HAL
open science

Premiers retours d'expérience sur l'utilisabilité et les usages de systèmes interactifs plastiques

Fatoumata Camara, Gaëlle Calvary, Rachel Demumieux, Vincent Ganneau

To cite this version:

Fatoumata Camara, Gaëlle Calvary, Rachel Demumieux, Vincent Ganneau. Premiers retours d'expérience sur l'utilisabilité et les usages de systèmes interactifs plastiques. Journal d'Interaction Personne-Système, 2014, Volume 1 (1), pp.1-37. 10.46298/jips.58 . hal-01058922

HAL Id: hal-01058922

https://inria.hal.science/hal-01058922v1

Submitted on 28 Aug 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Premiers retours d'expérience sur l'utilisabilité et les usages de systèmes interactifs plastiques

FATOUMATA CAMARA
[1] Laboratoire d'Informatique de Grenoble
[2] Orange Labs
GAELLE CALVARY[1]
RACHEL DEMUMIEUX[2]
VINCENT GANNEAU[1, 2]

Résumé : La propriété de plasticité a été introduite en 1999 en France en réponse à une variabilité nouvelle du contexte d'usage en intelligence ambiante. Une Interface Homme-Machine (IHM) est dite *plastique* lorsqu'elle est douée d'adaptation à son contexte d'usage dans le respect de propriétés centrées utilisateur. Dix ans plus tard, les avancées sont certaines. On recense de nombreux résultats aussi bien théoriques que pratiques. Cependant, l'utilisabilité et l'acceptabilité des IHM plastiques soulèvent de nombreuses questions qui nécessitent la mise en œuvre d'évaluations auprès d'utilisateurs. Cet article présente l'état de l'art en plasticité et en évaluation (méthodes et difficultés). Il décrit trois démonstrateurs et présente les premiers retours d'évaluations menées en laboratoire et sur le terrain.

Mots clés : Plasticité, Adaptation, Contexte d'usage, Evaluation, Expérimentation terrain.

Abstract: The *Plasticity* property has been introduced in 1999 in France to cope with the new variability of the context of use in ambient intelligence. A User Interface (UI) is said to be *plastic* if it is able to adapt to its context of use while preserving human-centered properties. Ten years later, theoretical as well as practical advances have been made. However, several questions related to usability and acceptability of plastic UIs need to be answered through evaluation with users. This paper presents a state of the art in plasticity and evaluation (methods and concerns). It describes three demonstrators and presents first results from evaluations led in laboratory and in the wild.

Key words: Plasticity, Adaptation, Context of use, Assessment, Field test.

Les articles de JIPS sont publiés sous licence Creative Commons Paternité 2.0 Générique.

1. INTRODUCTION

Avec les avancées des réseaux et la miniaturisation des dispositifs, l'utilisateur n'est plus immobile, assis devant un poste fixe. Il accède à des ressources variées (PC, téléphone, borne publique, services Web) qui apparaissent et disparaissent opportunément. L'interaction homme-machine devient possible en tout lieu, à tout instant, sur tout support. Dès lors, les Interfaces Homme-Machine (IHM) doivent pouvoir se conformer à un contexte d'usage désormais varié, variable et potentiellement imprévisible.

Au fil des années, les IHM ont progressivement évolué pour devenir adaptées, adaptables et adaptatives [Edmonds 1981] :

- *Adaptées :* dans une IHM adaptée, l'adaptation est mise en œuvre par le concepteur pour une cible prédéfinie d'utilisateurs [Benadi 2004]. Ceci constitue un premier niveau d'adaptation, qualifié de statique [Bobillier-Chaumon and al. 2005] ;
- *Adaptables* : une IHM adaptable est une IHM modifiable par l'utilisateur. Il choisit les adaptations qui lui conviennent le mieux en fonction de ses préférences et de ses habitudes (éléments graphiques, raccourcis, etc.) [Benadi 2004]. L'adaptation est dynamique, c'est-à-dire qu'elle a lieu pendant l'exécution du programme, mais demeure entièrement sous le contrôle de l'utilisateur [Bobillier-Chaumon and al. 2005] ;
- *Adaptatives* : dans une IHM adaptative, l'adaptation est prise en charge par le système sans intervention explicite de l'utilisateur. L'adaptation s'appuie sur un processus d'acquisition et d'exploitation d'un modèle utilisateur impliquant des formes d'apprentissage, d'inférence ou d'aide à la décision [Jameson 2003].

En 1999, face à l'intelligence ambiante, la communauté IHM a introduit la propriété de *plasticité,* définie comme étant la capacité d'adaptation d'une IHM à son contexte d'usage dans le respect de son utilisabilité [Thevenin and Coutaz 1999]. Dans cette définition, c'est une variation du contexte d'usage qui est à l'origine de l'adaptation. L'utilisateur peut être impliqué dans le processus d'adaptation couvrant ainsi les deux niveaux IHM *adaptables* et *adaptatives*. Les IHM utilisées peuvent être *adaptées* : plus elles le seront, plus probablement l'utilisabilité sera préservée. Cette préservation de propriétés centrées utilisateur est la raison d'être et le défi de la plasticité.

Dix ans après la définition de cette propriété, les avancées sont notables en conception d'IHM plastiques. Cependant, l'évaluation reste sous étudiée : peu de résultats d'évaluation sont publiés et validés par la communauté scientifique [Gena 2005]. De nombreuses questions restent encore sans réponse, notamment, à quel moment adapter, quelles données sont pertinentes pour le modèle utilisateur, quel niveau de contrôle laisser à l'utilisateur lors du processus d'adaptation ou encore comment intégrer les besoins d'une très large cible d'utilisateurs ?

Cet article apporte des éléments de réponse. Il énonce des critères ergonomiques formulés à partir d'enseignements issus d'évaluations de trois démonstrateurs. Ces démonstrateurs illustrent les leviers de l'adaptation : le remodelage et la redistribution d'IHM. Les plates-formes utilisées sont le téléphone mobile et l'ordinateur. Notre objectif est d'évaluer les performances et le ressenti des utilisateurs vis-à-vis du remodelage et de la redistribution d'IHM afin de déterminer l'intérêt et l'acceptabilité des adaptations. En particulier, il s'agit de comprendre dans quelle mesure les adaptations doivent être placées sous le contrôle des utilisateurs. Les observations

permettent d'alimenter les réflexions sur les méthodes et outils pour l'ingénierie de systèmes interactifs plastiques. Pour cela, nous relatons deux évaluations utilisateurs en laboratoire et une évaluation utilisateurs sur le terrain.

Après un bref état de l'art sur la plasticité et les méthodes et difficultés d'évaluation, cet article présente les démonstrateurs développés et détaille les principaux résultats des évaluations. L'article est structuré en quatre sections. La section 2 est consacrée aux concepts clé de la plasticité : le contexte d'usage, les leviers, le contrôle utilisateur et la valeur pour l'utilisateur. La section 3 traite de l'évaluation. Un ensemble de méthodes y sont présentées, ainsi que les difficultés liées à l'évaluation de systèmes adaptatifs. Les sections 4 et 5 décrivent respectivement les démonstrateurs implémentés ainsi que les évaluations menées (protocoles et résultats). Des recommandations ergonomiques sont, en particulier, formulées à partir des résultats d'évaluation. Enfin, des perspectives sont énoncées.

2. PLASTICITE: CONCEPTS CLE

Cette section présente les concepts clé de la plasticité au fil des évolutions de la définition. Aujourd'hui, la plasticité dénote la capacité d'adaptation d'une IHM à son contexte d'usage dans le respect de la valeur attendue par l'utilisateur [Dâassi 07]. Nous identifions trois mots clé : adaptation, contexte d'usage et valeur. Ils sont ici examinés avec pour l'adaptation deux centres d'intérêt : les moyens de l'adaptation (remodelage, redistribution, multimodalité) ainsi que le degré de contrôle du processus d'adaptation par l'utilisateur.

2.1 CONTEXTE D'USAGE

En plasticité, le contexte d'usage fait aujourd'hui référence au triplet <utilisateur, plate-forme, environnement>. L'utilisateur dénote l'utilisateur du système interactif ; la plate-forme représente l'ensemble des ressources d'interaction disponibles et l'environnement rassemble les caractéristiques physiques et sociales du lieu hébergeant l'interaction. Historiquement, l'effort a été porté sur la plate-forme [Grolaux and al. 2005]. L'environnement a été peu considéré. Il est en général réduit à la localisation de l'utilisateur. Au départ [Thevenin and Coutaz 1999], l'utilisateur n'était pas considéré. Il l'a été plus tardivement à travers la notion de modèle utilisateur.

En 2006, [Tarpin-Bernard 2006] souligne l'importance d'étendre la définition du contexte à la prise en compte de l'activité de l'utilisateur. L'argument donné par l'auteur est le suivant (page 35) : "un même utilisateur utilisant une même plate-forme d'interaction dans un environnement identique peut avoir des besoins très différents en fonction de l'activité qu'il mène à l'instant t". En fait, le besoin de prendre en compte l'activité de l'utilisateur avait été identifié auparavant mais [Tarpin-Bernard 2006] trouve impropre sa dilution dans la dimension utilisateur.

En 2007, la variabilité de l'objectif utilisateur a été intégrée à la réflexion. Jusqu'ici, l'objectif était supposé fixe, connu à la conception. [Gabillon and al. 2007] imagine des objectifs utilisateur émergents. Le cas d'étude est celui de Yoann, malade qui souhaiterait voir un médecin. Il lance son assistant personnel *Compose* et saisit son objectif "je voudrais voir un médecin" (Figure 1a). *Compose* détecte que Yoann est loin de chez lui. Il lui demande alors s'il souhaite consulter un médecin sur place ou rentrer chez lui (Figure 1b). Yoann préfère voir un médecin sur place. Compte tenu de l'heure, *Compose* lui propose quatre solutions (Figure 1c) parmi lesquelles il choisit le médecin de garde (le docteur Mabuse). *Compose* génère une IHM qui répond à l'objectif de

Yoann (Figure 1d). Elle fournit les éléments d'information et d'interaction nécessaires à l'atteinte de l'objectif exprimé : l'adresse du docteur Mabuse, son numéro de téléphone pré-composé, des boutons permettant d'appeler le médecin puis raccrocher, l'itinéraire pour se rendre au cabinet médical ainsi que les coordonnées de la pharmacie de garde la plus proche.

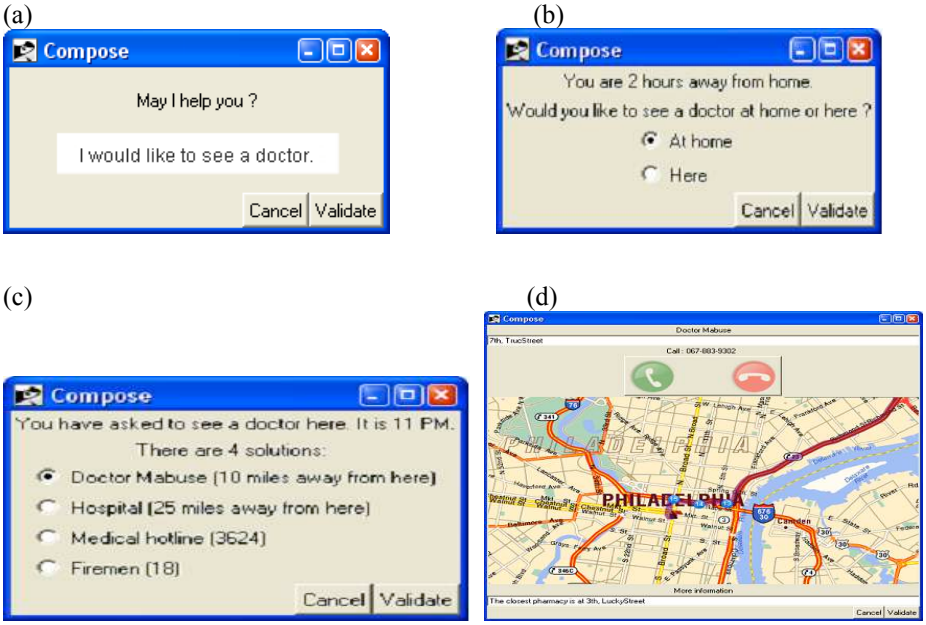


Figure 1: Compose, un système à objectif utilisateur émergent [Gabillon and al. 2007].

