

# Les environnements sonores à la rescousse de la cognition

Isabelle Viaud-Delmon

IRCAM CNRS UMR 9912

Équipe Espaces Acoustiques et Cognitifs

1, place Igor Stravinsky

75004 Paris

E-mail : [Isabelle.Viaud-Delmon@ircam.fr](mailto:Isabelle.Viaud-Delmon@ircam.fr)

## Résumé

*Les environnements sonores réalistes dans des scènes de réalité virtuelle offrent de nouveaux moyens thérapeutiques pour certains troubles psychiatriques et perceptifs. Leurs avantages reposent en particulier sur la sollicitation conjointe de plusieurs modalités sensorielles, permettant d'améliorer l'immersion du participant dans l'univers virtuel et de faciliter l'apprentissage. Après avoir précisé les modes de fonctionnement de la réalité virtuelle et ses applications en mode purement visuel, nous examinerons comment l'audition interagit avec la vision, et comment, dès lors, ces interactions permettent d'envisager de nouveaux traitements par réalité virtuelle.*

## Réalité virtuelle et thérapeutique

La réalité virtuelle, un oxymore introduit par l'Américain Jaron Lanier au début des années 1980, se définit comme l'ensemble des techniques et des interfaces qui immergent un utilisateur dans un environnement artificiel et lui permettent d'interagir avec lui. La réalité virtuelle a très tôt été utilisée en psychiatrie pour traiter certains troubles dans lesquels l'évitement de situations caractéristiques handicape les patients. En 1994, l'équipe de Barbara Rothbaum aux États-Unis [1], a publié la première étude suggérant l'intérêt d'exposer des malades phobiques à des environnements virtuels. Deux groupes d'étudiants sensibles à l'acrophobie, la peur des lieux situés en hauteur, étaient comparés : les uns commençaient un traitement, tandis que les autres restaient en attente. Le traitement proposait, après une session de familiarisation avec le matériel, sept séances d'expositions progressives à différents environnements virtuels représentant des scènes de vide : un pont de hauteur variable au-dessus de l'eau, un balcon au rez-de-chaussée ou au 20<sup>e</sup> étage, ou encore un ascenseur panoramique. On évaluait le degré d'anxiété, le comportement d'évitement et l'attitude des sujets avant et après le traitement. Ces mesures ont montré une nette amélioration pour le groupe traité.

Depuis cette étude pionnière, de nombreuses recherches ont mis en évidence l'efficacité de la réalité virtuelle pour traiter différents troubles anxieux. L'acrophobie, la phobie de l'avion ou de la conduite, la phobie sociale (peur du regard des autres), la claustrophobie (peur de lieux clos), l'agoraphobie (peur des lieux ouverts et des foules) et l'arachnophobie (peur des araignées) sont couramment traitées de cette façon dans des centres médicaux aux États-Unis et dans certains pays européens. D'autres troubles, comme le stress post-traumatique, certains troubles des conduites alimentaires et diverses addictions au tabac, à l'alcool, aux drogues et au jeu, sont également abordés par des thérapies en réalité virtuelle.

Il est notable que les applications de la réalité virtuelle, qu'elles relèvent du jeu, de l'apprentissage ou de l'art, stimulent principalement la vision, même si un certain nombre de sons sont restitués pour créer une ambiance sonore. Toutefois, les stimulations virtuelles qui associent vision et audition, commencent à se révéler utiles pour traiter des pathologies psychiatriques, mais aussi des troubles de la perception. En associant plusieurs modalités sensorielles, la réalité virtuelle stimule de façon plus efficace l'intégration multisensorielle, qui potentialise les apprentissages et accroît les réponses émotionnelles.

Une des complications de la réalité virtuelle intégrant des environnements sonores 3D est qu'elle demande, au niveau technique, une maîtrise de la correspondance spatiale et temporelle entre le rendu des différentes informations sensorielles, sans quoi l'intégration multisensorielle du sujet percevant risque d'être perturbée.

