

Serious game et apprentissage en réalité virtuelle : résultats d'une étude préliminaire sur la mémorisation en langue étrangère

Laurence Schmoll
Université de Strasbourg

Manuel Veit
Holo3

Mickaël Roy
École Supérieure de Pédagogie de Karlsruhe

Antonio Capobianco
Université de Strasbourg

À paraître in : *Synergies de pays germanophones* 7/2014.

Cette étude se propose d'analyser l'influence de l'interaction 3-D sur l'apprentissage d'une langue étrangère à partir d'une expérimentation menée auprès d'apprenants d'allemand langue étrangère. Les résultats doivent permettre de réfléchir à l'intégration de ce facteur technologique dans un scénario de *serious game*. Les résultats de l'étude empirique montrent une influence positive de l'interaction 3D sur la mémorisation lexicale à court terme, mais ne permettent pas d'observer une influence significative à long terme. L'étude souligne l'importance d'une phase de prise en main de la technologie d'interaction 3D en environnement virtuel, proposant quelques pistes de mise en œuvre de cette phase de découverte.

Jeux sérieux, réalité virtuelle, apprentissage, langues, interaction

Diese Studie untersucht den Einfluss der 3D-Interaktion auf das Fremdsprachenlernen. Sie basiert auf einer mit Lernern von Deutsch als Fremdsprache durchgeführten Untersuchung. Mit Hilfe der Ergebnisse soll die mögliche Integration dieser Technologie in die Konzeption von *Serious Games* erforscht werden. Die Ergebnisse der empirischen Studie zeigen einen positiven Einfluss der 3D-Interaktion auf das kurzfristige Behalten lexikalischer Einheiten. Die Studie erlaubt jedoch keine Rückschlüsse auf eine signifikante langfristige Wirkung. Die Studie unterstreicht außerdem die Wichtigkeit einer Einführungsphase, die den NutzerInnen erlaubt, sich mit der Technologie für 3D-Interaktion vertraut zu machen ; sie bietet außerdem einige Umsetzungsmöglichkeiten dieser Entdeckungsphase.

serious game, virtuelle Realität, Lernen, Sprachen, Interaktion

This study aims at analyzing the influence of 3-D interaction on learning a foreign language from an experiment conducted with German language learners. The objective is to draw empirical evidence on the usefulness of Virtual Reality technologies in the conception of a serious game. The results of the empirical study show a positive influence of 3D interaction on lexical short-term memorization, but did not revealed a significant influence on long-term memorization. The study also highlights the importance of the presence of a training phase of the use of the interaction techniques before any learning activities are performed. We propose some guidelines for the implementation of this training phase.

serious game, virtual reality, learning, languages, interaction

1. Introduction

Les jeux sérieux pour l'enseignement-apprentissage des langues revêtent des formes variées pour des objectifs pédagogiques différents. Une étude précédente (Schmoll 2011) montre que beaucoup exploitent des jeux déjà existants (jeux de l'oie, pendus, mots croisés, etc.), adaptés au support informatique, que le concepteur détourne dans une intention pédagogique. La plupart présentent également des objectifs essentiellement linguistiques : vocabulaire, point grammatical. Cette particularité s'explique par le fait que des objectifs de ce type n'appellent

généralement qu'une seule bonne réponse et entraînent par conséquent peu de marge d'erreurs dans la programmation et l'utilisation du jeu.

De telles applications présentent, à notre sens, un double défaut. D'une part, ces jeux ne sont en définitive que des exercices fermés cachés sous un habillage qui se veut ludique. Or, un des intérêts majeurs du jeu sérieux éducatif est de proposer au joueur une didactique la plus invisible possible (Ollivier 2012), une motivation intrinsèque découlant du plaisir immédiat du joueur qui, dans le meilleur des cas, expérimente le « flow » (1) (Csikszentmihalyi 1990). D'autre part, apprendre une langue ne se réduit pas à apprendre du vocabulaire et de la grammaire. Ce ne sont que les instruments d'un savoir-faire et d'un savoir-être à acquérir pour interagir en contexte (réception et production orales et écrites). Autrement dit, ces jeux ne permettent pas de développer la compétence de communication (Hymes 1984) chère aux méthodologies les plus récentes en didactique des langues : l'approche communicative (Coste 1976) et la perspective actionnelle (Conseil de l'Europe 2001).