En 2008, une dimension supplémentaire est donnée au sujet par l'introduction d'objectifs utilisateur latents, c'est-à-dire non explicitement exprimés par l'utilisateur. EMMA (Embedded Manager for Mobile Adaptation) [Ganneau and al. 2008] s'inscrit dans cette lignée. EMMA est un système sensible au contexte sur téléphone mobile (plate-forme Windows Mobile CE, SPV C100). EMMA maintient un modèle utilisateur bayésien à partir de données contextuelles recueillies en mobilité : la localisation géographique de l'utilisateur, le jour de la semaine, l'heure, l'activité de l'utilisateur sur le téléphone (gestion du profil silencieux/sonore, application utilisée, etc.). L'hypothèse est que l'activité de l'utilisateur est fonction du contexte. En conséquence, lorsqu'un contexte est reconnu et nommé par l'utilisateur comme étant celui correspondant par exemple à la maison, au bureau ou encore aux vacances, EMMA propose à l'utilisateur les activités qu'il effectue en général dans ce contexte. En figure 2a, EMMA reconnaît un contexte clé. Dans ce contexte clé, qui a été nommé "Maison" par l'utilisateur, EMMA propose d'adapter le téléphone : réorganiser les fonctionnalités (Figure 2d) pour rendre saillantes les applications les plus utilisées dans ce contexte ; commuter le profil du téléphone tel que le mode vibreur, silencieux ou encore normal (Figure 2c) ; changer les paramètres de personnalisation comme la couleur du fond d'écran. L'utilisateur peut accepter ou refuser les propositions. Les changements de contexte et les adaptations à l'initiative du système peuvent être placés sous le contrôle de l'utilisateur via des boîtes

de dialogue (Figure 3). L'utilisateur peut choisir entre trois niveaux de contrôle : contrôlé, négocié et observable.

- *Contrôlé* : le système laisse la liberté à l'utilisateur de contrôler certains aspects de l'adaptation. Il s'agit notamment de la gestion des contextes clé (ajout, modification, suppression) ;
- *Négocié* : l'utilisateur est associé aux décisions. Le système propose et l'utilisateur dispose ;
- *Observable* : l'adaptation est perceptible à l'utilisateur mais il ne peut pas la contrôler. Seul un retour d'information lui est donné.

Figure 2: EMMA, un système sensible au contexte à objectif utilisateur latent [Ganneau and al. 2008].

Figure 3: EMMA, illustration des options disponibles pour le contrôle utilisateur [Ganneau and al. 2008].

2.2 LEVIERS DE LA PLASTICITE

Les deux leviers de la plasticité sont le remodelage et la redistribution [Calvary and Coutaz 2002].

Le *remodelage* est une adaptation locale à la plate-forme de l'utilisateur. Elle consiste en un remaniement de l'IHM sans en changer la distribution sur les ressources d'interaction (écran, clavier, souris, etc.). Ce remaniement peut porter sur la présentation (un menu déroulant peut se transformer en un jeu de boutons radio) ou le dialogue (des tâches utilisateur peuvent être ajoutées, modifiées ou supprimées). *FlexClock* [Grolaux and al. 2002] est un exemple typique de remodelage : selon la taille de la fenêtre, la présentation de l'heure change passant d'un texte à un cadran. Un calendrier est éventuellement ajouté selon la place disponible.

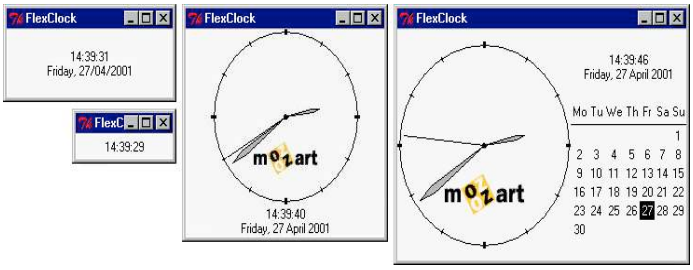
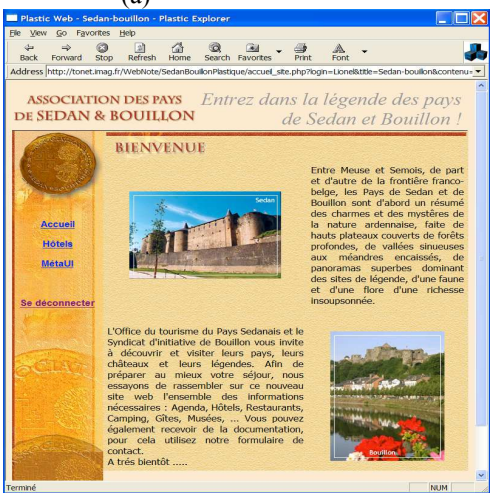


Figure 4: FlexClock, un exemple de remodelage [Grolaux and al. 2002].

Au contraire du remodelage, la *redistribution* change l'état de distribution de l'IHM sur l'ensemble des plates-formes en présence. La redistribution s'opère par migration totale ou partielle du système interactif. Sedan-Bouillon en est un exemple. C'est un site Web plastique pour la promotion des pays de Sedan et Bouillon. La version plastique est une version élaborée à partir d'une version simplifiée du site public (http://www.bouillon-sedan.com/). Elle est structurée en trois zones respectivement dédiées au titre, à la navigation et au contenu (Figure 5a). Dans sa version plastique, Sedan-Bouillon requiert l'identification de l'utilisateur (Lionel) (Figure 5b). Si Lionel se connecte à ce même site à partir d'un PDA, alors cette double connexion est détectée et une proposition de redistribution lui est faite : une IHM de contrôle utilisateur apparaît, expliquant à Lionel qu'il peut répartir le site, à son gré, entre le PC et le PDA (Figure 5c). La redistribution se fait au grain des espaces de dialogue. Lionel choisit, par un jeu de cases à cocher, sur quelle(s) plate(s)-forme(s) il souhaite placer ces trois espaces. La redistribution s'opère alors, composant l'IHM à façon.

(a)



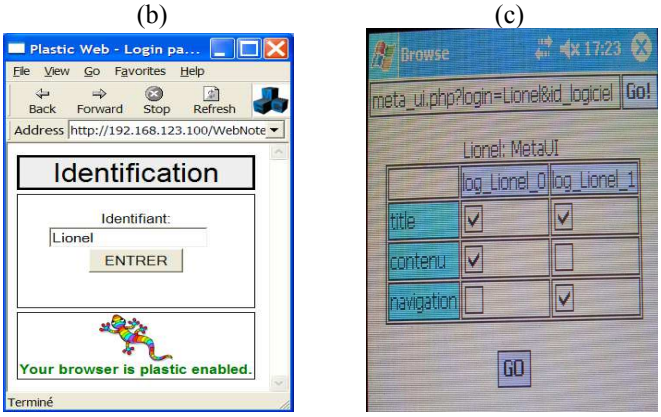


Figure 5: Sedan-Bouillon, un site web plastique (L. Balme et A. Demeure).

La redistribution peut requérir un remodelage. Le remodelage peut être *intramodal*, c'est-à-dire préserver la modalité courante. Il peut être *intermodal*, c'est-à-dire changer de modalité. Il peut être aussi *multimodal*, c'est-à-dire combiner plusieurs modalités. Même si l'adaptation n'est pas au cœur des recherches en multimodalité, elle suscite de l'intérêt dans ce domaine [Bellik 2006] [Huebsch and Kadner 2007] [Hesselman and al. 2008]. Dans tous ces travaux, la multimodalité est utilisée comme moyen d'adaptation. Pour exemple, le scénario suivant a été présenté dans [Hesselman and al. 2008] : Consuelo séjourne dans un hôtel et aimerait regarder un film avant de se rendre en ville pour dîner. Il est abonné au service de vidéo à la demande proposé par son opérateur mobile. Il utilise l'écran tactile disponible dans sa chambre d'hôtel pour sélectionner le film "Casino Royale". Comme une télévision haute définition est également disponible dans sa chambre et qu'il a à portée de main son portable, il reçoit le service en trois parties synchronisées : la partie vidéo du film sur la télévision haute définition, la partie audio sur le système hi-fi de la chambre et l'IHM de contrôle du service sur le mobile (Figure 6a). Lorsque l'heure de dîner approche, Consuelo quitte la chambre et se rend au parking pour prendre sa voiture. Il aimerait bien continuer à regarder son film. Aussi, il utilise les contrôles sur son téléphone mobile pour transférer la partie vidéo du film sur le téléphone (Figure 6b). Une fois installé dans la voiture, son téléphone mobile lui indique que son microphone, haut-parleurs et écran sont désactivés pour des raisons de sécurité. A partir de ce moment, seule la partie audio du film est disponible, et ce, grâce au système audio de la voiture. Comme il n'est pas très intéressant de regarder un film sans image, Consuelo décide de le suspendre à l'aide de commandes vocales. Pour ne pas perdre le fil à son retour, Consuelo annote le film avec quelques mots, vocalement, via le système audio de la voiture (Figure 6c). Enfin, avant de prendre la route, il utilise les commandes présentes sur le volant pour choisir une radio sur Internet qu'il écoutera tout le long du trajet.

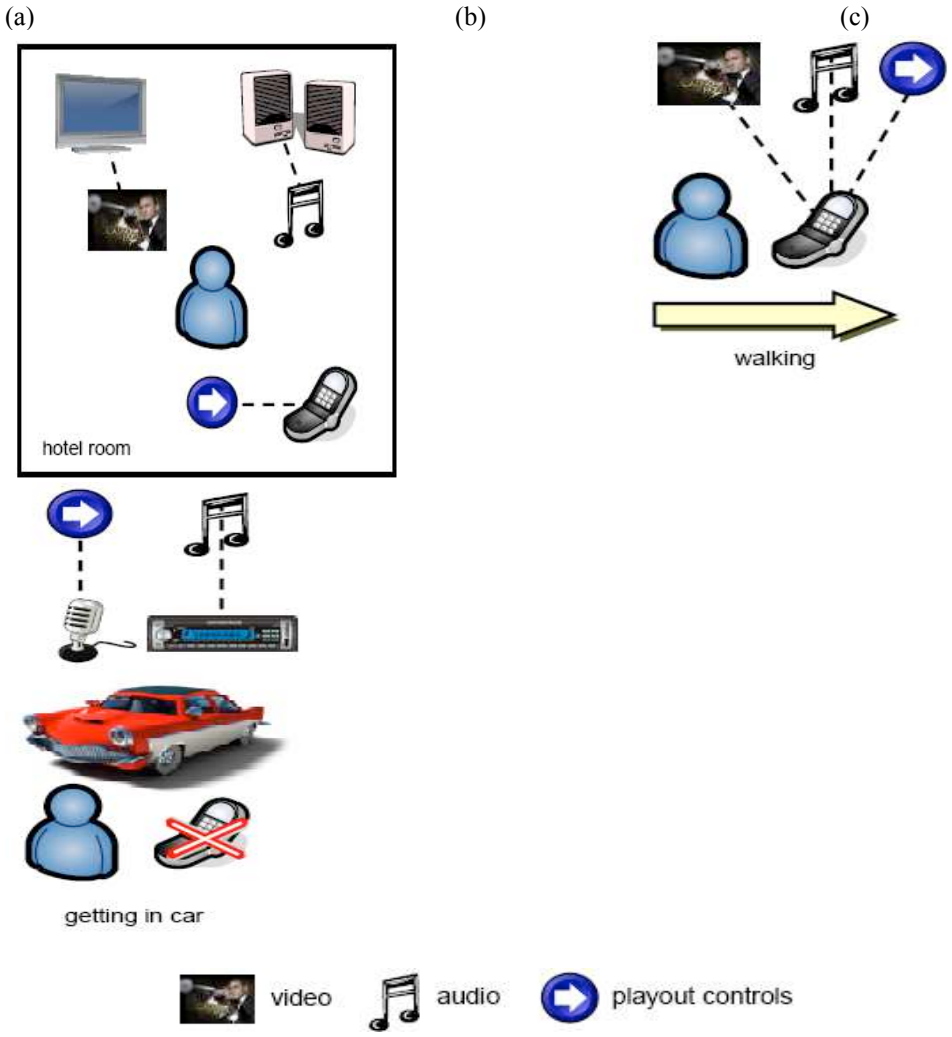


Figure 6: Scenario illustratif de l'adaptation par la multimodalité [Hesselman and al. 2008].

Explorer la multimodalité en plasticité est une perspective d'autant plus intéressante que les modalités se sont multipliées (vocale, tactile, gestuelle 2D, 3D, etc.), les technologies et infrastructures se sont développées et des résultats d'évaluation utilisateur sont disponibles. Par exemple, dans [Holly and al. 2002], les auteurs ont exploré différentes combinaisons de modalités. Le but est d'évaluer l'impact de telles combinaisons sur des sujets dotés d'une vue correcte. Les résultats montrent que :

- Les feedbacks uni-modaux audio, visuel et tactile sont équivalents en termes de charge mentale ;
- Les combinaisons de modalités sont intéressantes pour les retours d'information. Combiner l'audio, le visuel et le tactile réduit la charge mentale ;

- Un retour tactile est efficace seul ou en combinaison avec du visuel.

