

Evaluation de la sonification d'un système multimédia automobile

Sabine Langlois, Séverine Loiseau

Renault / Ergonomie cognitive
1, av du Golf
78288 Guyancourt, France
sabine.langlois@renault.com
severine.loiseau@renault.com

Julien Tardieu, Andrea Cera, Nicolas Misdariis

Ircam-CNRS STMS
1, place Igor Stravinsky
75004 Paris, France
j1tardieu@gmail.com, andreawax@yahoo.it
nicolas.misdariis@ircam.fr

RESUME

Equiper un système multimédia embarqué sur véhicule de guidage sonore doit permettre au conducteur de manipuler son système « sans les yeux », en restant concentré sur la route. Le guidage est en effet censé faciliter la découverte du système et la navigation dans l'architecture du menu.

Une maquette a été réalisée puis évaluée auprès d'utilisateurs afin d'éprouver l'efficacité et l'acceptabilité du guidage sonore mis en place (sons hiérarchiques, synthèse vocale à vitesse normale ou accélérée sous forme de spearcons, extraits musicaux, sons de confirmation d'action). Les résultats obtenus confirment l'intérêt du guidage sonore. Bien que la sonification du système ne permette pas de naviguer et d'accéder plus rapidement aux données, les participants souhaitent la conserver car ils estiment leur conduite plus sûre. Le guidage sonore leur donne d'ailleurs l'impression de devoir moins souvent détourner le regard de la route.

Les sonifications les plus utiles sont : la synthèse vocale et les sons hiérarchiques dans les premiers niveaux du menu, ainsi que les lettres dans les listes longues des sous menus. Les sons hybrides entre earcons et icônes auditives sont préférés aux sons orchestrés (musèmes) car le lien à la fonction représentée est plus intuitif. Les spearcons sont rejetés : incompréhensibles, ils ne permettent pas à l'utilisateur de se repérer dans les listes de données. A toutes les étapes de la recherche, les sons de défilement et de butée/liste vide sont appréciés.

MOTS CLES : IHM sonores, design sonore, spearcons, synthèse vocale, évaluation ergonomique.

ABSTRACT

An automotive multimedia system that would be sonified should offer to the driver the possibility to interact with this system without his eyes, having thus a better visual focus on the road. The sonification should indeed facilitate the discovery of the system and the navigation within the menu tree.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

IHM 2010, 20-23 Septembre 2010, Luxembourg, Luxembourg

We have prepared a prototype to test with users the efficiency and acceptability of the sonification, made by hierarchical sounds, vocal synthesis, spearcons (compressed vocal synthesis), musical samples and feedback sounds. Our results confirm that sonification is interesting. Even if it doesn't allow to navigate and to reach the data faster, the participants want to keep it as they find their driving safer.

Vocal synthesis and hierarchical sounds used on the first menu levels are the most helpful sounds with the alphabetical letters used in the long list of data. Hybrid sounds, composed of earcons and auditory icons, are preferred to orchestra sounds (musems) because the link to the fonction is more intuitive. Spearcons are not appreciated because they are not understood and do not help the positioning in the lists of data. Feedback sounds are appreciated all along the navigation, in particular the sound of scrolling and the sound « end of the list / empty list ».

CATEGORIES AND SUBJECT DESCRIPTORS: H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

GENERAL TERMS: Design, Human factors.

KEYWORDS: sonic human-machine interface, sound design, spearcons, vocal synthesis, ergonomics, experimentation.

INTRODUCTION

Renault a souhaité évaluer l'intérêt de l'utilisation du son pour faciliter la navigation et la recherche de contenu dans les grandes bases de données des systèmes multimédia embarqués sur véhicule. On s'intéresse ici aux systèmes pilotés par une commande à distance, et présentés sous forme d'arborescence hiérarchique.

Quatre catégories de retours sonores ont été définies et ont fait l'objet d'un maquetage puis d'une évaluation auprès d'utilisateurs. Ces retours sonores ont pour objectif de faciliter l'apprentissage de la structure d'un menu, le repérage du positionnement et l'identification des items dans les longues listes de données, et de confirmer les actions réalisées par le conducteur, en limitant le recours à l'attention visuelle.

Ce document décrit le guidage sonore défini et maqueté, la méthodologie adoptée pour évaluer son efficacité et

son acceptabilité ainsi que les résultats de cette évaluation et les perspectives du projet.

DESCRIPTION DU GUIDAGE SONORE

Nous avons réalisé une maquette interactive, pilotable au moyen d'un pad gestuel. Elle permet d'accéder en théo-

rie à différents menus d'un système multimédia : Communication, Navigation (au sens de fonction GPS), Véhicule (au sens d'informations sur le véhicule), Musique, Images et Vidéos. Dans le cadre de l'étude, seuls les menus Communication et Musique ont été développés, et rendus disponibles. L'arborescence est la suivante :

