



HAL
open science

Interagir dans la réalité virtuelle : au-delà de l'effet “ wow ”, étude des technologies immersives sur l'embodiment en situation d'utilisation d'une L2

Virginie Privas-Bréauté, Maud Ciekanski

► To cite this version:

Virginie Privas-Bréauté, Maud Ciekanski. Interagir dans la réalité virtuelle : au-delà de l'effet “ wow ”, étude des technologies immersives sur l'embodiment en situation d'utilisation d'une L2. Drôles d'objets: un nouvel art de faire, Oct 2021, La Rochelle, France. 10.5281/zenodo.6061164 . hal-03592646

HAL Id: hal-03592646

<https://hal.univ-lorraine.fr/hal-03592646>

Submitted on 1 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Interagir dans la réalité virtuelle : au-delà de l'effet « wow », étude des technologies immersives sur l'*embodiment* en situation d'utilisation d'une L2.

Virginie Privas-Bréauté et Maud Ciekanski

ATILF (UMR 7118) - Université de Lorraine – CNRS

Résumé. Le groupe de recherche VRALL (*Virtual Reality Assisted Language Learning*)¹ questionne le potentiel de la réalité virtuelle (RV) dans les apprentissages de langue, en situations formelles ou informelles. L'une des perspectives adoptées pour comprendre les apprentissages avec la RV est celle de l'*embodiment* (ou cognition incarnée) qui explique l'accès à la connaissance comme le résultat de relations d'influences réciproques entre le corps, les émotions et l'environnement. Elle s'appuie pour partie sur les neurosciences cognitives (le paradigme de l'enaction de Varela) et affectives et sociales (le paradigme de la pensée émotionnelle d'Immordino-Yang). Sont présentés les résultats d'une étude menée auprès d'étudiants de Master2 (master MEEF² et master FLE³) qui ont testé, dans le cadre d'un atelier de découverte, différentes immersions virtuelles avec des technologies 2 et 3D, impliquant divers types d'interaction en L2. La méthodologie compréhensive et qualitative mise en œuvre s'appuie sur une approche phénoménologique des expériences vécues, recueillies par divers instruments de recherche. Les résultats montrent une différence dans l'adhésion des participants à la RV, en fonction non seulement des technologies utilisées, des scénarii, des ressources mais aussi de la nature multimodale des ressources. Ils indiquent que la nature extrême-sensorielle de ces ressources peut provoquer une surcharge cognitive et émotionnelle entravant l'attention des apprenants et leur apprentissage.

Mots-clés. Réalité virtuelle, *embodiment*, avatar, technologies 2D et 3D, apprentissage des langues en situation informelle.

Introduction

Cette étude s'inscrit dans une réflexion générale sur le renouvellement des situations d'apprentissage des langues étrangères et cherche à comprendre les effets potentiels de la réalité virtuelle (RV) sur le développement des compétences langagières en langue-cible, à travers les interactions produites dans l'environnement virtuel. L'immersion est généralement perçue comme l'expérience la plus favorable au développement de la compétence de communication en langue, grâce à une exposition plus riche en quantité et en qualité. Les travaux sur l'acquisition des langues montrent notamment que de nombreuses dimensions de la compétence interactionnelle ne peuvent se développer que

¹ L'équipe est composée de 5 enseignants-chercheurs (4 en didactique de l'anglais et 1 en sciences du langage), rattachés à l'équipe Didactique des langues et sociolinguistique de l'ATILF (Université de Lorraine). Il s'agit de Maud Ciekanski, Carmenne Kalyaniwala, Nicolas Molle, Virginie Privas-Bréauté et Kossi Seto Yibokou.

² Métiers de l'Enseignement, Education et Formation, mention second degré parcours anglais et espagnol.

³ Français Langue Etrangère.

lors d'interactions authentiques (Pekarek Doehler & Berger, 2018). Les technologies immersives, de moins en moins coûteuses et de plus en plus ergonomiques, permettraient d'immerger les apprenants dans des situations de communication proches de situations authentiques, comme le proposent les ressources d'*edutainment*⁴ grand public, articulant visées ludiques et éducatives, et proposant la simulation de diverses pratiques sociales (Hine, 2005).

Le terme parapluie « réalité virtuelle » recouvre une diversité de ressources et de technologies. Les notions d'immersion, de téléprésence et de réalisme en constituent trois dimensions-clés. Elles permettraient d'apporter un gain en termes d'engagement cognitif, corporel et émotionnel dès lors que les utilisateurs-apprenants ressentent un sentiment de présence (Borona *et al.*, 2018). Des effets positifs pour les apprentissages de langue ont été identifiés sur les acquisitions lexicales (Alfadil, 2020), les connaissances linguistiques (Papin, 2019), la compétence interculturelle (Yeh *et al.*, 2020), les processus de production orale (Xie *et al.*, 2019) et de compréhension orale (Roy, 2017).

Notre étude concerne l'utilisation d'avatars dans des mondes virtuels et de vidéos immersives 360° qui permettent l'exploration d'un lieu filmé, selon différents degrés d'interactivité, par des technologies 2 et 3D. Ces ressources de divertissement, non-conçues à des fins didactiques, sont accessibles *via* des plateformes grand public. Dans ces situations virtuelles, l'utilisateur contrôle la direction dans laquelle il explore ce qui a été filmé, selon ses envies. Les deux types de ressources permettent des interactions avec l'interface numérique ou avec les objets en trois dimensions de l'environnement numérique. La présentation d'informations rendues saillantes grâce à la tridimensionnalité et à la multimodalité proposées en font des ressources prometteuses pour l'apprentissage. Pour autant, interagit-on avec l'interface numérique des objets connectés ou bien avec et dans la réalité projetée et co-construite par les interactions à l'écran ? Cette drôle de réalité (ou cette invention de la réalité, pour paraphraser Watzlawick, 1978) aurait-elle des effets sur les comportements interactionnels ?

Les expériences mises en place auprès de deux cohortes d'étudiants de master 2 (MEEF et FLE) questionnent (1) les effets comparés des artefacts choisis (outils et technologies 2 et 3D), en lien avec leur rejet ou adhésion comme prolongement de soi, sur l'immersion ressentie lors de l'utilisation des ressources et (2) l'utilisation d'un corps avatarisé ou ressenti pour le développement des compétences interactionnelles. Si interagir en L2 c'est aussi s'approprier un nouvel univers par tout un jeu d'interactions, la RV ferait-elle de nous des interactants plus conscients de ce processus de co-construction, et donc plus aptes à nous engager dans de nouvelles réalités ?

Après la présentation de notre cadrage théorique, nous discuterons des enjeux méthodologiques liés aux recherches écologiques avec la RV, puis discuterons des résultats obtenus.