L'environnement, peu exploré en plasticité, est bien intégré en multimodalité. Dans [Lemmelä and al. 2008], les auteurs développent deux prototypes dans le but d'évaluer différentes modalités dans deux types d'environnement : la rue en situation de marche et la voiture en situation de conduite. Les résultats de ces études démontrent que le contexte d'usage a une influence sur les préférences des utilisateurs en terme de modalité d'interaction. En voiture, par exemple, les utilisateurs préfèrent l'interaction vocale et seraient prêts à utiliser un système entièrement vocalisé. Par contre, en situation de marche, ils ne trouvent pas cette modalité d'interaction utile. Ils estiment plus appropriée l'interaction gestuelle (2D, et surtout, 3D).

2.3 CONTROLE UTILISATEUR

Le contrôle utilisateur caractérise le degré de liberté accordé à l'utilisateur dans le processus d'adaptation. Le processus d'adaptation est un processus de type Perception-Action. A gros grain, il s'agit d'un processus en deux étapes qui consiste successivement à percevoir le contexte (Perception) puis adapter l'IHM en fonction du contexte observé (Action).

Dans la taxonomie de [Dieterich and al. 1994] pour la classification des systèmes interactifs dotés de capacités d'adaptation à l'exécution, les auteurs proposent un découpage du processus d'adaptation en quatre étapes :

- *L'initiative* : elle représente la décision d'un acteur (système ou utilisateur) de suggérer une adaptation ;
- *La proposition* : elle correspond aux différentes possibilités d'adaptation ;
- *La décision* : elle reflète le choix d'une adaptation parmi celles proposées ;
- *L'exécution* : elle consiste en la mise en œuvre effective de l'adaptation choisie.

Ces étapes, augmentées par l'étape de perception du contexte, couvrent l'ensemble du processus d'adaptation (cf. section 2.2). La question du contrôle utilisateur se pose à chaque étape. L'utilisateur peut intervenir aussi bien au niveau de la perception pour fixer les aspects du contexte à observer qu'au niveau de l'action pour orienter les décisions. De façon générale, on distingue quatre niveaux de contrôle :

- *Non observable* : l'adaptation s'opère à l'insu de l'utilisateur ;
- *Observable* : l'adaptation est perceptible par l'utilisateur sans qu'il dispose par contre de moyen de contrôle ;
- *Négocié* : l'utilisateur est associé aux décisions ;
- *Contrôlé* : le système laisse la liberté à l'utilisateur de contrôler le processus d'adaptation.

2.4 VALEUR

Une évolution majeure de la définition de la plasticité a consisté, en 2007, en l'introduction de la notion de *valeur*, en remplacement de l'utilisabilité. Avec la valeur, la plasticité devient la capacité d'une IHM à s'adapter à son contexte d'usage dans le respect de la *valeur* attendue par l'utilisateur [Dâassi 2007].

La notion de valeur a été introduite en IHM en 2004 par Gilbert Cockton. L'auteur estime que l'utilisabilité n'est pas suffisante [Cockton 2004a] [Cockton 2005]. L'exemple donné est celui de la gestion du chauffage [Cockton 2004a]. Les systèmes sont peut-être utilisables au sens où ils sont faciles à utiliser et à apprendre mais aucun ne répond à la véritable motivation de l'utilisateur, à savoir économiser. Ce n'est pas par plaisir que l'usager programme son système de chauffage. C'est pour réduire sa facture. Un système de qualité devrait donc lui rendre observable le montant de l'économie faite.

En 2004, la *valeur* n'était pas définie de manière formelle. Gilbert Cockton disait qu'elle se cernait en discutant avec les gens [Cockton 2004b]. Le premier cadre de développement pour la *conception centrée valeur* (*Value Centered Design* (VCD)) [Cockton 2005] est apparu en 2005. La limite de ces travaux provenait principalement de la confusion autour du mot *value*. Son usage a, par exemple, créé la confusion avec la *conception sensible à la valeur* (*Value Sensitive Design* (VSD)). En effet, en anglais, le mot "*value*" peut avoir plusieurs sens. Il peut être dénombrable ou pas. Dans VSD, "*value*" est non dénombrable. Il est souvent employé au pluriel comme dans la phrase "*human values*". "*Value*" prend un autre sens lorsqu'il est employé au singulier. C'est le sens qui était recherché dans VCD. Avec cette acception, les valeurs font partie de la valeur.

Pour lever toute ambigüité, Gilbert Cockton propose d'adopter, en 2006, le mot "*worth*" (valeur) à la place de "*value*" [Cockton 2006]. Ainsi, au lieu de parler de "*Value Centered Design* ", on parlerait de "*Worth Centered Design*" (WCD). Ce changement de terminologie est allé au-delà d'une substitution de lettre. Par la même occasion, l'auteur a donné une définition de "*worth*". Le manque de définition du mot "*value*" constituait une faiblesse des précédents travaux et contribua certainement à empirer la confusion autour du mot.

La définition de "*worth*" adoptée est issue du dictionnaire "Concise Oxford English Dictionary" [Thompson 1996]. Selon Gilbert Cockton, pour faire court, la valeur ("*worth*") est un motivateur. Pour les concepteurs, il stipule ceci : concevoir pour la valeur consiste à concevoir des choses qui motivent soit à l'achat, l'apprentissage, l'utilisation ou la recommandation d'un produit interactif et idéalement de la plupart d'entre eux ("*designing worth means designing things that will motivate people to buy, learn, use or recommend an interactive product, and ideally most of all these*").

La préservation de la valeur étant le différentiateur entre adaptation et plasticité, il est important d'identifier et d'évaluer les facteurs de valeur.

3. EVALUATION DE SYSTEMES ADAPTATIFS

Cette section traite de l'évaluation des systèmes adaptatifs. Elle s'articule en trois parties. La première relate les difficultés liées à l'évaluation des systèmes adaptatifs. La deuxième présente les conditions possibles d'évaluation (avec ou sans utilisateur, en laboratoire ou sur le terrain). Enfin, la troisième partie expose deux méthodes d'évaluation spécifiques aux systèmes adaptatifs.

3.1 DIFFICULTES LIEES A L'EVALUATION DE SYSTEMES ADAPTATIFS

L'évaluation des systèmes adaptatifs est loin d'être une question triviale [Carbonell and Simonin 2004]. Si l'évaluation demeure encore aujourd'hui un des verrous en adaptation, c'est parce qu'elle est difficile par nature. A ce jour, peu de résultats d'évaluation sont validés par la communauté scientifique. A cela, plusieurs raisons :

- Vaste espace problème de l'adaptation

L'espace problème de l'adaptation est vaste. Les nombreux espaces problème [Thevenin 2001] [Vanderdonckt and al. 2005] [Tarpin-Bernard 2006] [Calvary 2007] en témoignent : ils ajoutent de nouvelles dimensions au fil du temps. En conséquence, l'évaluation peut porter sur différents aspects (contrôle utilisateur, pertinence des adaptations proposées, prévisibilité, transparence de la stratégie d'adaptation, moments de l'adaptation, etc.).

- Manque de critères et de métriques

Si des recommandations ergonomiques ont été formulées, il n'existe pas de référentiel de critères et de métriques pour l'évaluation des systèmes adaptatifs. L'absence de référentiel est en partie dûe à la spécialisation des systèmes par métiers ou domaines d'application. Très souvent, dans le monde académique, les recherches sont focalisées, nuisant en conséquence à une appréhension globale. Ces considérations cloisonnées complexifient le travail de standardisation. Si les chances de parution d'un tel référentiel dans un avenir proche sont faibles, on peut envisager un référentiel global couvrant les éléments communs de l'évaluation, appuyé de référentiels spécifiques.

- Collecte de données objectives en temps réel

Pour obtenir des résultats objectifs, les évaluations sont à conduire en situations réelles. Cependant, les opinions et ressentis des utilisateurs sont recueillis généralement lors de séances d'entretiens (individuels ou en groupe), postérieures à la phase d'utilisation des systèmes. Ces opinions et ressentis sont, d'une part, subjectifs et, d'autre part, ne constituent pas des données fraîches. Avec le temps qui s'écoule entre la fin de l'utilisation et les entretiens, d'importants « détails » peuvent échapper aux utilisateurs. Pour pallier à ces problèmes, il est nécessaire de mettre en place des outils "*standalone*" ou à embarquer dans le système interactif pour la collecte de données objectives. Le recueil d'avis utilisateurs en temps réel ou rapproché peut se faire par des journaux de bord accessibles via le Web. Cette solution peut s'avérer lourde pour les utilisateurs [Höök 1997].

- Coût de l'évaluation

L'évaluation de systèmes adaptatifs est coûteuse non seulement en ressources mais aussi en temps.

Les adaptations sont motivées par des changements de contexte d'usage. La multiplicité des contextes d'usage et des changements de contexte d'usage peut nécessiter de nombreuses passations de mêmes utilisateurs. La combinatoire augmente encore en faisant varier les participants.

Le temps requis pour mesurer l'impact de l'adaptation peut être long. Il est indispensable de laisser plusieurs jours, semaines, voire mois à l'utilisateur pour découvrir les différentes facettes du système adaptatif en situations réelles. En outre, bien souvent, un temps d'apprentissage est nécessaire au système afin d'acquérir des connaissances plus précises sur l'utilisateur. Cette étape, préalable aux premières adaptations, mais qui se poursuit pendant toute la période d'utilisation, peut être plus ou moins longue selon l'activité de l'utilisateur.

3.2 CONDITIONS POSSIBLES

3.2.1 METHODES D'EVALUATION AVEC *VERSUS SANS* UTILISATEURS

Dans l'évaluation de systèmes adaptatifs, et d'une manière générale, les méthodes d'évaluation peuvent être classées en deux catégories selon qu'elles impliquent ou non l'utilisateur pendant l'évaluation. Nous en citons ici les principales mais sans les développer :

- les tests utilisateurs [Nielsen 1993] ;
- les entretiens [Blanchet and al. 2000] ;
- les questionnaires [Blanchet and al. 2000] ;
- le paradigme du Magicien d'Oz [Amalberti andValot 1992] [Preece and al. 1994] ;
- les *focus group* [Greenbaum 1998] [Nielsen 1993].

Les méthodes n'impliquant pas l'utilisateur dans l'évaluation ont recours à un ou plusieurs experts dans des domaines tels que l'ergonomie, la sociologie ou la psychologie cognitive. Parmi ces méthodes, on trouve :

- l'évaluation heuristique [Nielsen and Molich 1990] ;
- l'évaluation experte [Gena 2006] ;
- l'analyse des tâches [Diaper 1990] ;
- l'inspection cognitive [Lewis and al. 1990] [Polson and al. 1992].

Pour évaluer nos démonstrateurs, des évaluations expertes ont tout d'abord été conduites en cours de conception. Puis, des évaluations avec des utilisateurs en laboratoire et sur le terrain ont été menées.

3.2.2 EN LABORATOIRE *VERSUS* SUR LE TERRAIN

Deux conditions d'évaluation sont possibles : l'évaluation en laboratoire et l'évaluation sur le terrain [Jambon and al. 2006]. Les termes anglo-saxons consacrés sont *field experimentation* [Goodman and al. 2004], *experimentation in the field* [Kellar and al. 2005] ou encore *experimentation in the wild*.

L'évaluation en laboratoire se décline généralement sous la forme de tests utilisateurs. Cette méthode peut donner des résultats fiables. Cependant, elle ne prend pas en compte tous les facteurs pouvant intervenir dans l'interaction homme-machine et ne permet pas de cerner les usages réels des dispositifs [Norman 1999] [Kjeldskov and Paay 2005].

L'évaluation sur le terrain présente l'avantage de fournir des données sur l'utilisation réelle des systèmes. Cependant, elle n'est pas envisageable pendant les premières phases de la conception car elle requiert un système fonctionnel du point de vue technique et pose les problèmes d'enregistrement des observations et de validité des analyses [Jambon et al. 2006].

Les outils utilisés pour l'évaluation des usages sur le terrain sont variés et dépendent fortement des contraintes d'utilisation du système (comme la mobilité). Sans être exhaustifs, nous pouvons citer les techniques d'auto-documentation utilisées par [Sohn et al. 2008] et [Hulkko et al. 2004]. Ces derniers ont recueilli l'activité des utilisateurs par l'intermédiaire de journaux de bord (diary) ou encore de photographies, de collages et d'échantillonnage d'expérience (ESM). Les utilisateurs devaient décrire leurs activités, leurs besoins et leurs attentes selon différents contextes d'usage, prendre des photographies illustrant leurs contextes ou encore répondre à des questionnaires en ligne servant à décrire leurs activités et leurs ressentis. En outre, l'analyse de logs (automatic logging) peut être aussi utilisée pour identifier les usages réels. Ces outils tracent

l'utilisation d'un dispositif en contexte réel (fréquence d'usage, temps d'utilisation, périodes d'utilisation ou de non utilisation, etc.) et ils sont généralement complétés par des questionnaires ou des entretiens avec des utilisateurs (explicitation de leurs usages).