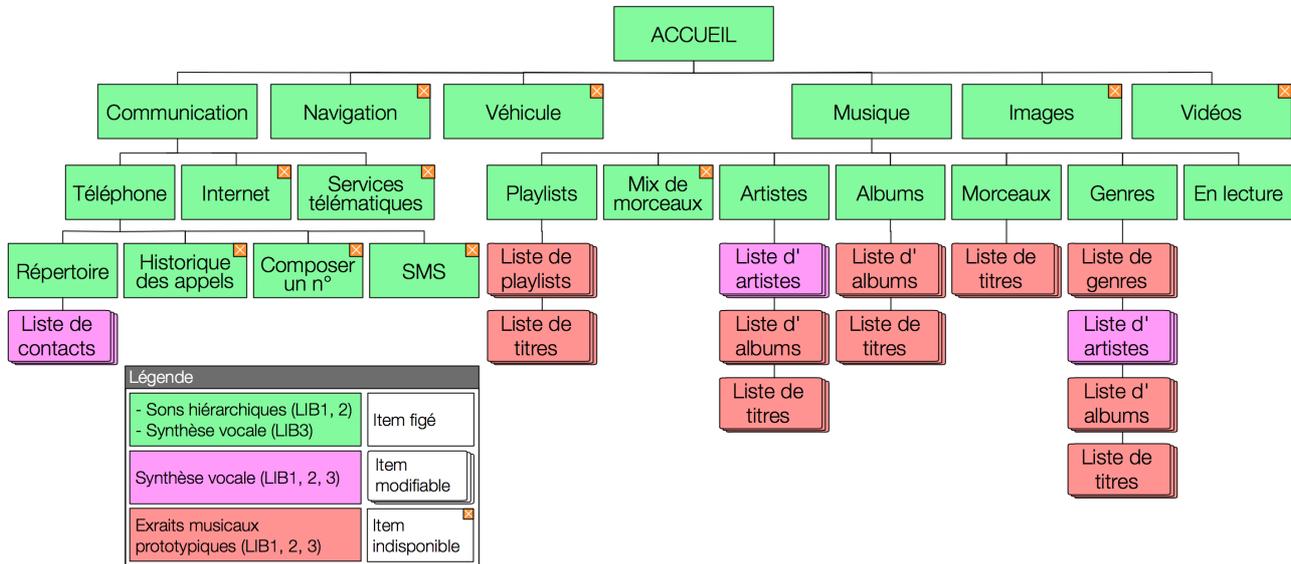


Figure 1 : Arborescence des menus et principe d'utilisation des 3 librairies (LIB1, 2, 3) de sons suivant ces menus

Trois librairies de sons sont activables sur la maquette. Nous les nommerons dans l'article LIB1, LIB2 et LIB3. Quatre catégories de sons sont utilisés dans les 3 librairies : (1) sons de confirmation d'action, (2) sons hiérarchiques (sons associés aux items du menu), (3) synthèse vocale et (4) échantillons prototypiques (extrait musical représentatif). Les sons de synthèse vocale ont trois variantes : une simple lettre (« Lettre »), le groupe de mots complet (par exemple « Services télématiques ») dit à vitesse normale (« Synthèse vocale »), ou dit à vitesse accélérée (« Spearcon »). Les 4 catégories de sons sont utilisées dans les 3 librairies comme indiqué dans la légende de la figure 1. Nous entendons par « item modifiable », un item que l'utilisateur de la version série de l'interface pourrait modifier. Ces items sont organisés en liste, dont la longueur nous a incités à prévoir les trois vitesses d'exploration évoquées ci-dessus : normale, accélérée (seulement dans les menus sonifiés avec la synthèse vocale) et rapide, sonifiée par la simple lettre dans toutes les listes.

Les sons de confirmation d'action ainsi que les sons hiérarchiques ont été créés dans le cadre d'une collaboration avec l'équipe Perception et Design Sonores de l'Ircam, en association avec le compositeur/designer sonore A. Cera.

Intégrée dans une démarche scientifique, la phase de création a été précédée d'une étude bibliographique qui a permis de faire un état de l'art des travaux relatifs à la sonification des menus hiérarchiques. D'un point de vue fonctionnel, ce problème soulève deux questions fondamentales par rapport à la modalité auditive : la représentation du niveau hiérarchique de l'arborescence et la représentation du contenu des items du menu. L'étude de la

littérature permet de distinguer trois principales approches :

- « *earcon* », son synthétique pour lesquels la relation son/signification est arbitraire et doit donc être apprise par l'utilisateur ; cette notion a été introduite par Blattner [1] puis, par la suite, étudiée et expérimentée notamment par Brewster [2], Leplâtre [3] et Helle [4]. L'intérêt des earcons est de pouvoir facilement représenter une structure hiérarchique en modulant les caractéristiques sonores (hauteur, timbre, rythme, durée, etc.).

- « *icône sonore* » (en anglais, « *auditory icon* »), introduite par Gaver [5], qui établit un lien direct ou métaphorique avec l'objet ou le concept qu'il représente, en faisant référence à un son de notre environnement, et donc de fait, quasi-directement compris par l'utilisateur. On peut trouver une mise en oeuvre de cette approche notamment dans les travaux de Barrass [6] ou Conversy [7].

- « *spearcon* » (contraction de « *speech* » et « *earcon* »), basé sur une synthèse vocale compressée dans le temps tout en maintenant une hauteur constante. La reconnaissance du mot ou groupe de mots n'est pas nécessairement assurée, mais le spearcon joue le rôle d'une empreinte univoque. Cette notion a été introduite par Walker [8] afin de pallier les limitations d'une sonification reposant sur des earcons ou des icônes auditives (difficulté de création et d'apprentissage, en particulier pour les earcons). Les travaux de Walker, entre autres ceux menés avec Palladino [9] sur un menu de téléphone portable mettent en avant les performances d'apprentissage de spearcons (identiques à celles de la synthèse vocale ; 1,1 présentation suffit), comparées à celles des earcons

(6,5 présentations sont nécessaires). Dingler [10] ajoute les icônes auditives à la comparaison, pour des catégories d'objets : il faut 8,5 cycles d'apprentissage pour les earcons, 3 pour les icônes auditives et 1,1 pour les spearcons ou le vocal.

