

Titre .

Influence d'images évocatrices et distractrices sur une tâche de jugement en acoustique des salles

Auteurs.

Aurore Defays¹, Stéphane Safin¹, Alexis Billon², Christine Decaestecker³, Nadine Warzée³

Adresses postales et électroniques des auteurs.

1) Université de Liège – LUCID-ULg (Lab for user cognition and innovative design)

1, chemin des chevreuils, Bât B52, 4000 Liège - Belgique

2) INTELSIG group - Département E.E.I.

Université de Liège B28

B-4000 Sart-Tilman (Liège 1) - Belgium

3) ULB - Faculté des Sciences Appliquées - LISA CP 165/57 (Laboratoire de l'Image : Synthèse et Analyse)

Service LIST Avenue F. Roosevelt, 50, B-1050-Bruxelles (Belgium)

aurore.defays@student.ulg.ac.be

stephane.safin@ulg.ac.be (auteur correspondant)

abillon@ulg.ac.be

cdecaes@ulb.ac.be

nwarzee@ulb.ac.be

Numéros de téléphone et télécopie des auteurs.

1) Tel : +32.4.366.95.15 Fax : +32.4.366.95.17

2) Tel +32 4 3662650 Fax +32 4 3662649

3) Tél : 32-2-650.27.61 Fax : 32-2-650.22.98

Résumé de la soumission.

Dans le cadre de la conception d'un espace immersif multimodal visant à soutenir la tâche d'évaluation de la qualité acoustique de salles sur base de simulations numériques, nous nous intéressons aux liens entre l'image et le son. 54 sujets ont été confrontés à une tâche de jugement du degré de réverbération d'un extrait sonore dans une situation contrôle (sans image) puis dans une situation expérimentale avec, pour la moitié des sujets, une image visant à soutenir le jugement (cohérente avec le son) et, pour l'autre moitié, une image « distractive ». Nos résultats montrent une influence positive des images évocatrices et pas d'effet des images distractrices.

Mots clés

multimodalité, intermodalité, acoustique des salles

Logiciel utilisé

MS Word 2004 pour Mac

Forme de participation

Article de recherche long

Influence d'images évocatrices et distractrices sur une tâche de jugement en acoustique des salles

Aurore Defays¹, Stéphane Safin¹, Alexis Billon², Christine Decaestecker³, Nadine Warzée³

¹LUCID-ULg (Lab for User Cognition and Innovative Design)
Université de Liège - Belgique
aurore.defays@student.ulg.ac.be
stephane.safin@ulg.ac.be

²INTELSIG (Laboratory for signal and Image exploitation)
Université de Liège - Belgique
abillon@ulg.ac.be

³LISA (laboratoire d'images : synthèse et analyse) - Université Libre de Bruxelles - Belgique
cdecaes@ulb.ac.be
nwarzee@ulb.ac.be

RESUME

Dans le cadre de la conception d'un espace immersif multimodal visant à soutenir la tâche d'évaluation de la qualité acoustique de salles sur base de simulations numériques, nous nous intéressons aux liens entre l'image et le son. 54 sujets ont été confrontés à une tâche de jugement du degré de réverbération d'un extrait sonore dans une situation contrôle (sans image) puis dans une situation expérimentale avec, pour la moitié des sujets, une image visant à soutenir le jugement (cohérente avec le son) et, pour l'autre moitié, une image « distractive ». Nos résultats montrent une influence positive des images évocatrices et pas d'effet des images distractrices.

MOTS CLES : multimodalité, expérimentation, acoustique des salles

ABSTRACT

In the context of the design of a multimodal environment for room acoustics evaluation, we investigate the link between a projected image and the sound of an acoustic simulation. 54 subjects have been tasked to judge the level of reverberation of a sound in a control situation (without image) and in an experimental situation. In this experimental condition, half the subjects were presented a coherent image aiming at supporting the evaluation task and the other half, a disturbing image. Our results show a positive influence from the evocative images, and no influence from the distractive images.

CATEGORIES AND SUBJECT DESCRIPTORS: H.1.2 User/Machine Systems (Human factors, Human information processing) ; H.5.1 : Multimedia Information Systems (audio output, virtual reality)

GENERAL TERMS: Experimentation, Human Factors

KEYWORDS: multimodality, experimentation, room acoustics.

INTRODUCTION

Cette étude se situe dans le cadre du développement d'un dispositif de réalité virtuelle dédié à l'acoustique des salles. Ce projet de recherche dénommé Auralias vise à développer un environnement de travail collaboratif acousticiens/architectes permettant le rendu acoustique d'un espace architectural en cours de conception à partir de simulations numériques, afin de permettre aux acteurs impliqués d'améliorer sa qualité acoustique de façon préventive, plutôt que corrective. La simulation comprend un rendu sonore en deux dimensions, couplé à une représentation 3D projetée en grande taille du modèle de la salle simulée et prend place dans un système immersif dédié.

