



Projet VERBATIM

Sous-projet 1 – Lot 3

Comparaison de méthodologies d'élaboration de recommandations ergonomiques pour les interfaces multimodales

Version : 2

Date : 15/12/2006

Auteurs : Joëlle Blanquet, Jérôme Gatefin (France Télécom R&D)



LSR



Table des matières

1. Introduction	3
2. Méthode 1 : Elaboration de recommandations ergonomiques pour les interfaces multimodales, modalité par modalité	4
2.1 Hypothèses de recommandations ergonomiques pour chaque modalité à partir de guides existants.....	4
1 ^{ère} conclusion.....	6
2.2 Recommandations élaborées par modalité à partir de tests utilisateurs.....	7
2 ^{ème} conclusion.....	8
3. Méthode 2 : Elaboration de recommandations ergonomiques pour les interfaces multimodales à partir de l'ensemble des modalités	9
3.1 Recommandations multimodales élaborées "globalement" à partir de tests utilisateurs.....	9
3 ^{ème} conclusion.....	11
3.2 Les critères d'ergonomie d'un point de vue multimodal.....	12
3.2.1 Guidage.....	12
3.2.2 Charge de travail.....	14
3.2.3 Contrôle explicite	15
3.2.4 Adaptabilité.....	16
3.2.5 Gestion des erreurs	17
3.2.6 Homogénéité/Cohérence.....	18
3.2.7 Signifiante des codes et dénominations	18
3.2.8 Compatibilité.....	19
4 ^{ème} conclusion.....	19
4. Comparaison des deux méthodes et discussion	20
5. Bibliographie.....	21

1. Introduction

Le présent document se situe dans le cadre du projet RNRT VERBATIM, dont l'objet est la VERification Biformelle et Automatisation du Test des Interfaces Multimodales.

Le Sous Projet 1 concerne les aspects utilisateurs. Les lots 2,3,4 étudient les recommandations ergonomiques dans le but d'évoluer vers des spécifications formelles.

L'objectif du sous-projet 1 était d'identifier des recommandations ergonomiques génériques d'un niveau " macro ", décrivant le bon comportement des interfaces suivant différents contextes et modalités, et susceptibles d'être formalisées en vue d'une validation automatisée. On entend par interfaces le support d'interaction entre l'humain et le service. On entend par contexte, ce qui est du domaine de l'utilisateur et ce qui est du domaine des contraintes environnementales. On entend par modalité, les moyens d'interaction entre le service et l'humain.

Les recommandations ergonomiques doivent permettre l'accommodation fluide et continue de l'interaction entre l'utilisateur (caractéristiques individuelles, activités,...), les interfaces (supports), les modalités (visuelle, vocale,...), et les contextes (environnement physique, privé/public,...).

Le domaine de recherche appliquée consistant à mettre au point des recommandations ergonomiques pour les services multimodaux est un domaine encore nouveau. Un état de l'art a été réalisé dans le cadre du lot 1.

Le présent document correspond au livrable du lot 3 : "Comparaison de méthodologies d'élaboration de recommandations ergonomiques pour les interfaces multimodales".

L'objectif de ce lot 3 était d'étudier la possibilité de réutilisation de recommandations ergonomiques existantes et leur couplage avec de nouvelles recommandations. Une première méthode consiste à effectuer l'identification des recommandations ergonomiques, modalité par modalité. Chaque modalité étant considérée séparément, dans le cas d'un service multimodal, cette méthode traite les modalités de manière incrémentale. Une seconde méthode consiste à prendre en compte le service multimodal avec plusieurs modalités en global, et à élaborer les recommandations ergonomiques à partir de l'ensemble des modalités. La problématique sous-jacente est d'étudier les possibilités de réutilisation des recommandations ergonomiques concernant les interfaces, les modalités, les contextes.

2. Méthode 1 : Elaboration de recommandations ergonomiques pour les interfaces multimodales, modalité par modalité

Pour élaborer des recommandations ergonomiques dans le domaine du multimodal, une première méthode consiste à effectuer l'identification des recommandations ergonomiques, modalité par modalité.

2.1 Hypothèses de recommandations ergonomiques pour chaque modalité à partir de guides existants

Cette identification peut être réalisée par une extraction de recommandations ergonomiques monomodales à partir des guides de recommandations existants pour chaque modalité. Il s'agit alors de repérer des recommandations ergonomiques susceptibles d'être pertinentes pour les services multimodaux, notamment pour les services que nous avons pu observer à travers différentes maquettes et prototypes (France Télécom, CLIPS IMAG).