L'analyse comparée de ces différentes approches ainsi que le respect du cahier des charges de l'étude ont permis d'établir une stratégie de design sonore. Pour la navigation 3 principales fonctions ont été identifiées :

- la différenciation des menus multimédia : chaque menu est associé à une famille de sons qui doit permettre une bonne distinction entre les menus ;

- le défilement vertical : le son associé à un item doit indiquer son niveau de profondeur dans la structure (par exemple, indiquer que l'item « Répertoire » se trouve dans le sous-menu « Téléphone » sur menu « Communication ») ; - le défilement horizontal : au sein d'un même niveau hiérarchique, les sons associés aux items doivent indiquer l'avancement dans la liste (par exemple, indiquer que « Répertoire » est le premier item de la liste du menu « Communication », tandis que « SMS » est le quatrième).

Pour l'ensemble des sons hiérarchiques, des solutions ont été proposées suivant deux directions.

D'une part, une approche conventionnelle, mélangeant les concepts d'earcons et d'icônes auditives (LIB1), chacun étant assigné à des fonctions spécifiques : des icônes sonores pour la différenciation, des earcons issus d'une synthèse sonore paramétrable par modélisation physique pour les défilements vertical et horizontal, avec en plus un procédé d'hybridation de ces 2 composantes qui favorise l'homogénéité du rendu. Nous appellerons par la suite cette librairie « hybride ».

D'autre part, une approche plus originale basée sur la notion de « musème » (LIB2), issue des travaux d'analyse musicologique de Philip Tagg [11] et qui renvoie à l'idée d'une unité de sens exprimée dans un contexte musical. Pour le compositeur, il s'agit ici de représenter de manière métaphorique – grâce à un motif musical – chaque menu principal (fonction différenciation), la durée et la nature mélodique de ce motif étant ensuite modifiées afin d'assurer respectivement les fonctions de défilement vertical et horizontal et de couvrir ainsi la surface complète de l'arborescence.

1. Sons de confirmation d'action (« feedback sounds »)

Ils servent à confirmer par le son l'action réalisée par l'utilisateur afin de lui permettre de savoir que son action est ou non correcte et prise en compte par le système. Trois sons de confirmation d'action sont proposés. Ils sont identiques pour les 3 librairies.

Son « butée » : ce son est joué dans les listes d'items pour indiquer à l'utilisateur que le focus (curseur) qui se déplace à l'écran sur les items de la liste est parvenu en butée.

Son « défilement » : ce son est joué dans les listes d'items qui n'utilisent pas de sons hiérarchiques (dans

les listes qui utilisent des sons hiérarchiques, aucun son de défilement n'est joué car le son hiérarchique a été créé de façon à avoir une attaque franche pour bien marquer le passage d'un item à un autre).

Son « liste vide » : ce son d'erreur est joué dans les listes d'items lorsque l'utilisateur sélectionne (action « ok ») un item qui n'ouvre aucun écran. Il s'agit du même son que le son « butée ». En complément de LIB1, 2 et 3, la librairie LIB4 est constituée de ces seuls sons de confirmation d'action, car ce sont ceux habituellement présents sur les systèmes multimédia usuels tels que téléphones portables. Nous appellerons par la suite cette librairie « Guidage sonore Off » ou « Sans son ».

2. Sons hiérarchiques

Ces sons sont associés aux différents items de menu de LIB1 et LIB2 afin de renseigner l'utilisateur sur l'univers fonctionnel visité (Communication, Navigation, Véhicule ...) et la profondeur verticale et horizontale de la structure du menu. L'objectif de ces sons est de faciliter l'apprentissage du contenu et de la structure du menu afin de permettre une navigation plus rapide dans le système.

Lorsqu'une liste de menu est ouverte, le son hiérarchique associé à l'item sur lequel est positionné le focus est joué. Lorsque l'utilisateur déplace le focus dans cette liste, le son hiérarchique de l'item atteint est joué. De plus, si l'utilisateur revient à la liste du niveau de menu précédent (action « Retour simple»), le focus se place automatiquement sur l'item précédemment sélectionné et le son associé à cet item est joué.

3. Spearcons et synthèse vocale

L'intérêt de la synthèse vocale est de ne nécessiter aucun apprentissage. Dans la LIB3, la synthèse vocale remplace les sons hiérarchiques des librairies 1 et 2 (cf. Légende - Figure 1). Dans les 3 librairies, nous utilisons la synthèse vocale dans les listes Répertoire et Artistes afin de faciliter le repérage dans les listes nominatives.

D'un point de vue pratique, la synthèse vocale est réalisée avec voix féminine française Alice de la synthèse TTS Acapela. Elle est précalculée dans le cadre de l'expérimentation, sachant que, sur véhicule de série, elle serait générée à la volée pour s'adapter à la nature dynamique du contenu de certains sous-menus (par exemple, la « Liste d'albums » du menu Musique). A partir de cette synthèse, nous avons produit les spearcons grâce au logiciel Audacity, avec une compression de tempo égale à 30% (<http://audacity.sourceforge.net/>). Un silence de 0,2 s. sépare le spearcon de la synthèse vocale.

4. Extraits musicaux prototypiques

Le but est de faciliter la recherche de données musicales aux utilisateurs qui souvent ne connaissent pas les titres des morceaux et des albums disponibles dans leur librairie musicale. La durée des échantillons varie de 2 à 5 secondes. Pour les listes Playlists, Genres et Albums, les échantillons sont construits à partir d'un morceau représentatif de la playlist, du genre et de l'album survolé. Pour la liste Morceaux, les échantillons sont construits à partir d'un moment caractéristique du morceau. Ces extraits contiennent de la voix si le morceau est chanté, et, si possible, le moment où le titre est chanté.

EVALUATION

L'évaluation répond à 3 objectifs : vérifier, hors situation de conduite, la qualité des 2 librairies de sons hiérarchiques ; vérifier, en situation de conduite simulée, l'acceptabilité des sons ainsi que leur efficacité ; recueillir l'avis des participants sur les différents sons du guidage sonore.