Le projet dans son ensemble utilise une démarche centrée utilisateurs. Sur base de l'analyse d'activités d'acteurs du milieu de l'acoustique des salles, nous tentons de proposer des solutions qui soient adaptés aux contextes d'usages particuliers. Cette implication de l'utilisateur nous incite aussi à avoir des réflexions sur les mécanismes cognitifs mis en œuvre dans l'évaluation de la qualité acoustique d'espaces architecturaux. C'est dans cette optique que cette étude se situe.

En effet, afin de comprendre quels types d'image doivent être affichés dans notre studio immersif en soutien à la tâche d'évaluation de la qualité acoustique de l'espace simulé, il nous est nécessaire de comprendre un peu plus en profondeur les liens entre les modalités visuelles et auditives. En particulier, nous essayons de comprendre en quoi une image de l'espace – et ses caractéristiques – peut fournir une ressource ou au contraire entraver la tâche d'écoute et de jugement. Pour ce faire, nous adoptons une approche expérimentale.

Dans un premier temps, nous décrivons le principe de l'acoustique des salles et rendons compte du travail de l'acousticien dans des projets de conception. Nous expliquons ensuite l'objet du projet Auralias, qui vise à soutenir cette activité. Nous abordons dans la section suivante les liens entre les modalités visuelles et auditives décrits dans la littérature pour conclure sur nos hypothèses de recherche. Dans les sections suivantes,

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 2009, 13-16 Octobre 2009, Grenoble, France

Copyright 2009 ACM 978-1-60558-461-4/09/10 ...\$5.00.

Dans ce contexte, le projet Auralias vise à aider l'activité de l'acousticien en proposant un système immersif permettant de présenter non plus uniquement la valeur des paramètres objectifs, mais aussi de soutenir son avis

à l'aide du rendu sonore de la salle en cours de conception.

Pour ce faire, un environnement est proposé dans lequel les utilisateurs (architectes et acousticiens) sont immergés dans un modèle de l'espace architectural. Cet environnement immersif est installé dans une salle d'environ 38 m² et est composé :

- d'un écran de projection dont la taille et la forme ont été spécialement choisies pour permettre de couvrir pratiquement l'intégralité de la vision périphérique et d'augmenter le sentiment d'immersion ;
- d'un dispositif d'auralisation (rendu sonore en deux dimensions) sur base de 8 haut-parleurs ;
- d'un système de tracking permettant de localiser les utilisateurs pour adapter le rendu sonore à leur positionnement dans l'espace immersif ;
- d'une interface de commandes sur UMPC ;
- de deux PC en réseau pour la gestion du son et de l'image.

Cet environnement immersif permettra aux acteurs de la conception de se baser sur un rendu sonore et visuel pour juger de la qualité acoustique de la salle. La tâche principale des utilisateurs sera donc une tâche d'écoute, et une vue synthétisée de la salle interviendra en support de l'immersion sonore. Si le rendu sonore se veut aussi fidèle que possible, reproduisant les caractéristiques acoustiques, bonnes ou mauvaises, d'une salle [7], il nous semble intéressant, dans le cadre de la conception de cet environnement, d'étudier l'impact que pourraient avoir les caractéristiques du modèle visuel sur la tâche.

En effet, les modèles 3D proposés en architecture et en acoustique (voir figure 2) diffèrent sur plusieurs caractéristiques [3], notamment sur leur complexité (nombre de faces) et leur niveau de réalisme (textures, ombrages...). On peut s'imaginer qu'un modèle trop réaliste ou trop fini va capter l'attention visuelle de l'auditeur au détriment de la tâche d'évaluation sonore. Mais à l'inverse, on peut aussi faire l'hypothèse qu'une image plus réaliste renforcera l'immersion et permettra une écoute plus attentive.



Figure 2 : modélisation acoustique (nécessaire à la simulation et à l'auralisation – à gauche) et architecturale (à droite) d'un même amphithéâtre.

Pour savoir quel type de modèle afficher pour soutenir l'écoute, il nous apparaît nécessaire de comprendre, avant d'adresser spécifiquement la question de ses caractéristiques (niveau de réalisme), si l'utilisation d'une image cohérente ou incohérente avec le son va

induire une amélioration ou une diminution de la performance à une tâche d'écoute.