## Intégration multisensorielle

Le cerveau perçoit généralement un même objet par plusieurs canaux sensoriels. La fusion de deux informations peut donner lieu à une interprétation perceptive qui ne se résume pas à une simple sommation des informations, mais qui produit des illusions, comme dans le cas de l'effet McGurk [2] (décrit par Grimault & Devergie page 35) : lorsque l'on voit une bouche prononcer la syllabe /ga/ en même temps qu'on entend un stimulus composé de la syllabe /ba/, on perçoit la syllabe /da/. Il s'agit d'un son intermédiaire du point de vue phonétique. Cette illusion reflète une stratégie employée par le cerveau pour combiner les informations visuelles et auditives afin d'aboutir à une perception unique et cohérente.

Imaginons maintenant une salle de conférence. Vous écoutez un séminaire donné par un conférencier équipé d'un micro. Alors qu'il n'y a pas de correspondance spatiale entre la direction de sa voix, émise par des haut-parleurs derrière vous ou sur le côté, et sa position dans l'espace, vous savez qu'il est bien l'orateur. Les informations auditives sont couplées automatiquement par votre cerveau aux informations visuelles. C'est l'effet du ventriloque : la voix du marionnettiste et la poupée en mouvement, bien que spatialement disjoints, sont fusionnées. Vient le temps des questions, diffusées par les mêmes haut-parleurs. Une personne dans la salle prend la parole, mais vous ne la voyez pas ; vous êtes incapable d'identifier qui parle puisque votre cerveau ne peut établir de correspondance spatiale entre la direction de la voix et un indice visuel. La personne se lève et c'est alors seulement que vous l'identifiez comme étant l'orateur, grâce aux informations visuelles liées à son mouvement ; vos informations auditives vous indiquent pourtant une autre direction.

Comment le cerveau trouve-t-il la correspondance auditive et visuelle appropriée pour déterminer qu'un son et une image proviennent de la même source ?

Plusieurs études ont éclairé les mécanismes en jeu. En particulier, Stein et Meredith [3] ont observé que les neurones d'une région cérébrale où se projettent certains axones du nerf optique, le colliculus supérieur, répondaient aussi bien à des informations visuelles qu'à des informations auditives. L'activité électrique de ces neurones, qualifiés de multisensoriels, devient beaucoup plus importante si les deux modalités sont stimulées à peu près au même moment et selon la même direction de l'espace. L'amplitude de cet effet dépend du degré de coïncidence spatiale et temporelle entre les informations.

Comme le système visuel est bien mieux équipé que le système auditif pour juger de la direction d'un stimulus, on pense généralement que la position d'un stimulus visuel dicte souvent la position apparente d'un stimulus auditif correspondant.

Cependant, plusieurs études suggèrent maintenant que l'estimation visuelle ne détermine pas systématiquement la perception résultante. Le cerveau serait capable de pondérer les informations sensorielles selon leur qualité. Pour parvenir à cette conclusion, Alais et Burr [4] ont demandé à des sujets d'estimer la localisation de stimuli unimodaux sonores (des clicks de 1,5 ms dont la perception de la localisation spatiale était contrôlée par les différences interaurales de temps entre les deux haut-parleurs de part et d'autre de l'écran) et visuels (des blobs Gaussiens). Ils ont ensuite proposé des stimuli bimodaux visuels et auditifs localisés à la même position, ou étant spatialement disparates. Les stimuli visuels variaient dans leur taille, rendant leur localisation plus incertaine. Quand le stimulus visuel était très précis, les chercheurs observaient une dominance visuelle ; lorsqu'il était moins précis, la perception auditive l'emportait. De plus, quand les stimuli auditifs et visuels étaient présentés conjointement et sans conflit spatial, les sujets discriminaient plus finement leurs positions qu'avec un seul sens. Cela indique que l'intégration des informations spatiales visuelles et auditives améliore l'estimation de la position du stimulus. Ces observations s'accordent avec l'hypothèse que le cerveau utilise une règle statistique de combinaison optimale, reposant sur la fiabilité relative des informations sensorielles, pour combiner les signaux provenant des différents sens et pour percevoir la localisation d'un objet multisensoriel.