Nos hypothèses à tester sont les suivantes : (1) la sonification permet de réaliser plus rapidement les tâches de navigation dans le système multimédia, tout en conduisant ; (2) elle limite des détournements de regard hors de la scène de conduite ; (2 bis) elle permet une meilleure focalisation de l'attention sur la conduite ; (3) les sons hiérarchiques sont plus acceptables sur le long terme que la synthèse vocale ; (4) les sons hiérarchiques de type musème sont plus acceptables sur le long terme que ceux de type hybride, car ils font moins directement référence à des sons environnementaux.

1. Protocole d'évaluation

Participants. 24 personnes ont participé à l'évaluation du système (11 F., 13 H., entre 21 et 48 ans, tous salariés de Renault).

Installation. L'évaluation de la qualité des librairies hiérarchiques et le recueil de l'avis de participants sur le guidage sonore a été réalisé sur table. L'évaluation de l'acceptabilité et de l'efficacité des sons a été elle réalisée sur un simulateur afin de placer les participants en situation de conduite (roulage sur voie rapide).

Déroulement général. L'évaluation s'est déroulée en 7 étapes : {1} une présentation du test ; {2} une phase d'évocation des 2 librairies hiérarchiques (LIB1, 2) ; {3} une évaluation, à l'aide de la maquette (présentation sonore et visuelle), de la cohérence entre les sons hiérarchiques et les fonctions associées ; {4} un choix entre LIB1 et LIB2 ; {5} une familiarisation avec la maquette (+ pad gestuel) associée à la librairie hiérarchique choisie dans la phase précédente ; {6} la réalisation de 6 tâches courantes (e.g, appeler un correspondant) sur la maquette, en situation de conduite simulée, suivant 3 conditions expérimentales : avec la librairie choisie à l'étape 4 (LIB1 ou 2), avec la librairie vocale (LIB3), et avec le guidage sonore Off à l'exception des sons « défilement », « liste vide » et « butée » (LIB4) ; {7} un questionnaire permettant aux participants de verbaliser librement sur leurs préférences et leurs stratégies de jugement.

La durée totale du test est d'environ 1h30 ; l'étape 6 dure en moyenne 30 minutes (10 minutes par condition).

RESULTATS

1. Qualité des librairies hybrides et musèmes

Phase d'évocation libre (étape 2). Dans la librairie hybride, seul le son Téléphone est correctement associé à la fonction par la majorité des participants (74%). Le son Images est compris par 43% des participants, tandis que le son Communication évoque des fonctions telles que Téléphone et Internet liées au même univers fonctionnel Communication. Les autres sons sont eux rarement associés à la bonne fonction, et le son Navigation n'est identifié par aucun des participants. La majorité des sons évoquent des fonctions diverses. En revanche, près de la

moitié des participants associe le son Musique à un son d'erreur (type Windows).

Dans la librairie musème, aucun son n'est correctement associé à la fonction qu'il représente. Le son Musique, le mieux identifié, l'est par 30% des participants. La plupart des sons sont associés spontanément à des sons d'alerte ou d'erreur.

Cohérence son-fonction (étape 3). Cette étape correspond davantage à la manière dont les utilisateurs découvrirait le système en situation réelle que l'étape d'évocation libre. Les sons sont ici globalement jugés cohérents avec la fonction qu'ils représentent, en particulier les sons hybrides Communication, Véhicule et Images et les sons musèmes Musique et Vidéos. En revanche, dans la librairie hybride, la moitié des participants considèrent que le son Musique n'est pas du tout adapté à la fonction. Dans la librairie musème, ce sont les sons Images et Communication qui sont considérés comme non adaptés à la fonction par la moitié des participants. Le tableau ci-dessous reprend ces résultats :

Librairie hybride (LIB1)	Librairie musème (LIB2)
☺ Communication (63%)	☺ Musique (54%)
☺ Véhicule (42%)	☺ Vidéos (54%)
☺ Images (38%)	
☹ Navigation (☺ 25% / ☹ 33%)	☹ Navigation (☺ 21% / ☹ 33%)
☹ Vidéos (☺ 21% / ☹ 21%)	☹ Véhicule (☺ 25% / ☹ 33%)
☹ Musique (63%)	☹ Images (63%)
	☹ Communication (50%)

Tableau 2 : Evaluation de la cohérence son/fonction

Continuité du thème sonore dans les menus Communication et Musique (étape 3). La continuité du thème sonore est jugée satisfaisante par la grande majorité des participants dans les menus Musique (83%), Communication (96% entre Communication et ses items, 92% entre Téléphone et ses items) de la librairie musème.

En revanche, dans la librairie hybride, cette continuité du thème sonore n'est jugée satisfaisante par la majorité des participants que dans le menu Musique (96%). Si le 1^{er} niveau du menu Communication satisfait quant à lui 96% des participants, 42% des participants trouvent que le sous-menu Téléphone n'est pas homogène.

Dans les commentaires généraux formulés par les participants, il est intéressant de souligner que certains déplorent la perte du caractère évocateur des sons du dernier niveau de menu, impliquant une mémorisation difficile. Néanmoins, 29% des participants soulignent et apprécient que la profondeur du menu soit donnée par la simplification des sons à mesure que l'on s'enfonce dans la structure.

Préférence entre les librairies hybride et musème (étape 4). La majorité des participants (79%) disent préférer la librairie hybride à la librairie musème (13%). La librairie hybride est préférée car elle est jugée plus explicite, suggestive, intuitive, ce qui facilite la mémorisation, l'apprentissage et la familiarisation avec le système. Notons que 8% des participants aimeraient toutefois faire un mix entre ce qu'ils ont préféré dans les deux librairies, retenant les sons Musique et Vidéos de la librairie musème et les sons restants de la librairie hybride.

