

# *L'étrangeté perceptive en réalité virtuelle*

**Résidence en recherche artistique à l'IRCAM en 2019-2020**

**Vincent Isnard & Trami Nguyen**

## **Résumé**

Les vidéos à 360° en réalité virtuelle (RV) présentent certaines limites, comme l'impossibilité d'effectuer des déplacements dans l'environnement par la volonté propre du spectateur ni des interactions physiques directes avec des objets virtuels. La voix du spectateur pourrait être un moyen d'augmenter cette interactivité puisqu'elle ne nécessite pas de jeu physique mais pourrait suffire à produire des conséquences physiques dans le monde virtuel. Dans cette recherche artistique élaborée dans le cadre d'une résidence à l'IRCAM en 2019-2020, nous avons cherché à évaluer la valeur ajoutée d'interactivité lorsque le spectateur utilise sa propre voix, en incluant des transformations en temps réel de timbre et de spatialisation pour l'intégrer dans un scénario au contexte futuriste à travers un dialogue avec une intelligence artificielle ayant pris forme humaine. En effet, un test de Turing inversé fictif sert de prétexte à cette interaction dialoguée et doit permettre d'évaluer de manière fictive notre propre degré d'humanité. Un test scientifique sur la perception en RV est réalisé en parallèle à cette interaction afin d'évaluer si la qualité de la voix du spectateur transformée en temps réel lui permet d'incarner davantage son personnage dans la fiction en RV, en jouant sur l'effet d'étrangeté que ces transformations peuvent générer.

## **Préambule**

Cet article détaille les enjeux scientifiques soulevés dans le cadre d'une expérience de perception multisensorielle en RV, les propositions artistiques relevant d'un contexte futuriste et anthropocénique insérées dans une narration en RV, ainsi que les choix technologiques (vidéo à 360° immersive, son 3D ambisonique et binaural) effectués lors de notre résidence en recherche artistique à l'IRCAM en 2019-2020. L'installation développée en RV à l'issue de la résidence et présentée lors du Forum IRCAM du 4 au 6 mars 2020 est un premier aboutissement du projet. Celui-ci continue à être développé actuellement, d'une part dans un cadre scientifique (test perceptif), d'autre part dans un cadre artistique (installation interactive et film autonome en RV), non pas de manière cloisonnée mais par une émulation forte et inspirante entre science et art.

## **Le test de Pieter Musk**

### *Enjeux scientifiques et technologiques*

Trois grands enjeux à la fois scientifiques et technologiques ont nourri l'expérience perceptive développée en RV : l'étrangeté perceptive, l'adaptation perceptive de contenus sonores et visuels en RV et l'utilisation de la voix pour

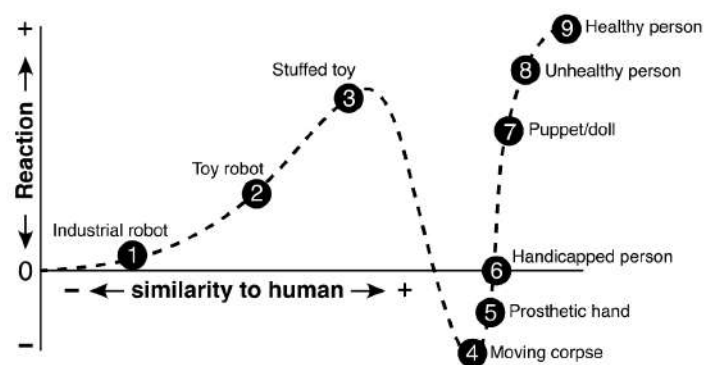
interagir en RV. Trois axes de recherche en ont découlé et nous ont inspiré alternativement ou simultanément lors du développement de notre installation.

### L'étrangeté perceptive :

L'étrangeté perceptive est un concept dont les enjeux se font de plus en plus prégnants à l'heure actuelle de l'avènement des assistants vocaux, des agents conversationnels incarnés ou encore des robots.

Pour saisir facilement ce concept, on peut par exemple s'appuyer sur la littérature fantastique : dans *Le marchand de sable* (1815) d'E.T.A. Hoffmann, un jeune garçon est épris d'une jeune fille. Cependant, il la trouve troublante sous bien des aspects (froideur au contact de sa peau, visage impassible...). A la fin de cette nouvelle, il réalisera que la jeune fille n'était pas un être humain mais un robot développé à la perfection par son "père" physicien. L'étrangeté ou l'inquiétude que peut provoquer cette familiarité humaine a été étudiée par la suite notamment en psychologie (e.g. Freud, 1919).