Au sein de France Télécom, on dispose de nombreux travaux relatifs à l'ergonomie des usages et concernant les services de télécommunications mobiles, fixes, internet, vocaux, etc. Ces travaux ont permis de rédiger un certain nombre de guides de recommandations ergonomiques qui ont été utilisés pour repérer et proposer des recommandations ergonomiques susceptibles d'être pertinentes pour les interfaces multimodales.

Tous les services multimodaux que nous avons pu manipuler proposaient au moins la modalité "graphique" et la modalité "vocale". Les interfaces graphiques étaient de type web ou de type WIMP (Windows Icons Menus Pointing). Les interfaces vocales pouvaient comporter plusieurs voix et dans certains services, la voix était "véhiculée" graphiquement par un avatar.

L'examen des guides existants, en particulier ceux produits à France Télécom à partir de tests utilisateurs et d'expertises ergonomiques, nous a permis de repérer des recommandations ergonomiques monomodales "incontournables" pour la modalité graphique comme pour la modalité vocale.

Voyons dans quelle mesure ces recommandations monomodales restent pertinentes pour un service multimodal. Plusieurs cas se présentent :

- Des recommandations existant pour chacune des modalités, ou encore existant dans les guides pour une des modalités, mais pertinentes dans une autre modalité.

Par exemples :

Reco. *Identifier le début d'une session par un message d'accueil personnalisé*

- Reco.** *Avertir l'utilisateur de temporisation(s) en cours et éditer un message indiquant l'état de cette temporisation (exemple : déconnexion dans N secondes)*
- Reco.** *Réduire au minimum le nombre d'étapes obligatoires pour optimiser l'atteinte d'un objectif (exemple : proposer des « séquences » de choix prédéfinis pour les cas les plus fréquents).*

Pour un service monomodal, graphique ou vocal, ces recommandations sont chacune fortement recommandée, voire incontournable.

Pour un service multimodal, la question se pose différemment car leur mise en œuvre va dépendre de la gestion de la coopération entre les modalités, autrement dit : redondance ou complémentarité entre le vocal et le graphique.

La réutilisation de ces recommandations reste hypothétique et demande à être vérifiée selon le type de coopération mis en œuvre entre modalités et avec des retours de tests utilisateurs.

- Des recommandations relatives à une modalité, ne pouvant pas être transposées à une autre modalité, mais qui peuvent rester pertinentes dans un service multimodal.

Par exemples :

- Reco.** *Guider l'utilisateur dans le repérage visuel et l'identification des actions possibles (nommage, localisation, codage graphique) de façon cohérente dans toute l'application*
- Reco.** *Veiller à maintenir la lisibilité par un contraste visuel suffisant entre le fond et le texte*
- Reco.** *Opter pour un indicateur de hiérarchie dès que la structure comporte plus de deux niveaux de profondeur*

Certaines recommandations associées à la perception et à la lisibilité des éléments visuels restent applicables à la modalité graphique dans un service multimodal.

De même, certaines recommandations associées à la perception et à l'audibilité des éléments sonores restent applicables à la modalité vocale dans un service multimodal.

Cependant, la question de la coopération entre les modalités et de la cohérence d'ensemble se pose une fois encore. Des vérifications de cohérence doivent être faites entre modalités. En effet, une interface graphique homogène, couplée à une interface vocale homogène, pourraient néanmoins être incohérentes entre

elles. Par exemple, dans le cas d'une interface graphique qui afficherait une fonction "téléphone" en vert alors que l'interface vocale énoncerait "sélectionner la fonction *communiquer* en jaune".

La réutilisation de ces recommandations est probable pour ce qui concerne en particulier les aspects perceptifs, mais doit néanmoins être examinée du point de vue de la cohérence d'ensemble entre modalités.

- Des hypothèses de recommandations multimodales, établies à partir de recommandations monomodales similaires pour deux modalités.

Par exemples :

Reco. S'assurer que les procédures de validation soient cohérentes dans toute l'application

Reco. Veillez à l'homogénéité : la même commande déclenche la même action et se formule toujours de la même manière

Certaines recommandations sont identiques dans plusieurs modalités. Cependant, la coopération entre modalités est à examiner de près dans un service multimodal. Par exemple, si un même terme doit être utilisé pour désigner la même action en graphique comme en vocal, il reste que la modalité vocale peut proposer des synonymes afin de lever d'éventuels blocages.

Ces recommandations sont hypothétiques et demandent à être vérifiées selon le type de coopération mis en œuvre entre modalités et avec des retours de tests utilisateurs.