Par ailleurs, nous avons pu montrer que pour reconnaître un objet réaliste, le son conjugué à la vision permettait un temps de réaction bien plus rapide que pour des objets abstraits.

L'expérience précédemment décrite et la plupart des expériences en neurosciences ont été conduites avec des stimuli n'ayant aucune signification. Cependant, la réalité virtuelle est utilisée afin de permettre l'interaction en temps réel avec des environnements présentant des caractéristiques sémantiques. Nous avons donc mené des expériences s'intéressant à la potentielle modification de l'intégration sensorielle par la valeur sémantique de l'objet sur lequel est testé l'intégration [5]. Celles-ci suggèrent que les sons provenant d'objets biologiques (animaux) sont traités plus rapidement que les sons d'objets non biologiques (moyens de transport). Enfin, en utilisant des objets significatifs, nous avons observé des temps de réaction extrêmement rapides signalant une importante facilitation due à une intégration visuoauditive. Cet effet d'intégration était nettement plus important que celui rapporté dans la littérature, et pourrait être dû au réalisme des stimuli et à l'immersion des sujets dans un environnement de réalité virtuelle [6].

Peut-on utiliser ces connaissances pour améliorer les applications médicales de la réalité virtuelle ?

La variété des systèmes sensoriels est actuellement peu représentée dans les mondes virtuels utilisés en médecine. Les sons y sont peu présents, et dépourvus des indices qui permettraient leur localisation à 360° et optimiseraient l'intégration cérébrale des informations visuelles et auditives. Cependant, des problèmes peuvent surgir si la cohérence subjective des signaux sonores et visuels est difficile à obtenir.

Dans ce cas, la nouvelle modalité sensorielle, au lieu d'enrichir l'intégration sensorielle, perturbe les sujets, car des informations non cohérentes créent une « surcharge cognitive ». De plus, les patients psychiatriques présentent souvent une hypersensibilité au bruit, alors que leur audition est généralement normale. Il est ainsi nécessaire de comprendre de quelles façons la modalité auditive peut être utilisée efficacement.

## Le traitement des phobies

Puisque la modalité auditive fournit constamment des informations sur le monde environnant et la façon dont nous y évoluons, son ajout aux mondes virtuels ne peut être que bénéfique. Effectivement, lors du travail avec des patients anxieux, cet apport permet d'améliorer leur sentiment d'être présent dans l'environnement virtuel et non plus dans l'environnement réel [7]. Pour cela, on crée des scènes sonores virtuelles en 3D qui utilisent la technique binaurale et ambisonique (voir le numéro 71 d'A&T).

Nous avons mis en place des environnements visuels et auditifs permettant d'étudier la phobie des chiens et d'en suggérer des traitements [8]. Nous avons pu observer que l'utilisation d'informations visuelles et auditives dans un environnement virtuel permet d'exposer un sujet de façon progressive au stimulus qu'il redoute. Le sujet doit naviguer dans plusieurs scènes virtuelles, où il peut voir ou entendre surgir des chiens.

Les sources sonores fournissent des indices précis sur la localisation des stimuli. Il peut par exemple percevoir des chiens arriver en courant derrière lui. Les personnes exposées à ces chiens virtuels en ont autant peur que des chiens réels, alors qu'avant l'expérience elles pensaient que des chiens virtuels ne pourraient pas les effrayer. De plus, la peur est plus intense lorsque le chien virtuel est vu et entendu que lorsque le chien est seulement vu ou entendu, démontrant l'effet de potentialisation de l'information lorsque celle-ci est présentée dans plusieurs modalités sensorielles simultanément.