⁴ Ce mot valise construit sur *education* et *entertainment* renvoie à tout média conçu pour éduquer par le divertissement et a été utilisé dès 1954 par Walt Disney ("Educational Values in Factual Nature Pictures". *Educational Horizons*, 33, n°2, pp. 82-84 (<https://www.jstor.org/stable/42922993>)).

1. Le paradigme de l'enaction : repenser le corps dans les apprentissages ?

1.1. Les neurosciences cognitives

Dans les années 1980, des linguistes ont commencé à explorer la façon dont les concepts abstraits pouvaient être basés sur des métaphores de concepts corporels et physiques. Pour Lakoff et Johnson (1999), l'esprit est intrinsèquement incarné, la pensée est principalement inconsciente et les concepts abstraits, largement métaphoriques. Ce sont là trois découvertes majeures pour les sciences cognitives⁵. L'idée selon laquelle l'esprit doit être compris dans le contexte de sa relation avec un corps physique qui interagit avec le monde suscite alors un intérêt croissant. Certains soutiennent que l'Homme a évolué à partir de créatures dont les ressources neurales étaient principalement consacrées au traitement perceptif et moteur, et dont l'activité cognitive consistait en grande partie en une interaction immédiate dans un environnement précis. Par conséquent, la cognition humaine, plutôt que d'être centralisée, abstraite et nettement distincte des modules d'entrée et de sortie périphériques, pourrait avoir des racines profondes dans le traitement sensori-moteur.

Le paradigme de l'enaction (Varela, 1993 ; 1996) s'inscrit dans le champ des savoirs incorporés et met l'accent sur le rôle des processus sensori-moteurs dans l'émergence de la connaissance. L'enaction repose sur trois principes-clés : le couplage participatif, l'émergence du sens et le chemin d'apprentissage individuel. Pour Varela, les langues servent aux hommes à se rencontrer (« se coupler », selon son propre terme) pour faire émerger le sens, et cette construction du sens prend forme de manière individuelle en fonction de l'environnement dans lequel l'individu évolue. La place du langage dans ce processus est donc primordiale et l'environnement ainsi que le contexte de l'utilisation du langage sont les conditions de l'émergence du sens. Le chercheur replace ainsi le corps, et les émotions, au cœur du processus d'accès à la connaissance :

L'intervention de l'implication du corporel et de l'émotionnel dans les processus d'apprentissage qui fait justement émerger du nouveau, c'est essentiel parce que sinon ça ne s'enracine pas dans un travail à dimension structurelle du corps et donc c'est un apprentissage superficiel (qui) passe très vite à une espèce de classement d'oubli d'abstraction (Varela cité par Trocmé-Fabre, 1994).

Eschenauer s'appuie sur ce paradigme pour expliquer les liens étroits entre émergence du langage et processus neural, biologique, dans des contextes précis de communication :

Le langage est complexe et comprend tous les canaux sensori-moteurs (langage non verbal : sensoriel, émotionnel, kinesthésique) et verbaux. Le besoin de « langager » est vital biologiquement pour l'homme et pour les êtres vivants. Il résulte d'actes de création langagière, c'est-à-dire de la capacité à s'adapter dans des situations de communication, et de combiner des structures ou symboles connus pour en faire émerger de nouveaux qui font sens dans l'interaction. S'appuyer sur cette nature créative du langage permet aux élèves d'atteindre une forme d'autonomie dans la capacité à communiquer (Eschenauer, 2018).

⁵ « The mind is inherently embodied. Thought is mostly unconscious. Abstract concepts are largely metaphorical. These are three major findings of cognitive science » (1999 : 3).

Dans l'enaction, la pensée n'est pas séparée du corps, la pensée et la parole sont incarnées, la cognition est incarnée et émerge en situation de communication.

1.2. Les neurosciences affectives et la pensée émotionnelle

Pour Damasio, les émotions font partie intégrante du processus d'apprentissage puisque « les sites d'induction de l'émotion déclenchent un certain nombre de signaux vers d'autres sites du cerveau [...] et vers le corps » (Damasio, 2002 : 76). Ces émotions (en plus de rappeler l'hypothèse du filtre affectif de Krashen, 1981) impactent précisément la conscience et rendent possible l'accès à la cognition :

Les aspects de la cognition qui sont les plus sollicités dans l'éducation, notamment l'apprentissage, l'attention, la mémoire, la prise de décision, la motivation et le fonctionnement social, sont à la fois profondément affectés par l'émotion et, en fait, subsumés dans les processus de l'émotion. Les émotions impliquent la perception d'un déclencheur émotionnellement compétent, une situation réelle ou imaginaire qui a le pouvoir d'induire une émotion, ainsi qu'une chaîne d'événements physiologiques qui vont permettre des changements à la fois dans le corps et dans l'esprit⁶ (Damasio, 1994 : 35, notre traduction).

Immordino Yang précise que l'émotion et la cognition sont soutenues par des processus neuronaux interdépendants. Pour elle, il est impossible sur le plan neurobiologique de construire des souvenirs, de s'engager dans des pensées complexes ou de prendre des décisions significatives sans émotion (Immordino Yang, 2016). Ses recherches l'ont amenée à proposer l'expression « *emotional thought* », pensée émotionnelle, qu'elle définit comme un important chevauchement entre la cognition et l'émotion (Immordino Yang, 2016 : 37). La pensée émotionnelle peut, selon elle, être consciente ou non et constitue le moyen par lequel les sensations corporelles liées à l'émotion entrent dans notre conscience. Ainsi, les aspects corporels des émotions sont représentés comme une boucle allant de la pensée émotionnelle au corps et inversement. Ici, les pensées émotionnelles peuvent modifier l'état du corps de manière caractéristique, par exemple en contractant ou en relâchant les muscles ou en modifiant le rythme cardiaque. À leur tour, les sensations corporelles de ces changements, réelles ou simulées, contribuent, consciemment ou non, aux sentiments, qui peuvent influencer la pensée. Le ressenti de ces sensations corporelles, qu'elles soient conscientes ou non, peut alors biaiser les processus cognitifs tels que l'attention et la mémoire.

1.3 L'*embodiment*, une notion féconde pour comprendre la RV ?

La notion d'*embodiment* (ou cognition incarnée) cherche à réévaluer la place et le rôle du corps dans le processus d'apprentissage. Tout processus cognitif s'inscrit dans un corps situé dans un environnement qu'il soit naturel, culturel, social ou technique. Les travaux

⁶ « The aspects of cognition that are recruited most heavily in education, including learning, attention, memory, decision-making, motivation and social functioning, are both profoundly affected by emotion and in fact **subsumed** within the processes of emotion. Emotions entail the perception of an emotionally competent trigger, a situation either real or imagined that as the power to induce an emotion, as well as a chain of physiological events that will enable changes in both the body and the mind ». (Damasio, 1994 : 35)

sur les enfants bilingues ont notamment souligné le rôle de la cognition incarnée dans leurs apprentissages de langue en situation naturelle et ses effets sur l'engagement et l'attention, la mémorisation du lexique en contexte, l'appropriation sémantique et la compréhension de la langue-culture.