À la lumière de ces deux constats, le projet EVEIL-3D s'intéresse à la conception de jeux sérieux plus ambitieux exploitant les ressorts du jeu vidéo de rôles dans un environnement immersif tridimensionnel. En effet, permettre à l'apprenant d'incarner un personnage dans un monde virtuel semble ouvrir les perspectives du point de vue des activités et des interactions (verbales, gestuelles et corporelles) situées (Mondada 2000) que nous lui proposons. Les objectifs du projet sont d'analyser comment le jeu sérieux en environnements immersifs permet un apprentissage de la langue cible, de fournir une solution matérielle (2) et logicielle « clé en main », à un coût abordable, pour les établissements du secondaire de la région tri nationale du Rhin supérieur, et enfin de sensibiliser les enseignants à ces nouvelles technologies.

Dans ce but, nous nous demanderons plus précisément dans le cadre de cet article, quelle peut être l'influence d'un facteur technologique (ici l'interaction 3-D) sur l'apprentissage en langue. Après avoir présenté un état de l'art, nous décrivons une étude menée sur l'influence de l'interaction gestuelle et corporelle, en environnements immersifs, sur l'apprentissage des langues. Nous présenterons ensuite les résultats obtenus et les analyserons dans le but de les intégrer dans le scénario de jeu vidéo sérieux.

2. Etat de l'art

2.1 Serious game

Le jeu vidéo sérieux ou *serious game* est défini comme une « application informatique, dont l'intention initiale est de combiner, avec cohérence, (...) des aspects sérieux (Serious) [...] avec des ressorts ludiques issus du jeu vidéo (Game) » (Alvarez 2007). Sa conception a pour ambition, dans le cas des serious games éducatifs, de développer la compréhension et l'acquisition de savoirs, savoir-faire et/ou savoir-être.

Djaouti (2011) propose le modèle DICE qui formalise les étapes de la conception des jeux vidéo sérieux. Cet outil théorique permet d'aboutir à un modèle générique itératif (conception / test / retour à la réalisation, etc.) pour les quatre phases : « Définir », « Imaginer », « Créer » et « Évaluer ». L'expérimentation que nous présentons dans cet article s'inscrit dans l'étape « Évaluer » puisqu'il s'agit d'évaluer l'influence du degré d'interaction gestuelle et corporelle, en environnement immersif tridimensionnel, sur l'apprentissage. Les résultats permettront de revenir aux trois premières étapes dans une perspective d'amélioration.

2.2 Réalité virtuelle et apprentissage

Les applications de la réalité virtuelle sont nombreuses. Elles vont de la revue de projet industrielle aux thérapies cognitivo-comportementales, en passant par l'apprentissage de procédures et de gestes techniques. L'utilisation de la réalité virtuelle pour l'apprentissage au sens large est apparue rapidement, dès les années 1995. Nous pouvons citer notamment ScienceSpace (Dede 1996). Ce travail suggère des gains intéressants associés à l'introduction de la sensorimotricité (i.e. l'activité motrice et les sensations) ou de la redondance entre les modalités sensorielles, par rapport à des environnements d'apprentissage fondés sur le visuel seul. Plus récemment, des études essaient de comprendre le rôle de la motricité dans le processus de mémorisation (Larrue 2012). Toutefois, cette dernière étude se focalise sur la mémorisation d'information spatiale, et non pas sur l'apprentissage.

La réalité virtuelle permet de rendre l'utilisateur acteur du monde virtuel dans lequel il est immergé. Cette « immersion corporelle » de l'apprenant est vue comme un agent catalyseur de l'acte d'apprentissage. C'est ce que

nous avons voulu vérifier dans cette étude, en nous focalisant sur l'apprentissage du vocabulaire lié aux informations spatiales.

3. Matériel et méthode

Étant donné que cette expérimentation a pour but d'évaluer l'influence de l'interaction sensorimotrice sur l'apprentissage, nous n'avons pas mis en œuvre une solution pouvant s'apparenter à un serious game, mais nous avons plutôt mis en place une activité d'apprentissage. Cette dernière se veut suffisamment générique pour apporter des éléments probants pour l'élaboration de scénarii de jeu vidéo sérieux en environnements immersifs. Dans cette activité, nous nous intéressons au vocabulaire de la spatialité intégré dans des actes de langage (consignes pour se repérer dans l'espace) contextualisés et associés à une activité sensorimotrice. Cela conduit l'apprenant non pas à simplement entendre et répéter du vocabulaire, mais à comprendre des énoncés et agir en conséquence.