3.3 METHODES POUR L'EVALUATION DES SYSTEMES ADAPTATIFS

Dans [Chin 2001], l'auteur passe en revue trente-deux publications contenant des évaluations empiriques pour divers systèmes adaptatifs. Il réalise une synthèse des méthodes utilisées pour les évaluer. Il note que les critères d'évaluation peuvent inclure des mesures tant quantitatives (par exemple, le temps de réalisation d'une tâche) que qualitatives (par exemple, le feedback de l'utilisateur sur l'utilité du système). Les données ainsi obtenues sont traitées à l'aide d'outils statistiques comme par exemple l'analyse de la variance (ANOVA). Ce type d'évaluation vise à donner des résultats généralisables car statistiquement valides et fidèles. Deux approches classiques sont utilisées pour l'évaluation empirique des systèmes adaptatifs [Tobar 2003] :

- L'évaluation *comparative* : elle consiste à évaluer, par des tests utilisateurs généralement réalisés en laboratoire, deux versions du système : l'une avec adaptation et l'autre sans. La valeur de l'adaptation est la différence d'utilité entre les deux versions [Höök 1997] [Höök 2000] [Lieberman and al. 2004].
- L'évaluation *structurée* ou *layered evaluation* : cette approche a été introduite par [Totterdell and Boyle 1990]. Elle consiste à évaluer séparément les éléments adaptatifs du système pour identifier l'intérêt et la pertinence des adaptations proposées [Brusilovsky and al. 2001] [Tobar 2003] [Brusilovsky and al. 2004]. La valeur de l'adaptation est évaluée à deux niveaux [Totterdell and Rautenbach 1990] [Benyon and Murray 1993] [Brusilovsky 1996] : au niveau du modèle utilisateur (représentation correcte des caractéristiques, buts, préférences, etc. de l'utilisateur [Brusilovsky 1996]) et au niveau des choix d'adaptation (validité et pertinence des adaptations proposées). L'évaluation structurée n'a pas pour objectif de remplacer les pratiques d'évaluation existantes : chaque adaptation peut précisément être évaluée avec ces techniques.

4. DEMONSTRATEURS

Cette section présente deux démonstrateurs de plasticité : PlasticSMS et MobilePlastic. Le troisième est EMMA : il a fait l'objet de la section 2.1. Les démonstrateurs illustrent les concepts clé de la plasticité : le remodelage, la redistribution et le contrôle utilisateur. La redistribution n'est pas assurée dans EMMA qui, par contre, existe en deux versions : avec ou sans IHM de contrôle. A l'inverse, PlasticSMS et MobilePlastic couvrent la redistribution et intègrent une IHM de contrôle.

4.1 PLASTICSMS

PlasticSMS est un prototype développé en C++. Il offre un service de messagerie plastique permettant de communiquer via SMS depuis un téléphone mobile ou un ordinateur. En situation de mobilité, le téléphone est utilisé de manière traditionnelle. Lorsque l'environnement s'augmente de plates-formes telles qu'un ordinateur, le service offre à l'utilisateur la possibilité de redistribuer l'IHM vers une seconde plate-forme lorsque la proximité de celle-ci est détectée (connexion Bluetooth) (Figures 7 et 8).

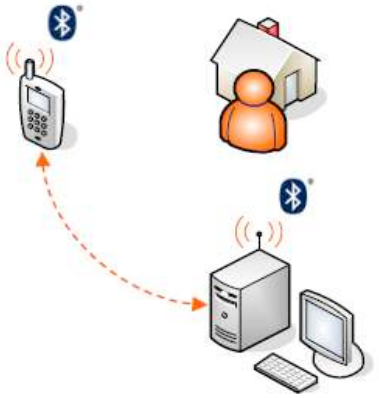


Figure 7: Redistribution de l'IHM de PlasticSMS.

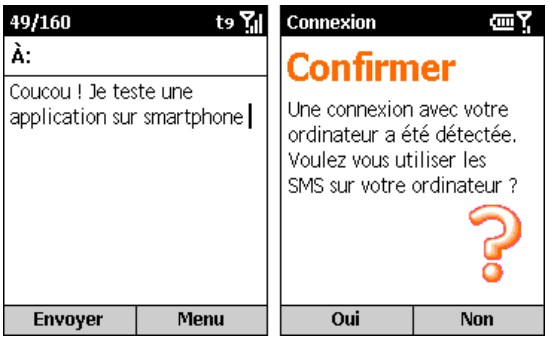


Figure 8: IHM de PlasticSMS sur téléphone.

L'IHM de saisie est alors reproduite sur l'ordinateur, le message pouvant être saisi indifféremment sur le mobile ou l'ordinateur. Les avantages supposés pour l'utilisateur sont un confort visuel accru par la taille de l'écran et une saisie facilitée par la disponibilité d'un clavier en entrée. De plus, l'IHM est dotée de capacité de remodelage. Lorsque l'utilisateur agrandit (respectivement réduit) la taille de la fenêtre, des fonctionnalités supplémentaires apparaissent ou disparaissent (Figure 9). Ainsi, l'utilisateur peut modifier la taille des caractères, envoyer le message ou encore accéder à ses contacts enregistrés dans le répertoire du téléphone.

Figure 9: IHM de PlasticSMS sur PC.

L'observabilité et le contrôle des adaptations sont accessibles par l'intermédiaire d'une IHM dite « de contrôle utilisateur » (Figure 10). L'utilisateur dispose d'options permettant d'accepter ou de refuser la redistribution, de demander une confirmation à chaque adaptation proposée par le système, d'adapter automatiquement sans être consulté ou encore de refuser toute adaptation (utilisation classique du SMS).

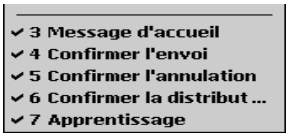


Figure 10: Un exemple d'options accessibles à travers l'IHM de contrôle.

4.2 MOBILEPLASTIC

MobilePlastic est une application développée en C#. Elle propose des fonctions complémentaires à PlasticSMS. Elle permet de gérer les contacts du téléphone mobile, de consulter l'historique des appels entrants et sortants, d'enregistrer ses SMS (fichiers textes) et de consulter le niveau de batterie du téléphone (Figure 11). En terme de plasticité, elle est dotée de possibilités de redistribution (les actions se réalisant sur l'ordinateur) mais pas de remodelage.

Figure 11: MobilePlastic sur PC.

Tout comme dans PlasticSMS, les utilisateurs peuvent contrôler l'adaptation. Des messages de confirmation apparaissent sur le téléphone mobile lorsque l'utilisateur réalise des actions critiques telles que l'ajout, la modification, la suppression d'un contact, l'envoi ou l'annulation d'un SMS. En outre, la liste des contacts du répertoire est réorganisée en fonction de la fréquence des appels.

5. EVALUATION DES DEMONSTRATEURS ET RESULTATS

Cette section décrit le déroulement des évaluations et présente les principaux résultats. Elle s'organise en deux parties. La première est dédiée aux évaluations effectuées en laboratoire ; la deuxième relate les évaluations sur le terrain.

5.1 PLASTICSMS ET MOBILEPLASTIC : EVALUATIONS EN LABORATOIRE

Les évaluations de PlasticSMS et MobilePlastic se sont déroulées en laboratoire. Elles ont donné lieu à un focus group et des tests utilisateurs.

5.1.1 OBJECTIFS

Les principaux objectifs du focus group sont les suivants :

- cerner les attentes des usagers (identifier les principaux besoins pouvant être satisfaits par ce type de services, recueillir des idées nouvelles en termes de fonctionnalités, des propositions de services, etc.),
- appréhender l'intérêt perçu de la plasticité (le ressenti des participants face à un système qui capitalise leur préférences, modifie l'IHM en fonction de leurs actions, etc.),
- appréhender l'intérêt de la négociation avec le système (contrôle utilisateur nécessaire ou non, contraignant ou non, etc.),
- appréhender le ressenti des utilisateurs vis-à-vis du caractère potentiellement intrusif de ce type de système.

Les principaux objectifs des tests utilisateurs sont les suivants :

- évaluer la compréhension du concept de plasticité,
- déterminer l'intérêt d'IHM plastiques (avantages et inconvénients).

Nous indiquons ci-dessous le profil des différents participants (Tableau 1). La section suivante précise les protocoles.

	Focus Group	**Tests utilisateurs**
Effectif		
Nombre de participants	5 participants	20 participants
Caractéristiques		
Sexe	2 hommes 3 femmes	11 hommes 9 femmes
Age	Moyenne d'âge : 39 ans	Moyenne d'âge : 36 ans

	Age minimum : 20 ans Age maximum : 57 ans	Age minimum : 16 ans Age maximum : 71 ans
Expérience mobile	Aucun participant ne devait avoir ou avoir eu un mobile de type SPV. Ainsi, tous avaient le même niveau d'expérience.	
Echantillon des utilisateurs	Deux catégories d'âge ont été ciblées : un groupe de 16 à 25 ans et un groupe de 45 ans et plus.	
Usage du mobile	Un usage régulier du mobile était demandé.	
Usage de l'ordinateur	Un usage régulier de l'ordinateur était demandé.	

Tableau 1: Profil des participants du focus group et des tests utilisateurs.

5.1.2 PROTOCOLES

Le focus group a duré quatre-vingt dix minutes. La séance a été enregistrée en audio et vidéo et a donc impliqué, en plus des cinq participants, un modérateur et un observateur en régie. La séance s'est déroulée de la façon suivante :

1. Accueil des participants
2. Questionnaire de caractérisation des participants (profils, habitudes, usages des mobiles et de l'ordinateur, etc.)
3. Présentation retro-projetée de la plasticité (illustration de différents exemples de prototypes)
4. Thèmes abordés lors de la discussion :
   - avantages et inconvénients de l'adaptativité,
   - intérêt de l'adaptation aux plates-formes, à l'utilisateur, à l'environnement, pour soi et pour les autres,
   - avantages et inconvénients des démonstrateurs.
5. Clôture et synthèse de la séance.

La passation des tests était individuelle et durait environ deux heures. Les tests utilisateurs ont regroupé vingt utilisateurs. Les séances se sont déroulées de la façon suivante :

1. Accueil des participants
2. Questionnaire de caractérisation des participants (profils, habitudes, usages des mobiles et de l'ordinateur, etc.)
3. Illustration d'exemples de différentes formes d'adaptation
4. Premier entretien permettant d'appréhender l'intérêt perçu des IHM plastiques
5. Tâches de familiarisation avec les démonstrateurs
6. Scénarios d'utilisation sur PlasticSMS :
   - envoi d'un SMS à partir du téléphone avec modification de la taille de la police,
   - envoi d'un SMS à partir de l'ordinateur.
7. Second entretien permettant de recueillir l'avis des utilisateurs sur l'utilisation de PlasticSMS
8. Scénarios d'utilisation sur MobilePlastic :
   - consultation de l'état de la batterie,
   - enregistrement et modification d'un contact,
   - consultation d'un SMS enregistré,
   - suppression d'un contact,
   - émission d'un appel depuis le téléphone mobile.

Le contrôle des adaptations et l'apprentissage des préférences des utilisateurs ont été simulés. Pour cela, tous les participants ont réalisé les scénarios dans le même ordre :

- Pour les premiers scénarios, toutes les adaptations proposées par le système étaient négociées avec l'utilisateur.
- Pour les derniers scénarios, les adaptations étaient réalisées sans négociation auprès de l'utilisateur. De plus, l'évolution des scénarios illustrait une réorganisation des contacts après avoir envoyé des messages et appelé les mêmes contacts.

5.1.3 LIMITES

Nous identifions quelques limites aux protocoles :

- Concernant le focus group, les retours utilisateurs sont basés uniquement sur le déclaratif des participants. Par conséquent, ces résultats sont à utiliser avec précaution, les participants ayant tendance à surestimer ou sous-estimer certains usages.
- La mise en situation d'un panel réduit d'utilisateurs (vingt pour les tests utilisateurs) peut paraître restreinte. Toutefois, cette méthode permet d'obtenir une tendance qualitative en mettant en évidence la perception des utilisateurs et les principales difficultés d'utilisation.
- En laboratoire, les utilisateurs sont mis en situations d'utilisation les plus réalistes possible. Néanmoins il est difficile d'approcher un large éventail de contextes d'usage (comme les usages en mobilité par exemple).
- Les vingt utilisateurs du test ont tous passé les scénarios dans le même ordre (étant donné notre volonté d'illustrer l'apprentissage de leurs actions par le système). Un risque sur l'apprentissage système doit être souligné.

5.1.4 PRINCIPAUX RESULTATS

Tout d'abord, nous présentons la perception des utilisateurs vis-à-vis des concepts clé de la plasticité. Puis, nous abordons les performances observées et les difficultés rencontrées par les participants lors des tests utilisateurs.