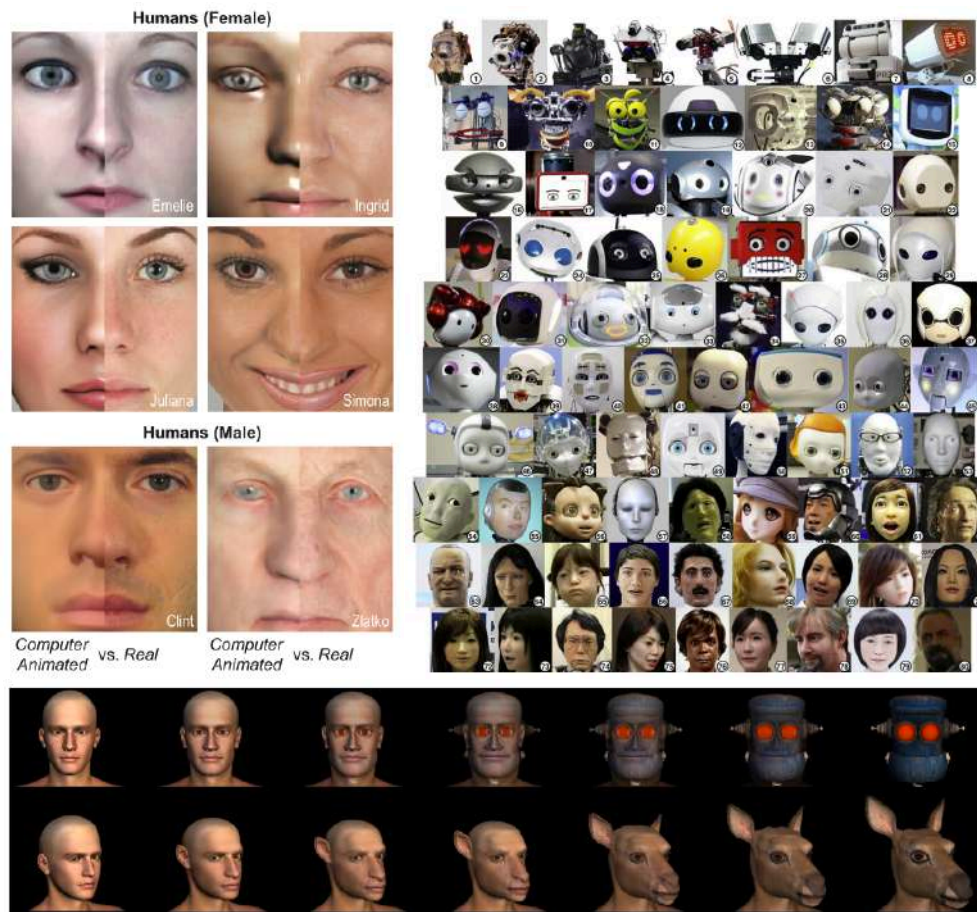
En 1970, le roboticien Mori propose l'hypothèse de la vallée de l'étrangeté (*uncanny valley*). Selon son hypothèse, moins une entité est similaire à un être humain, moins la réaction qu'elle provoque est forte (e.g. réaction émotionnelle...), et inversement, plus la similarité augmente, plus la réaction est forte. Mais cette relation n'est pas linéaire : elle présente un creux appelé "vallée de l'étrangeté" (Fig. 1). Ce creux indique que lorsque la similarité est assez forte mais néanmoins imparfaite, l'entité quasi-humaine peut provoquer une réaction très négative (e.g. aversion). Les interprétations pour expliquer une telle réaction négative sont multiples : on peut se demander si la personne est morte, sujette à un pathogène, à une absence perturbante de défauts, ou encore d'un point de vue cognitif, il pourrait s'agir d'un conflit entre les indices perceptifs mis en jeu, on croirait reconnaître un être humain mais on n'en serait pas sûr et notre système perceptif s'en trouverait perturbé.



**Figure 1. Vallée de l'étrangeté.** D'après Mathur & Reichling (2016).

Un certain nombre d'auteurs ont testé cette hypothèse dans un cadre scientifique, en particulier pour évaluer si des robots androïdes, généralement trop lisses pour être véritablement humains, pouvaient générer ce type de réactions négatives, ce qui s'avèrerait très dommageable pour l'industrie robotique (Fig. 2).

Le scénario artistique de notre installation inclut une conversation avec une intelligence artificielle incarnée. Il s'agira de trouver la meilleure représentation visuelle et auditive de cette entité pour la rendre la plus crédible et pertinente possible dans le cadre de notre fiction, potentiellement en jouant sur un effet d'étrangeté chez le participant.