3.1 Déroulement de l'expérimentation

Dans le cadre de cette étude expérimentale, nous avons fait passer 30 sujets (15 femmes et 15 hommes, tous âgés de 15 ou 16 ans) tous issus du même lycée et de la même formation. L'ensemble des sujets n'ont aucune expérience particulière des environnements immersifs.

Phase 1

Les sujets répondent à un premier questionnaire (Q1) visant à évaluer les compétences lexicales, en allemand langue étrangère, liées aux notions de placement dans l'espace. La tâche consiste à positionner, sur l'image d'un espace de vente, des objets suivant des consignes orales se focalisant sur le positionnement spatial d'objets (e.g. « Pose la lampe verte sur l'étagère du haut, sur la gauche » ; « Mets le tableau, sur la table, au milieu à l'arrière »).

Phase 2

Deux semaines après, ces sujets sont exposés à une phase d'immersion virtuelle. Durant cette phase, des consignes en allemand sont énoncées afin de permettre de réaliser l'aménagement d'un espace de vente. Pour ce faire, le sujet doit placer les objets virtuels à la bonne position dans l'espace de vente. Initialement, les objets sont placés sur une table devant le sujet. Lorsque le positionnement d'un objet est juste, la consigne suivante est énoncée. Si l'objet sélectionné ou la position cible sont fausses, l'objet retourne à sa position initiale sur la table. Ce processus continue jusqu'au positionnement des 17 objets ou jusqu'à la fin du temps imparti (20 min). À la fin de cette phase, les sujets remplissent le même questionnaire que lors de la phase 1 (Q3).

Les sujets sont divisés en deux groupes : un groupe exposé à un niveau d'interaction faible et un groupe exposé à un niveau d'interaction fort.

- Niveau d'interaction faible : l'environnement présenté au sujet est statique, et les actions (choix de l'objet et choix de la position) sont entièrement réalisées avec le pavé directionnel du périphérique. Une première étape permet de sélectionner l'objet à positionner. Une seconde étape permet alors de sélectionner la position cible. La sélection de l'objet ainsi que la sélection de la position cible se fait en utilisant les touches gauche et droite du pavé directionnel de la Wiimote© dans une liste circulaire (i.e. que lorsque l'utilisateur atteint la fin de la liste, la sélection est déplacée au début de cette liste). Tout au long de la manipulation, il est demandé au sujet de maintenir les deux bras le long du corps et de maintenir une position fixe debout ;
- Niveau d'interaction fort : un dispositif de capture de mouvement permet non seulement au sujet d'avoir du head-tracking (recalcule de l'image en fonction du point de vue de l'utilisateur, ce qui donne l'illusion de mouvoir sa tête dans l'environnement virtuel et de l'explorer de façon naturelle), mais également de positionner les objets en les saisissant et en les plaçant à l'aide d'une métaphore d'interaction répandue en réalité virtuelle, le ray-casting (Bowman 1997). Cette métaphore consiste à attacher un rayon au périphérique d'interaction. En pointant sur un objet à positionner et en appuyant sur la gâchette de la Wiimote©, l'objet est alors comme attaché au rayon. Tout mouvement en translation ou en rotation du rayon implique les mêmes transformations à l'objet sélectionné.

Phase 3

Deux semaines après la phase d'immersion, les sujets répondent à un questionnaire similaire à celui de la Phase 1, afin d'évaluer les compétences acquises (Q4).