1^{ère} conclusion

Pour conclure, il apparaît que l'examen des guides de recommandations ergonomiques existant par modalité peut être une aide dans la conception des services multimodaux. En particulier :

- pour reprendre les recommandations touchant tout ce qui est de l'ordre de la bonne perception des informations fournies par une modalité (par exemple, taille de police minimale pour pouvoir être lisible, minima de contraste, etc).
- pour extrapoler des recommandations d'une modalité à une autre
- pour mettre en évidence les choix à faire au niveau de la redondance ou de la complémentarité.

Cependant, ces reprises de recommandations monomodales doivent être ajustées selon le type de coopération entre modalités et vérifiées par des tests utilisateurs sur le service multimodal.

2.2 Recommandations élaborées par modalité à partir de tests utilisateurs

Cette méthode ne peut être réalisée que dans le cas où le service multimodal offre une totale redondance des modalités.

Ce type de service peut être le résultat d'une combinaison en un seul service dit multimodal, de deux services monomodaux à l'origine. Historiquement, les deux services ont pu être développés pour permettre à l'utilisateur de choisir l'une ou l'autre modalité selon son contexte.

Dans ce cas particulier, le service peut être utilisé exclusivement dans chaque modalité et l'ajustement de la coopération entre modalité n'est pas réellement traité.

Ce fut le contexte de l'expérimentation utilisateurs du service multimodal "Plan Resto" (cf. SP5.3), où nous avons pu tester ce service d'une part en monomodal, modalité graphique uniquement et modalité vocale uniquement, et d'autre part en multimodal (avec et sans avatar virtuel).

Les tests effectués sur les versions monomodales du service Plan Resto avaient comme objectif principal de permettre la comparaison avec les versions multimodales en termes d'efficacité, d'efficacité, de charge mentale et de satisfaction.

Les tests utilisateurs du service sur les deux versions monomodales nous ont permis :

- pour la version vocale, de mettre en évidence d'une part la lourdeur du dialogue pour un service de ce type : "rechercher un restaurant à Paris" peut constituer une tâche relativement fastidieuse lorsqu'elle est réalisée uniquement en vocal, et cela d'autant plus quand l'utilisateur n'a pas une idée précise de ce qu'il cherche. D'autre part, l'importance de la prosodie de la synthèse vocale car certaines questions ont été interprétées par les utilisateurs comme des affirmations occasionnant des blocages dans l'interaction. Notons pour ces deux exemples, qu'en multimodal, la modalité graphique constitue un support pertinent pour ce type de service et permet par ailleurs, de lever certaines ambiguïtés dues aux problèmes de prosodie de la synthèse vocale.
- pour la version graphique, de mettre en évidence des problèmes liés à l'interface graphique (par exemple : mauvaise reconnaissance des boutons, surcharge informationnelle, problèmes de navigation, etc.)

Dans ce cas particulier de services multimodaux permettant une utilisation exhaustive dans chacune des modalités, il est nécessaire de tester le service modalité par modalité, pour vérifier qu'il peut effectivement être utilisé même lors de la désactivation d'une des modalités.

Pourtant, si les recommandations ergonomiques pour chaque modalité peuvent être reprises, elles ne doivent pas être appliquées sans prendre en compte les autres modalités. La cohérence entre modalités doit être examinée de manière à ne pas susciter des erreurs ou des contradictions dans la présentation des informations ou dans les interactions proposées à l'utilisateur. L'exemple d'une interface graphique homogène, couplée à une interface vocale homogène, décrit plus haut illustre la nécessité de vérifier la cohérence entre modalités : cas d'une interface graphique qui afficherait une fonction "téléphone" en vert alors que l'interface vocale énoncerait "sélectionner la fonction *communiquer* en jaune".

2^{ème} conclusion

Pour conclure, les tests utilisateurs modalité par modalité, uniquement réalisables dans le cas où il est possible d'interagir intégralement avec le service quelque soit la modalité peuvent éventuellement constituer une aide. Dans notre étude, ils ont permis de mettre en lumière certains problèmes ergonomiques toujours valables pour les versions multimodales. En aucun cas, ces tests ne pourraient être suffisants étant donné l'importance de prendre en compte la coopération et la cohérence d'ensemble des modalités.

3. Méthode 2 : Elaboration de recommandations ergonomiques pour les interfaces multimodales à partir de l'ensemble des modalités

Une seconde méthode consiste à prendre en compte le service multimodal avec plusieurs modalités en global, et à élaborer les recommandations ergonomiques à partir de l'ensemble des modalités.

Nous avons abordé cette question selon plusieurs approches :

- Réaliser des tests utilisateurs de services multimodaux
- Reprendre les critères "classiques" d'ergonomie d'un point de vue multimodal

3.1 Recommandations multimodales élaborées "globalement" à partir de tests utilisateurs

Une première approche consiste à élaborer des recommandations ergonomiques pour les interfaces multimodales suite à l'analyse des résultats d'une expérimentation d'application multimodale avec des utilisateurs.