Selvam reprend les grands principes des savoirs incorporés, de l'enaction et de la pensée émotionnelle afin de représenter l'*embodiment* au cœur d'interrelations entre cognition, comportement et émotion (Fig. 1).

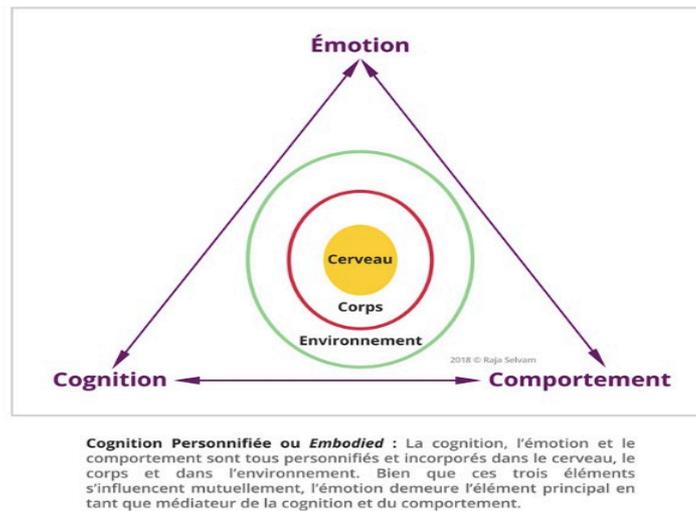


Figure 1. Représentation schématisée de l'embodiment (Selvam, 2018, n.p.)

Bien que ces trois éléments s'influencent mutuellement, l'émotion y joue un rôle principal. En outre, dans la théorie de la cognition incarnée, les connaissances à acquérir sont reliées aux expériences physiques de l'apprenant. Il existerait ainsi un lien très fort entre le corps, la perception et l'action dans l'apprentissage (Gadille & Impedovo, 2020). Cette compréhension des processus cognitifs est particulièrement intéressante pour l'apprentissage des langues. Pour Kramersch, les apprenants de langues ne sont pas seulement des communicateurs et des résolveurs de problèmes, mais des personnes à part entière pour lesquelles : « les formes symboliques ne sont pas seulement des éléments de vocabulaire ou des stratégies de communication, mais aussi des expériences incarnées, des résonances émotionnelles et des imaginations morales » (Kramersch, 2006 : 251).

La notion d'*embodiment* semble particulièrement féconde pour appréhender les apprentissages avec la RV. La RV génère chez les utilisateurs des expériences fortement sensorielles et émotionnelles (le fameux « effet wow ») suscitées par les interactions avec les outils, les acteurs/personnages, l'environnement virtuel. Elle provoque une expérience médiatisée qui plonge un ou plusieurs utilisateurs au cœur d'un environnement artificiel dans lequel l'utilisateur peut ressentir et interagir en temps réel *via* des interfaces sensori-motrices. L'utilisateur juge crédible l'expérience, accepte de se prendre au jeu et en réponse ressent un sentiment de présence (Bouvier, 2009). Pour Méran-Martinez & Spanghero-Gaillard (2021), l'interaction est la première caractéristique de la RV ; l'immersion et le sentiment de présence en sont deux autres

subséquentes (Ciekanski *et al.*, 2020). L'immersion renvoie « aux stimuli produits par le dispositif de réalité virtuelle et à leur degré de fidélité par rapport à la réalité » (Roy, 2014 : 3). Le sentiment de présence désigne l'effet de percevoir comme réels ou vivants les objets, événements ou personnages avec lequel l'utilisateur interagit dans la RV. Le degré de présence ressenti est souvent assimilé au degré de réalisme (social, perceptif, cognitif) des ressources. Un fort sentiment de présence peut générer une illusion d'immédiateté (d'absence de médiation) chez l'utilisateur de RV. Burkhardt *et al.* (2006) ont notamment détaillé les apports des interactions avec la RV sur l'apprentissage : l'amplification de la motivation par l'aspect ludique et attractif de l'outil ; la création d'une immersion au moyen de rétroactions simultanées sur les modalités visuelles, auditives et haptiques, qui agit comme un catalyseur de l'apprentissage et la proposition de perspectives multiples sur une situation ou un objet, soutenant diverses stratégies au cours de l'interaction avec l'environnement virtuel.

Nous nous intéressons dans cette étude à la place du corps, comme vecteur d'expériences et d'émotions dans les interactions suscitées par la RV et aux effets de l'*embodiment* sur l'apprentissage d'une langue en situation écologique (c'est-à-dire non-didactisée). Dans quelle mesure le corps virtuel, ressenti ou avatarisé, peut-il être un vecteur pertinent pour un apprentissage de langue en immersion ? Quels sont les effets des technologies utilisées sur le sentiment de présence et la perception de l'environnement ? Dans quelle mesure cette mise en saillance des sensations et émotions des utilisateurs favoriserait-elle un apprentissage ?

2. Méthodologie

Dans le cadre de cette étude, une soixantaine d'étudiants⁷ ont utilisé des ressources de RV lors d'ateliers organisés au Centre de Ressources Yves Châlon (CLYC) de l'Université de Lorraine, *via* des technologies 2D (ordinateur et smartphone) et 3D (casque 3D), en anglais ou en espagnol. Les ressources de la RV sont caractérisées par un fort degré de multimodalité : à chaque mode communicationnel correspond une modalité, c'est-à-dire une forme particulière de communication (Fig.2) :

Modes	Modalités
Textuel	Sous-titre, texte
Parole	Audio
Visuel	Image statique, image dynamique, modification de la focale
Iconique	Boutons de commande
Graphique	Schéma
Haptique	Actions tactiles et manuelles sur l'interface et dans l'environnement
Spatial	Déplacement d'un point à l'autre

Figure 2 : Correspondances entre modes et modalités dans les ressources sélectionnées.

⁷ Les ateliers ont eu lieu lors de cours portant sur les pédagogies innovantes en apprentissage des langues auprès de deux promotions d'étudiants de Master Métiers de l'Enseignement, de l'Éducation et de la Formation (MEEF) anglais/ espagnol en 2018-2019 puis en 2019-2020, et auprès d'une promotion d'étudiants de Master FLE en 2020-2021.