2. Efficacité du guidage sonore (étape 6)

Les 3 variables recueillies lors de cette étape sont : la durée chronométrée de réalisation de la tâche, le niveau de confort de conduite ressenti (échelle en 6 points allant de « très inconfortable » à « très confortable »), la fréquence de détournement de regard de la route ressentie (évaluée directement par le participant sur une échelle en 4 points allant de « rarement » à « très souvent »).

Des analyses de la variance, avec un intervalle de confiance fixé à 95%, nous permettent d'étudier l'influence qu'ont sur ces 3 variables numériques dépendantes les variables explicatives catégorisées suivantes : « condition de sonification » (3 conditions : librairie vocale, librairie hiérarchique et sons OFF – sauf sons de confirmation d'action) et « tâches » (T1 : Appeler le coiffeur, T2 : Appeler Pizza Rapido, T3 : Lancer le morceau Divine idylle, T4 : Lancer le morceau Superstition, T5 : Lancer l'artiste Gare du Nord, T6 : Lancer l'artiste Ray Charles). Il s'agit d'un plan expérimental à mesures répétées, dont les variables inter sujets identifiées sont le genre et la classe d'âge. Des tests de Fisher ont permis de comparer les différentes modalités des variables (par exemple : sons hiérarchiques vs. vocal vs. sons OFF pour la variable « condition de sonification »).

L'analyse réalisée porte sur la dernière des 3 conditions testées par les participants afin de s'affranchir de l'influence de l'ordre de présentation et de se concentrer sur la situation où le sujet a essayé les divers guidages proposés par le système (dans un ordre contrebalancé entre les sujets) et a désormais bien en main le système. L'étude a été effectuée sur 22 des 24 participants (un participant n'a pu conduire sur simulateur, un autre avait une mauvaise vision).

Effet du guidage sonore sur la durée des tâches. Pour les 3 conditions de guidage, le graphe ci-dessous indique la moyenne et l'écart-type de la durée de chaque tâche, ainsi que pour l'ensemble des tâches (« T1-T6 »).

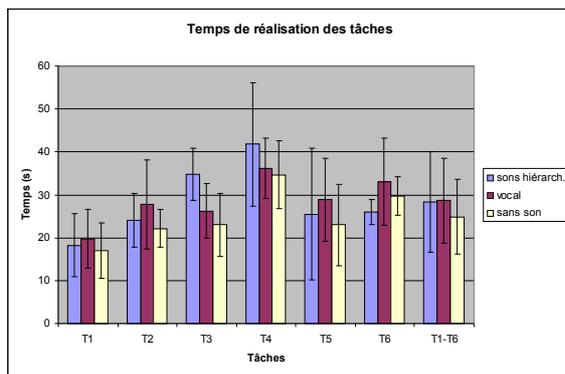


Figure 2 : Temps de réalisation des tâches (moyenne +/- un écart type), par tâche (« T1 » à « T6 »), et en moyenne (« T1-T6 »).

Un effet tout juste significatif du guidage sonore sur la durée de réalisation des tâches est constaté ($p=0,047$). Le test de Fisher indique que la seule différence porte sur le vocal plus long que sons OFF ($p=0,02$). Les sons hiérarchiques sont similaires au vocal ($p=0,9$) et à sons OFF ($p=0,08$).

Si notre étude porte plus particulièrement sur l'influence de la condition de sonification, il est néanmoins intéressant d'examiner l'influence des autres variables. Celle de la tâche est significative ($p<0,0001$). La tâche 1 (appeler coiffeur) se révèle significativement plus rapide que les autres tâches ($p=0,003$ pour T1 vs. T2), et T4 (lancer le morceau Superstition) significativement plus longue ($p=0,009$ pour T5 vs. T8). La rapidité de réalisation de la tâche 1 peut s'expliquer par le fait que le chemin d'accès à la donnée recherchée (coiffeur) est le plus court des chemins d'accès parcourus.

Le genre n'a pas d'influence ($p=0,4$), en revanche la classe d'âge en a ($p<0,0001$), ce qui est une des explications de l'importance des écarts type.

Effet du guidage sonore sur le confort perçu. La condition de sonification a une influence significative ($p<0,0001$) sur le niveau de confort de conduite perçu.

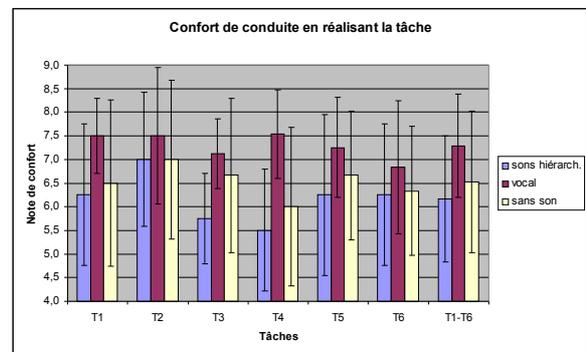


Figure 3 : Confort perçu (moyenne +/- un écart type)

Les participants jugent en effet leur conduite plus confortable avec la librairie vocale que lorsque le guidage sonore est désactivé ($p<0,0001$). En revanche, aucun effet significatif de la librairie hiérarchique n'est observé sur la perception de confort de conduite : le confort de conduite ressenti est similaire entre les conditions librairie hiérarchique et Sons OFF ($p=0,2$), tandis que $p=0,002$ pour sons hiérarchiques vs. vocal).

Il n'y a ici pas d'influence de la variable « tâche » ($p=0,5$). Le genre reste non influent ($p=0,6$), et inversement pour l'âge ($p<0,0001$).