LA MULTIMODALITE

En psychologie cognitive, une modalité est liée à l'activité d'un organe sensoriel spécifique. Par exemple, la vue est la modalité tandis que l'œil est l'organe sensoriel. Aujourd'hui, il est unanimement reconnu qu'une modalité peut avoir de l'influence sur une autre. La psychologie cognitive appelle ce phénomène l'intermodalité : toutes les activités de notre vie quotidienne reposent sur la participation simultanée et l'interaction de nos sens, même si nous prêtons peu d'attention à ces interactions [13]. La plupart des objets et des événements de notre environnement nécessitent la mobilisation de plusieurs sens pour être perçus mais, dans la plupart des cas, la convergence des systèmes existe peu voire pas du tout, chaque sens prélevant les informations qui lui sont propres. Cet ensemble d'informations recueillies doit être harmonisé pour que l'organisme réponde adéquatement à la situation et perçoive l'environnement comme une unité, c'est-à-dire attribue une *signification* à ce qui est perçu : les informations perçues de l'environnement par les différentes modalités se voient ainsi attribuer des « valeurs psychologiques » (*évaluation*) qui vont être combinées (*intégration*) pour permettre d'effectuer un jugement (*décision*) [9].

Il existe quatre situations différentes d'intermodalité, variant sur leur rapport temporel et sur les propriétés de l'objet perçues par chacune des modalités [5] qui vont permettre des types d'intégration différents.

Situation	Propriétés de l'objet	Mode d'exploration
Transfert intermodal	Mêmes propriétés	En mode successif
Intégration complémentaire	Différentes propriétés	En mode simultané
Intégration redondante	Mêmes propriétés	En mode simultané
Conflits perceptifs	Propriétés incompatibles	En mode simultané

Tableau 1 : Situations d'intermodalité [5]

- Le transfert intermodal est une situation où une information prélevée par une modalité sensorielle (phase d'encodage) est utilisée dans une autre (phase de décodage). Dans ce cas, les différents sens reçoivent les mêmes propriétés sur l'objet mais successivement.
- Dans l'intégration complémentaire, les modalités reçoivent simultanément des propriétés distinctes d'un même objet. Ensuite, c'est l'intégration entre les différentes informations qui permet de créer la signification de l'objet.
- L'intégration redondante concerne le fait que plusieurs modalités accèdent simultanément aux mêmes propriétés de l'objet. Sachant que ces diverses modalités traitent celles-ci différemment,

l'individu doit coordonner ces informations pour créer la signification.

- Dans les conflits perceptifs, les différents sens reçoivent des propriétés incompatibles simultanément sur un même objet, comme c'est le cas dans certaines illusions auditives ou visuelles.

Dans le cas de notre projet, nous cherchons à connaître les propriétés de l'objet véhiculées par la modalité visuelle, en complément de la modalité auditive. Il s'agit de voir si ces informations sont complémentaires ou redondantes. Il s'agit aussi, en plaçant les personnes en situation de conflits perceptifs, de voir si une modalité prend le pas sur l'autre.

En ingénierie de l'Interaction Homme-Machine, la notion de multimodalité renvoie à des liens entre modalités d'interaction, qui diffèrent sensiblement des modalités au sens psychologique du terme. Une modalité d'interaction caractérise les échanges entre le système et l'utilisateur. Elle est définie comme la relation qui existe entre deux niveaux d'abstraction distincts : le dispositif physique (niveau physique) et langage d'interaction (niveau logique). Dans le cas de notre espace d'immersion, les deux modalités d'interaction en sortie du système sont un écran et des hauts parleurs, qui renvoient tous deux des informations sur la salle de spectacle modélisée et sur son acoustique.

Coutaz et Nigay [2], dans le modèle CARE, identifient quatre types de liens entre modalités dans les interfaces multimodales. Si ces propriétés sont souvent utilisées pour décrire des interactions en entrée des systèmes, elles peuvent aussi caractériser les flux d'informations en sortie.

- L'assignation : chaque modalité véhicule des informations spécifiques, que l'on ne peut retrouver dans les autres modalités.
- La redondance : les modalités sont utilisées en complément les unes des autres. L'ajout d'une modalité va permettre de clarifier, de désambiguïser le message.
- La complémentarité : il est nécessaire d'utiliser l'ensemble des modalités définies pour parvenir à obtenir l'information souhaitée.
- L'équivalence : les différentes modalités peuvent être utilisées de manière équivalente. Elles apportent les mêmes informations.
- En outre Martin [10] ajoute le lien de concurrence : plusieurs modalités apportent de l'information indépendante en parallèle, simultanément, ce qui peut mener à des conflits dans le cas où ces informations indépendantes sont contradictoires.