Enfin, nos études ont démontré qu'une seule séance d'exposition à un environnement dans lequel les participants rencontrent des chiens virtuels qui émettent des aboiements et grognements améliorerait le contrôle des émotions négatives puisqu'à l'issue de la séance, les participants s'approchent plus des chiens virtuels qu'au début de la séance [9].

## Les acouphènes

Une autre application de l'intégration multisensorielle concerne le traitement de troubles perceptifs comme les acouphènes. Un acouphène est une perception sonore permanente ou intermittente (sifflement, bourdonnement) qui surgit indépendamment d'un son extérieur. Le plus souvent, on en ignore la cause et on sait mal le traiter. On considère que de 8 à 10% de la population adulte des pays développés vit avec des acouphènes.

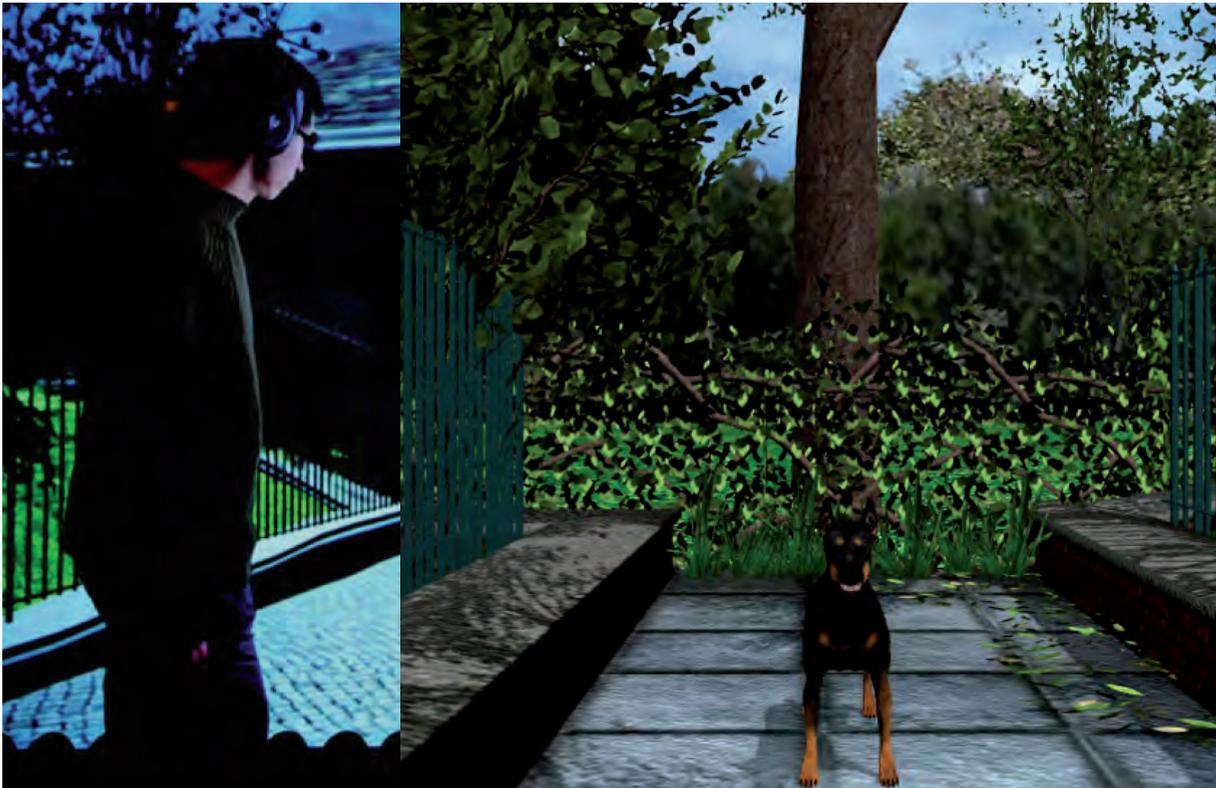


Fig. 1. Traiter la phobie des chiens. Le sujet est équipé de lunettes polarisantes pour voir en trois dimensions et d'un capteur de position de la tête. Les informations de position de sa tête permettent d'ajuster la stéréo visuelle et les indices de localisation des sons en fonction de ses mouvements. Le sujet doit naviguer dans plusieurs scènes virtuelles, où il peut voir ou entendre surgir des chiens.