Perception des utilisateurs

Pour évaluer le ressenti de tous les participants au focus group et aux tests utilisateurs, un premier entretien a été réalisé sur la base d'une présentation rétro-projetée illustrant différents exemples de services et d'applications adaptatives :

- l'adaptation à la plate-forme a été perçue comme étant la plus utile (vingt-sept occurrences) ;
- l'avantage est une optimisation des tâches et une utilisation plus conviviale (vingt-huit) lorsque le dispositif sur lequel l'IHM est déportée possède des caractéristiques plus confortables (écran plus grand ou encore présence d'un clavier) ;
- les principaux inconvénients cités par les participants sont la confidentialité des données (huit) et la crainte de perdre le contrôle de ses actions (douze) ;
- neuf participants ont déclaré qu'ils utiliseraient ce type de services illustrés par les démonstrateurs dans la sphère professionnelle et familiale ;
- les idées de fonctionnalités adaptatives les plus fréquemment citées seraient de pouvoir disposer d'un agenda adaptatif (quatorze occurrences) sur mobile (rappel du contrôle technique de la voiture, de la liste des courses, etc.), une

localisation GPS donnant des informations sur la localisation de son interlocuteur (huit) ou encore des informations circonstanciées en voiture ou en faisant ses courses (quinze).

Performance des participants (tests utilisateurs)

Tous les résultats d'utilisation des deux démonstrateurs ne sont pas exposés dans cet article. Nous avons extrait les performances de certains scénarios et essentiellement ceux relatifs aux leviers de l'adaptation (redistribution, remodelage) ainsi qu'au contrôle utilisateur.

Tâche	Réussite	Temps moyen		Nombre d'assistances
		Jeunes	Plus de 45 ans	
Elargir la fenêtre de l'application	13/20	13,14	22,5	6
Saisir le SMS	20/20	21,1	21,9	1
Envoyer le SMS	20/20	36,9	63,5	13
Consulter le niveau de batterie du téléphone	20/20	3,1	14,8	11
Ajouter un contact	20/20	3,1	44,6	15
Modifier un contact	20/20	34,3	41,2	15

Tableau 2: Performance des utilisateurs (PlasticSMS et Mobile Plastic)

Les utilisateurs ont majoritairement réussi les différents scénarios. En revanche, pour certaines tâches, l'intervention de l'expérimentateur a été nécessaire. De plus, nous avons essentiellement constaté des différences de temps entre le groupe des jeunes et celui des plus de 45 ans. A titre d'exemple, la consultation de l'état de la batterie sur l'ordinateur a posé plus de problèmes aux seniors. Le temps moyen d'exécution des utilisateurs de ce groupe (M = 14,8) est considérablement plus élevé que celui des juniors (M = 3,1). Les informations ne semblaient pas assez visibles selon les commentaires des testeurs.

Remodelage de l'IHM

Pour le premier scénario d'utilisation de PlasticSMS, nous avons demandé aux testeurs d'envoyer un message texte à partir de l'ordinateur. L'écran était identique à celui du mobile. L'utilisateur devait ensuite élargir la fenêtre pour voir apparaître les options du répertoire et de l'envoi.

Figure 12: Illustration du remodelage de l'IHM.

Nous avons observé des échecs uniquement au niveau du remodelage. En effet, la plupart des participants ont bien vu la possibilité d'agrandir la fenêtre, mais n'ont pas perçu le lien avec l'existence de fonctions supplémentaires. Treize utilisateurs ont jugé l'indication de l'agrandissement de la fenêtre incompréhensible. Notons que parmi les participants qui ont réussi la tâche (treize utilisateurs), sept sont ceux qui ont adopté la stratégie d'agrandir la fenêtre en utilisant l'icône d'agrandissement de Windows et six (sur treize) ont utilisé la flèche.

Contrôle utilisateur

Pour le premier scénario, la confirmation des actions de redistribution de l'IHM sur l'ordinateur et d'envoi du message étaient accessibles sur le téléphone. Douze utilisateurs les ont jugées inutiles et préfèraient qu'elles soient visibles uniquement sur l'ordinateur. Les avis sont identiques pour les autres scénarios ayant nécessité de confirmer les actions sur le mobile.

Perception des éléments adaptatifs

Le dernier scénario d'utilisation consistait à émettre un appel. Les utilisateurs pouvaient alors constater la réorganisation de la liste des contacts en fonction de la fréquence. Notre objectif était de recueillir leur avis sur cet élément adaptatif. Dix-huit participants l'ont jugé utile car cela permet d'éviter des manipulations supplémentaires (gain de temps).

5.1.5 DISCUSSION

Les évaluations en laboratoire nous ont permis de collecter un certain nombre d'observations objectives et déclaratives sur l'utilisation d'IHM plastiques. Néanmoins, les limites des méthodologies que nous avons détaillées plus haut nous conduisent à établir un certain nombre de constats :

- En laboratoire, il est difficile de simuler les adaptations à l'utilisateur et notamment à leurs préférences et à leurs usages. En effet, l'utilisation des dispositifs sur une très courte période (deux heures) rend difficile une réelle prise en compte de ces caractéristiques.
- Pour mettre en évidence le remodelage des IHM du téléphone sur l'ordinateur, nous avons opté pour une IHM similaire au mobile et pouvant s'agrandir pour proposer des fonctionnalités supplémentaires et des modes d'interaction différents. Globalement, tous les participants n'ont pas perçu la pertinence de conserver dans un premier temps un affichage cohérent avec l'IHM du mobile. Il semble préférable d'afficher directement l'IHM dans sa globalité sur l'ordinateur.
- Evaluer la redistribution des IHM sur le terrain auprès d'utilisateurs nécessite de développer des prototypes stables pouvant être installés et utilisés facilement par les utilisateurs chez eux ou en mobilité.
- Nos évaluations et le retour des utilisateurs ont été relativement biaisés par les différences des IHM en termes de design. En effet, les participants ont exprimé des améliorations en termes d'organisation des informations entre les deux démonstrateurs. Il apparaît donc nécessaire de prendre en compte dès la conception les objectifs de l'évaluation et la diversité des scénarios d'usage qui pourront être soumis aux testeurs.
- En laboratoire, il n'est pas possible d'identifier les différences d'usage pour une même application selon les plates-formes utilisées. Ainsi, il serait intéressant d'observer si la fréquence d'envoi des messages varie selon l'utilisation du téléphone mobile ou de l'ordinateur. De la même manière, les usages sont-ils reproduits à l'identique selon le dispositif utilisé ?
- Concernant le contrôle utilisateur (IHM de contrôle), nous avons observé un consensus sur la nécessité d'intégrer cette IHM sur l'ordinateur plutôt que sur le mobile. Nous avons été relativement surpris par ce résultat car il nous semblait important de notifier certaines actions sur le mobile pour rendre observable la répercussion des actions sur ce dispositif.

5.2 EMMA : EVALUATION SUR LE TERRAIN

L'évaluation d'EMMA s'est produite sur le terrain. En effet, EMMA impose des conditions réelles de mobilité pour collecter des données d'usage sur une période de plusieurs semaines. En outre, des évaluations sur le terrain sont nécessaires pour l'apprentissage des contextes clés de l'utilisateur.

5.2.1 PROTOCOLE

Dans le but de recueillir les usages et les impressions sur les deux versions d'EMMA, l'évaluation a été menée auprès de deux groupes différents de participants. Dans la première version, les utilisateurs avaient la possibilité, grâce à l'IHM de contrôle, de paramétrer les adaptations. Ils pouvaient ajouter, modifier et supprimer des contextes d'usage clés à tout moment. Dans la deuxième version, ils étaient tributaires des propositions du système. Le contrôle qu'ils exerçaient était limité à la négociation au fil des adaptations (ils ne disposaient pas de l'IHM de contrôle). Les tableaux ci-après récapitulent le protocole d'évaluation et le profil des participants.

Étapes	Objectifs
Lancement de l'étude	Remise des mobiles équipés d'EMMA aux participants Explication du déroulement de l'étude Remise d'un support expliquant le fonctionnement de l'étude Passation d'un questionnaire de caractérisation
Suivi	Envoi régulier de données d'usage collectées et stockées par e-mail par les participants
Conclusion de l'étude	Récupération des mobiles auprès des participants Entretien individuel avec chaque participant

Tableau 3: Etapes du protocole d'évaluation

Effectifs	
Nombre de participants initial	12 participants (divisés en 2 groupes)
Nombre d'abandons / échecs	1 abandon / 1 échec
Nombre de participants retenus	10 participants
Caractéristiques	
Sexe	5 hommes 5 femmes
Age	Moyenne d'âge : 19 ans Age minimum : 18 ans Age maximum : 23 ans
Equipement	Tous les participants utilisent un téléphone mobile. Ils ont en moyenne eu 2,9 mobiles avant leur mobile actuel.
Expérience mobile	Tous les participants utilisent un mobile (depuis au moins un an à plus de trois ans)
Usage du mobile	
Nombre de contacts enregistrés	8 participant ont entre 50 à plus de 100 contacts 2 participants ont entre 20 et 50 contacts
Usage des SMS	5 participants envoient entre 30 et 200 SMS par mois 5 participants envoient entre 15 et 30 SMS par mois
Usage des MMS	3 participants n'ont jamais envoyé de MMS 7 participants envoient de 2 à 10 MMS par mois
Usage des emails	3 participants envoient occasionnellement des emails
Usage de l'Internet mobile	4 participants utilisent occasionnellement cette fonction

Tableau 4: Profil des participants

5.2.2 LIMITES

Nous identifions quelques limites dans l'évaluation :

- Pour cette évaluation, une démarche qualitative a été mise en place avec un panel réduit d'utilisateurs. En effet, mettre en place une démarche quantitative sans tout d'abord tester l'application avec un petit panel d'utilisateurs est risqué et coûteux. Ainsi, les résultats développés dans cet article s'appliquent uniquement dans le cadre de cette étude, de l'application testée et de la population sollicitée (à savoir des jeunes plutôt technophiles).

- Les résultats relatifs à l'acceptabilité des adaptations proposées s'appliquent uniquement aux adaptations testées au cours de cette étude : réorganisation des fonctionnalités et automatisation de certaines tâches comme le changement de profil de sonnerie.
- Mener une étude sur le long terme est nécessaire pour évaluer des applications adaptatives. Les six semaines de test se sont avérées suffisantes pour le recueil de données, l'apprentissage, les propositions d'adaptation et les révisions des adaptations selon les usages des utilisateurs. Néanmoins, les participants ont eu tendance à oublier certaines informations comme à quel moment les premières adaptations ont été proposées, dans quels contextes étaient-ils ou encore comment ont-ils perçu les adaptations proposées ?
- Pour pallier aux difficultés de recrutement des utilisateurs ayant un mobile bien spécifique comme le SPV C100 et qui par ailleurs n'était plus en vente au moment de l'étude, nous avons prêté les téléphones aux participants. En outre, cela nous permettait de contrôler le niveau d'expérience des participants avec cette gamme de téléphone. Néanmoins, nous avons constaté que, par exemple, les utilisateurs ont très peu personnalisé le mobile (fond d'écran, etc.) alors qu'ils le font en règle générale. Ceci s'explique probablement par le fait qu'ils devaient restituer le matériel en fin d'expérience. Certains usages sont donc à prendre avec précaution car ils peuvent être dus à l'utilisation d'un mobile prêté.

5.2.3 PRINCIPAUX RESULTATS

Les résultats obtenus sont issus d'une part des fichiers de logs et d'autre part des verbatim des participants lors des entretiens individuels. Un programme d'exploitation des logs a été implémenté afin de faciliter le traitement des résultats. Tous les entretiens ont été retranscrits et une analyse du contenu a été conduite.

Fonctionnement de l'application

Les dix participants ont tous utilisé EMMA durant au moins six semaines. Nous avons ainsi collecté vingt cinq mille cinq cent quatre-vingt onze (25591) données tout au long de la durée de cette expérimentation terrain. Ce nombre permet de constater la stabilité fonctionnelle et opérationnelle de l'application mais aussi les usages effectifs de l'application par les testeurs.

Les phases d'apprentissage sont en moyenne au nombre de quatre-vingt trois par participant. Ce nombre nous indique en moyenne la fréquence a priori des sollicitations du système. Malheureusement, nous n'avons pas tracé en parallèle des phases d'apprentissage celles qui donnaient lieu à des propositions système (si le système détectait de nouveaux usages et donc des adaptations) ou celles pour lesquelles le système ne détectait aucun nouvel usage (donc aucune proposition). Or, ces données seraient pertinentes pour mesurer objectivement la fréquence des dialogues entre le système et l'utilisateur. Au cours des entretiens, tous les participants ont jugé les propositions système trop fréquentes mais aucune donnée ne nous permet de quantifier dans quelle mesure.

Contextes clés

Le nombre moyen de contextes clés proposés par le système et/ou ajoutés (pour les participants disposant de l'IHM de contrôle) était de 4,7.

Une participante n'a pas eu de proposition de contexte tout au long de l'étude. En effet, les algorithmes de "clustering" étaient conditionnés à l'usage du changement de profil de sonnerie pour identifier les contextes clés des utilisateurs. Or, cette participante a activé le mode vibreur ou silencieux en passant par le menu "Paramètres" et "Sons" mais jamais en utilisant les profils de sonnerie. Néanmoins le système lui a proposé des adaptations telles que la réorganisation des menus.