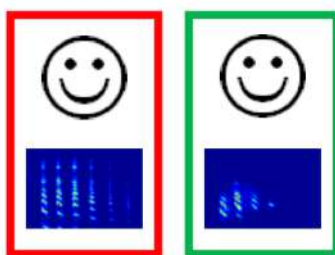
**Figure 2. Exemples d'études récentes sur la perception de l'étrangeté.** Dans chacune de ces études, il s'agit d'évaluer la réaction que provoque une entité androïde comparée à un être humain. D'après : à gauche : Chattopadhyay & MacDorman (2016) ; à droite : Mathur & Reichling (2016) ; en bas : Ferrey et al. (2015).

### Cohérence sonore et visuelle des contenus artistiques en RV :

La RV nécessite à l'heure actuelle des ordinateurs puissants et coûteux. En particulier, comme pour les jeux vidéo, les rendus visuels par synthèse d'environnements virtuels sont complexes à mettre en œuvre (c'est moins vrai pour les captations en 360° immersive), tandis que le son est généralement beaucoup plus simple à générer, transformer et diffuser en RV pour obtenir un rendu cohérent avec l'intention artistique. Le défaut de réalisme de l'environnement virtuel n'est pas forcément un problème en soi, car la narration suffit souvent à être pleinement immergé dans l'expérience en RV. L'enjeu réside donc surtout dans la cohérence de l'environnement virtuel visuel et auditif avec l'intention artistique.

Cependant, du fait de la forte divergence des processus de créations en image et son, on risque de générer une incongruence perceptive gênante en RV (le cas limite étant un rendu visuel non-réaliste, car complexe à mettre en œuvre, associé à un contenu sonore réaliste, car obtenu et manipulé facilement). C'est pourquoi, au lieu de proposer une complexification du processus de création de l'image, souvent aux dépens de l'intention artistique, on peut proposer une "dégradation" du contenu sonore pour tendre vers une meilleure congruence perceptive visuo-auditive et une meilleure expérience immersive en RV.

Dans un travail antérieur, nous avons proposé des "esquisses auditives" de sons complexes comme pendant d'esquisses visuelles, créées sur la base des indices de reconnaissance auditive (Fig. 3 ; cf. Isnard, 2016 ; Isnard et al., 2016).



**Figure 3. Représentation schématique de l'incongruence (à gauche) et de la congruence (à droite) visuo-auditive qui peuvent être générées en fonction du niveau d'adaptation des contenus image et son.** Peu de traits visuels suffisent à reconnaître un objet visuel (ici un visage). De la même manière, le système auditif humain peut se contenter de peu de traits auditifs pour reconnaître un son. On fait l'hypothèse que la congruence visuo-auditive est meilleure dans le cas où les complexités des objets visuel (e.g. visage) et auditif (e.g. voix) sont adaptées.

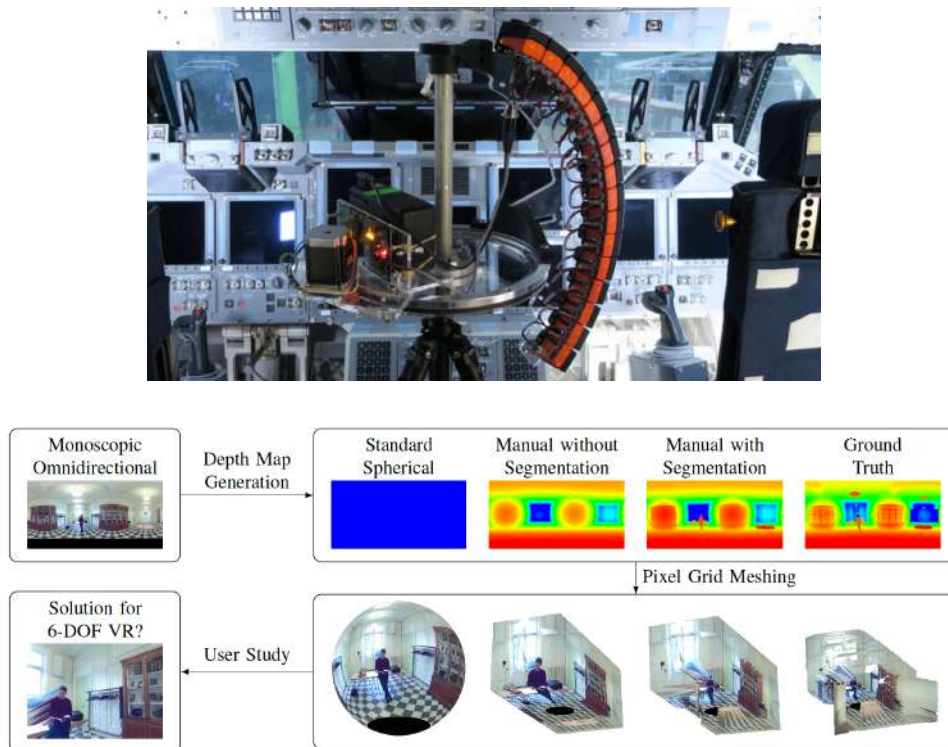
Dans le cadre de notre installation en RV, les contenus image et son seront transformés pour correspondre à l'intention artistique et au contexte futuriste et anthropocénique du scénario. Il s'agira d'abord d'adapter l'image et le son en correspondance pour favoriser la congruence visuo-auditive et une meilleure expérience d'immersion en RV.