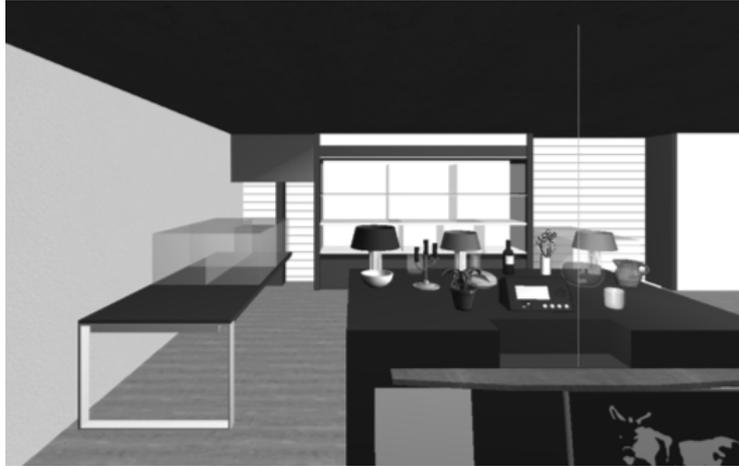


Fig. 1 L'environnement virtuel à aménager. Les objets à positionner se trouvent sur la table, les positions cibles sont matérialisées par les cubes transparents.

3.2 Matériel

Pour la phase d'immersion, nous avons utilisé un visiocasque Sony HMZ-T1 (cf. Figure 2). Les sujets ont bénéficié d'une vision stéréoscopique, avec une résolution de 1920 x 1080 pixels (soit 960 x 1080 pixels par œil). La capture de mouvement a été réalisée en utilisant l'Optitrack V120:Duo. La capture de mouvement permet de suivre la position de la tête de l'utilisateur, mais également le périphérique d'interaction (une Wiimote© sur laquelle nous avons fixé un capteur 6 degrés de liberté). L'ensemble fonctionnant sur un ordinateur portable (équipé d'un processeur Intel Core i7 Q740 1.73 Ghz, de 8 Go de RAM et d'une carte graphique NVidia GTX 460M)

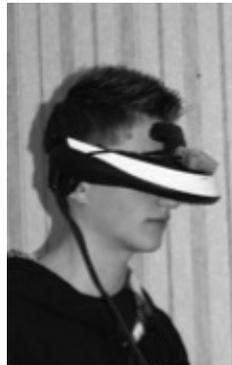


Fig. 2 Le visiocasque est relié à l'ordinateur portable et au système de capture de mouvement.

4. Résultats

Le Tableau 1 indique les moyennes des taux de réussite aux différents questionnaires. Nous avons réalisé un test de Student comparant les deux groupes (pour chaque questionnaire). Le seuil de significativité est fixé à 5%. Nous remarquons qu'il n'existe pas de différences significatives entre les deux groupes. Cela semble indiquer que le niveau d'interaction n'a pas d'influence sur l'apprentissage du vocabulaire lié à la spatialisation.

Tableau 1: Les moyennes des taux de réussite aux questionnaires Q1, Q3 et Q4.			
	taux de réussite à Q1	taux de réussite à Q3 (immédiatement après)	taux de réussite à Q4 (deux semaines après)

	(avant immersion)	l'immersion)	l'immersion)
interaction faible	38.96	45.91	42.68
interaction forte	37.37	57.26	53.23
<i>p</i> -value	<i>p</i> = 0.55	<i>p</i> = 0.19	<i>p</i> = 0.24

Dans le Tableau 2, nous indiquons les valeurs de *p* (toujours calculé à l'aide d'un test de Student) entre les taux de réussite aux questionnaires Q1, Q3 et Q4 pour les deux groupes. On remarque que seule la différence entre le questionnaire Q1 et Q3, pour le groupe "Interaction forte", indique une tendance. Ce qui semble indiquer que l'interaction forte peut mener à un meilleur apprentissage du vocabulaire lié à la spatialité.

Tableau 2 : Comparaison entre les taux de réponses justes aux différents questionnaires.		
	Q1 vs. Q3	Q1 vs. Q4
interaction faible	<i>p</i> = 0.29	<i>p</i> = 0.39
interaction forte	<i>p</i> = 0.07	<i>p</i> = 0.14

5. Discussion et conclusion

5.1 Discussion

Notre hypothèse de départ considérait qu'une interaction forte pouvait permettre à l'apprenant de se sentir davantage engagé (Brougère 2007) dans une action nécessitant un certain niveau de compréhension langagier. L'objectif attendu est d'estomper la situation d'apprentissage et de donner du sens à l'acquisition de la langue cible à travers une contextualisation par l'immersion visuelle. Par ailleurs, la dimension constructiviste de l'activité avait pour ambition de démontrer une amélioration de l'acquisition du vocabulaire de la spatialisation à travers le « faire » et la sensorimotricité (Dede 1996).