L'utilisation de la multimodalité dépend fortement du rôle de l'utilisateur prévu par le scénario de la ressource. Le scénario peut prévoir différents degrés d'interaction et d'interactivité. Les ressources sélectionnées dans l'atelier invitaient à expérimenter plusieurs comportements — observateur ou participant, actif ou passif — selon les affordances de l'environnement et la situation scénarisée qui offre plus ou moins de liberté aux utilisateurs.

La méthodologie de recherche a combiné approches quantitative et qualitative dans une perspective émique⁸. L'un des enjeux méthodologiques liés à la compréhension de l'*embodiment* dans la RV était de recueillir des données sous la forme d'une description phénoménologique des expériences subjectives vécues par les participants. L'approche phénoménologique cherche en effet à décrire les phénomènes, tout en mettant l'accent sur l'expérience telle qu'elle est vécue à travers les stratégies d'utilisation et d'interaction. Merleau-Ponty (1964) fait ainsi de l'expérience corporelle le prisme de l'expérience. Cela a des effets sur la façon de rentrer en relation avec le monde impliquant une *compréhension* par le corps et le mouvement qui permet de penser l'expérience dans un espace social et environnemental. L'approche phénoménologique vise en outre à « mettre le monde entre parenthèses » (Husserl, 1947/2014) pour mieux conscientiser notre intentionnalité, mieux le voir et agir.

Sur ces principes, nous avons conçu un rapport d'étonnement, inspiré du journal d'étonnement de Develotte (2006 : n.p.) : « il [s'agit] de faire jouer un rôle cathartique au journal d'étonnement, sur le plan des émotions et un rôle incitatif à la réflexion intellectuelle sur le plan rationnel ». Rédigé « à chaud », après chaque immersion virtuelle, il comprenait huit questions permettant aux étudiants de rendre compte, d'une part, de leur expérience, leur ressenti, leurs pratiques d'utilisation et, d'autre part, de ce qu'ils avaient compris, appris, retenus à l'issue de l'immersion en anglais. Le rapport d'étonnement cherchait à mettre en exergue tout à la fois les comportements, les émotions et les effets potentiels sur les processus cognitifs. Les étudiants ont pu tester au maximum quatre ressources lors des ateliers. Ces données ont été complétées par des questionnaires et des entretiens de type *Stimulated Recall* (Gass & Mackey, 2000).

3. Résultats

3.1 Expériences sensorielles et émotionnelles de la RV 3D selon la nature des ressources et leurs effets sur les interactions

Dans le cadre de trois ateliers proposés aux étudiants de master MEEF, ces derniers devaient tester différents genres de ressources de RV avec un casque 3D pour ensuite réfléchir à des applications concrètes en classe de langue visant le développement de compétences langagières et transversales (Privas-Bréauté, 2021). Nous rendons compte

⁸ Le choix de la perspective émique nous a semblé essentiel dans cette recherche dans la mesure où il s'agit de capter l'exploration subjective des participants, en nous concentrant sur la signification de l'expérience, du contexte, du processus participatif dans lequel ils se sont ou non engagés, au gré des circonstances et des opportunités de l'environnement.

ici des réponses obtenues à deux questions du rapport d'étonnement : quelles ont été vos premières impressions ? Quelles émotions avez-vous ressenties ? L'analyse des réponses a fait l'objet d'une analyse de contenu (Bardin, 2013) à partir des occurrences du terme « interaction » dans leurs propos. Les réponses obtenues ont été classées en trois catégories et selon des valences positives ou négatives, comme rapporté dans les tableaux 1, 2 et 3.

3.1.1. Ressources ludiques 3D

Le premier atelier de RV correspondait pour certains à une première prise en main du casque 3D et a consisté en l'utilisation d'applications de loisirs telles que :

- « Mission : ISS » dans laquelle les utilisateurs écoutent et lisent des instructions les invitant à réaliser des micro-tâches dans la station spatiale internationale ;
- « Disney » qui plonge les étudiants dans l'univers féérique du groupe international.

Les utilisateurs peuvent interagir avec le monde virtuel pour réaliser des tâches. Ils sont uniquement en situation de réception de la langue-cible (orale et écrite).

	Réponses à valence positive	Réponses à valence négative
Interactions avec la RV	Intéressant (x2) pour aller dans des lieux où on ne pourrait pas aller physiquement. Très réaliste, très bien fait (graphisme, dessins), très beau. Interactif Ludique Très immersif	Difficile d'utilisation Difficultés dans le contrôle Difficultés à manier les manettes, à se mouvoir dans l'espace, à bien placer le casque pour que l'image ne soit pas floue Difficultés de prise en main Manque d'interaction avec l'environnement (x3)
Emotions ressenties	Surprise Étonnement Émerveillement (x3) Impressionnant (x2) Enthousiasme, joie Euphorie Excitation Divertissement Amusement	Déception quand cela ne fonctionne pas Perturbé, déstabilisé
Sensations ressenties lors de l'expérience	Se sentir à l'aise Intégré au jeu Curiosité, envie de découvrir de nouveaux environnements « Nous sommes stimulés » Possibilité d'interagir avec le décor (saisir des objets) « On se sent dans un univers à part »	Perte d'équilibre, vertige Apesanteur dans la station ISS déstabilisante

Tableau 1. Réponses issues du rapport d'étonnement du premier atelier.

3.1.2. Visites virtuelles 3D

Lors du deuxième atelier, les étudiants ont testé des applications des vidéos 360° de type touristique telles que « Travel VR », « Google 360 », « The People's House » et « The Smithsonian's Venice » qui proposent des visites de lieux, commentés ou non, en langue cible. Les utilisateurs peuvent explorer l'univers graphique autour d'eux par la vision 360° et sont uniquement en situation de réception de la langue-cible (orale et écrite).

Les réponses obtenues sont présentées dans le Tableau 2.

	Réponses à valence positive	Réponses à valence négative
Interactions avec la RV	« On est immergé, on a l'impression que l'on peut toucher, communiquer, se déplacer » Attrayant Très impressionnant avec le son ambiant	Manque d'interaction avec l'environnement (x3) Difficile d'utilisation, difficulté de contrôler les mouvements
Emotions ressenties	Émerveillement (x3) Impressionnant (x2) Divertissement, amusement Relaxant pour certaines vidéos (grâce à la musique, la voix accompagnant la visite était calme) Calme et joie Curiosité	Déception Scepticisme
Sensations ressenties lors de l'expérience	A l'aise Envie de découvrir de nouveaux environnements Visiter de manière amusante Application calme qui permet de faire des voyages touristiques plus accessibles	Tournois, malaise Se sent inactif et passif, dans une vidéo 360 (x2)

Tableau 2. Réponses issues du rapport d'étonnement du deuxième atelier.