Actuellement, nous ne disposons pas d'une grande expérience "utilisateur" dans le domaine des interfaces multimodales, ni d'une grande expérience "ergonome" dans le domaine des tests d'utilisabilité de services multimodaux. C'est pourquoi il nous semble important de préciser que les tests utilisateurs de service multimodaux ne peuvent être réellement pertinents et maîtrisés qu'à certaines conditions :

- Le service multimodal doit être suffisamment robuste pour être testable par des utilisateurs en situation réaliste. Si cette condition ne peut pas être satisfaite, il est toujours possible d'utiliser la technique du Magicien d'Oz (simulation à l'aide de compères et à l'insu de l'utilisateur pour certaines parties de l'interface), pour étudier des points précis de l'interaction multimodale.
- Le service multimodal ne doit pas, autant que possible, proposer d'autres nouvelles technologies, sous peine de biaiser les résultats des tests. Par exemple, il n'est pas toujours possible alors de savoir si les résultats obtenus sont imputables à une technologie (la multimodalité), plutôt qu'une autre (la réalité virtuelle).

Ces conditions se trouvaient réunies dans le service multimodal "Plan Resto" et nous avons pu élaborer un certain nombre de recommandations ergonomiques pour les interfaces multimodales à partir de l'expérimentation utilisateur de ce service. Les résultats de cette étude sont présentés dans le livrable du SP5 Lot3 (L5.3).

Les recommandations élaborées recouvrent plusieurs aspects de la multimodalité :

- Des recommandations pouvant concerner une interaction dans une modalité particulière mais en contexte multimodal.

Par exemple :

Reco. *Prévoir un bouton push-to-talk (l'utilisateur maintient le bouton appuyé jusqu'à ce qu'il ait fini de parler) afin que l'utilisateur puisse contrôler le temps qui lui est imparti pour parler. Cette recommandation est d'ailleurs appuyée par d'autres études réalisées à France Télécom R&D.*

Ces recommandations sont propres à la multimodalité dans la mesure où elles pointent des dispositifs inhérents aux interfaces multimodales. L'exemple proposé ci-dessus traite en particulier de l'accès à une modalité à partir d'une autre modalité (ici, prise de parole à partir de l'interface graphique).

- Des recommandations qui concernent en particulier la multimodalité en entrée et en particulier en sortie.

Par exemple en entrée :

Reco. *Avoir une équivalence des modalités en termes de performances* et d'actions possibles. En effet, cela permettra à l'utilisateur de choisir la modalité avec laquelle il se sentira le plus à l'aise et cela permettra une utilisation plus aisée dans différents contextes, notamment en mobilité (forte luminosité, ambiance bruyante, etc.).*

** Il s'agit ici, autant que possible, de faire en sorte que les performances, en termes de temps de traitement de la requête par le système, soient à peu près équivalentes d'une modalité à l'autre.*

Par exemple en sortie :

Reco. *Favoriser la complémentarité modale plutôt que la redondance totale et réserver la redondance pour des messages d'alerte ou de haute importance par exemple.*

Reco. *Eviter la redondance totale pour les énoncés longs. En effet, cela risque d'entraîner une surcharge mentale pour les utilisateurs.*

Ces recommandations mettent en évidence l'importance de toujours resituer le service multimodal par rapport à ses types d'interaction avec l'utilisateur : en entrée pour désigner l'interaction de l'utilisateur vers le système et en sortie pour désigner l'interaction du système vers l'utilisateur.

- Des recommandations qui concernent en particulier la coopération entre modalités.

Par exemple :

Reco. *Choisir un ou plusieurs types de coopérations entre les modalités de façon pertinente. En effet, la redondance ne saurait se justifier pour tous les énoncés en sortie du système, elle est à privilégier pour les messages d'erreurs et/ou pour les feedbacks informant que sans une modification de la requête de l'utilisateur, celle-ci ne pourra pas aboutir.*

Reco. *Permettre à l'utilisateur de désactiver la synthèse vocale. Le système doit lui permettre d'avoir un contrôle permanent sur la synthèse vocale à l'aide d'une icône dédiée et d'une commande vocale (exemple : "couper le son !"). L'utilisateur doit également pouvoir faire une pause et reprendre à l'endroit où il avait arrêté la synthèse vocale. Il doit aussi pouvoir faire répéter le dernier énoncé de la synthèse vocale.*

Ces recommandations constituent le cœur de la problématique de la multimodalité d'un point de vue ergonomique et ne peuvent émerger, à l'heure actuelle, que suite à l'analyse de résultats de tests utilisateurs.