HYPOTHESES

Notre étude tente d'investiguer l'influence de la modalité visuelle sur une tâche d'écoute. Nous nous interrogeons donc sur les effets d'interférences potentiels qui pourraient survenir entre ces deux modalités. En d'autres

mots, nous essayons d'étudier dans quelle mesure la présence d'une vue synthétisée de la salle va affecter la performance de jugement de l'acoustique réalisé par les utilisateurs de notre système. Quatre questions nous intéressent particulièrement :

- Les sujets non entraînés sont-ils capables de discriminer des extraits sonores aux propriétés différentes, c'est-à-dire d'évaluer la qualité acoustique d'un son ?
- Un modèle visuel cohérent avec un son, affiché simultanément à ce dernier, affecte-t-il l'identification des propriétés de ce son ?
- Un modèle visuel incohérent avec le son détériore-t-il la performance, en plaçant les sujets dans une situation de conflit perceptif, ou la modalité auditive prend-elle le pas sur la visuelle ?
- Les effets du modèle visuel sont-ils identiques suivant le profil de l'auditeur (formation en ingénierie du bâtiment) ? En effet, l'expertise joue un rôle essentiel dans les processus de perception et de classification [11].

Ces premières hypothèses nous permettront de jauger de l'importance de la modalité visuelle pour soutenir la tâche des acousticiens et architectes dans notre système. En particulier, nous voulons investiguer le type de rapport qu'entretiennent les modalités auditives et visuelles en sortie de notre système pour la tâche visée :

- Elles peuvent être assignées, auquel cas seule l'information auditive sera importante pour réaliser la tâche. Les images ne devraient pas modifier la performance.
- Dans le même ordre d'idées, elles peuvent être équivalentes, auquel cas l'information provenant d'une seule modalité suffira à effectuer la tâche : il n'y a pas de pouvoir expressif supplémentaire induit par la combinaison de deux modalités.
- Elles peuvent être redondantes, auquel cas les images ajoutent de l'information pertinente pour la tâche et désambiguïsent le message. Ainsi les informations auditives suffiront à mener la tâche à bien, mais les informations visuelles permettront d'améliorer la performance, si toutefois elles sont cohérentes avec le son.
- Elles peuvent être complémentaires, auquel cas la tâche est impossible (ou quasi-impossible) sans la modalité visuelle. La performance sera très réduite sans la présence de la modalité visuelle, et devrait être affectée par des informations contradictoires présentées dans les deux modalités.
- Elles peuvent être concurrentes, c'est-à-dire que la modalité visuelle apporte des informations différentes et indépendantes de l'audio. Cela pourrait améliorer la performance à la tâche si les informations sont congruentes, mais aussi la détériorer si les elles sont contradictoires.

METHODOLOGIE

Pour répondre à ces questions, nous mettons en place un dispositif expérimental. Dans cette expérience, la tâche de jugement acoustique proposée aux sujets consiste à classer des extraits sonores selon un paramètre acoustique objectif, la réverbération, en s'aidant ou non d'images. Nous plaçons la moitié des sujets dans une situation où le couple son/image est cohérent, et l'autre moitié dans une situation de *conflit perceptif* [4]. Nous cherchons à investiguer si la présence d'une image associée à un son facilite et/ou perturbe la performance d'une tâche d'évaluation comparative du degré de réverbération d'extraits sonores.

Extraits sonores. Dans les extraits, les 80 premières millisecondes de la réponse acoustique sont fixes et sa décroissance est ensuite variée pour obtenir cinq temps de réverbération :

- TR1=1,60 s
- TR2=1,80 s
- TR3=2,00 s (durée idéale pour une salle de concert)
- TR4=2,20 s
- TR5=2,40 s

Un échantillon sonore anéchoïque, des percussions, est convolué avec ces réponses acoustiques. L'auditeur doit alors juger les extraits sonores obtenues selon leur durée de réverbération et les classer dans l'ordre de réverbération croissante.

Le temps de réverbération caractérise la capacité d'une salle à maintenir longtemps une excitation acoustique. Ce critère rend compte de la profondeur, la « présence » du son [1]. Ce paramètre constitue une des dimensions principales du jugement affectif d'une salle [12]. De plus, le concept de réverbération est simple à appréhender et l'association d'un lieu réverbérant avec certaines caractéristiques visuelles de l'édifice, comme sa taille, est relativement aisé. Un pré-test effectué préalablement à la tâche d'écoute a été mis en place pour valider cette assertion (voir ci-dessous).

Plan expérimental. Pour répondre à nos questions de recherche, le plan expérimental suivant est proposé (tableau 2).