3.1.3. Interactions orales dans un monde virtuel 3D

Lors du troisième atelier, les étudiants ont interagi dans le monde virtuel *Sansar*, version 3D du monde synthétique *Second Life*, via un avatar pour communiquer avec d'autres personnes. Ce monde virtuel 3D est ouvert à tout public. Il arrive donc fréquemment de visiter un monde où d'autres avatars inconnus se promènent. Les réponses obtenues sont présentées dans le Tableau 3.

	Réponses à valence positive	Réponses à valence négative
Interactions avec la RV	Très réaliste, très bien fait (graphisme, dessins), très beau Interactif Monde futuriste Monde riche	Difficile de reconnaître son avatar à l'écran Difficile d'utilisation, Difficulté de contrôler les mouvements
Emotions ressenties	Fascination Joie Étonnement (x3)	Déception quand cela ne fonctionne pas

	Surprise Amusement	Sentiment de gêne à parler avec des inconnus
Sensations ressenties lors de l'expérience	Impressionné devant les nombreuses personnes présentes à l'écran « J'ai pu discuter en anglais avec des gens ouverts à la discussion » « Je ne m'attendais pas à rentrer en contact aussi rapidement avec des inconnus » « Parfait pour un échange direct avec un natif » Déconnexion du monde réel Très immersif (x3)	Maux de tête, léger vertige Oppressant Mal à l'aise de parler et interagir avec des inconnus dans un environnement inconnu Rencontres très artificielles « Au bout d'un moment on se retrouve entouré d'autres avatars inconnus et on devient un objet de curiosité » L'enseignant peut perdre le contrôle, danger pour les élèves

Tableau 3. Réponses issues du rapport d'étonnement du troisième atelier.

Il en ressort que ces ressources 3D provoquent des expériences à fort degré sensoriel et émotionnel. De façon globale, on dénote davantage d'éléments à valence positive que négative dans les rapports d'étonnement (ce qui tendrait à confirmer le fameux « effet wow » de la RV). Toutefois, chaque ressource a engendré pour certains utilisateurs des émotions et des sensations négatives. En outre, le casque 3D provoque chez certains un malaise physique qui perturbe l'expérience. Enfin, l'envie d'interagir dans ou avec l'environnement 3D semble conditionnée par trois facteurs : le réalisme perçu de la situation, l'esthétique perçue de l'environnement et l'agentivité des participants. Un dernier aspect nous semble particulièrement intéressant : en l'absence de difficultés techniques de manipulation qui rendent manifeste les outils utilisés, les participants témoignent d'une expérience quasi immédiate dans le monde virtuel. Les expériences réussies avec les casques 3D tendraient à invisibiliser les objets du réel.

3.2 Expériences sensorielles et émotionnelles de la RV 2D et 3D et leurs effets sur la cognition

Nous avons étudié⁹, dans une perspective écologique et compréhensive, les pratiques d'apprentissage d'un groupe d'étudiants de master FLE de niveau hétérogène en anglais, lors de deux ateliers de découverte des vidéos immersives 360° 2D et 3D. Les ressources ont été proposées sans didactisation. Les étudiants n'ont reçu aucune instruction pour leurs essais, ni d'objectif d'écoute, mais ont pu bénéficier d'un accompagnement technique. Les étudiants ont testé *via* trois supports (casque 3D, ordinateur et smartphone) au moins trois vidéos immersives 360° différentes, à leur convenance, parmi une sélection de six ressources proposant des visites guidées d'une vingtaine de minutes¹⁰. 66 rapports d'étonnement ont été recueillis ainsi que les captations des quatre étudiants volontaires pendant leur immersion virtuelle, complétées par des interviews individuelles de type *stimulated recall*, qui ont été retranscrites.

⁹ Cette étude a été réalisée en collaboration avec Kossi Seto Yibokou (Université de Lorraine).

¹⁰ Les vidéos immersives 360° testées sur smartphone et ordinateur sont en accès libre sur *Youtube* et les vidéos 3D sur la plateforme *Oculus Rift*.

Les ressources sélectionnées ont permis aux participants de recourir à une variété de modes écologiques (parole, visuel, iconique, haptique et spatial) et non-écologiques (textuel, graphique) pour explorer la ressource. Dans ces ressources, la communication y est fortement multimodale puisqu'elle mobilise entre cinq et sept modes selon la vidéo. Le guidage des visites guidées virtuelles encadre l'exploration des utilisateurs par divers dispositifs d'orientation (par exemple, suivre des instructions orales ou écrites) ou d'information (le discours explicatif d'une voix-off) qui peuvent être proposés de façon explicite (le guide est visible et oriente les regards et la compréhension) ou à découvrir (les instructions sont cachées dans l'environnement numérique et le but est de les retrouver en explorant les lieux) (Annexe 1). Comme pour toute utilisation de ressources numériques, une certaine compétence technique est requise en deçà de laquelle les affordances de l'environnement ne sont pas perçues par les utilisateurs.

3.2.1. Perception de la multimodalité

Les rapports d'étonnement indiquent que les étudiants ont été attentifs à la multimodalité des ressources, quelle que soit la technologie 2D ou 3D utilisée, lors de leur exploration virtuelle¹¹. 70% des explorations ont mobilisées de trois à quatre modes ; regarder une vidéo immersive est donc une activité fortement multimodale. Parmi les sept modes composant la multimodalité des vidéos immersives 360° testées, quatre ont plus particulièrement favorisé l'attention et la concentration des utilisateurs. Il s'agit des modes visuels, parole, haptique et spatial (Fig.3).

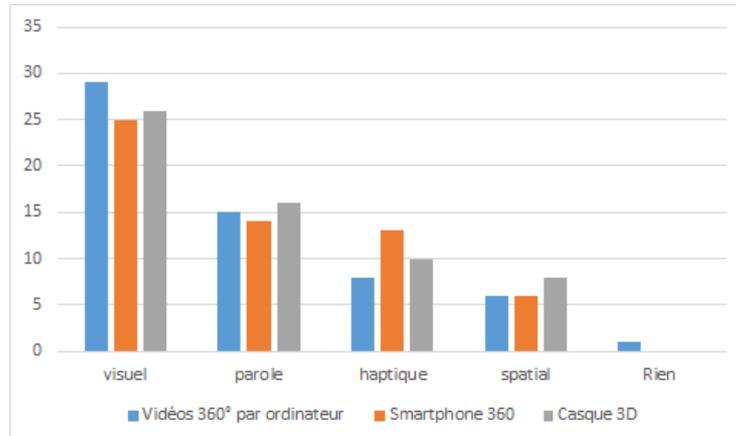


Figure 3. Perception de la multimodalité par support ayant favorisé l'attention des participants.

