processus peut nous aider à reconsidérer le design des interfaces et la structuration des contenus d'un service.

On peut en effet considérer que le design d'une application est d'autant plus en phase avec les exigences du nomadisme et des situations de mobilité qu'ils optimisent les désengagements visuels temporaires de l'activité d'écran en projetant des phases de pauses. De la fréquence des pauses qu'un service projette va donc dépendre l'aptitude de l'utilisateur à réinvestir visuellement l'environnement proximal et, bien évidemment, ceux avec qui il coexiste. Il est donc possible d'optimiser les utilisations et, par là même, d'atténuer les effets néfastes de l'absorption en créant des services et des interfaces qui sont structurés autour de rythmes moins absorbants, en permettant des micro-pauses. Par exemple, un jeu vidéo sur smartphone sera plus ou moins absorbant et addictif en fonction du nombre de pauses qui viennent rythmer le déroulement des parties. L'étude détaillée des usages des technologies de la communication en situation de mobilité montrent donc qu'elles participent à des formes de multi-activité complexes qui nous demandent de mieux étudier le déroulement temporel des utilisations pour comprendre les effets de ces technologies sur les processus attentionnels et, plus généralement, sur la participation la plus ordinaire des utilisateurs à la vie publique telle qu'elle s'opère bien souvent sous la forme d'échange de regards.

BIBLIOGRAPHIE

Bell, G., & Dourish, P. (2007). Yesterday's tomorrows: notes on ubiquitous computing's dominant vision. *Personal & Ubiquitous Computing, 11*(2), 133-143.

Brown, B., & Laurier, E. (2012). The normal natural troubles of driving with *GPS*. In *Proceedings CHI 2012*, ACM Press, pp. 1621-1630.

Circella, G., Mokhtarian, P., & Poff, K. (2012). A Conceptual Typology of Multitasking Behavior and Polychronicity Preferences. *Electronic International Journal of Time Use Research, 9*(1), 59-107.

Datchary, C., & Licoppe, C. (2007). La multi-activité et ses appuis : l'exemple de la « présence

obstinée » des messages dans l'environnement de travail. *Activités*, *4*(1), 4-29, http://www.activites.org

Drewes, H., De Luca, A., & Schmidt, A. (2007). Eye-gaze interaction for mobile phones. In *Proceedings Mobility 2007*, ACM Press, pp. 364-371.

Esbjornsson, M., Juhlin, O., & Weilenmann, A. (2007). Drivers Using Mobile Phones in traffic: An Ethnographic Study of Interactional Adaptation. *International Journal of Human Computer Interaction, 22*(1-2), 37-58.

Figeac, J. (2009). L'appropriation de la Mobile TV autour des réseaux de communication, *Réseaux*, 156, 81-111.

Figeac, J. (2010). La participation des usagers de la Mobile TV aux formes visuelles de la coordination sociale. *Les Cahiers du Numérique, 6*(2), 39-62.

Gaver, W.W. (1991). Technology affordances. In *Proceedings CHI 1991*, ACM Press, pp. 79-84.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton-Mifflin.

González V. M., & Mark G. (2005) Managing currents of work: Multi-tasking among multiple collaborations. In H. Gellersen, K. Schmidt, M. Beaudouin-Lafon, & W. Mackay (Eds.) *Proceedings of the Ninth European Conference on Computer-Supported Cooperative Work (ECSCW 2005),* (pp.143-162). Dordrecht: Springer.

Goodwin, C. (1984). Notes on story structure and the organization of participation. In M. Atkinson, & J. Heritage (Eds.), *Structures of social action* (pp. 225-246). Cambridge: Cambridge University Press.

Goodwin, C. (1996). Transparent vision. In E. Ochs, E.A. Schegloff, E., & S.A. Thompson S. (Eds.), *Interaction and grammar* (pp. 370–404). Cambridge: Cambridge University Press.

Goodwin, C. (2002). Multi-modality in Girls' Game Disputes. *Journal of Pragmatics, 24*(10-11), 1621-1649.

Haddington, P., & Rauniomaa, M. (2011). Technologies, Multitasking and Driving: Attending to and Preparing for a Mobile Phone Conversation in a Car. *Human Communication Research, 37*, 223-254.

Harrison, C., Amento, B., Kuznetsov, S., & Bell, R. (2007). Rethinking the progress bar. In *Proceedings. of the ACM Symposium on User interface Software and Technology*. ACM, Press, pp. 115-118.