Contrôle Son uniquement	Expérimental Son + image	
	Image évocatrice	Image distractive
G1 :10 archi G2 : 17 divers	G1 :10 archi G2 : 17 divers	
G3 :10 archi G4 : 17 divers		G3 :10 archi G4 : 17 divers

Tableau 2 : plan expérimental.

Chaque sujet est amené à faire la tâche d'évaluation de la réverbération deux fois : une situation contrôle et une situation expérimentale. Dans la situation contrôle, aucune image n'est présentée aux sujets afin de vérifier qu'ils sont effectivement capables de réaliser cette tâche

de jugement acoustique sur base du temps de réverbération et d'obtenir, pour chacun, une valeur de performance de référence.

Sujets. Nous avons décidé de réaliser cette expérience avec deux types de sujets cibles : d'une part, des étudiants en fin de parcours en architecture (G1 et G3) et, d'autre part, des étudiants terminant également leurs études mais dans d'autres filières (G2 et G4). De cette manière, nous espérons déceler des différences de sensibilité acoustique entre un public averti et non averti en matière de conception architecturale des salles.

Images. Dans la seconde condition, expérimentale, la lecture de chaque extrait sonore est accompagnée par l'affichage d'une image (photo) représentant un lieu. Ces images ont été soigneusement choisies pour évoquer chacune, par ses caractéristiques visuelles, une réverbération différente.



Figure 3 : images proposées par ordre croissant de réverbération évoquée.

Le caractère évocateur de ces images par rapport à la réverbération de la salle a été validé par un pré-test. Partant du principe que la réverbération perçue d'un espace est fonction de sa taille et de son revêtement, le choix des images a été fait sur ces deux critères : la taille de la salle (du plus petit au plus grand) et les matériaux utilisés (du plus absorbant – tapis, rideaux, etc. – au plus réverbérant – marbre par exemple). Les caractéristiques géométriques des salles étaient comparables pour toutes les images proposées (salles rectangulaires). Treize images ont été proposées aux pré-testeurs (14 sujets) et il leur a été demandé de classer les salles par ordre de réverbération. À l'analyse des résultats, un certain consensus sur le classement des images apparaît. Pour

valider plus précisément quelles images différaient les unes des autres dans le classement, nous avons fait des comparaisons paires des images (test de Wilcoxon). Grâce à ce pré-test, les images les plus « ambiguës » ont été supprimées de notre schéma expérimental. Nous avons sélectionné et classé cinq images (figure 4) différenciant les unes des autres dans l'ordre adéquat et avec un niveau de signification suffisant ($p < 0.01$).

Conditions. Les participants répartis en deux groupes.

- Dans le premier groupe (regroupant G1 et G2), l'image associée au son dans la condition expérimentale est dite « évocatrice », c'est-à-dire qu'il y a une correspondance entre la réverbération de l'extrait sonore et la réverbération évoquée par l'image. Dans cette condition, les images visent à soutenir la tâche de jugement de la réverbération des sons.
- Le second groupe (G3 et G4) est confronté à des images « distractrices » visant à perturber le jugement. Les extraits sonores et les images ont été appariés aléatoirement, de sorte qu'il n'existe plus aucun lien plausible entre la réverbération du son et la résonance perçue de l'espace.

Entre la situation contrôle et la situation expérimentale, les extraits sonores sont comparables mais différents : il s'agit du même instrument, mais l'extrait et son rythme sont différents.

Déroulement. Pour le déroulement de l'expérience, les sujets utilisent une interface développée avec MatLab pour lire les extraits sonores, visualiser les images, encoder les résultats. Cette interface affiche les consignes, standardisées pour limiter les biais liés à l'attitude de l'expérimentateur. Ces consignes ont été formulées de façon à ce qu'elles n'incitent pas trop les participants à regarder les images afin d'éviter que la tâche d'origine auditive se transforme en tâche purement visuelle (ou de classement d'images).

- « Bonjour, dans la première partie de cette expérience, vous allez entendre cinq extraits sonores. Nous vous demandons de classer ces différents extraits par ordre de réverbération (1 pour le moins résonnant et 5 pour le plus résonnant). »
- « Maintenant, nous allons vous présenter cinq autres extraits sonores. Ces différents extraits seront accompagnés de l'image de l'endroit où l'extrait sonore a été enregistré. Nous vous demandons également de classer ces extraits, par ordre de réverbération, du moins résonnant (1) au plus résonnant (5). »

Les cinq extraits sonores, présentés en ordre aléatoire pour chaque sujet, varient uniquement selon leur temps de réverbération. Les sujets peuvent les écouter plusieurs fois si souhaité.