3^{ème} conclusion

Pour conclure, grâce aux tests utilisateurs nous avons pu élaborer des recommandations qui concernent purement la multimodalité. Nous avons également cerné des problèmes qui nous paraissent fondamentaux lorsque l'on étudie l'ergonomie des interfaces multimodales, comme la question de la coopération et de la cohérence entre les modalités. Par ailleurs, ces tests ont permis de pointer le fait

que les utilisateurs n'ont pas d'expérience de la multimodalité et qu'il est nécessaire de placer les tests utilisateurs dans une dimension plus large que la simple prise en main du service. Seule une approche longitudinale (usage inscrit dans la durée et dans un contexte réel d'utilisation) permettrait de recueillir des données plus riches et plus réalistes concernant l'usage de ce type d'interface. Un autre constat émanant de cette étude est que du point de vue de l'ergonomie, l'expérience des tests d'utilisabilité sur des interfaces multimodales est en cours de constitution.

3.2 Les critères d'ergonomie d'un point de vue multimodal

Etant donné l'expérience encore jeune de la multimodalité aussi bien pour les utilisateurs que pour les ergonomes, il apparaît utile de se questionner sur ce domaine. Une des voies possibles de cette réflexion est de reprendre les critères ergonomiques [Bastien-Scapin1993] qui constituent une base de questionnement intéressante car largement utilisés et reconnus en ergonomie.

Nous allons donc reprendre chacun de ces critères en se questionnant systématiquement sur leur pertinence, leur élargissement et leur transposition au domaine de la multimodalité :

3.2.1 Guidage

"L'ensemble des moyens mis en œuvre pour conseiller, orienter, informer et conduire l'utilisateur lors de ses interactions avec l'ordinateur."

Incitation

"Inciter l'utilisateur à effectuer des actions spécifiques en lui fournissant des indices. Par exemple, guider les entrées de données en indiquant le format adéquat et les valeurs acceptables: Date (jj/mm/aa) __ / __ / __."

Le potentiel offert par la multimodalité ouvre différentes possibilités d'incitation. En effet, une interface peut inciter l'utilisateur à effectuer certaines actions selon différentes modalités. Elles peuvent être exclusives : pour certaines actions, l'incitation se fera dans une modalité comme le vocal et pour d'autres actions, elle se fera dans une autre modalité comme le graphique.

La question reste de savoir comment présenter les informations de façon pertinentes, en choisissant (ou en laissant le choix à l'utilisateur) les modalités

adaptées aux types de tâches à réaliser, au contexte ou aux préférences de l'utilisateur.

Groupement/Distinction entre items

"Groupement des différents éléments visuels de façon cohérente et ordonnée."

"Groupement/Distinction par la Localisation : Positionner les items les uns par rapport aux autres afin d'indiquer leur appartenance, ou non, à une classe donnée d'objets. Par exemple, grouper les options de menus en fonction des objets sur lesquels elles s'appliquent."

"Groupement/Distinction par le Format : Donner aux éléments des caractéristiques graphiques particulières afin d'indiquer leur appartenance, ou non, à une classe donnée d'objets. Par exemple: utiliser un symbole et la couleur rouge pour les boîtes de dialogue d'alerte ou d'erreur."

Ce critère reste pertinent quand le service multimodal propose la modalité graphique.

Feed-back immédiat

"Dans tous les cas, l'ordinateur doit répondre à l'utilisateur en fonction des actions et des requêtes de ce dernier. Par exemple, dans les cas où les traitements sont longs, une information indiquant à l'utilisateur que les traitements sont en cours devrait lui être fournie."

Pour un service multimodal graphique/vocal (comme Plan resto) la particularité de la reconnaissance vocale nous conduit à préconiser un feedback textuel de ce que la reconnaissance vocale a pris en compte. Ceci vient du fait que cette modalité est jugée par les utilisateurs comme étant généralement moins fiable que le graphique.

Néanmoins, il est probable qu'une utilisation prolongée d'un service multimodal amène l'utilisateur à une certaine expertise de ce que la reconnaissance vocale peut comprendre, ainsi, l'utilisateur peut s'ajuster par rapport au système.

Il se peut également que les technologies en matière de reconnaissance vocale progressent encore de façon significative dans les prochaines années à venir et que le taux d'erreur soit tellement infime que cette modalité d'interaction sera alors jugée par les utilisateurs plus sûre qu'elle ne l'est actuellement.