Effet du guidage sonore sur le détournement du regard perçu. La condition de sonification influence la perception de détournement de regard ($p<0,0001$). Les participants ont l'impression de détourner plus souvent leur regard de la scène de conduite lorsque le guidage sonore est coupé que lorsque le guidage vocal est activé ($p<0,0001$).

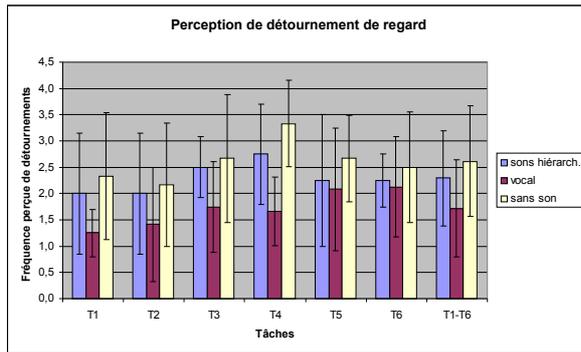


Figure 4 : Détournement de regard perçu (moy. +/- un écart type)

En revanche, aucune différence significative n'est observée entre Sons OFF et la condition hiérarchique ($p=0,2$, tandis que $p=0,006$ pour les sons hiérarchiques vs. vocal). La tâche a un effet significatif ($p=0,04$) : les tâches de téléphonie donnent aux participants l'impression de moins requérir de détournements de regard tandis que les tâches T5, T7 et T8 ($p=0,05$ pour T2 vs. T8).

L'influence des facteurs inter sujets est similaire à celle observée sur les deux précédentes variables dépendantes ($p=0,01$ pour l'âge, $p=0,7$ pour le genre).

Commentaires associés au confort de conduite et aux détournements de regard. Lors de l'étape 6, les participants ont expliqué les notes de confort et de détournement de regard attribuées après chaque tâche en invoquant plusieurs éléments du guidage sonore.

90% des participants soulignent leur intérêt pour les lettres alphabétiques associées au défilement lettre à lettre dans les listes. Ces lettres sont jugées utiles et limitent les détournements de regard, au profit d'une conduite plus sûre.

71% des participants disent préférer utiliser l'interface visuelle plutôt que les extraits musicaux. Ces derniers sont jugés inutiles car impliquent de connaître la musique disponible, nécessitent du temps pour les écouter alors qu'un coup d'œil rapide vers l'écran permet de trouver la musique recherchée.

52% des participants disent ne pas apprécier les spearcons : incompréhensibles et donnant l'impression d'un bégaiement, ils sont source de confusion et de stress et ne constituent pas une aide pour l'utilisateur.

3. Avis post conduite des participants {étape 7}

Satisfaction relative aux sons. Les bibliothèques vocale et hiérarchique évaluées sont jugées de bonne qualité.

L'appréciation des sons hiérarchiques porte essentiellement sur les sons les mieux identifiés lors de l'étape 3 (Téléphone et Communication de la bibliothèque hybride, Communication et Musique de la bibliothèque musée). Inversement, les sons peu appréciés sont les moins évocateurs, hormis le son Véhicule de la bibliothèque hybride, bien identifié mais mal jugé car trop familier (bruit de moteur).

Concernant les sons de confirmation d'action, les sons de défilement et d'erreur / butée sont jugés utiles et évocateurs et sont largement appréciés (respectivement 92 et 75%).

Les extraits musicaux satisfont 67% des participants car ils permettent notamment de pallier l'oubli du titre des morceaux et remplissent une fonction de confirmation.

Enfin les spearcons ne sont pas du tout appréciés (13% seulement de participants satisfaits) car, jugés incompréhensibles, ils perturbent la concentration sur la conduite.

Utilisation du guidage sonore sur le long terme. Une forte majorité (92%) indique que la sonification rend la conduite plus sûre, mais qu'elle ne permet pas de supprimer le visuel (83%).

Par ailleurs, 71% des participants pensent qu'ils continueraient d'utiliser le guidage sonore dans leur véhicule sur le long terme. De plus, 71% des participants préféreraient continuer d'utiliser la même bibliothèque sonore plutôt que d'en changer pour conserver le bénéfice de l'apprentissage.

Quelques participants utiliseraient la bibliothèque vocale lors de la phase de découverte du système (bibliothèque plus explicite, ne requérant aucun apprentissage) puis utiliseraient sur le long terme la bibliothèque hiérarchique, moins lassante, plus courte, plus agréable et plus personnalisée. Après l'appropriation du système, la bibliothèque hiérarchique serait ainsi utilisée par 35% des participants, la bibliothèque vocale par 30,5% des participants, et 30,5% des participants opteraient pour une bibliothèque mixte vocale et hiérarchique, le vocal étant utilisé pour remplacer les sons non explicites du dernier niveau de menu. Seulement 1 personne préfère la condition non sonifiée.

DISCUSSION

Notre hypothèse de réduction de la durée de réalisation des tâches de navigation dans le système multimédia grâce à la sonification n'est pas validée. Les sons hiérarchiques utilisés dans le haut du menu, joints à des sons vocaux dans les listes, n'apportent pas de gain par rapport au cas de la simple sonification par les sons de butée, d'erreur et de défilement. Le vocal présent dès le haut du menu entraîne quant à lui un allongement de la durée de réalisation des tâches. Sur les sons hiérarchiques, notre résultat est similaire à celui de Leplâtre [3] qui a appliqué le principe des earcons hiérarchiques pour la sonification d'un menu de téléphone portable. Cette étude compare deux groupes (l'un dispose du menu visuel seul, l'autre du visuel et des sons) qui doivent effectuer des tâches de navigation sur le téléphone. Elle montre que la sonification diminue significativement le nombre d'appuis de touche et d'erreurs mais qu'en revanche elle ne change pas la durée entre deux appuis sur les touches (et donc le temps moyen pour effectuer la tâche). Notre résultat sur le vocal contredit celui de Palladino [12], qui a comparé le temps de recherche dans une liste de 50 contacts téléphoniques, avec visuel, dans trois conditions : sans son, avec vocal, avec spearcon + vocal ; elle montre que la condition n'influe pas sur le temps de réalisation de la tâche.