Questionnaire. Avant de clôturer cette expérience, les participants doivent remplir un questionnaire. Celui-ci investigate plusieurs variables

- Données démographiques (âge et sexe)
- Pratique de la musique
- Usage de différentes salles (amphithéâtres, salles de concert...). Il s'agit de variables que nous jugeons pertinentes pour évaluer l'expertise en qualité acoustique de salles.
- Habitude à lire des plans. Ce critère a pour but de vérifier la distinction entre le groupe architecte et autres.
- Informations subjectives sur le déroulement de l'expérience.

Nous nous attendons à ce que la performance à la tâche d'écoute dans la situation d'images « évocatrices » soit meilleure que dans la condition contrôle. En outre, nous pensons qu'en présence d'images « distractrices », cette performance sera moins bonne que lors de la condition contrôle.

RESULTATS

Participants. 54 participants (27 hommes et 27 femmes) ont réalisé cette tâche d'écoute et de classement d'extraits sonores. La moyenne d'âge était de 25,7 ans avec un écart type de 6,49 ans.

Analyses. L'analyse des résultats porte sur deux variables : le temps d'exécution de la tâche et la performance. La performance à cette tâche d'écoute est évaluée à l'aide du coefficient de corrélation de Spearman. Celui-ci exprime l'adéquation entre l'ordre théorique (les réponses correctes) et l'ordre observé (les réponses du sujet). Ce coefficient varie de -1 (inverse parfait de l'ordre correct) à +1 (ordre correct). Les différences entre groupes indépendants ont été éprouvées par des tests (non-paramétriques) de Mann-Whitney et les différences intra-participants (au sein d'un même groupe) par le test des Signes. Tous les tests ont été éprouvés au seuil de 5% en bilatéral.

Profil. Dans un premier temps, nous nous demandons si les étudiants qui manipulent fréquemment des plans, des vues 3D, etc. présentaient des différences par rapport à un public non averti en ce qui concerne l'acoustique des salles. Nous ne trouvons pas de différence significative entre les groupes (architecture vs autre filière) ni sur le temps, ni sur la performance. De même, la variable « usage des plans » n'est pas discriminante. Nous pouvons donc considérer que ces deux types de publics sont équivalents et fusionner ces groupes G1 et G2 ainsi que les groupes G3 et G4 pour les analyses ultérieures. Nous voulons cependant nuancer ces résultats puisque l'échantillon servant à cette expérience est relativement jeune. En effet, il est constitué majoritairement d'étudiants. En ciblant des publics de professionnels,

nous pourrions éventuellement trouver des différences entre ces deux groupes.

Caractéristiques personnelles. Différentes données complémentaires sur les participants ont été récoltées afin de définir certains profils de participants et voir si ceux-ci avaient un impact sur la performance. Ces facteurs sont : le fait que le participant joue un instrument de musique fréquemment ou pas, son degré de fréquentation des auditoriums, son degré de fréquentation des théâtres, et enfin son degré de familiarisation avec les concerts de musique non amplifiée. La seule influence significative identifiée concerne le fait de jouer assez fréquemment un instrument. En effet, les participants sont significativement plus rapides pour classer les sons extraits lorsqu'ils font fréquemment de la musique par rapport à ceux qui en font peu ou pas ($p=0.002$). Il n'y a, en revanche, aucun impact significatif sur la performance. Ce résultat peut être aisément expliqué par le fait que les sujets sont habitués à des tâches d'écoute, la phase d'apprentissage n'étant plus nécessaire. En outre, il est à noter que la variable « habitude de fréquentation des concerts de musique non amplifiée », sans doute très importante pour ce genre de tâche, ne donne pas de résultats. Cependant, la quasi-totalité des sujets ne sont pas des habitués, ne permettant pas de distinguer deux groupes différents pour ce facteur.

Impact des images. Une comparaison globale des performances des participants placés en situation évocatrice (G1, G2) avec celles des participants placés en situation distractive (G3, G4) ne révèle aucune différence significative. Au sein de chacun de ces deux groupes de sujets, nous avons ensuite analysé l'impact des images évocatrices, d'une part, et celui des images « distractives », d'autre part. Nous avons alors constaté un impact positif des images évocatrices sur la performance ($p=0.003$) et sur le temps de réponse ($p<0.0001$) (figure 4). Les participants (G1, G2) sont plus rapides et plus efficaces pour classer les sons en présence de ce type d'images, par rapport à la situation contrôle (sans image).