#### L'utilisation de la voix pour interagir en RV :

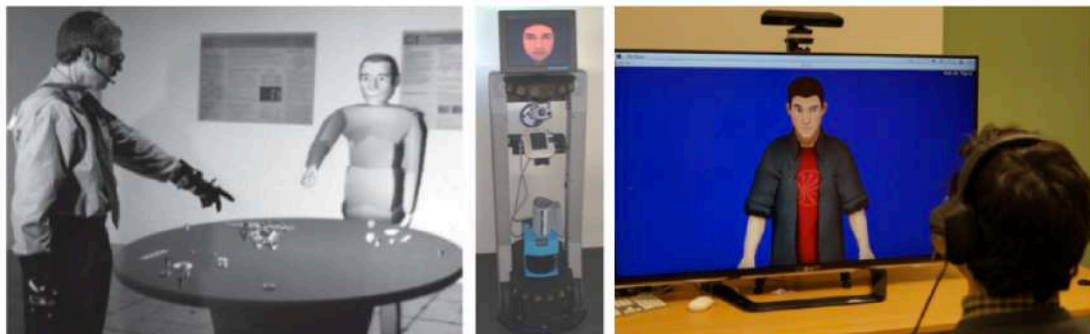
Les vidéos à 360° en RV ne permettent ni déplacement ni interaction. L'image est captée en 360° et le participant qui la visualise peut seulement tourner sur 360° pour observer toute la scène immersive. Son intérêt est qu'elle est relativement simple à mettre en œuvre (en comparaison au développement d'un environnement de synthèse en 3D) et qu'elle permet d'obtenir une qualité parfaitement réaliste comme matière brute avant traitements.

Certaines propositions sont en développement pour pallier à ces limites. Par exemple des réseaux de caméras *light fields*, ou encore une reconstitution de l'effet de parallaxe sur une image initialement monoscopique à l'aide de techniques computationnelles (Fig. 4). On peut cependant également s'inspirer de l'interactivité proposée par les "agents virtuels incarnés", pour lesquels la voix, les gestes ou le regard du participant peuvent être impliqués dans une interaction avec ce type d'agents virtuels (Fig. 5).

Dans notre installation, nous avons opté pour l'utilisation de la voix du participant, qui représente selon nous une manière peu coûteuse, simple et efficace pour améliorer l'interactivité en RV à 360°. De plus, nous proposons d'ajouter des transformations en temps réel de la voix propre du participant pour lui permettre d'incarner au mieux un personnage de fiction (par exemple, si le participant devait incarner un monstre dans un jeu vidéo, on lui proposerait de transformer sa propre voix en temps réel en voix de monstre pour qu'il incarne au mieux son personnage). Cependant, on pourra se demander si de telles transformations ne seront pas susceptibles de générer un effet d'étrangeté chez le participant et si la congruence visuo-auditive sera toujours respectée.



**Figure 4. Exemples de techniques pour proposer un déplacement du corps en RV à 360°.** En haut : réseau de caméras *light fields* (Google) ; en bas : processus de synthèse de l'effet de parallaxe sur une image monoscopique (d'après Dinechin & Paljic, 2018).



**Figure 5. Exemples d'agents conversationnels incarnés.** D'après : à gauche : Kopp et al. (2003) ; au milieu : Tamagawa et al. (2011) ; à droite : Baur et al. (2013).

### *Scénario artistique*

A l'origine, le test de Turing a été proposé par le célèbre mathématicien pour déterminer si une entité donnée est un être humain ou une machine (du type intelligence artificielle), en supposant qu'avec l'amélioration des technologies la frontière entre les deux deviendrait d'autant plus ténue et que les machines pourraient finir par se faire passer pour des humains en imitant certaines de nos capacités cognitives. Le test consiste essentiellement à poser des questions à cette entité inconnue via une interface. En fonction des réponses données, on doit en déduire si cette entité est un être humain ou une machine.