Lisibilité

"Les caractéristiques lexicales de présentation des informations sur l'écran doivent faciliter la lecture de ces informations. Par exemple, il est préférable de présenter un texte avec quelques lignes longues plutôt que de nombreuses lignes courtes."

Ce critère reste pertinent quand le service multimodal propose la modalité graphique.

3.2.2 Charge de travail

"L'ensemble des éléments de l'interface qui a un rôle dans la réduction de la charge perceptive ou mnésique des utilisateurs, de même que dans l'augmentation de l'efficacité du dialogue."

Brièveté

"Limiter le travail de lecture, d'entrée et les étapes par lesquelles doivent passer les usagers."

"Concision : Réduire la charge de travail au niveau perceptif et mnésique pour ce qui est des éléments individuels d'entrée ou de sortie. Par exemple, lorsqu'une unité de mesure est associée à un champ de données, celle-ci doit faire partie du label du champ plutôt qu'être saisie par l'utilisateur."

"Action Minimales : Limiter les étapes par lesquelles doivent passer les utilisateurs. Par exemple, ne pas demander aux utilisateurs d'entrer des données qui peuvent être déduites par le système."

Ce critère est particulièrement fondamental dans le cadre de la multimodalité dont on attend qu'elle permette de faire mieux qu'une interface graphique.

La multimodalité peut théoriquement améliorer la brièveté, dans le sens où le vocal, notamment en langage naturel, peut permettre d'outrepasser une navigation parfois longue et fastidieuse. En effet, un des avantages de la modalité vocale est de pouvoir accéder directement à un item cible dans une longue liste ou dans une grande base de données, ou encore dans une interface où la profondeur des menus est grande.

En revanche, l'ajout de la modalité vocal dans une interface graphique peut aussi être une source d'erreurs supplémentaires faisant perdre en brièveté. La gestion de la coopération entre les modalités est cruciale pour satisfaire le critère de brièveté.

Densité informationnelle

"Réduire la charge de travail du point de vue perceptif et mnésique, pour des ensembles d'éléments et non pour des items. Par exemple, limiter la densité informationnelle de l'écran, en affichant seulement les informations nécessaires."

C'est également un des critères dont on peut attendre que la multimodalité puisse proposer un avantage certain par rapport à une interface graphique.

En effet, la multimodalité en sortie du système peut théoriquement permettre d'alléger la charge perceptive visuelle (par exemple : une partie de ce qui devait être présenté graphiquement peut être lu par une synthèse vocale) ou la charge perceptuelle auditive (par exemple : la possibilité d'outrepasser les limites du vocal en termes de nombres d'items maximum, grâce à l'affichage graphique).

Néanmoins, la multimodalité peut aussi augmenter la densité informationnelle, notamment lorsqu'il y a redondance totale entre ce qui est écrit et ce qui est lu par la synthèse vocale. La gestion des modalités en sortie (du système vers l'utilisateur) paraît essentielle pour satisfaire le critère de densité informationnelle.

3.2.3 Contrôle explicite

"Prise en compte par le système des actions explicites des utilisateurs et le contrôle qu'ont les utilisateurs sur le traitement de leurs actions."

Actions explicites

"Expliciter la relation entre le fonctionnement de l'application et les actions des utilisateurs. Par exemple, l'entrée de commandes doit se terminer par une indication de fin (« Enter », « OK ») à laquelle des possibilités d'édition doivent être préalables."

Ce critère constitue un enjeu pour la prise de parole dans une interface multimodale dont le fonctionnement n'est pas toujours clair selon le dispositif employé :

- push-to-talk

- appui-début/parole/appui-fin
- appui-début/fenêtre temporelle pour la parole/fin automatique
- reconnaissance permanente sans action spécifique de l'utilisateur

Contrôle utilisateur

"L'utilisateur doit pouvoir contrôler le déroulement des traitements informatiques en cours. Par exemple, autoriser l'utilisateur à interrompre tout traitement en cours."

Ce critère est tout à fait pertinent pour la multimodalité. Par exemple : Laisser la possibilité à l'utilisateur d'interrompre la synthèse vocale à tout moment.

3.2.4 Adaptabilité

"Capacité à réagir selon le contexte et selon les besoins et les préférences des utilisateurs."

Flexibilité

"Mettre à la disposition des utilisateurs des moyens pour personnaliser l'interface afin de rendre compte de leurs stratégies ou habitudes de travail et des exigences de la tâche. Par exemple, les utilisateurs doivent pouvoir désactiver des affichages inutiles."

Ce critère constituerait un des principaux avantages de la multimodalité, décrit notamment dans [Oviatt2002] ou encore [Reeves2004].