La répartition du nombre de contextes clés moyen selon la version d'EMMA (avec ou sans contrôle utilisateur) est la suivante :

	Nombre de contextes clés
Participants avec IHM de contrôle (N=5)	6,4
Participants sans IHM de contrôle (N=5)	3
Ensemble des participants (N=9)	**4,7**

Tableau 5 : Nombre moyen de contextes selon la version utilisée d'EMMA

Les participants qui disposaient des options permettant de paramétrer et de personnaliser les contextes clés les ont effectivement utilisées. Ils ont en effet ajouté eux-mêmes des contextes en plus de ceux proposés par le système. Pour ce groupe, on observe une répartition du nombre de contextes allant de quatre à dix contextes selon les participants. Quant aux groupes ayant EMMA sans IHM de contrôle, ce nombre varie de zéro à cinq contextes.

Nombre de contextes réellement utiles

En analysant les usages des participants, nous avons constaté que le nombre de contextes réellement utiles, à savoir pour lesquels on observait des usages, pouvait être différent du nombre total de contextes. Deux participants sur neuf ont des contextes vides. Six participants ont des contextes quasiment vides (c'est-à-dire représentant entre 1 et 2% de leurs usages). La répartition des contextes utiles (représentant 80% de leurs usages) est la suivante :

Nombre de contextes utiles	**Nombre de participants (N=9)**
2 contextes utiles	5
3 contextes utiles	2
1 contexte utile	2
Nombre moyen de contextes utiles	**2,11**

Tableau 6 : Nombre de contextes réellement utiles (avec des usages constatés)

On obtient ainsi une moyenne de contextes utiles de 2,11 par participant. La majorité des participants a eu des usages distincts dans deux voire trois contextes d'usage. Pour compléter ces données, nous avons calculé le pourcentage de temps d'activité (avec des usages) dans les différents contextes. Il apparaît bien un usage bipolaire voire tripolaire pour la grande majorité des participants (80% des usages sont réalisés dans pas plus de trois contextes).

Ajout et modification de contextes

Pour déterminer l'intérêt de l'IHM de contrôle et son usage réel par les participants, nous avons tracé les actions réalisées dans ce menu et nous avons recueilli l'avis de ces utilisateurs quant aux actions réalisées lors des entretiens.

Tout d'abord, nous constatons des écarts entre ce que déclarent avoir fait les utilisateurs et les données recueillies dans les logs. Etant donné la longueur de l'étude, les participants ont eu dû mal à se souvenir précisément des événements. Ensuite, en analysant les logs des utilisateurs, certaines informations sont manquantes car nous avons tracé au global le nombre de contextes proposés par le système et ceux ajoutés par les utilisateurs et à quel moment (date en cours). Une extraction d'un fichier log illustre le format et les données recueillies concernant les propositions du système :

Ajout de contextes :
10/04/2008 = 2
13/04/2008 = 1
26/04/2008 = 1
Nombre d'ajouts de contextes : 4

Modification de contextes :
10/04/2008 = 1
28/04/2008 = 2
Nombre de modifications de contextes : 3

Suppression de contextes :
28/04/2008 =2
Nombre de suppressions de contextes : 2

Il aurait été intéressant de collecter ces données sous forme de session indiquant les actions proposées par le système, celles réalisées par l'utilisateur, le moment et son contexte (sa localisation géographique, le jour, la date et l'heure). Ainsi, notre compréhension pourrait être meilleure. En outre, il convient aussi de tracer finement les paramétrages modifiés dans l'IHM de contrôle et ce par contexte clé (cf. illustration de certaines options disponibles à partir de ce menu et que l'on pourrait tracer).

En effet, nous avons tracé le nombre d'occurrences de consultation du menu "EMMA" mais nous n'avons pas collecté les options modifiées dans ce menu.

En terme d'usage, nous avons donc recoupé les verbatim des utilisateurs avec les données objectives des fichiers logs. Parmi les cinq participants disposant de l'IHM de contrôle, on recueille :

- trois ont eux-mêmes ajouté des contextes clés (l'un pour s'amuser et les deux autres ont surtout utilisé ces contextes comme moyen de commutation rapide de profil de sonnerie à partir du plug-in EMMA) ;
- un seul participant a refusé un contexte proposé par le système car il ne comprenait pas pourquoi EMMA lui proposait cet ajout ;
- deux utilisateurs ont tenté de supprimer des contextes qu'ils avaient eux-mêmes ajoutés (mais la suppression des contextes ne fonctionnait pas) ;
- tous les participants ont modifié certaines options accessibles dans le menu comme la suppression de la vocalisation, les fonds d'écran ou encore les sonneries pour un contexte donné.

Tous les utilisateurs étaient satisfaits de disposer de ce menu mais tous ont estimé que certaines propositions étaient difficilement compréhensibles. Ils auraient souhaité disposer de plus d'informations.

Intitulé des contextes

Tous les contextes pour lesquels nous observons des usages ont été nommés par des lieux (comme "maison", "école", "stage" ou encore "Quimper") ou encore par des périodes temporelles comme "week-end" ou "vacances". Il semble donc que, selon le point de vue utilisateur, la localisation géographique ait un poids prépondérant dans la constitution de contextes clés. Cette hypothèse reste quand même à vérifier auprès d'utilisateurs n'ayant reçu préalablement aucun exemple d'intitulé de contexte. En effet, lors de la séance d'explication à propos du déroulement de l'expérimentation, un descriptif contenant des exemples d'intitulés de contexte avait été fourni aux utilisateurs.

Usages des applications par contexte clé

En terme d'utilité de l'application, une des questions qui se pose est la pertinence de la réorganisation des fonctionnalités selon les contextes clés. En effet, il est légitime de vérifier si les utilisateurs ont bien des usages distincts selon les contextes clés. Ainsi, l'application peut-elle faciliter l'accès aux fonctionnalités les plus utilisées dans un contexte donné (ou du moins augmenter la rapidité d'accès) ? Pour répondre à cette question, nous avons relevé pour les neuf participants (ayant eu des contextes clés) quels étaient globalement la composition et l'ordre des applications utilisées selon les différents contextes clés.

Nous avons centré cette analyse sur les neuf premières fonctionnalités du téléphone. Prenons l'exemple d'un testeur, ce dernier a eu trois contextes clés. Pour chaque contexte, la réorganisation est notée de la façon suivante :

- (0) si elle est identique,
- (1) si un menu est différent,
- (2) si deux menus sont différents.

Contexte 1	Contexte 2	Contexte 3	Notation
Paramètres	Paramètres	Paramètres	0
Messages	Messages	Messages	0
Contacts	Contacts	Enr.Video	1
App.photo	App.photo	Photos Vidéos	1
Réveil	Photos Vidéos	App.photo	2
Enr.Video	Enr.Video	WindowsMedia	1
Photos Vidéos	WindowsMedia	Contacts	2
ActiveSync	Journal des appels	Journal des appels	1
Journal des appels	ActiveSync	ActiveSync	1

Tableau 7 : Procédure de calcul des différences de réorganisation entre les contextes clés

Le diagramme ci-dessous résume les différences observées pour l'ensemble des participants sur tous leurs contextes clés utiles.

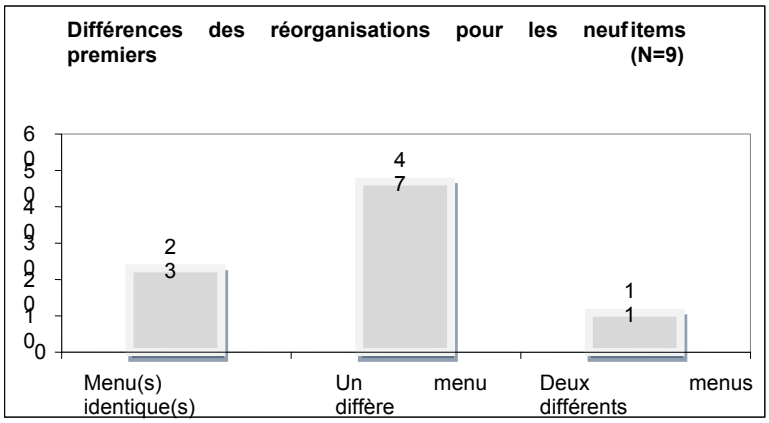


Figure 13 : Différence de réorganisation des neuf premières fonctionnalités (tous les testeurs confondus)

Au travers de ces résultats, nous constatons que globalement les utilisateurs ont bien des usages différents selon les contextes puisque l'ordre des menus varie (dans cinquante huit cas). Dans uniquement vingt trois cas, les menus sont identiques pour différents contextes.

En revanche, si nous regardons plus précisément la réorganisation par contexte clé utile des quatre premières fonctionnalités les plus utilisées, nous observons que dans dix-neuf cas les fonctionnalités sont réordonnées dans le même ordre quel que soit le contexte et dans vingt-deux cas l'ordre des fonctionnalités varie selon le contexte. Il est donc difficile de trancher sur l'utilité de cette adaptation proposée par EMMA. Dans la moitié des cas, la réorganisation de l'ordre des menus n'est pas pertinente, puisque les principaux usages sont les mêmes. Dans l'autre moitié, la réorganisation a permis de remonter des fonctionnalités qui ne sont utilisées que dans certains contextes. Une solution à investiguer consisterait à proposer une partie du menu statique pour les quatre premières fonctionnalités et une partie dynamique pour les autres. Ainsi, les applications a priori plus secondaires pourraient être accessibles plus rapidement.

Décisions utilisateurs

Afin de proposer des adaptations a priori pertinentes du point de vue des utilisateurs, il est nécessaire de les négocier. Pour cela, non seulement l'utilisateur doit pouvoir prendre connaissance des actions réalisées par le système et les comprendre. Mais il doit pouvoir aussi agir sur les adaptations soit en les refusant soit en les acceptant. C'est dans cette optique que nous avons intégré un certain nombre de messages systèmes permettant à l'utilisateur de manifester son accord ou son refus aux propositions d'EMMA. En outre, ces décisions utilisateurs sont tracées par le système et apprises afin soit de ne plus proposer les adaptations toujours refusées soit de les automatiser au-delà d'un certain nombre d'acceptations. Pour finir, ces données nous éclairent aussi sur la pertinence des adaptations proposées par le système.

Nous avons collecté globalement les décisions utilisateurs pour toutes les adaptations proposées par EMMA, c'est-à-dire : les propositions d'ajouts de contextes clés, les réorganisations des menus, les changements de profils de sonnerie, les changements de contextes clés, de sonneries, de fonds d'écran, de page d'accueil et de couleurs et l'automatisation des adaptations.

	Acceptation	Refus	Automatique	Pas pris en compte (time-out)
Toutes adaptations confondues	26%	32%	26%	16%
Ajout de contextes clés	40%	19%	0%	41%
Changement de contextes clés	18%	50%	11%	21%
Réorganisation des menus	46%	19%	22%	13%
Changements de profils de sonnerie	51%	30%	14%	5%

Tableau 9 : Décisions utilisateurs selon le type d'adaptation proposé

Globalement, les participants ont, d'une manière générale, un peu plus rejeté les adaptations proposées qu'ils ne les ont acceptées (32 versus 26%). Par ailleurs, 16% des adaptations proposées n'ont pas été vues par les utilisateurs au moment où elles étaient proposées. En revanche, ils ont automatisé beaucoup d'adaptations proposées, et ce, certainement pour éviter d'être très régulièrement sollicités par le système.

Les propositions d'ajout de contextes clé ont été majoritairement acceptées par les participants. Seulement 19% de ces propositions ont été rejetées.

Les propositions de changements de contexte ont été refusées très majoritairement par les participants. Tous les testeurs ont déclaré que le système les proposait alors qu'ils étaient toujours dans le même contexte. A titre d'exemple, deux participantes ont été obligées d'ajouter un contexte clé chez elle car le système leur proposait d'en ajouter un nouveau alors qu'elles étaient toujours à leur domicile. Pour ne plus être perturbées par les messages, elles ont fini par ajouter un contexte "ordinateur" ou "maison2". Précisons que ces deux participantes ne disposaient pas de l'IHM de contrôle. La position géographique CELL-ID LAC est très sensible sur ce téléphone mobile ce qui explique ces dysfonctionnements constatés par les utilisateurs.

Les participants ont majoritairement accepté les réorganisations des fonctionnalités pour un contexte donné (46% versus 19%). Cependant, les participants n'avaient pas l'impression que leurs usages avaient évolué et ils se demandaient pourquoi le système proposait des réorganisations (verbatim issus des entretiens).

Le changement de mode de profil de sonnerie d'un contexte à un autre a été plébiscité par les utilisateurs. Ils ont jugé cette adaptation comme étant très utile car elle évite un certain nombre de manipulations sur le téléphone pour modifier le mode de sonnerie.

Perception utilisateur

Les entretiens individuels ont permis d'identifier les points forts et les points faibles de l'application, la pertinence ressentie des adaptations ou encore les difficultés rencontrées par les participants.