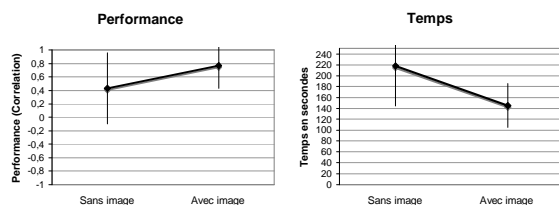


Figure 4 : résultats pour le groupe soumis à une image évocatrice. Les deux tendances sont significatives.

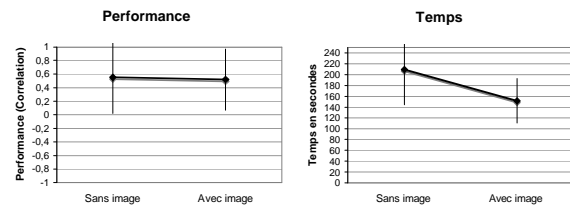


Figure 5 : résultats pour le groupe soumis à une image distractive. Seule la différence de temps est significative.

Par contre, au sein du groupe placé en situation distractive (G3, G4), la présence de ces images n'a pas d'effet sur la performance ($p=0.7$), mais continue d'influer sur le temps de réponse ($p=0.0002$) par rapport à la situation contrôle (figure 5). A l'instar de la situation précédente, les participants sont plus rapides pour classer les sons en présence d'images.

Ces résultats suggèrent un effet d'apprentissage pour les deux conditions qui agit sur la rapidité de réponse mais pas nécessairement sur la performance. Au niveau de la performance, il semblerait que la présence d'une image cohérente améliore les résultats. Par contre, la présence d'une image incohérente ne détériore pas la performance des sujets.

Paramètres subjectifs de l'utilisation des images.

Nous avons aussi évalué dans quelle mesure les participants avaient le sentiment que les images les avaient aidés. Nous n'avons pas trouvé de différence significative quant à l'apport des images perçu entre les groupes en situations évocatrice et distractive.

DISCUSSIONS

Notre projet dans sa globalité cherche à investiguer ce qu'il convient d'afficher à l'écran de notre espace de simulation. Nous cherchons à savoir quel type de représentations présenter (représentations symboliques, niveaux de rendu des textures, types d'éclairages, ...) pour favoriser la tâche d'écoute ou de jugement acoustique précédemment décrite. Cette première étude nous permet de répondre à une question préalable mais fondamentale, à savoir si une image peut interférer, dans un sens ou dans l'autre, avec le jugement acoustique.

Globalement, nous pouvons affirmer qu'en situation évocatrice, c'est-à-dire quand les caractéristiques visuelles sont cohérentes avec le son, on retrouve un effet de facilitation. Pour les situations distractives, où les images et le son n'ont pas de lien, nous n'avons pas trouvé d'impact préjudiciable des images sur la performance à la tâche d'écoute. Les modalités auditives et visuelles, en sortie de notre système, peuvent donc être qualifiées de *redondantes* [2] : la tâche est réalisable avec la modalité auditive uniquement, mais l'information visuelle permet de désambigüiser le message et d'apporter des informations complémentaires. Par contre, la modalité visuelle peut être outrepassée si elle apporte des informations contradictoires et place le sujet en

situation de conflit perceptif. Ces constatations sont d'une importance cruciale pour le développement de notre projet. Elles suggèrent qu'il est dans tous les cas intéressant d'afficher une image en soutien à l'évaluation auditive : dans le pire des cas, la performance ne sera pas détériorée et dans le meilleur des cas, elle pourra être soutenue. Elles suggèrent aussi qu'il faut privilégier des images qui apportent des informations complémentaires au son : des informations visuelles sur les matériaux utilisés dans la salle (notamment les matériaux absorbants) ainsi qu'une vision claire de la taille de l'espace sont des indices pertinents pour soutenir l'évaluation des propriétés acoustiques d'un espace. Si ces résultats sont très encourageants, ils ne permettent cependant pas encore d'investiguer de manière systématique le lien entre les caractéristiques du modèle visuel et la cohérence entre l'image et le son. Les textures, les lumières et la précision du modèle fournissent-ils des informations qui permettront aux utilisateurs de se faire une idée plus précise du comportement de la salle ? Cette question fera l'objet de nos prochaines recherches.