Cependant cette flexibilité "théorique" - avoir plusieurs modalités permet de changer de modalité d'interaction quand : erreur, changement de contexte, préférences utilisateurs... - se doit d'être accompagné d'une certaine fluidité dans l'interaction. Il ne faut pas, a priori, que le changement de modalité soit plus coûteux que de rester sur la modalité initiale.

Prise en compte de l'expérience de l'utilisateur

"Le système doit respecter le niveau d'expérience de l'utilisateur. Par exemple, prévoir des choix d'entrées pas-à-pas ou multiples selon l'expérience des utilisateurs."

Ce critère reste tout à fait pertinent pour la multimodalité. D'autant plus que les interfaces multimodales proposent parfois des modalités peu communes (haptique, reconnaissance de geste ou ne serait-ce que le fait de parler à une machine...).

3.2.5 Gestion des erreurs

"Moyens permettant d'une part d'éviter ou de réduire les erreurs, d'autre part de les corriger lorsqu'elles surviennent."

Protection contre les erreurs

"Mettre en place des moyens pour détecter et prévenir les erreurs. Par exemple, toutes les actions possibles sur une interface doivent être envisagées et plus particulièrement les appuis accidentels des touches du clavier afin que les entrées non-attendues soient détectées."

Ce critère reste pertinent pour la multimodalité et l'est d'autant plus qu'il y a plusieurs modalités à disposition de l'utilisateur, donc autant d'erreurs potentielles à prévoir et à anticiper.

Qualité des messages d'erreur

"S'assurer que l'information donnée aux utilisateurs sur la nature des erreurs commises (syntaxe, format, etc.) et sur les actions à entreprendre pour les corriger, soit pertinente, facile à lire et exacte. Par exemple, utiliser un vocabulaire neutre, non-personnalisé, non réprobateur dans les messages d'erreurs; éviter l'humour."

Ce critère reste pertinent quand le service multimodal propose la modalité graphique et la modalité vocale.

Correction des erreurs

"Mettre à la disposition des utilisateurs des moyens pour corriger leurs erreurs. Par exemple, fournir la possibilité de modifier les commandes lors de leur saisie."

Ce critère constitue un des avantages attendus de la multimodalité, dans la mesure il est possible de récupérer une erreur réalisée dans une modalité sur une autre modalité.

Cependant, ce point précis correspond à l'un des dix mythes concernant la multimodalité qu'Oviatt a relevé dans un de ses articles [Oviatt1999] : *"Les erreurs sur un mode sont compensées par un autre mode"*. Ainsi, les insuffisances d'un mode (par exemple les erreurs de reconnaissance de la parole) ne sont pas toujours compensées par un autre mode, souvent, les erreurs se cumulent d'un mode à l'autre.

3.2.6 Homogénéité/Cohérence

"Les choix de conception d'interface doivent être conservés pour des contextes identiques, et doivent être différents pour des contextes différents. Par exemple, toujours afficher au même endroit l'incitation pour la saisie des données ou des commandes."

Ce critère reste pertinent pour la multimodalité quelque soit les modalités implémentées.

3.2.7 Signifiante des codes et dénominations

"Il doit y avoir adéquation entre l'objet ou l'information affichée ou entrée, et son référent. Par exemple, rendre les règles d'abréviation explicites."

Ce critère reste pertinent pour la multimodalité.

Concernant la dénomination :

En entrée (de l'utilisateur vers le système), quand des commandes sont réalisables en graphique (bouton) ou en vocal (en prononçant le nom de la commande), la reconnaissance vocale doit au moins accepter le libellé exact de la commande qui est affiché à l'écran, sachant qu'elle peut également accepter des synonymes de cette même commande.

En sortie, les commandes présentées à la fois vocalement et graphiquement doivent être les mêmes et présentées dans le même ordre, d'une modalité à l'autre.

3.2.8 Compatibilité

"Il faut qu'il y ait accord entre les caractéristiques des utilisateurs et des tâches, d'une part, et l'organisation des sorties, des entrées et du dialogue d'une application donnée, d'autre part. Par exemple, les termes employés doivent être familiers aux utilisateurs, et relatifs à la tâche à réaliser."

Ce critère représente un des enjeux de la multimodalité : il est attendu de la multimodalité qu'elle permette un meilleur ajustement entre l'utilisateur et le système. En effet, la multimodalité peut présenter "théoriquement" l'avantage de s'adapter à l'utilisateur selon son contexte, selon ses préférences ou encore selon la nature de la tâche qu'il a à réaliser.