Avis général

Aucun participant n'a clairement rejeté l'application. Ils ont pour la plupart estimé que le concept était intéressant et pratique (cinq participants). En revanche, les propositions système sont gênantes car trop nombreuses et deux participants ont déclaré que ces propositions semblaient trop rapides (un jour après l'installation d'EMMA voire le jour même). Deux testeurs ont aussi jugé que cette application serait plus utile pour des personnes ayant des usages très variés ou pour des professionnels.

Concernant la réorganisation des menus, cinq participants ont un avis positif pour cette adaptation ("pratique", "bien", "plutôt intéressant"). En revanche une personne a été gênée de ne pas retrouver toujours la même organisation des menus. Trois ont estimé que les réorganisations étaient trop fréquentes et pas toujours justifiées et une personne ne se rendait pas compte des réorganisations car ses usages semblaient trop proches d'un contexte à un autre.

Le basculement de profils de sonnerie a été jugé positivement par six participants. Une seule a clairement indiqué qu'elle préfère le faire elle-même.

Messages systèmes

Comme nous l'avons développé un peu plus haut, tous les participants ont estimé que les sollicitations du système étaient trop nombreuses et certaines incompréhensibles. En effet, les participants ne comprenaient pas toujours pourquoi EMMA proposait de réorganiser les menus par exemple ou d'ajouter un contexte clé. Il semble important voire nécessaire de proposer des messages plus explicites sur les raisons des adaptations proposées par le système.

Disposer ou non de l'IHM de contrôle

Parmi les cinq participants qui utilisaient EMMA sans l'IHM de contrôle, quatre ont jugé qu'il était gênant de ne pas pouvoir contrôler les adaptations. Ils auraient souhaité pouvoir paramétrer certains contextes ou tout simplement les supprimer. Il semble donc nécessaire pour ce type d'application d'ajouter un menu dédié permettant de personnaliser et modifier les adaptations.

5.2.4 ENSEIGNEMENTS

Cette évaluation nous a permis de tirer un certain nombre d'enseignements qui peuvent donner un éclairage sur les bonnes pratiques à respecter au regard de l'évaluation terrain.

- Le recueil des retours utilisateurs doit être pris en compte au plus près du moment de l'adaptation. Au-delà de plusieurs semaines, les participants ont du mal à se souvenir précisément de certaines informations et de leurs impressions. Des méthodes de recueil en ligne sont à investiguer pour la suite des études (journal de bord, SMS de "sondage", etc.).
- Pour éviter toutes pertes de données objectives qui sont précieuses pour l'analyse des comportements du système et de l'utilisateur, il convient de les stocker sur le terminal (si elles ne sont pas trop volumineuses) et de les envoyer automatiquement par GPRS ou GSM à un serveur.
- Le moment de l'adaptation semble être réellement un point critique dans le ressenti et l'acceptation du système. Pour EMMA, les adaptations étaient proposées uniquement lors de la reprise en main du téléphone après une période d'inactivité. Il s'avère que cette limite n'est pas suffisante puisque

majoritairement tous les participants ont estimé que les sollicitations du système étaient trop nombreuses et gênantes. De plus, le clustering et les propositions d'adaptations ont peut-être été réalisés de manière trop systématique. Tester différentes périodes d'adaptation et la fréquence des apprentissages et clustering semble nécessaire pour tenter de déterminer les bons seuils.

- Pour déterminer les options nécessaires à intégrer dans l'IHM de contrôle et les usages de ces fonctionnalités, il convient de tracer précisément les options sélectionnées ou non et à quel moment.
- L'observabilité et la raison des propositions du système semblent aussi être des éléments très importants pour favoriser l'acceptabilité des systèmes adaptatifs. Durant cette étude, tous les utilisateurs se sont plaints de ne pas toujours comprendre pourquoi le système proposait telle modification du menu ou tel ajout de contexte. Il convient donc de travailler sur des modes de dialogue avec l'utilisateur lui permettant de comprendre les actions du système et les propositions faites avant de les accepter ou de les refuser.
- La réorganisation des menus semble être une adaptation pertinente pour certaines cibles d'utilisateurs. Néanmoins, comme nous avons pu le constater, les usages peuvent être très peu variés entre différents contextes clés, les premiers menus étant dans l'ensemble plutôt stables et liés aux fonctions de communication. Une approche consisterait à proposer un menu statique pour les quatre voire cinq premières fonctionnalités et un menu dynamique (adaptatif) pour les fonctionnalités secondaires. Ainsi, les utilisateurs retrouveraient une certaine stabilité des informations et des accès plus rapides pour des tâches réalisées de manière moins fréquente.

5.3 RECOMMANDATIONS ERGONOMIQUES

Les résultats des évaluations nous ont permis de formuler un ensemble de recommandations ergonomiques pour la conception de systèmes interactifs plastiques. Etant donné que ces recommandations sont le fruit des premières évaluations, elles doivent être considérées avec prudence. Elles mettent en évidence quatre critères du référentiel de [Bastien et Scapin 1993] qui méritent une attention particulière. Les quatre critères qui semblent incontournables sont les suivants : contrôle explicite, charge de travail, adaptabilité et guidage.

Ces critères identifient les conditions dans lesquelles la plasticité peut constituer un facteur facilitateur d'usage en téléphonie mobile. Parmi eux, le contrôle explicite est identifié comme un requis par l'ensemble des utilisateurs. Les recommandations énoncées s'articulent autour des deux leviers de la plasticité (remodelage et redistribution) et du contrôle utilisateur.

- Pour le remodelage, les recommandations sont les suivantes :
  - Réduire la navigation (charge de travail).
  - Déléguer au système les tâches répétitives (charge de travail).
  - Prendre en compte les préférences et les besoins de l'utilisateur (adaptabilité).
- Pour la redistribution, les recommandations sont les suivantes :
  - Privilégier la charge de travail à l'homogénéité/cohérence (charge de travail).
  - Améliorer le confort et la convivialité (adaptabilité).

  - Réduire la charge de travail (charge de travail).
- Pour le contrôle utilisateur, les recommandations sont les suivantes :
  - Expliquer la raison d'être des propositions (guidage)
  - Limiter les propositions du système (charge de travail).
  - Placer l'adaptation sous le contrôle de l'utilisateur (contrôle explicite).

6. CONCLUSION ET PERSPECTIVES

L'année 2009 marque la première décennie des travaux en plasticité. Autour de ce sujet, traité aujourd'hui aux plans national et international, les avancées sont importantes. Cependant, l'évaluation reste un point faible où les résultats sont globalement pauvres.

Cet article tente d'apporter des réponses à des questions, jusqu'ici sans réponse, sur l'acceptabilité et l'utilisabilité des IHM adaptatives. Pour ce faire, trois démonstrateurs (PlasticSMS, MobileSMS et EMMA) ont été développés et évalués dans le but de recueillir le ressenti des utilisateurs vis-à-vis du remodelage et de la redistribution ainsi que de comprendre l'intérêt du contrôle utilisateur. Les méthodes d'évaluation utilisées impliquent, pour la plupart, l'utilisateur (focus group, tests utilisateur, entretiens). Les évaluations de PlasticSMS et MobileSMS se sont déroulées en laboratoire. EMMA qui nécessite la collecte de données sur des usages réels a été évaluée sur le terrain.

A notre connaissance, les résultats exposés dans cet article figurent parmi les premiers retours utilisateurs sur les IHM plastiques. Nous avons recueilli le ressenti des utilisateurs par rapport aux concepts qui ont fait l'intérêt de ces travaux. L'adaptation à la plate-forme par la redistribution sur une plate-forme plus confortable a été perçue comme la plus utile. Par ailleurs, nous avons observé des échecs uniquement au niveau du remodelage. En effet, il semble préférable d'afficher directement l'IHM, remodelée, dans sa globalité. Nous avons également appris qu'il est préférable de placer l'adaptation sous le contrôle de l'utilisateur. Les enseignements tirés des évaluations ont permis de formuler de premières recommandations ergonomiques sur la base du référentiel de [Bastien et Scapin 1993]. Les critères qui sont apparus comme pertinents au cours des évaluations sont : le contrôle explicite, la charge de travail, l'adaptabilité et le guidage. La charge de travail semble être le critère prépondérant. Cinq des neuf recommandations ergonomiques données dans cet article sont relatives à ce critère.

Toutefois, les résultats présentés dans cet article sont à manipuler avec précaution car les protocoles adoptés présentent des limites. Nous nous sommes inscrits dans une démarche qualitative. Les évaluations ont donc été menées auprès d'un panel réduit d'utilisateurs. Les retours du focus group sont uniquement basés sur le déclaratif. En laboratoire, il est difficile d'approcher un large éventail de contextes d'usage. Pour pallier aux difficultés de recrutement dans l'évaluation d'EMMA, nous avons prêté les téléphones aux participants. Avec ce nouvel équipement, les participants n'ont pas tout à fait adopté leurs pratiques usuelles. Nous avons constaté que, par exemple, ils ont très peu personnalisé le mobile alors qu'ils le font en règle générale. Après six semaines, les participants ont eu tendance à oublier certaines informations sur les propositions d'EMMA lors des premiers jours d'utilisation.

Pour le futur, nous pensons investiguer d'autres concepts clé de la plasticité. La valeur qui vise à aller au-delà de l'utilisabilité est une notion à laquelle nous souhaitons particulièrement nous intéresser. Nous souhaitons, tout d'abord, comprendre ce que représente la valeur d'un système adaptatif aux yeux des utilisateurs. Nous pensons mener cette étude par entretiens puis focus groups avec élaboration de maquettes

visuelles, démonstrateurs et/ou prototypes. Il s'agira ensuite, sur la base de la conception centrée valeur [Cockton 2006], de mettre en place un prototype stable qui sera évalué auprès d'utilisateurs en situations réelles.

REMERCIEMENTS
Nous remercions chaleureusement Dominique Scapin et Jean Vanderdonckt pour l'intérêt porté aux travaux. Nous associons à ces remerciements les relecteurs de cet article qui, par leurs remarques constructives, nous ont permis d'en améliorer l'écriture.

RÉFÉRENCES

Amalberti, R., Valot, C. Le Magicien d'Oz. CERMA, journée du PRC Rhône-Alpes.1992.

Bastien, J.M.C., Scapin D. Ergonomic Criteria for the Evaluation of Human-Computer Interfaces, Rapport technique INRIA, N°156, Juin 1993.

Bellik, Y. Présentation Multimodale de l'Information. Habilitation à Diriger des Recherches. Université d'Orsay Paris-Sud, 2006, 90 pages.

Benadi, S. Structuration des données et des services pour le télé-enseignement. Thèse de Doctorat de l'Institut National des Sciences Appliquées de Lyon, Septembre 2004.

Benyon, D., Murray, D. Adaptive Systems : From Intelligent Tutoring to Autonomous Agents. In Knowledge-Based Systems, 197–219, 1993.

Bobillier-Chaumon, M.E., Carvallo, S., Tarpin-Bernard, F., Vacherand-Revel, J. Adapter ou uniformiser les interactions personnes-systèmes? Revue d'Interaction Homme-Machine, Europia, Vol. 6, 2, 2005, 91-129.

Blanchet, A., Ghiglione, R., Massonat, J., Trognon, A. Les techniques d'enquête en sciences sociales : observer, interviewer, qustionner. Dunod, Paris, France, 2000.

Brusilovsky, P. Methods and Techniques of Adaptive Hypermedia. In User Modeling and User-Adapted Interaction, 87–129, juillet 1996.

Brusilovsky,P., Karagiannidis, C., Sampson, D. The Benefits of Layered Evaluation of Adaptive Applications and Services. In Proceedings of the Workshop on Empirical Evaluations of Adaptive Systems, 8th International Conference on User Modeling (UM'01), S. Weibelzahl, D. N., G. Weber, éditeurs, 1–8, juillet 2001.

Brusilovsky, P., Karagiannidis, C., Sampson, D. Layered Evaluation of Adaptive Learning Systems. In International Journal of Continuing Engineering Education and Life-Long Learning, 14(4), 402–421, 2004.

Calvary, G., Coutaz, J. Plasticité des Interfaces : une nécessité ! Information-Interaction-Intelligence, Actes des deuxièmes Assises nationales du GDR I3, Cépaduès Editions, J. Le Maître (Ed), Nancy, France, Décembre 2002, 247-261.

Calvary, G. Plasticité des Interfaces Homme-Machine. Habilitation à Diriger des Recherches. Université Joseph Fourier, Novembre 2007, 97 pages.

Carbonell, N., Simonin, J. Interfaces adaptatives – Modèles de l'utilisateur. Journée d'étude Interfaces adaptatives, Laboratoire Paragraphe, Université Paris 8, juin 2004.

Chin, D. N. Empirical evaluation of user models and user-adapted systems. User Modeling and User-Adapted Interaction 11(1-2), 2001, 181–194.