**Fig. 2 : Acouphènes et réalité virtuelle.** Un sujet est équipé d'un visiocasque, d'écouteurs et d'une baguette équipée de capteurs. En bougeant la baguette, elle contrôle le bouquet d'étincelles (visible sur l'écran) auquel est relié un son qui reproduit celui de l'acouphène. La scène comporte d'autres objets sonores dont la localisation dans l'espace est asservie aux mouvements de la tête du sujet. Cela enrichit son environnement virtuel auditif et favorise son impression d'être immergé dans l'espace. En déplaçant la baguette, le patient apprend progressivement à maîtriser la localisation de son acouphène. Cette impression de le maîtriser fait qu'il perd son caractère agressif.

L'implication des structures centrales auditives, mais également extra-auditives, rend compte de la perception consciente de l'acouphène mais surtout de la gêne, parfois invalidante, ressentie par certains patients (hypersensibilité aux sons, troubles du sommeil, troubles attentionnels, réactions anxieuses et dépressives).

Si les acouphènes font le plus souvent suite à une lésion auditive périphérique responsable d'une hypoacousie, l'entretien du symptôme et ses caractéristiques psychoacoustiques peuvent être expliqués par une modification tonotopique de l'activité des voies auditives cortico-sous corticales [10]. Cette réorganisation résultant de phénomènes de plasticité neuronale (voir l'article de Schön et Tillmann page 18), a été mise en évidence expérimentalement. Cette situation clinique et son équivalent neurophysiologique sont à rapprocher de ce qui est rencontré dans le cadre des syndromes douloureux chroniques intervenant en particulier après amputation [11] où des thérapies utilisant les techniques de réalité virtuelle ont déjà été employées avec succès afin de diminuer les douleurs liées aux membres fantômes. Les acouphènes sont une sensation fantôme et à ce titre il est intéressant d'étudier si les patients souffrant d'acouphènes pourraient bénéficier de techniques de réalité virtuelle grâce auxquelles on leur permettrait de contrôler cette sensation, comme il est proposé pour les membres fantômes.

Dans ce but, une étude clinique a été menée à l'Hôpital européen Georges Pompidou pour la prise en charge des acouphènes unilatéraux.

La thérapie proposée consiste à favoriser la dissociation entre la perception de l'acouphène et sa représentation mentale en travaillant sur l'appropriation progressive par le patient de la localisation de l'acouphène dans l'espace de sorte à permettre, à terme, sa maîtrise au niveau émotionnel. En pratique, la méthode implique dans un premier temps de synthétiser un son reproduisant les caractéristiques auditives de l'acouphène perçu par le patient. Dans un second temps, l'acouphène synthétique est inséré dans un environnement de réalité virtuelle interactif où il est représenté visuellement par un bouquet d'étincelles que le patient manipule en temps réel de sorte à en contrôler la localisation dans l'espace. Les différentes composantes sonores de l'environnement virtuel (acouphène synthétique, ainsi que diverses sources ou ambiances sonores) sont spatialisées au moyen de techniques binaurales [11]. Nous pensons que le contrôle actif de l'acouphène synthétique rendu dans les modalités visuelles et auditives favorise l'apparition de la maîtrise de la perception acouphénique, lui enlevant son caractère aversif.

Le succès de la méthode repose sur la similitude spectrale perçue entre l'acouphène synthétique et l'acouphène du patient. Par ailleurs, la possibilité de contrôler la localisation de l'acouphène synthétique suppose qu'un phénomène de fusion intervienne entre ce dernier, présenté à l'oreille controlatérale du patient, et l'acouphène. Ce phénomène de fusion est facilité par l'intégration des différentes informations sensorielles sollicitées lors de l'immersion en réalité virtuelle.