Les résultats que nous avançons ici permettent de mettre en lumière le fait que l'interaction 3-D (et l'engagement corporel qu'elle induit) ne semble pas toujours avoir d'influence majeure sur l'apprentissage d'un lexique pourtant adapté aux environnements immersifs tridimensionnels.

À court terme, nous constatons que l'apport d'une interaction forte permet une meilleure mémorisation du vocabulaire (différence significative entre le taux de réussite à Q1 et Q3 pour l'interaction forte), mais sur le plus long terme, nous n'observons pas d'acquisition (différence non significative entre le taux de réussite à Q1 et Q4 pour l'interaction forte). Ce résultat contradictoire nous paraît s'expliquer par le fait que les sujets ont abordé l'activité sans aucune expérience particulière des environnements immersifs. En effet, nous avons souvent observé des difficultés liées à l'utilisation de la capture de mouvement, ne rendant pas transparente l'utilisation de la technologie et gommant ainsi potentiellement les gains attendus par l'interaction gestuelle et corporelle sur l'apprentissage.

Lors de l'expérimentation nous n'avons pas accordé un temps suffisamment important à la prise en main de l'environnement immersif, prenant certainement inconsciemment le parti pris que cette catégorie d'âge présente une plus grande adaptabilité aux nouvelles technologies. Or, les résultats mitigés nous incitent à penser, qu'avant toute chose, il faudrait proposer une phase d'apprentissage de cette dite technologie. L'objectif serait ici de permettre à l'utilisateur de prendre en main les différentes facettes de la manipulation en environnements immersifs afin de réduire au maximum la charge cognitive liée à la manipulation et ainsi de maximiser les gains potentiels liés à l'engagement corporel et au constructivisme de l'activité.

La question se pose alors, dans le cadre d'un jeu vidéo sérieux, de savoir comment intégrer cette étape supplémentaire dans un scénario dont les « ressorts ludiques » (Alvarez 2007) risquent d'être mis à mal par les objectifs sérieux à atteindre. Pour ce faire, il faut tout d'abord que le matériel et son utilisation soient justifiés par le scénario. Pour le projet EVEIL-3D, nous avons fait le choix d'un scénario à réalité alternée expliquant l'entrée dans le dispositif (mentionné, en note de bas de page, dans la partie 1) par la recherche, dans l'environnement

virtuel, d'un architecte de la Cathédrale Notre-Dame de Strasbourg. Ce dernier a été inexplicablement aspiré dans la maquette 3-D de cette Cathédrale. Mandaté par la fille de l'architecte dans la réalité (réception d'une lettre, correspondance par courriel, recherches initiales sur des sites Internet qui permettent une sensibilisation au contexte culturel et au champ lexical de l'architecture attendant), le joueur-apprenant, qui a la particularité d'appartenir à une génération ayant la faculté d'agir sur le monde virtuel, intègre ainsi le dispositif. Il permet alors à la fille (qui est un allié radio du joueur) de suivre l'évolution du joueur à distance et le smartphone, couplé à un capteur 6 degrés de liberté, représente le moyen d'interagir avec l'environnement virtuel.

Par ailleurs, pour l'apprentissage de la technologie en tant que telle, différentes solutions peuvent être avancées avant l'entrée dans le scénario. Une phase de découverte peut, par exemple, être proposée sous la forme d'un tutoriel non scénarisé. Toutefois, cette étape retarde l'entrée dans le jeu et peut paraître ennuyeuse et dirigiste, ce qui contrevient à notre volonté de didactique invisible. Une autre solution est celle de l'espace freeplay. Ce choix présuppose néanmoins que les joueurs ont déjà en partie une certaine connaissance, voire compétence, des commandes et des interactions avec le dispositif. Or, nous avons pu constater que ce n'était pas le cas, notamment dans le cas d'une immersion tridimensionnelle. La solution qui nous paraît la plus appropriée est donc, dans la partie du scénario qui se passe dans l'univers virtuel, une première phase qui intègre uniquement la manipulation des moyens technologiques, ainsi que l'ensemble des interactions corporelles basiques à maîtriser dans un univers immersif tridimensionnel : navigation, sélection, manipulation et contrôle d'applications, qui sont les interactions fondamentales en interaction 3-D (Bowman 2004). Durant cette étape, il s'agira pour le joueur d'accomplir un certain nombre d'actions nécessitant ces interactions ; actions qui n'auront rien à voir avec un quelconque autre apprentissage et pour lesquelles les consignes auront été données en dehors du dispositif, durant la partie du scénario qui se déroule en classe. L'enseignant pourra même s'assurer de la compréhension de ces consignes par un usage de la langue maternelle de l'apprenant. Nous proposons ainsi une progression qui limite, autant que faire se peut, la surcharge cognitive, en demandant au joueur-apprenant :