Il ressort des rapports d'étonnement que le visuel oriente la compréhension orale des étudiants. En outre, les modes sont utilisés par les étudiants dans la réalisation d'actions intermodales (visuel + parole + haptique, par exemple) qui accompagnent le processus de compréhension. Toutefois, ces stratégies intermodales peuvent générer chez les étudiants une surcharge cognitive due à un trop plein attentionnel, au détriment de la compréhension du message (« *je ne retiens rien parce que je me concentre sur l'environnement* »).

¹¹Les étudiants mentionnent avoir eu recours en moyenne à 3 modes avec les casques 3D, 2,56 avec les ordinateurs et 2,51 avec les smartphones.

3.2.2. Perception des opportunités d'apprentissage selon la scénarisation des visites

L'analyse des quatre interviews de type *stimulated recall* a porté sur l'identification des « situations potentielles de développement » (Mayen, 1999) rencontrées par les utilisateurs avec les vidéos immersives. Le dispositif de participation des ressources et sa scénarisation ont eu un effet sur l'attention des utilisateurs. Certains lexiques liés à des connaissances historiques et architecturales sont apparus de façon saillante et en redondance à l'écrit et à l'oral dans la vidéo. Ce sont les seules opportunités d'apprentissage qui ont été conscientisées et relatées lors des interviews. Le réalisme de la situation perçue et le sentiment de présence ressentis ont déclenché une envie d'interagir (« *j'avais tout le temps envie de poser des questions* »). L'impossibilité de le faire a cependant limité les opportunités d'apprentissage. La faible agentivité générée par le scénario a été ressentie comme un frein à l'envie d'apprendre. La possibilité de mettre en place un rôle actif dans la découverte du lieu a généré de nombreuses manipulations et déplacements, en fonction de la compréhension et de la curiosité des étudiants. Les instructions orales ont également stimulé le sentiment d'une présence sociale provoquant une envie d'interagir oralement avec le système (« *j'avais envie de réagir, de poser des questions quand j'étais bloqué* »). Dans ces situations, les opportunités d'apprentissage rapportées étaient liées aux difficultés rencontrées dans la réalisation des tâches qui ont stimulé l'exploration de l'environnement et aux éléments rendus saillants dans l'environnement (à l'oral par répétition ou à l'écrit et à l'oral par redondance).

3.2.3. Nature des éléments retenus par outil

Les rapports d'étonnement invitaient les étudiants à rendre compte de ce qu'ils avaient compris et retenu à l'issue de chaque visite immersive et à donner des exemples précis de connaissances nouvelles apprises sur le sujet traité ou d'éléments de langue-culture saisis pendant l'écoute. 93% des rapports d'étonnement rendent compte d'un apprentissage d'éléments langagiers ou de connaissances nouvelles. 84% des rapports d'étonnement mentionnent des exemples précis de ce qui a été retenu. Il en ressort que 49% des éléments retenus sont langagiers, 34% extra-langagiers et 17% culturels (Fig. 4).

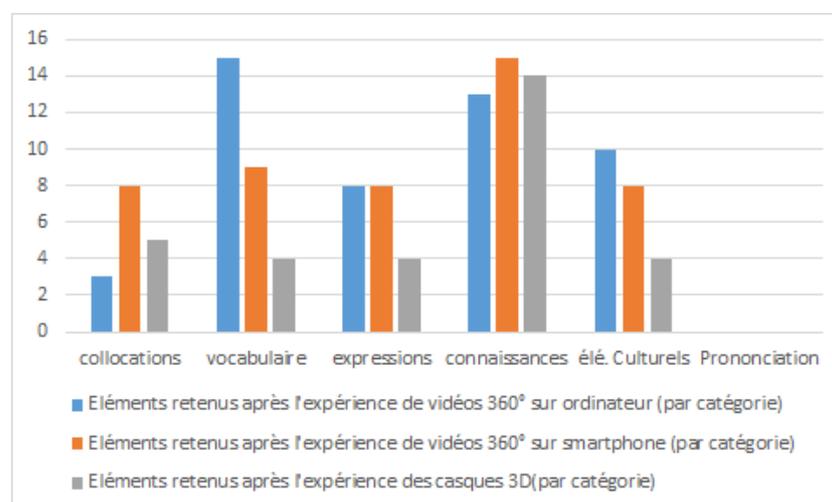


Figure 4. Nature des éléments retenus par outil.

La nature des éléments retenus varie selon la technologie 2D ou 3D (Fig.4). 49 éléments ont été mentionnés suite aux visites immersives avec un ordinateur, 48 avec un smartphone et 31 avec un casque 3D. Par leurs affordances, les technologies génèrent des situations immersives différentes qui ont des effets sur l'attention des utilisateurs et sur les stratégies d'écoute et de compréhension mobilisées. Les casques 3D permettent l'immersion la plus proche d'une situation écologique. Les résultats indiquent qu'ils favorisent l'attention des utilisateurs sur les connaissances extra-langagières par rapport aux éléments langagiers ou culturels des vidéos immersives. Le recours aux modes textuels permis par les technologies 2D avec l'activation du sous-titrage semble soutenir la rétention d'éléments langagiers. La présence d'une médiation textuelle (sous-titrage ou transcription ponctuelle de mots techniques comme sur la vidéo 1, etc.) semble favoriser les saisies lexicales. Les données des rapports d'étonnement montrent les effets du mode textuel sur la compréhension orale de vidéos, les effets de l'intermodalité texte et parole sur l'apprentissage du lexique et les effets des informations visuelles sur les saisies langagières en compréhension orale. Enfin, les rapports d'étonnement indiquent que les casques 3D permettent l'immersion mobilisant le plus grand nombre d'affordances multimodales. Pour autant, cette forte stimulation attentionnelle semble moins soutenir le processus de compréhension orale. Cela peut être dû à une plus forte distraction ressentie ou à davantage d'effets d'attention divisée. En outre, plus de problèmes techniques ont été rapportés avec les casques 3D (7/20 contre 5/23 avec smartphone et 3/23 avec l'ordinateur), ce qui a pu affecter la qualité de la concentration et de l'attention des participants.

3.2.4. Le rôle des émotions : facilitateur ou frein aux apprentissages immersifs ?

Les rapports d'étonnement rendent compte d'une forte expérience émotionnelle. Les substantifs et les verbes traduisant une émotion dans les discours des étudiants, ainsi que les adjectifs traduisant une intensité, ont été analysés selon la catégorisation de Kerbrat-Orecchioni (1980). 18/20 rapports d'étonnements indiquent une expérience émotionnelle avec les casques 3D, 20/23 avec les smartphones et 18/23 avec les ordinateurs. Les discours des étudiants montrent que les émotions ressenties dépendent du sentiment d'immersion vécu, de leur adhésion au sujet traité ainsi qu'à l'esthétique visuel et sonore de la ressource. Pour autant, l'intensité des émotions ressenties diffère selon les outils : 85% des émotions positives ont été ressenties avec un casque 3D ou un smartphone et seulement 37% avec un ordinateur. À la surcharge cognitive s'ajouterait ainsi une surcharge émotionnelle, selon les technologies. Les rapports d'étonnement montrent que plus l'émotion ressentie est forte, moins les éléments retenus sont précis, tant au niveau de la langue que des connaissances extra-langagières.