Notre deuxième hypothèse de moindre perception de détournement de regard en présence de sonification, et son corollaire de meilleure focalisation sur la conduite, est validée, tant par les résultats recueillis à l'étape 6 (influence significative de la sonification vocale) qu'après la conduite, où 92% des participants estiment leur conduite plus sûre avec le son. Il serait intéressant, sur une maquette disposant d'une interface moins expérimentale et plus aboutie que celle utilisée ici, de collecter des données quantitatives de détournement de regard, à l'aide d'un oculomètre, afin de savoir si l'attention visuelle du conducteur est réellement moins sollicitée lorsque le guidage sonore est activé.

Si les sons hiérarchiques n'induisent pas en conduite un ressenti plus sécuritaire, les personnes indiquent a posteriori les considérer comme plus acceptables que le vocal : certains les adopteraient d'emblée, d'autres après une phase d'apprentissage de la structure du menu avec le vocal, si bien qu'à terme la majorité utiliserait les sons hiérarchiques (35% emploieraient la librairie hiérarchique et 30,5% une librairie mixte hiérarchique et vocale). Si la librairie vocale est préférée parce qu'elle facilite la découverte du système, aucune interprétation des prompts vocaux n'étant nécessaire, la librairie hiérarchique est in fine retenue car jugée plus agréable et moins lassante. Une librairie mixte, alliant sons hiérarchiques sur les premiers niveaux de menu et sons vocaux sur le dernier niveau de menu serait également appréciée car elle permettrait de pallier les difficultés de compréhension et de mémorisation des sons hiérarchiques du dernier niveau de menu. Notre hypothèse selon laquelle les sons hiérarchiques sont plus acceptables que le vocal est donc validée, pour les premiers niveaux du menu.

La majorité des participants (71%) affiche son intention de conserver à long terme le guidage sonore. Ce résultat, obtenu après une courte utilisation des sons (30 minutes en conduite environ) est à tempérer par le résultat recueilli par Helle [4] sur un téléphone portable sur lequel des earcons bien conçus ont été installés : au bout de trois semaines, 9 participants sur 17 avaient enlevé les sons, les jugeant gênants dans les lieux publics, trop longs et trop aigus. Certes le contexte est ici différent : le milieu automobile est moins ouvert sur l'extérieur, et la technologie sonore disponible sur un système multimédia embarqué est visiblement très supérieure à celle du téléphone portable utilisé par Helle. Une évaluation longitudinale (prêt d'un véhicule équipé pendant une période définie) permettrait d'observer l'effet du guidage sonore sur l'appropriation du système et d'étudier l'évolution de son utilisation sur le long terme.

Par ailleurs, soulignons le souhait des participants de disposer de sons courts : nous avons créé les sons hiérarchiques avec ce principe de design, nous sommes donc incités à poursuivre dans cette voie.

Les participants n'apprécient pas les sons musèmes, sauf les mieux identifiés, même s'ils reconnaissent leur meilleure aptitude que les sons hybrides à assurer la continuité du thème sonore. Notre hypothèse selon laquelle ces sons seraient plus acceptables que les sons hybrides n'est donc pas validée. Plusieurs sons de la librairie hybride doivent néanmoins être retravaillés pour rendre la fonction qu'ils représentent plus évidente. C'est le cas notamment du son Musique qui évoque davantage

un son d'erreur. Nous obtenons grâce à la comparaison entre ces deux librairies sonores le résultat suivant : l'acceptabilité des sons est directement liée à leur taux de reconnaissance, élevé lorsque le son entretient une relation forte avec l'item à représenter. Ce point justifie notre principe d'utiliser des composantes iconiques dans les sons hiérarchiques, et nous conduit à abandonner les sons hiérarchiques au profit du vocal dès que la relation son / item se relâche. Notre conclusion sur l'intérêt de la composante iconique rejoint les travaux de Dingler [10] qui comparent la facilité d'apprentissage du vocal, des icônes auditives et des earcons.

Nous avons intégré les spearcons à notre système afin de proposer une augmentation de la vitesse de production sonore en cohérence avec la vitesse de défilement dans la liste. Cependant les spearcons sont rejetés car ils ne sont pas compris. Notre résultat contredit ceux de Palladino [12] et de Dingler [10] selon lesquels quasiment une seule présentation du spearcon suffit pour qu'il soit appris. Si les spearcons de Palladino sont d'un type différent des nôtres (compression logarithmique chez Palladino vs. linéaire), ceux de Dingler sont similaires, la qualité des spearcons que nous avons utilisés ne semble donc pas en cause. Quant aux travaux sur l'acceptabilité des spearcons, la seule étude portant sur ce sujet, celle de Walker [13] n'apporte pas d'élément clair.