Hudson, J., Christensen, J., Kellogg, W., & Erickson, T. (2002). "I'd be overwhelmed, but it's just one more thing to do" : Availability and interruption in research management. In *Proceedings of ACM CHI 2002 Conference on Human Factors in Computing Systems*, pp. 97-104.

Isomorsu, M., Marika, T. Soili, V., & Kuutti, K. (2007). Experimental evaluation of five methods for collecting emotions in field settings with mobile applications. *International Journal of Human-Computer Studies, 65*(4), 404-418.

Kane, S.K.**,** Wobbrock, J.O., & Smith, I.E. (2008). Getting off the treadmill: Evaluating walking user interfaces for mobile devices in public spaces. In *Proceedings of MobileHCI '08,* ACM Press, pp. 109-118.

Kaptelinin, V., & Nardi, B. (2012). Affordances in HCI: toward a mediated action perspective. In *Proceedings CHI 2012,* ACM Press, pp. 967-976.

Kaufman-Scarborough, C. (2003). Two perspectives on the tyranny of time: Polychronicity and monochronicity as depicted in *Cast Away*." *The Journal of American Culture, 26*(1), 87-95.

Kenyon, S. (2010). What do we mean by multitasking? Exploring the need for methodological clarification in time use research. *International Journal of Time Use Research, 7*(1), 42-60.

Lahlou, S. (2006). L'activité du point de vue de l'acteur et la question de l'inter-subjectivité : huit années d'expériences avec des caméras miniaturisées fixées au front des acteurs (subcam). *Communications, 80*, 209–234.

Lahlou, S. (2011). How can we capture the subject's perspective? An evidence-based approach for the social scientist. *Social Science Information, 50*(3-4), 607-655.

Latour, B. (1999). *Pandora's hope. Essays on the reality of science studies*. Cambridge: Harvard University Press.

Laurier, E. (2000). Why people say where they are during mobile phone calls. *Environment and*

Planning D: Society and Space, 19, 485-504.

Laurier, E, Lorimer, H., Brown, B., Jones, O., Juhlin, O., Noble, A., Perry, M., Pica, D., Sormani, P., Strebel, I., Swan, L, Taylor, A., Watts, L., & Weilenmann, A. (2008). Driving and passengering. Notes on the ordinary organization of car travel. *Mobilities, 3*(1), 3-23.

Lee, I., Kim, J., & Kim, J. (2005). Use contexts for the mobile internet: A longitudinal study monitoring actual use of mobile internet services. *International Journal of human-Computer Interaction, 18*(3), 269-292.

Leont'ev, A. N. (1974). The Problem of Activity in Psychology. *Soviet Psychology, 13*(2), 4-30.

Licoppe, C. (2008). « Logiques d'innovation, multiactivité et zapping au travail ». *Hermès, 50*, 171-178.

Livingston, E. (2008). *Ethnographies of Reason*. Farnham: Ashgate.

Mark, G., Christensen, U., & Shafae, M. (2001). A methodology using a microcamera for studying mobile IT usage and person mobility. *CHI 2001, Workshop on Mobile Communication: Understanding Users, Adoption and Design*.

Mintzberg, H. (1973). *The Nature of Managerial Work*. New York: Prentice Hall.

Mondada, L. (2012). "Talking and driving". Multiactivity in the car. *Semiotica*, *191*, 223-256.

Morel, J., & Licoppe, C. (2011). Studying mobile video telephony. In M. Büscher, J. Urry, & K. Witchger, *Mobile Methods* (pp. 164-182). Routledge.

Nevile, M. (2012) Interaction as distraction in driving: A body of evidence. *Semiotica, 191*, 169-196.

O'Connaill, B., & Frohlich, D. (1995). Timespace in the Workspace. Dealing with Interruptions. *Proceedings of Human factors in Computing (CHI'95)*, Pittsburgh, PA: ACM Press.

Oh, Y.S., & Lee, K.P. The development of user Observation method and analysis tool in Mobile environment - With emphasis on the development of observation method at user's point of view by wearable micro-camera and analysis tool. http://dpl.kaist.ac.kr/web_wiki/images/1/18/Ohys2005.pdf

Omodei, M. M., McLennan, J., & Whitford, P. (1998). Improving decision making in complex natural settings: Using a head-mounted video camera to improve performance of competitive orienteers. *International Journal of Sport Psychology, 29*(2), 115-131.