4. Discussion

4.1. Interagit-on avec l'interface numérique des objets connectés ou bien avec et dans la réalité co-construite par les interactions à l'écran ?

Les résultats montrent que les étudiants se sentent interagir dans la RV une fois les difficultés de manipulation dépassées. Les outils 2D ou 3D deviennent alors invisibles aux utilisateurs (qu'il s'agisse d'un smartphone ou d'un casque 3D). Les ressources immersives sont perçues comme un média d'exploration et d'expérimentation. Les participants ont fait l'expérience de différents degrés d'agentivité et d'implication. Placer le corps au centre de l'apprentissage a eu des effets sur l'exploration de l'environnement et sur la rétention des connaissances qui en découlent (observer à sa guise un bâtiment tout en écoutant des explications sur son architecture, son histoire, se déplacer tout en écoutant des instructions, réaliser des tâches dans des environnements spécifiques). Tout cela semble avoir des effets sur la mémorisation des apprenants et sur leur compréhension en contexte.

4.2. Cette drôle de réalité aurait-elle des effets sur les comportements interactionnels ?

Il semble que les effets de sidération et le mal de RV (à l'image du mal de transport) freinent les interactions et l'apprentissage. Tout l'enjeu est de comprendre comment les canaliser pour faciliter interaction, rétention et apprentissage.

Les résultats montrent par ailleurs que les étudiants interagissent dans la situation d'immersion et non avec un écran. Ils ont ainsi l'impression de tomber dans les escaliers, qu'on leur parle directement et plusieurs ont ressenti l'envie de répondre aux interlocuteurs virtuels (qu'il s'agisse d'une voix-off ou d'une représentation à l'écran). Cela a donc eu un effet sur leur comportement interactionnel, que l'on peut rapprocher de l'échoïsation empathique, soutenue par les effets d'*embodiment* vécus dans ces environnements. Les étudiants ont été particulièrement sensibles à l'intention communicative développée dans les ressources. On note également que le processus d'empathie est fortement sollicité dans les interactions en RV. Ces phénomènes sont particulièrement intéressants pour les apprentissages de langue.

4.3. La réalité virtuelle ferait-elle de nous des interactants plus conscients des processus de co-construction du sens, et donc plus aptes à nous engager dans de nouvelles réalités ?

En rendant saillante la multimodalité des environnements, la RV semble solliciter une attention plus complexe aux paramètres de la situation de communication. Les étudiants utilisent en moyenne 3 à 4 modes dans les actions intermodales pour construire leur compréhension. En outre, la multimodalité des ressources immersives décuple les sensations et cela peut être un frein aux saisies langagières et extra-langagières. Enfin, les modes haptique et spatial peuvent agir comme un distracteur à la compréhension lorsque les sensations ressenties sont fortes. Cette hyperstimulation sensorielle semble renforcée avec les casques 3D (*se sentir happer/ avoir besoin d'arracher son casque 3D*) qui favorise

la mise à disposition de l'utilisateur à l'expérience, par le corps et par le regard (*se sentir coupé du monde par le casque 3D*). L'expérience corporelle concourt fortement au sentiment de présence (les étudiants ont l'impression qu'on leur parle réellement et souhaitent répondre...) et génère de l'empathie pour les sujets traités. Ces éléments participent du processus de saisie langagier et extra-langagier.

Enfin, nous avons pu observer des comportements interactionnels spontanés mais très incongrus, sur le mode du jeu, entre certains de nos étudiants et des personnes inconnues rencontrées sur *Sansar*. Ainsi, l'un de nos étudiants a d'abord échangé verbalement quelques banalités avec un avatar apparaissant sous la forme d'un robot, avant de se mettre à danser avec lui. S'engager dans de nouvelles réalités, c'est aussi peut-être réinventer les normes sociales qui sous-tendent nos interactions.

Conclusion

La réalité virtuelle présente de nombreux avantages en tant que dispositif d'apprentissage des langues. La caractéristique sensorielle et émotionnelle de la RV déclenche l'émerveillement et peut favoriser la motivation et l'engagement des apprenants. Les affordances des outils technologiques favorisent l'exploration multi-sensorielle et multimodale en situation de communication contextualisée et permettent à l'expérience d'être vécue comme étant incarnée. Toutefois, les résultats obtenus sont très contrastés (indépendamment du niveau des étudiants et de leurs compétences techniques) en ce qui concerne les processus cognitifs mobilisés.

L'étude questionnait le potentiel heuristique de l'*embodiment* pour mieux appréhender les expériences avec la RV. L'approche méthodologique a tenté de saisir les expériences subjectives vécues par une soixantaine d'étudiants. Chaque mode de recueil les invitait à une conscientisation de leur expérience sensorielle, émotionnelle et cognitive en langue-cible. Ce mode de recueil nous a semblé particulièrement pertinent pour aider les étudiants à dépasser le caractère inouï et inédit des expériences immersives (notamment celles vécues en 3D).

Bibliographie

- Alfadil, M. (2020). Effectiveness of virtual reality game in foreign language vocabulary acquisition. *Computers and Education*, 153, 1-13. DOI: 10.1016/j.compedu.2020.103893.
- Bardin, L. (2013). *L'analyse de contenu*. Presses Universitaires de France.
- Borona S., Tambouris, E. et Tarabanis, K. (2018). The Use of 3D Multi-user Virtual Environments in Computer Assisted Second Language Learning: A Systematic Literature Review, *International Journal of Learning Technology* 13(3), 249-274.
- Bouvier, P. (2009). *La présence en réalité virtuelle, une approche centrée sur l'utilisateur*. Thèse de doctorat, Université Paris-Est.