## Le mot de la fin

L'ajout du son 3D aux environnements virtuels visuels permet d'agrandir le champ des applications thérapeutiques de la réalité virtuelle en stimulant adéquatement l'intégration multisensorielle. Cependant, il ne faut pas négliger le potentiel d'une réalité virtuelle purement auditive, où un rendu sonore est fait en temps réel selon des données obtenues par un capteur posé sur le corps. L'utilisation de sons 3D couplés à la proprioception offre des perspectives encore plus intéressantes sans doute que celles de la réalité virtuelle visuelle et auditive, car le rendu visuel impose de nombreuses limitations technologiques que l'on a déjà dépassées pour le rendu audio. Ajouter un retour sonore à ses actions permet de renforcer le fait de se sentir l'auteur de ses actions et le fait de se les attribuer renforce le sens de la possession du corps, qui est atteint dans de nombreux troubles comme la schizophrénie mais également suite à certaines lésions cérébrales. De plus, le retour sonore d'une action est un aspect important de l'apprentissage moteur. L'avenir des applications de réalité virtuelle pour la santé pourrait donc relever principalement de réalité virtuelle acoustique.

## Références bibliographiques

- [1] Rothbaum B. O., Hodges L. F., Kooper R., Opdyke D., Williford J. S., North, M. «Effectiveness of computer-generated (virtual reality) graded exposure in the treatment of acrophobia.» *American Journal of Psychiatry*, 152 (4), pp 626-628, 1995.
- [2] McGurk H. & MacDonald J. «Hearing lips and seeing voices». *Nature*, 264 (5588), pp 746-748, 1976.
- [3] Stein B. & Stanford T. «Multisensory integration : current issues from the perspective of the single neuron». *Nature Rev. Neurol.*, 9 :255-266, 2008.
- [4] Alais D. & Burr D. «The ventriloquist effect results from near-optimal bimodal integration.» *Curr Biol*. 14(3) pp 257-262, 2004.
- [5] Sued C. & Viaud-Delmon I. «Auditory-visual object recognition time suggests specific processing for animal sounds.» *PLoS One*. 2009;4(4):e5256. doi: 10.1371/journal.pone.0005256, 2009.
- [6] Sued C., Bonneel N. & Viaud-Delmon I. «Integration of auditory and visual information in the recognition of realistic objects.» *Experimental Brain Research*, 194(1) pp 91-102, 2009.
- [7] Viaud-Delmon I., Warusfel O., Seguelas A., Rio, E., Jouvent R. «High sensitivity to multisensory conflicts in agoraphobia exhibited by virtual reality.» *European Psychiatry*, 21 (7), pp 501-508, 2006.
- [8] Sued C., Drettakis G., Warusfel O., Viaud-Delmon I. «Auditory-visual virtual reality as a diagnostic and therapeutic tool for cynophobia.» *Cyberpsychol Behav Soc Netw.*;16(2) pp 145-152, 2013.
- [9] Taffou M., Guerchouche R., Drettakis G., Viaud-Delmon I. «Auditory-visual aversive stimuli modulate the conscious experience of fear.» *Multisensory Research*, (2013).
- [10] Eggermont J. J. & Roberts L. E. «The neuroscience of tinnitus.» *Trends in Neurosciences* 27(11) pp 676-682, 2004.
- [11] Møller A. R. «Tinnitus and pain.» *Prog Brain Res*. 166 pp. 46-53, 2007.
- [12] Londero A., Viaud-Delmon I., Baskind A., Delerue O., Bertet S., Bonfils P., Warusfel O. «Auditory and visual 3D virtual reality therapy for chronic subjective tinnitus: theoretical Framework.» *Virtual Reality*, 14 (2), pp 143-151, 2010.