Oulasvirta, A., Tamminen, S., Roto, V. & Kuolahti, J. (2005). Interaction in 4-Second Bursts: The Fragmented Nature of Attentional Resources in Mobile HCI. In *Proceedings of CHI 2005*, ACM Press, pp. 919-928.

Pink, S. (2007). *Visual Anthropology*. London: Sage.

Relieu, M. (2002). The "glasscam" as an observational tool for studying screen-based mobile phone uses and management of parallel activities. *International Conference on Conversation Analysis (ICCA-02)*, Copenhagen, 17-21 May.

Relieu, M. (2009). Mobile phone "work": disengaging and engaging mobile phone activities with concurrent activities. In R., Ling, & S., Campbell (Eds.). *The reconstruction of space and time: Mobile communication practices* (pp. 215-230). New Brunswick (NJ): Transaction Publishers.

Roto, V., Oulasvirta, A., Haikarainen, T., Kuorelahti, J., Lehmuskallio, H., & Nyyssönen, T. (2004). Examining Mobile Phone Use in the Wild with Quasi-Experimentation. In *HIIT Technical reports*.

Stivers, T., & Sidnell, J. (2005). Introduction: multi-modal interaction. *Semiotica, 156*(1/4), 1-20.

Streeck, J., Goodwin, C., & LeBaron, C. (2011). *Embodied Interaction*. Cambridge: Cambridge University Press.

Tamminen, S., Oulasvirta, A., Toiskallo, K., & Kankainen, A. (2004). Understanding mobile contexts. *Personal and Ubiquitous Computing, 8*, 135-143.

Theureau, J. (2006). *Le cours d'action. Méthode développée*. Toulouse: Octarès Editions.

Tossell, C., Kortum, P., Rahmati, A., Shepard, C., & Lin, Z. (2012). Characterizing web use on smartphones. In *Proceedings CHI 2012,* ACM Press, pp. 2769-2778.

Wajcman, J., & Rose, K. (2011). Constant Connectivity: Rethinking Interruptions at Work. *Organization Studies*, *32*(7), 941-961.

Yilmaz, A., Javed, O., & Shah, M. (2006). Object tracking: A survey. *ACM Computing Surveys, 38*(4), 1-45.

Zouinar, M., Relieu, M., Salembier, P., & Calvet, G. (2004). Observation and capture of multimodal interaction in mobile situations. In *Proceedings UbiMob 2004*, ACM Press, pp. 5-8.

RÉSUMÉ

Pour comprendre empiriquement la manière dont est organisée la gestion temporelle des activités qui caractérise les situations de multi-activité, nous avons développé une méthode de captation vidéo des usages des téléphones mobiles en situation de mobilité. Elle est basée sur la synchronisation de deux types d'enregistrements vidéo : une prise de vue contextuelle recueillie à l'aide de lunettes caméra et la captation de l'activité d'écran des smartphones réalisées avec un enregistreur audio-vidéo. Elle permet notamment d'analyser comment les utilisateurs s'approprient les applications mobiles en situation de mobilité. Ces données révèlent l'organisation spatio-temporelle des basculements visuels des utilisateurs (autour de l'écran des terminaux mobiles) et délivrent des informations précieuses pour comprendre comment les utilisateurs gèrent actuellement leurs usages en parallèle des autres activités qui composent les situations ordinaires de multi-activité. Parler de multi-activité reflète une posture particulière des acteurs envers les implications séquentielles potentielles, qui émergent de l'imbrication temporelle des événementsliés à la fois à la navigation dans les services mobiles et à l'environnement de la mobilité ; ces événements formant des occasions d'opérer un basculement visuel d'un foyer d'activité à l'autre. Nous discutons alors du fait que la « texture pragmatique » des interfaces devrait être structurée pour proposer davantage d'opportunités séquentielles de ce type afin qu'elles projettent une organisation temporelle des usages plus en phase avec les paramètres de la multi-activité.

MOTS-CLES

mobilité, téléphone portable, multi-activité, lunettes caméra, basculements visuels

REFERENCEMENT

Licoppe, Ch., & Figeac, J. (2014). L'organisation temporelle des engagements visuels dans des situations de multi-activité équipée en milieu urbain. *Activités*, *11*(1), 65-85, http://www.activites.org/v11n1/v11n1.pdf

Article soumis le 12 avril 2013, accepté pour publication le 26 septembre 2013.