- Burkhardt, J. M., Lourdeaux, D., & Mellet-d'Huart, D. (2006). *Le traité de la réalité virtuelle*. École des Mines de Paris.
- Chateau, A., Ciekanski M., Molle, N., Paris J. & Privas-Bréauté, V. (2019), Adding Virtual Reality to the university self-access language centre: Brave new world or passing fad? *European Journal of Language policy*, « Virtual exchanges ».
- Ciekanski, M., Kalyaniwala, C., Molle, N., & Privas-Bréauté, V. (2021). Real and perceived affordances of Immersive Virtual Environments in a language teacher-training context: effects on the design of learning tasks, *ReDoc*, 1(1). <https://hal.archives-ouvertes.fr/hal-02938908>
- Damasio, A. (1994). *Descartes' error: Emotion, reason, and the human brain*. Penguin Books.
- Damasio, A. (2002). *Le Sentiment même de soi - Corps, émotions, conscience*. Odile Jacob.
- Develotte, D. (2006). Le Journal d'étonnement : aspects méthodologiques d'un journal visant à développer la compétence interculturelle. *LIDIL*, 105-124.
- Eschenauer, S. (2018). Créativité et empathie dans les apprentissages performatifs : vivre et incorporer ses langues, *Recherches & éducations*, Hors-série "Corps, gestes, paroles dans la situation d'enseignement".
- Gadille, M. & Impedovo, M-A. (2020). The Embodiment Dimension while Learning and Teaching in a Virtual World. *Journal of Systemics, Cybernetics and Informatics: JSCI*, 18 (2), 1-8.
- Gass, S. & Mackey, A. (2000). *Stimulated Recall Methodology in Applied Linguistics and L2 research*. Routledge.
- Hine, C. (ed.) (2005). *Virtual Methods: Issues in Social Research on the Internet*. Berg Publishers.
- Husserl, E. (1947/2014). *Méditations Cartésiennes*. Editions Vrin.
- Immordino-Yang, M.H. (2016). *Emotions, Learning and the Brain. Exploring the Educational Implications of Affective Neuroscience*. Norton & Company
- Kramsch, C. (2006). From communicative competence to symbolic competence. *Modern Language Journal*, 90 (2), 249-252
- Krashen, S. D. (1981). *Second Language Acquisition and Second Language Learning*. Pergamon.
- Kerbrat-Orecchioni, C. (1980). *L'énonciation de la subjectivité dans le langage*. Armand Colin.
- Lakoff G. & Johnson M. 1999. *Philosophy in the flesh. The embodied mind and its challenge to Western Thought*. Basic Books.
- Lan, Y. J. (2020). Immersion, Interaction and Experience-oriented Learning: Bringing Virtual Reality into FL Learning. *Language Learning & Technology*, 24(1), 1-15.
- Merleau-Ponty, M. (1964). *L'Œil et l'Esprit*, Editions Gallimard.
- Mayen, P. (1999). Des situations potentielles de développement. *Education Permanente*, 139(2), 66-86.
- Meyran-Martinez, C. & Spanghero-Gaillard, N. (2021). L'enseignement de l'anglais avec la réalité virtuelle. *Mélanges CRAPEL*, 42 (2). https://www.atilf.fr/wp-content/uploads/publications/MelangesCrapel/Melanges_42_2_5_Meyran-Martinez_Spanghero-Gaillard.pdf
- Molle, N., Privas-Bréauté, V. et Ciekanski, M. (2020). La réalité virtuelle comme vecteur d'immersion pour apprendre les langues. *EDL*, 34, 69-80.

- Papin, K. (2019). *La contribution des tâches de réalité virtuelle au désir de communiquer en français langue seconde à l'extérieur de la salle de classe à Montréal*. Thèse de Doctorat, Université de Montréal. <http://hdl.handle.net/1866/23558>
- Pekarek Doehler, S. et Berger, E. (2018). L2 Interactional Competence as Increased Ability for Context-sensitive Conduct: A Longitudinal Study of Story-openings, *Applied Linguistics*, 39(4), 555–578.
- Privas-Bréauté, V. (2021). Intégrer la réalité virtuelle dans les formations d'enseignants en langues : Dispositif innovant immersif inscrit dans un paradigme enactif. *The Langscape Journal*, 3, 103-116.
- Roy, Michael. (2017). *La réalité virtuelle pour l'apprentissage des langues*. Peter Lang.
- Selvam, R. (2018). La science de l'Embodiment ou personnification de la cognition et des émotions énelles: Implications pour de meilleurs résultats, quelle que soit la modalité thérapeutique employée *Integral Somatic Psychology*, n.p.
- Steuer, J. (1992). Defining virtual reality: dimensions determining telepresence. *Journal of Communication*, 42(4), 73-93.
- Trocmé-Fabre, H. (1994). Né pour créer du sens avec Francisco Varela, <https://www.canal-u.tv/chaines/cerimes/serie-ne-pour-apprendre/ne-pour-creer-du-sens-avec-francisco-varela>
- Varela, F. (1996). *Invitation aux sciences cognitives*, traduction de Pierre Lavoie. Éditions du Seuil.
- Varela F., Thompson E. et Rosch E. (1993). *L'inscription corporelle de l'esprit, Sciences cognitives et expérience humaine*, traduction de Véronique Havelange. Editions du Seuil.
- Xie, Y., Chen, Y. & Ryder, L. H. (2019). Effects of using mobile-based virtual reality on Chinese L2 students' oral proficiency. *Computer Assisted Language Learning*. 1 (21). DOI: <https://doi.org/10.1080/09588221.2019.1604551>
- Yeh, H. C., Tseng, S. S. & Heng, L. (2020). Enhancing EFL students' intracultural learning experiences through virtual reality. *Interactive Learning Environments*, 1-10. DOI: <https://doi.org/10.1080/10494820.2020.1734625>
- Watzlawick, P. (1978). *La réalité de la réalité*. Editions du Seuil.

Annexes 1

Description des six vidéos sélectionnées pour l'atelier des étudiants de master FLE.

Ressource	Type	Intitulé	Courte description du contenu et du mode de guidage
vidéo 1	360°	« <i>Take a Virtual Reality tour of six real exoplanets</i> » (4K, 360° VR experience)	L'exploration de six planètes par des vues rapprochées et des encarts scientifiques, guidée par une voix-off.
vidéo 2	360°	« <i>Discover Bangkok: A Guided City Tour - 360 VR Video</i> »	La visite des endroits les plus célèbres de la ville de Bangkok, avec un guide présent à l'image.
vidéo 3	360°	« <i>360° Underwater National Park - National Geographic</i> »	L'exploration sous-marine de Buck Island dans les Caraïbes, guidée par un guide photographe présent à l'image.

vidéo 4	RV 3D	« <i>The People's House - Inside the White House with Barack and Michelle Obama</i> »	Découverte des lieux emblématiques de la Maison Blanche, guidée par le couple Obama présent à l'image.
vidéo 5	RV 3D	« <i>Smithsonian Journey: Venice</i> »	Visite guidée des lieux emblématiques de la ville de Venise par un historien présent à l'image.
vidéo 6	RV 3D	« <i>Mission ISS</i> »	Monde synthétique proposant une exploration ludique et heuristique de la navette spatiale ISS, avec un tutoriel de prise en main.