Il est aussi à noter que notre étude possède de nombreuses incomplétudes : le public sollicité ne représente que partiellement les utilisateurs finaux du dispositif, les conditions expérimentales ne reflètent pas réellement la complexité de la tâche attendue et les extraits sonores ne varient que sur un seul paramètre. Pour tirer des conclusions définitives, il sera nécessaire de reproduire ce genre d'expérience dans des milieux plus écologiques. En outre, il subsiste une inconnue concernant l'interprétation de nos résultats. Nous avons montré un effet d'habituation sur le temps de réponse (les sujets sont plus rapides au second test) sans que cet effet soit identifiable pour la performance. Il est néanmoins possible qu'un effet d'apprentissage existe concernant la performance. Dans ce cas, on arriverait à des conclusions inverses : les images évocatrices ne diminuent pas (voire augmentent) la performance, mais les images distrayantes au contraire diminuent la performance en annulant l'effet d'apprentissage.

Par ailleurs, les résultats de cette étude nous invitent à une réflexion supplémentaire, liée à l'expertise des sujets. Nous n'avons pas en effet distingué d'effet de l'expertise, entre autres car le public est composé d'étudiants. Il serait intéressant d'étendre cette étude à d'autres types de publics. Des professionnels de l'architecture et de l'acoustique, mais aussi à des musicologues, des musiciens, chefs d'orchestre, ingénieurs du son ou toute autre catégorie socio-professionnelle nécessitant le développement d'une « oreille ». L'influence des images pourrait être toute différente pour ces publics.

CONCLUSIONS

Cette étude est une première étape dans nos investigations liées au projet Auralias. Elle fournit des résultats intéressants pour le développement de notre application, ainsi que pour toute application multimodale visant à soutenir une tâche d'écoute. Nos prochaines expérimentations, à la suite de cette étude préalable, auront pour but d'investiguer spécifiquement la question du degré de réalisme des modèles visuels. Nous pensons que celui-ci est directement lié au sentiment d'immersion perçu dans la salle de simulation.

Cette étude, malgré ses limites, apporte un éclairage intéressant et expérimentalement fondé pour la suite du développement. Nous estimons qu'elle illustre bien la nécessité d'avoir certaines réflexions d'ordre fondamental dans des projets de nature appliquée.

REMERCIEMENTS

Le projet Auralias est un partenariat entre l'Université de Liège et l'Université Libre de Bruxelles, subventionné par la Région wallonne de Belgique (Direction Générale des Technologies, de la Recherche et de l'Energie), programme WIST 2 (Wallonie, informations, Société, Technologies), convention n°616416.

BIBLIOGRAPHIE

1. Beranek L. (1996). *Concert and opera halls : How they sound*. AIP Press.
2. J. Coutaz, L. Nigay, Les propriétés « CARE » dans les interfaces multimodales. *Actes de la conférence IHM'94*, Lille, 1994.
3. Dawans, A., Demaret, J.N., Safin, S. (2009) Principes de modélisation pour la simulation acoustique. *Actes de la conférence SCAN 2009*, Nancy
4. De Gelder, B. & Bertelson, P. (2003). Multisensory integration, perception and ecological validity. *Trends in Cognitive Sciences*, 7(10), 460-467.
5. Delorme, A. (2003). L'organisation perceptive In A. Delorme & M. Flückiger (Eds). *Perception et réalité. Introduction à la psychologie des perceptions*. De Boeck Université.
6. Embrechts J.J. (2000). Broad spectrum diffusion model for room acoustics ray-tracing algorithms. *Journal Acoustical Society of America* 107, 2068-2081.
7. Embrechts, J.J., Werner, N. and Lesoinne, S. (2005). Auralization in Room Acoustics using Directional Impulse Responses Computed by Sound Ray Techniques. *Proceedings of the 4th Forum Acusticum Congress*, Budapest, 2535-2538.
8. Krokstad A., Strom S., Sordal S. (1968). Calculating the acoustical room response by the use of a ray tracing technique. *Journal of Sound and Vibration* 8, 118-125.

9. Massaro, D. W. (2004). From multisensory integration to talking heads and language learning. In G. Calvert, C. Spence & B.E. Stein (Eds.), *The Handbook of Multisensory Processes*. Cambridge : MIT Press
10. Martin, J. C., 1997, TYCOON: Theoretical Framework and Software Tools for Multimodal Interfaces. Intelligence and Multimodality in Multimedia Interfaces, AAAI Press.
11. Reason, J. (1993). *L'erreur humaine*. Paris : Presses Universitaires de France.
12. Schroeder M.R., Gottlob D., Siebrasse K.F. (1974). Comparative study of european concert halls: correlation of subjective preference with geometric and acoustic parameters. Journal Acousical Society of America 56, 1192-1201.
13. Streri, A. (2003). L'intermodalité. In A. Delorme & M. Flückiger (Eds). *Perception et réalité. Introduction à la psychologie des perceptions*. De Boeck Université.
14. Vorländer, M. (2007). *Auralization : Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*. Berlin : Springer.