C'est pourquoi un des enjeux restant est de satisfaire ce critère de façon pertinente : doit-on laisser un contrôle total à l'utilisateur sur les modalités qu'il peut utiliser ? Le système doit-il prendre en compte automatiquement le contexte de l'utilisateur (notamment à l'aide de capteur) et adapter les modalités disponibles en fonction de ce contexte ? Ou doit-on adopter une position intermédiaire entre ces deux extrêmes pour la conception ?

4^{ème} conclusion

Pour conclure, l'idée de reprendre les critères ergonomiques comme base de questionnement sur leur pertinence, leur élargissement et leur transposition au domaine de la multimodalité s'avère intéressante pour amener des éléments de réflexion concernant la conception et les tests des services multimodaux. Précisons que ce questionnement paraît difficile sans un minimum d'expérience "ergonome" sur la multimodalité et qu'enfin, il n'est pas figé et évoluera probablement avec les données disponibles dans l'avenir sur ce domaine.

4. Comparaison des deux méthodes et discussion

Le domaine de la multimodalité est encore très jeune du point de vue de l'expérience utilisateur comme du point de vue de l'expérience "ergonome". La question de la comparaison des deux méthodologies se positionne dans le contexte de la constitution de cette expérience.

La première méthode qui consiste à identifier des recommandations ergonomiques modalité par modalité peut constituer une aide pour les conceptions des services multimodaux. En particulier, pour reprendre les recommandations qui touchent les aspects perceptifs de l'interaction pour chaque modalité, pour extrapoler les recommandations d'une modalité à une autre et pour mettre en évidence certains choix à faire au niveau de la coopération entre modalités. Cependant, les hypothèses de recommandations ainsi élaborées doivent être examinées et ajustées selon le type de coopération entre modalités et vérifiées par des utilisateurs. Autrement dit, cette méthode ne saurait suffire pour élaborer des recommandations ergonomiques pour les interfaces multimodales.

La deuxième méthode qui consiste à élaborer des recommandations ergonomiques en considérant directement l'ensemble des modalités est indispensable à la conception et à l'évaluation des services multimodaux. Cependant, cette méthode se heurte à un certain nombre de difficultés : en premier lieu parce que les services multimodaux ne répondent pas tous aux conditions requises pour des tests utilisateurs réellement faisables et valides (cf. section 3.1 de ce document) ; en second lieu, parce qu'il apparaît nécessaire de placer l'utilisateur dans une interaction inscrite dans la durée et plus large que la simple prise en main d'un service ; et en troisième lieu, parce qu'on ne dispose pas encore suffisamment de recul sur les tests d'utilisabilité et sur la façon dont les critères d'ergonomie peuvent se développer du point de vue de la multimodalité.

La comparaison de ces deux méthodes nous amène à établir que la deuxième est indispensable pour l'élaboration de futures recommandations ergonomiques pour les interfaces multimodales, tandis que la première constitue plutôt une aide à la réflexion.

Enfin, au regard de notre expérience "ergonome" dans le domaine de la multimodalité, on peut légitimement penser que les futurs tests utilisateurs devront être envisagés sous un angle longitudinal et en prenant soin de s'assurer au préalable que les services testés seront motivants et auront du sens pour l'utilisateur. Une fois l'utilité du service vérifiée, il est possible de mieux s'interroger sur l'utilité de la multimodalité, dans un contexte réel d'utilisation. Il conviendra également de s'interroger plus précisément sur la nature des tâches demandées et des contenus présentés aux sujets afin de mieux répondre aux questions suivantes : quelles modalités pour quels types de tâches ? et quelles modalités pour quels contenus ?

5. Bibliographie

- [Bastien-Scapin1993] Bastien Christian JM and Scapin Dominique, Ergonomic criteria for the evaluation of human-computer interfaces, rapport de recherche INRIA, mai 93.
- [Oviatt1999] Oviatt, S.L. (1999). *Ten myths of multimodal interaction*, Communication of the ACM, Vol. 42, No. 11, November, 1999, pp. 74-81.
- [Oviatt2002] Oviatt, S.L. (2002). *Multimodal interfaces*. In the Human-Computer Interaction Handbook : Fundamentals, Evolving Technologies and Emerging Applications, J. JACKO AND A. SEARS, Eds. Lawrence Erlbaum Assoc., Mahwah, NJ, 2002, chap.14, 286-304.
- [Reeves2004] Reeves, L.M., Lai, J., Larson, J.A., Oviatt, S., Balaji, T.S., Buisine, S., Collings, P., Cohen, P., Kraal, B., Martin, J.C., McTear, M., Raman, T.V., Stanney, K.M., Su, H., Wang, Q.Y. (2004). **Guidelines for multimodal user interface design**. *Communications of the ACM – Special Issue on Multimodal Interfaces*, vol. 47(1), pp. 57-59.