En partant de cette source d'inspiration, l'écrivain de science-fiction Philip K. Dick, dans *Les androïdes rêvent-ils de moutons électriques ?* (1966 ; repris au cinéma

sous le titre de *Blade runner* par Ridley Scott en 1982, cf. Fig. 6), a imaginé le test fictif de Voight-Kampff conçu pour permettre à la police de retrouver et démasquer des androïdes évadés (les "répliants"), indiscernables sans cela de la population humaine. Ce test consiste à poser des questions dérangementantes pour examiner si elles provoquent des réactions émotionnelles chez le sujet, qui n'existent pas chez les répliants.



**Figure 6. Affiche du film "Blade runner".**

Aujourd'hui, c'est cependant moins la question des machines qui se feraient passer pour des humains qui semble vraiment problématique, dans la mesure où c'est toujours l'humain (pour l'instant) qui contrôle la machine et qui cherche à améliorer cette imitation pour le meilleur et pour le pire, que la problématique des humains qui se transformeraient progressivement en machines par le biais de diverses augmentations technologiques. On pense par exemple à l'ensemble de nos assistants électroniques (à commencer par Internet) jusqu'aux prothèses augmentées (cf. Frischmann & Selinger, 2018).

Pour notre installation, nous avons donc imaginé un test subversif pour déterminer si le participant ne serait pas lui-même, à un certain degré, une machine. Dans l'expérience en RV, le test est proposé par notre personnage de fiction Pieter Musk, une intelligence artificielle qui se présente comme le fils d'Elon Musk, le célèbre entrepreneur. À l'aide de son test de Turing inversé, Pieter Musk cherche à identifier le spectateur, en déterminant son degré d'humanité, pour lui permettre ou non d'accéder aux données confidentielles de son père. Les questions sont volontairement dérangementantes pour susciter une réaction émotionnelle. Un exemple : "Votre enfant de 7 ans rentre à la maison avec un bocal rempli de grenouilles mortes [...]. Il vous tend également le couteau encore ensanglanté qui lui a servi à découper les grenouilles [...]. Que lui dites-vous ? Réponse A : merveilleux ! Je te débarrasse de tout ça [...]. Réponse B : vous faites comme si de rien n'était [...]. Réponse C : vous roulez des yeux, pris de vertige [...]."

Dans l'installation, les réponses données par le participant font partie de la fiction et ne rentrent pas en compte dans l'analyse scientifique.

### ***Protocole expérimental du test scientifique***

Pour tester l'interactivité et l'immersion du participant dans l'installation en RV, plusieurs paramètres sont modifiés successivement au cours du test :

- le son traité en temps réel en timbre (voix humaine ou voix robotique) et en spatialisation (voix colocalisée avec la source vocale ou délocalisée) ;
- l'image traitée de manière correspondante, respectivement en distorsion et en dissociation RGB.

Pour chaque question fictive que lui pose Pieter Musk, le participant doit répondre à voix haute. A la suite de quoi, il effectue une évaluation perceptive sur une échelle visuelle pour déterminer si les traitements sur sa propre voix favorisent ou non son interaction en RV.

Nous faisons l'hypothèse que des traitements congruents entre l'image et le son, et entre l'environnement de fiction (futuriste) et la voix du participant (rendue robotique), favorisent cette interaction et améliorent l'interactivité.

L'ensemble de l'expérience en RV dure environ 30 min.

### ***Conception de l'installation en RV***

L'acteur choisit pour cette expérience est Piersten Leirom, performeur et danseur. Le tournage a eu lieu dans un studio à l'IRCAM en 2018. Le matériel de tournage a été le suivant (cf. Fig. 7) :

- pour l'image, nous avons utilisé une caméra 360° Insta Pro 2 (en location) qui présente l'intérêt d'avoir une résolution 8k, une gestion à distance du ventilateur (à couper pour limiter le bruit dans la prise de son) et de la capture vidéo, ou encore un stitching automatisé avant importation des images 360° sur PC ;

- le son a été enregistré en ambisonique à l'aide d'un microphone Eigenmike 32 capsules (appartenant à l'IRCAM). A noter que des microphones plus accessibles existent comme le Zoom H3-VR ou le Zylia ZM-1 ; de même pour l'image avec une large gamme de caméras grand public.