Enfin, la grande majorité (83%) des sujets n'envisagent pas de se passer du visuel. Ce résultat est en cohérence avec les résultats de Palladino [12], confirmés par Walker [9], qui indiquent que le temps de recherche d'un item dans une liste est sensiblement plus court avec visuel que sans, et ce même après dix répétitions. L'analyse des verbatim recueillis en conduite nous permet de voir que le guidage sonore est principalement utilisé dans la phase de recherche initiale. Dans cette phase, les sons associés au menu (en particulier les sons hiérarchiques hybrides et la synthèse vocale, explicités) et les lettres alphabétiques permettent à l'utilisateur de s'approcher au plus près de l'information recherchée (par exemple, atteinte de la lettre D lors de la recherche du morceau « Divine Idylle »). Puis, lors de la phase d'affinement de la recherche (recherche de l'item « Divine Idylle » dans la lettre D) l'utilisateur privilégie souvent l'interface visuelle pour repérer et atteindre l'information recherchée : l'information est repérée d'un simple coup d'œil. Dans cette phase, la synthèse vocale et les extraits musicaux se révèlent peu utiles. L'association d'un prompt vocal à l'ensemble des items survolés peut d'ailleurs être jugée lassante et donner l'impression d'un bégaiement du système lors du défilement rapide dans les listes (car la synthèse vocale est coupée lorsque le focus passe rapidement d'un item à l'autre). Néanmoins, la synthèse vocale et les extraits musicaux sont utilisés dans la phase finale, pour confirmer à l'utilisateur que l'item sur lequel il s'est arrêté est bien celui recherché.

CONCLUSIONS ET PERSPECTIVES

Les résultats obtenus confirment l'intérêt du guidage sonore sur un système multimédia automobile. Bien que la sonification du système ne permette pas de naviguer et d'accéder plus rapidement aux données, les participants préfèrent la conserver car, au final, leur conduite leur apparaît plus sûre. Les sons proposés doivent cependant être évocateurs, et ce à tous les niveaux de l'arborescence. Le guidage sonore vocal leur donne l'impression

d'avoir une conduite plus confortable et de devoir moins souvent détourner le regard de la route.

L'utilisation des lettres alphabétiques lors du défilement rapide dans les listes recueille une forte adhésion. Des travaux récents menés sur l'utilisation de lettres alphabétiques en synthèse vocale (Spindex) lors du défilement dans les listes, confirment l'efficacité de ce type de retour sonore. Mais la comparaison de différentes variantes de Spindex, effectuée par Jeon [14], révèle que celle utilisée dans notre maquette (Spindex de type « Minimal », à savoir « A » prononcé sur le premier item commençant par A puis « B » sur le premier item commençant par B, sans son sur les items intermédiaires de A) est la moins appréciée. Il serait alors intéressant de nous orienter vers l'utilisation des variantes de Spindex les plus appréciées (comme celle de type « atténué » où le son A est prononcé à un niveau restreint sur les items intermédiaires de A) afin d'allier efficacité et acceptabilité.

Enfin, alors que l'intervention d'un designer sonore s'est révélée et demeure indispensable pour la création des sons hiérarchiques associés aux items du menu, la création des extraits musicaux et des prompts en synthèse vocale doit, elle, être automatisée pour que l'intégration de ces sons soit envisageable dans un véhicule. En effet, autant les items du menu n'évolueront pas (ils sont figés par le constructeur), autant le contenu des listes des menus Musique et Communication va être régulièrement modifié par les utilisateurs du véhicule.

REMERCIEMENTS

Nous tenons à remercier A. Picot, R. Heim et D. Blumenthal (Renault) pour leur contribution à l'étude.

BIBLIOGRAPHIE

1. M. Blattner, D. Sumikawa, R. M. Greenberg. Earcons and icons: their structure and common design principles. *Human-Computer Interaction*, 4:11-44, 1989.
2. S. Brewster, P. Wright, A. Edwards. Experimentally derived guidelines for the creation of earcons. *Human Computer Interaction*, Huddersfield, UK, 1995.
3. G. Leplâtre, S. Brewster. Designing Non-Speech Sounds to Support Navigation in Mobile Phone Menus. *Proc. of Intern. Conf. on Auditory Display (ICAD)*, Atlanta, USA, pp 190-199, 2000.
4. S. Helle, G. Leplâtre. Menu sonification in a mobile phone - a prototype study. *Proc. of Intern. Conf. on Auditory Display (ICAD)*, Espoo, Finland, July 2001.
5. W. Gaver. Auditory Icons: Using Sound in Computer Interfaces. *Human-Computer Interaction*, 2 (2) pp. 167-177, 1986.
6. S. Barrass. Auditory information design. PhD thesis, The Australian National University, 1997.
7. S. Conversy. Conception d'icônes auditives paramétrées pour les interfaces homme-machine. PhD thesis, Université de Paris-Sud XI, 2000.
8. B. N. Walker, A. Nance, J. Lindsay. Spearcons: speech-based earcons improve navigation performance in auditory menus. *Proc. of Intern. Conf. on Auditory Display (ICAD)*, 2006.
9. D. K. Palladino, B.N. Walker. Learning rates for auditory menus enhanced with spearcons versus earcons. *Proc. of Intern. Conf. on Auditory Display (ICAD)*, 2007.
10. T. Dingler, J. Lindsay, B. N. Walker. Learnability of sound cues for environmental features: Auditory icons, earcons, spearcons, and speech. *Proc. of Intern. Conf. on Auditory Displays (ICAD)*, Paris, 2008.
11. P. Tagg. Teoria degli affetti in « Analysing popular music : theory, method and practice », *Popular Music*, 2 (1982), pp 37-67.
12. D. K. Palladino, B. N. Walker. Efficiency of spearcon-enhanced navigation of one dimensional electronic menus. *Proc. of Intern. Conf. on Auditory Display (ICAD)*, 2008.
13. B.N. Walker, A. Kogan. Spearcon performance and preference for auditory menus on a mobile phone. *Proc. of International Conference on Universal Access in Human-Computer Interaction (UAHCI)*, 2009
14. M. Jeon, B. N. Walker "Spindex": accelerated initial speech sounds improve navigation performance in auditory menus. *Proc. of HFES conference*, San Antonio, 2009