Cockton, G. From Quality in Use to Value in the World. In ACM Proc. CHI 2004, Late Breaking Results, 2004a, 1287-1290.

Cockton, G. Value-centred HCI. Proceedings of the third Nordic conference on Human-computer interaction, 2004b, 149-160.

Cockton, G. A development Framework for Value-Centred Design. In ACM Proc. CHI 2005, Late Breaking Results, 2005, 1292-1295.

Cockton, G. Designing worth is worth designing. In Proceedings of the 4th Nordic conference on Human-computer interaction: changing roles, 2006, 165-174.

Dâassi, O. Les comets : une nouvelle génération d'Interacteurs pour la Plasticité des Interfaces Homme-Machine. Thèse de Doctorat de l'Université Joseph Fourier, Grenoble I, Janvier 2007.

Demumieux, R., Ganneau, V., Calvary, G., Gegovska, E. Les interfaces plastiques premiers retours utilisateurs : évaluations en laboratoire. In Actes du colloque Ergo-IA 2008, L'humain au cœur des systèmes et de leur développement, 2008, 117-124.

Diaper, D. Task Analysis for Human-Computer Interaction. Ellis Horwood, Chichester, UK, 1990.

Dieterich, H., Malinowski, U., Kühme, T., Schneider-Hufschmidt, M. State of the Art in Adaptive User Interfaces. In Adaptive User Interfaces: Principles and Practices, Schneider-Hufschmidt & Al. (ed.), 1994, 13-48.

Dittrich, K., Gatziu, S., Geppert, A. The Active Database Management System Manifesto: A Rulebase of ADBMS Features. In Proceedings of the Second International Workshop on Rules in Database Systems (RIDS'95), pages 3–20. Springer-Verlag, London, UK, 1995.

Edmonds, E.A. Adaptive Man-Computer Interfaces. In Coombs, M. J. and Alty, J. L. (Eds), Computing Skills and the User Interface. London: Academic Press, 1981.

Florins, M. Graceful Degradation, A Method for Designing Multiplatform Graphical User Interfaces, PhD Université Catholique de Louvain, Canada, Juillet 2006.

Gabillon, Y., Calvary, G., Fiorino, H. Composing interactive systems by planning. In Proceedings of the 4th French-speaking conference on Mobility and ubiquity computing, ACM International Conference Proceeding Series, 2007, 37-40.

Ganneau, V., Demumieux, R., Calvary, G. EMMA: modèle utilisateur pour la plasticité des interfaces homme-machine en mobilité. In Proceedings of the 4th French-speaking conference on Mobility and ubiquity computing, ACM International Conference Proceeding Series, 2008, 1-8.

Gena, C. Methods and techniques for the evaluation of user-adaptive systems. In The Knowledge Engineering Review, Volume 20, 2005, 1-37.

Gena, C. A User-Centered Approach for Adaptive Systems Evaluation. InProceedings of the 5th Workshop on User-Centred Design and Evaluation of Adaptive Systems, 5th International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH'06), juin 2006, 430–439.

Goodman, J., Brewster, S., Gray, P. Using Field Experiments to Evaluate Mobile Guides. In Proceedings of the 3rd Annual Workshop on HCI in Mobile Guides. 2004.

Greenbaum, T. The handbook for Focus Group Research. Seconde edition, 1998.

Grolaux, D., Vanderdonckt, J., Van Roy, P. FlexClock, a Plastic Clock Written in Oz with the QTk toolkit. In Proceedings of the First International Workshop on Task Models and Diagrams for User Interface Design, 2002, 135 – 142.

Grolaux, D., Vanderdonckt, J., Van Roy, P. Attach Me, Detach Me, Assemble Me Like You Work, In proceedings of INTERACT 2005, M.F. Costabile and F. Paternò (Eds.), 2005, 98 – 212.

Hesselman, C., Cesar, P., Vaishnavi, I., Boussard, M., Kernchen, R., Meissner, S., Spedalieri, A., Sinfreu, A., Räck, C. Delivering interactive multimedia services in dynamic pervasive computing environments. In Proceedings of the 1st international conference on Ambient media and systems, Ambi-sys, 2008.

Holly, S.V., Jacko, J.A., Emery, V.K. Multimodal feedback: establishing a performance baseline for improved access by individuals with visual impairments. In Proceedings of the fifth international ACM conference on Assistive technologies, ACM SIGACCESS Conference on Assistive Technologies, 2002, 49-56.

Höök, K. Evaluating the Utility and Usability of an Adaptive Hypermedia System. In Proceedings of the 2nd International Conference on Intelligent User Interfaces, IUI 97.

Höök, K. Steps to take before intelligent user interfaces become real. Interacting with Computers (2000) 12(4), 409-426.

Huebsch, G., Kadner, K. Multimodal information access across multiple devices. In Proceedings of the 4th international conference on mobile technology, applications, and systems and the 1st international symposium on Computer human interaction in mobile technology, International Conference On Mobile Technology, Applications, And Systems, 2007, 736-742.

Hulkko, S., Mattelmäki, T., Virtanen, K., Keinonen, T. Mobile probes. In Proceedings of the third Nordic conference on Human-computer interaction, 43 – 51, 2004.

Jameson, A. Adaptive Interfaces and Agents. Human-Computer Interface Handbook, J.A. Jacko & A. Sears (Eds), Chapter 15, 2003, 305-330.

Jambon, F., Golanski, C., Pommier, P.J. Évaluation des dispositifs mobiles : sur le terrain ou en laboratoire ? In Actes de la 18ème Conférence Francophone sur l'Interaction Homme-Machine (IHM'06), ACM, New York, NY, USA, 35-42, 2006.

Kellar, M., Reilly, D., Hawkey, K., Rodgers, M., MacKay, B., Dearman, D., Ha, V., MacInnes, J., Nunes, M., Parker, K., Whalen, T., Inkpen, K. It's a Jungle Out There : Practical Considerations for Evaluation in the City. In CHI'05 extended abstracts on Human Factors in Computing Systems (CHI'05), ACM, New York, NY, USA, 1533–1536, avril 2005.

Kjeldskov, J., Paay, J. Just-For-Us : A Context-Aware Mobile Information System Facilitating Sociality. In Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services (MOBILEHCI'05), ACM, New York, NY, USA, 23-30, 2005.

Lemmelä, S., Vetek, A., Mäkelä, K., Trendafilov, D. Designing and Evaluating Multimodal Interaction for Mobile Contexts. In Proceedings of the 10th international conference on Multimodal Interfaces, ACM International Conference on Multimodal Interfaces, 2008, 265-272.

Lewis, C., Polson, P.G., Wharton, C., Rieman, J. Testing a Walkthrough Methodology for Theory-Based Design of Walk-Up-and-Use Interfaces. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'90), ACM, New York, NY, USA, 235–242, 1990.

Lieberman, H., Faaborg, A., Espinosa, J., Stocky, T. Common Sense on the Go: Giving Mobile Applications an Understanding of Everyday Life. In British Telecom Technology Journal, Vol. 22, No. 4, October 2004

Mariage, C. Espace de classification des méthodes et outils d'évaluation de l'ergonomie des systèmes à technologie Web. Mémoire en vue de l'obtention du Diplôme Interuniversitaire d'Etudes Approfondies en Informatique (DEA). Université Catholique de Louvain, 2001, 59 pages.

Nielsen, J., Molich, R. Heuristic Evaluation of User Interfaces. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'90), ACM, New York, NY, USA, 1990, 249–256.

Nielsen, J. Usability Engineering. Morgan Kaufmann, San Francisco, CA, USA, 1993.

Evaluation des EIAH : une nécessaire diversité des méthodes. In proceedings of TICE 2004, Compiène, France, 2004, 265-271.

Norman, D.A. Affordance, Conventions, and Design. In Interactions, 38–43, 1999.

Polson, P.G., Lewis, C., Rieman, J., Wharton, C. Cognitive Walkthroughs : A Method for Theory-Based Evaluation of User Interfaces. In International Journal of Man-Machine Studies, 741–773, mai 1992.

Preece, J., Rogers, I., Sharp, H., Benyon, D., Holland, S., Carey, T. Human-Computer Interaction. Addison-Wesley, Essex, UK, 1994.

Senach, B. L'évaluation ergonomique des interfaces homme-machine : une revue de la littérature. In L'ergonomie dans la conception des projets informatiques, J.-C. Sperandio, éditeur. Octares Éditions, 69–122, 1993.

Sohn, T., Li, K.A., Griswold, W.G., Hollan, J.D. A diary study of mobile information needs. In Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems, 433 – 442, 2008.

Tarpin-Bernard, F. Interaction Homme-Machine adaptative. Habilitation à Diriger des Recherches. Institut National des Sciences Appliquées de Lyon, Université Claude Bernard Lyon I, Décembre 2006, 99 pages.

Thevenin, D. and Coutaz, J. Plasticity of User Interfaces: Framework and Research Agenda. In: Edinburgh, A.S., Johnson, C. (eds.) Proc. Interact 1999, IFIP IOS Press Publ., Amsterdam, 110–117.

Thevenin, D. Adaptation en Interaction Homme-Machine : le cas de la Plasticité. Thèse de l'Université Joseph Fourier, Grenoble I, 2001, 234 pages.

Thompson,D. Concise Oxford English Dictionary, 1996.

Tobar, C. Yet Another Evaluation Framework. In Proceedings of the 2nd Workshop on Empirical Evaluation of Adaptive Systems, 9th International Conference on UserModeling (UM'03), S.Weibelzahl, A. Paramythis, éditeurs, 15–24, juin 2003.

Totterdell, P., Boyle, E. The Evaluation of Adaptive Systems. In Adaptive User Interfaces, D. Browne, P. Totterdell, M. Norman, éditeurs. Academic Press, 161–194, 1990.

Totterdell, P., Rautenbach, P. Adaptation as a Problem Design. In Adaptive User Interfaces, D. Browne, P. Totterdell, M. Norman, éditeurs, Academic Press, 59-84, 1990.

Vanderdonckt, J., Grolaux, D., Van Roy, P., Limbourg, Q., Macq, B., Michel, 93 B., A Design Space for Context-Sensitive User Interfaces, Proc. of ISCA 14th Int. Conf. on Intelligent and Adaptive Systems and Software Engineering IASSE'2005.

Fatoumata Camara est actuellement étudiante en deuxième année de doctorat. Son travail se déroule dans les Orange Labs, centre de Lannion et dans le Laboratoire d'Informatique de Grenoble (LIG), équipe IIHM (Ingénierie de l'Interaction Homme-Machine). Sa thèse s'inscrit dans le domaine de l'Informatique, plus précisément, en Interaction Homme-Machine (IHM). Elle s'intéresse au problème de l'adaptation, selon deux axes : Information et Interaction. Il s'agit d'étudier, conjointement, ces deux formes d'adaptation avec un barycentre porté sur la qualité de vie.

Gaëlle Calvary est Professeur en Informatique à Grenoble INP. Elle effectue ses recherches au Laboratoire d'Informatique de Grenoble. Son thème de recherche est la Plasticité des Interfaces Homme-Machine (IHM), c'est-à-dire l'étude de la capacité d'adaptation des IHM à leur contexte d'usage dans le respect de la valeur attendue par l'utilisateur cible. L'approche qu'elle a le plus explorée est l'Ingénierie Dirigée par les Modèles. Elle étudie la conservation et l'exploitation à l'exécution des modèles de conception pour la plasticité des IHM et, de façon, plus générale, leur qualité. Elle est membre du projet européen ITEA UsiXML et de l'ANR MyCitizSpace. Elle est membre du comité directeur du GDR I3 et co-fondatrice de l'atelier CESAME (Conception et Exécution de Systèmes interactifs doués d'Adaptation dans un monde Mixte en Evolution).

Rachel Demumieux est ergonome senior à Orange Labs (Direction Recherche et Développement de France Télécom). Elle travaille dans le domaine de la conception et l'évaluation d'interfaces de dispositifs utilisés en mobilité. Ses travaux de recherche portent sur les problématiques de collecte de données d'usage en mobilité et sur l'adaptation des interfaces en fonction des profils utilisateurs et des contextes d'utilisation. Elle a co-encadré la thèse de Vincent Ganneau et actuellement celle de Fatoumata Camara traitant de l'adaptation des interfaces et des modes d'interaction selon les contextes d'usage.

Vincent Ganneau est Ingénieur de Recherche et Développement chez Haploid, une start-up spécialisée dans la conception et la réalisation d'applications pour téléphones mobiles. Il est titulaire d'un Doctorat en Informatique sur l'Ingénierie de l'Interaction Homme-Machine. Sa thèse traite d'un Modèle Utilisateur pour la Plasticité des Interfaces Homme-Machine (IHM) en Mobilité. Ses travaux examinent l'adaptation des IHM du point de vue de l'usage. Ils explorent les Réseaux Bayésiens pour modéliser et exploiter les données collectées en mobilité afin d'identifier les contextes d'usage clés et proposer la meilleure adaptation.