Pour rappel, le son ambisonique (obtenu par enregistrement ou synthèse) peut être décodé sur tout type de système de restitution sonore (dôme ambisonique, système 5.1, etc.), et notamment en binaural, c'est-à-dire en son 3D dans un casque audio quelconque, en conservant toute l'information de spatialisation de la scène sonore originale.

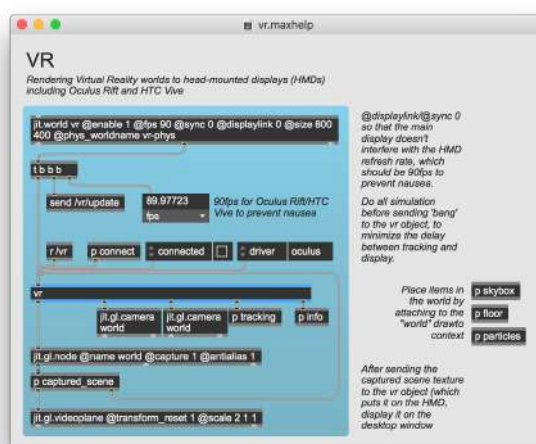


**Figure 7. Matériel de tournage.** A gauche : caméra 360° Insta Pro 2 ; à droite : microphone ambisonique Eigenmike 32 capsules.

Pour le montage, nous avons utilisé Adobe Premiere Pro qui prend en charge les images de l'Insta Pro 2. Pour le son, nous avons utilisé Reaper qui gère facilement des fichiers comportant un grand nombre de canaux. Les cuts de début et de fin de chaque plan ont été ajustés grâce aux claps effectués au tournage et aux timecodes.

A noter qu'il existe la panoplie d'outils (gratuits) Facebook 360 Spatial Workstation permettant de produire de la RV. Le Spatialiser permet de gérer du son spatialisé dans Reaper (ou d'autres DAW) avec un monitoring vidéo. L'Encoder permet de combiner une image de RV avec un son spatialisé en un seul fichier vidéo. Nous avons tout de même opté pour effectuer cet "encodage", du moins la lecture simultanée de l'image et du son, dans Max 8 (cf. paragraphe suivant), notamment car ces outils ne permettent pas actuellement de gérer des fichiers de haute-définition spatiale (8k pour l'image, 32 canaux pour le son).

La lecture et les traitements temps réel de l'image et du son ont donc été effectués dans Max 8. La connexion avec notre casque de RV, un Oculus Rift, a été rendue possible grâce à la bibliothèque "vr" développée par Graham Wakefield (Fig. 8 ; l'Oculus Rift n'est pas le seul casque pris en charge par cette bibliothèque). Cette bibliothèque est extrêmement pratique et efficace car elle permet de récupérer toutes les données spatiales du casque Oculus mais également des manettes. Et elle permet évidemment l'affichage d'une image en RV dans le casque de RV.



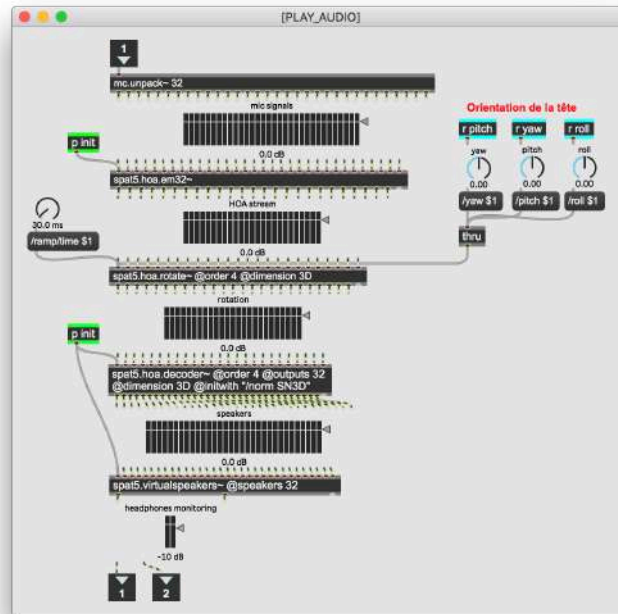
**Figure 8. Affichage de l'aide de la bibliothèque "vr" dans Max 8 permettant la connexion à l'Oculus Rift et l'affichage d'une image de RV dans le casque de RV.**

Pour le son spatialisé, nous avons utilisé la bibliothèque "spat" développée au sein de l'équipe Espaces Acoustiques et Cognitifs de l'IRCAM. Cette bibliothèque est particulièrement complète et flexible à utiliser pour tous les aspects de son 3D qui existent à l'heure actuelle (Fig. 9).

Au bout du compte, la restitution de la vidéo est effectuée à l'aide du casque de RV dans lequel on affiche l'image à 360°, tandis que le son est restitué sur le casque audio en binaural. Le participant équipé du matériel de RV peut observer la scène immersive tout autour de lui en tournant la tête et le corps, et les rendus visuels et auditifs sont alors restitués de manière cohérente et actualisés en temps réel (rotations simultanées de l'image 360° et de la scène sonore 3D).

Enfin, pour améliorer l'interactivité, la voix propre du participant est captée à l'aide d'un microphone serre-tête pour être traitée en temps réel à travers des vocoders pour générer un timbre robotique, ainsi qu'en spatialisation à nouveau à l'aide du Spat. Cependant, nous avons préféré ne pas modifier l'évolution du scénario de la fiction en fonction des réponses prononcées par le participant pour ne pas perturber ni alourdir le test scientifique. Les questions et réponses s'enchaînent donc suivant un ordre pré-établi.





**Figure 9. Conversion ambisonique vers binaural à l'aide de la bibliothèque "spat" dans Max 8 pour une lecture en RV du son spatialisé sur casque audio.** Les coordonnées spatiales du casque de RV sont transmises à l'objet du Spat pour la rotation de la scène sonore captée en ambisonique 32 canaux avant conversion en binaural.

A noter, pour ce dernier aspect, que lorsqu'on parle dans un microphone et qu'on s'écoute dans un casque audio, on entend notre voix telle que peuvent l'entendre des personnes autour de nous. Cependant, cela ne correspond pas au timbre que nous entendons nous-même, car nous entendons deux flux sonores distincts du timbre qu'entendent des personnes autour de nous (ou tel qu'il est capté par le microphone) : d'une part le son qui sort de notre bouche est filtré par notre tête avant d'atteindre nos oreilles, d'autre part le son qui est produit par nos cordes vocales passe également, par un deuxième trajet, directement par conduction osseuse dans nos oreilles avec un filtrage spécifique. Nous avons donc effectué un filtrage global simulant ces deux filtrages concomitants, tel que proposé par Porschmann (2000). Il s'agit globalement d'un filtrage des hautes-fréquences au-dessus de 5 kHz, qui permet donc de s'entendre à travers le microphone et le casque audio comme lorsqu'on s'entend soi-même parler naturellement sans tout le matériel sollicité ici pour la RV.

### *Premiers résultats du test scientifique*

Treize participants ont accepté de passer l'expérience et d'effectuer l'évaluation perceptive sur l'interaction et l'immersion en RV avec un dispositif simulant un dialogue. Globalement, sur l'ensemble des conditions expérimentales, les participants ont noté une bonne interaction avec leur propre voix (notée environ 4.5/7 en moyenne sur toutes les conditions confondues).

Cependant, les résultats étaient peu variables en fonction des conditions expérimentales, qui correspondaient à une variation des transformations temps réel. Pour la suite de l'étude, il s'agira donc d'élargir la gamme des effets pour tenter d'observer des résultats davantage différenciés en fonction des conditions.

Par ailleurs, les résultats étaient très variables en fonction des participants. Il semblerait que les participants plus familiers avec la RV appréciaient davantage l'expérience, probablement car maîtrisant mieux le système ils pouvaient se concentrer d'autant plus sur l'interaction vocale. Il s'agira donc pour la suite de l'étude de prendre en compte plus systématiquement les différents profils des participants (naïfs, joueurs de jeux vidéo, etc.) et de proposer une étape de familiarisation ou une comparaison voix naturelle vs. voix transformée pour mieux rendre compte de l'intérêt du dispositif le cas échéant.

Enfin, les participants nous ont indiqué en commentaires généraux qu'ils avaient globalement beaucoup apprécié l'expérience, l'originalité du scénario et l'interaction vocale. Cette expérience a généré relativement peu de symptômes de cybercinétose malgré sa durée assez longue (30 min ; il est généralement recommandé de limiter une expérience en RV à environ 20 min maximum), sans doute car les mouvements de l'image et du son étaient assez limités.

### ***Bilan***

Les retours sur cette installation en RV ont été très positifs et les premiers résultats du test scientifique très encourageants. Un tel dispositif est complètement fonctionnel pour effectuer des tests perceptifs avec des composantes artistiques qui fournissent de la matière aux problématiques scientifiques. Plusieurs améliorations sont néanmoins envisagées, en particulier : les conditions de transformations temps réel proposées, la variété des interactions vocales pour évaluer à partir de quel moment et jusqu'à quel point on apprécie l'effet rendu par les transformations sur notre propre voix, tout en remaniant le protocole expérimental pour restreindre le temps de passage en RV. De plus, l'ensemble du dispositif et des résultats perceptifs obtenus permettront de nourrir notre réflexion pour la suite de notre travail sur ce dispositif : d'une part le prolongement du test scientifique, d'autre part la production d'un film autonome en RV fortement inspiré de cette installation.

### **Remerciements**

Nous tenons à remercier l'ensemble des personnes qui ont été impliquées dans ce projet : Piersten Leirom ; Isabelle Viaud-Delmon, Olivier Warusfel et toute l'équipe Espaces Acoustiques et Cognitifs de l'IRCAM ; Jérémie Bourgogne, Cyril Claverie et l'ensemble de la Production de l'IRCAM ; Greg Beller, Markus Noisternig, Paola Palumbo et l'ensemble du département IRC de l'IRCAM ; Sebastian Rivas, Anouck Avisse et l'équipe du GRAME pour une résidence de travail effectuée au GRAME en 2019 et complémentaire à celle de l'IRCAM.

### **Références**

- Baur, T., Damian, I., Gebhard, P., Porayska-Pomsta, K., & André, E. (2013). A job interview simulation: Social cue-based interaction with a virtual character. In *2013 International Conference on Social Computing* (pp. 220-227). IEEE.
- Chattopadhyay, D., & MacDorman, K. F. (2016). Familiar faces rendered strange: Why inconsistent realism drives characters into the uncanny valley. *Journal of Vision*, 16(11), 7-7.

- Dick, P. K. (1979). *Les Androïdes rêvent-ils de moutons électriques?*. JC Lattès.
- de Dinechin, G. D., & Paljic, A. (2018). Cinematic virtual reality with motion parallax from a single monoscopic omnidirectional image. In *2018 3rd Digital Heritage International Congress (DigitalHERITAGE) held jointly with 2018 24th International Conference on Virtual Systems & Multimedia (VSMM 2018)* (pp. 1-8). IEEE.
- Ferrey, A. E., Burleigh, T. J., & Fenske, M. J. (2015). Stimulus-category competition, inhibition, and affective devaluation: a novel account of the uncanny valley. *Frontiers in psychology*, 6, 249.
- Freud, S. (1919). *L'inquiétante étrangeté et autres essais* ([1985] éd.). Paris: Folio.
- Frischmann, B., & Selinger, E. (2018). *Re-engineering humanity*. Cambridge University Press.
- Hoffmann, E. T. A. (1815). *Le marchand de sable*.
- Isnard, V. (2016). *L'efficacité du système auditif humain pour la reconnaissance de sons naturels* (Doctoral dissertation, Paris 6).
- Isnard, V., Taffou, M., Viaud-Delmon, I., & Suied, C. (2016). Auditory sketches: very sparse representations of sounds are still recognizable. *PloS one*, 11(3).
- Kopp, S., Jung, B., Lessmann, N., & Wachsmuth, I. (2003). Max-a multimodal assistant in virtual reality construction. *KI*, 17(4), 11.
- Mathur, M. B., & Reichling, D. B. (2016). Navigating a social world with robot partners: A quantitative cartography of the Uncanny Valley. *Cognition*, 146, 22-32.
- Pörschmann, C. (2000). Influences of bone conduction and air conduction on the sound of one's own voice. *Acta Acustica united with Acustica*, 86(6), 1038-1045.
- Tamagawa, R., Watson, C. I., Kuo, I. H., MacDonald, B. A., & Broadbent, E. (2011). The effects of synthesized voice accents on user perceptions of robots. *International Journal of Social Robotics*, 3(3), 253-262.

## Présentation des auteurs

### *Vincent Isnard*

Docteur de l'IRCAM en neurosciences spécialisé en perception auditive et titulaire de trois Masters en technologies du son et d'une Licence en philosophie, Vincent Isnard est chercheur, ingénieur du son et réalisateur en informatique musicale. Ses pratiques musicales contemporaines se sont également développées dans les classes de Laurent Durupt et Denis Dufour au Conservatoire.

### *Trami Nguyen*

Pianiste, performeuse et artiste visuelle, Trami Nguyen est diplômée d'un master de la HEM de Genève. Co-fondatrice de l'Ensemble Links, elle défend des répertoires contemporains et des créations de concerts participatifs, scénographiés, immersifs et/ou multidisciplinaires. Ses projets visuels s'articulent autour de performances réalisées en Europe et s'étendent au domaine de la réalisation en réalité virtuelle.