1. D'observer, de se déplacer et d'agir / interagir corporellement et gestuellement dans le dispositif afin de lui laisser le temps de prendre en main les moyens technologiques à sa disposition ;
2. De comprendre des énoncés langagiers et d'agir / interagir en conséquence ;
3. De réinvestir les notions rencontrées dans la phase précédente dans des interactions verbales.

5.2 Conclusion

Dans cet article, nous avons présenté les résultats d'une étude préliminaire visant à évaluer l'influence d'une interaction gestuelle et corporelle sur l'apprentissage d'une langue en environnements immersifs. Les résultats obtenus semblent indiquer que cette immersion corporelle a une influence positive sur l'apprentissage. Toutefois, afin de maximiser ce gain, l'interface entre l'Homme et la Machine doit être maîtrisée par le joueur. Dans le cadre d'un serious game immersif et interactif, nous proposons plusieurs solutions pour tenir compte de cette phase de prise en main. Dans une prochaine étude, nous évaluerons l'influence des différentes solutions proposées, d'une part sur la maîtrise de l'environnement immersif et de l'apprentissage de la langue, et d'autre part sur l'intégration dans un *serious game*.

Remerciements

Le projet EVEIL-3D est cofinancé par l'Union Européenne via le Fonds Européen de Développement Régional (FEDER), dans le cadre du programme INTERREG IV - Rhin Supérieur.

Bibliographie

- Alvarez J., 2007, *Du jeu vidéo au Serious game. Approches culturelle, pragmatique et formelle*, Thèse de Doctorat en Science de la communication et de l'information, Université de Toulouse II et III, France.
- Bowman, D., Hodges, L. 1997, « An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments », In the *Proceedings of the 1997 Symposium on Interactive 3D Graphics*.
- Brougère G., Bézille H. 2007, « Des usages de la notion d'informel dans le champ de l'éducation » In *Revue Française de Pédagogie*, n° 158, 117-160.

- Conseil de l'Europe - Division des Politiques Linguistiques 2001, *Un cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer*, Paris : Didier.
- Coste, D., Courtillon, J., Ferenczi, V., Martins-Baltar, M., Papo, E. 1976, *Un niveau-seuil*, Strasbourg : Conseil de la coopération culturelle du Conseil de l'Europe.
- Csikszentmihalyi, M. 1990, *Flow, the Psychology of Optimal Experience.*, New York : Harper and Row.
- Dede, C. a. 1996, « ScienceSpace: virtual realities for learning complex and abstract scientific concepts », *Proceedings of the 1996 Virtual Reality Annual International Symposium (VRAIS 96)*, 246-254.
- Djaouti D. 2011, *Serious Game Design. Considérations théoriques et techniques sur la création de jeux vidéo à vocation utilitaire*. Thèse de Doctorat en Informatique, Université de Toulouse III, France.
- Hymes, D. H. 1984, *Vers la compétence de communication*, Paris : Hatier.
- Lane, G. 1970, *How to do things with Words*, éd. The William James Lectures delivered at Harvard University.
- Larrue, F. a. 2012, « Brain Computer Interface vs Walking Interface in VR : The Impact of Motor Activity on Spatial Transfer ». *ACM symposium on Virtual Reality Software and Technology (VRST'2012)* , 113-120.
- Mondada L., Pekarek Doehler S. 2000, « Interaction sociale et cognition située : quels modèles pour la recherche sur l'acquisition des langues ? », *Acquisition et interaction en langue étrangère* [En ligne], 12 | 2000, mis en ligne le 09 novembre 2010, consulté le 21 novembre 2013. URL : <http://aile.revues.org/947>
- Ollivier, C. 2012, « Approche interactionnelle et didactique invisible – Deux concepts pour la conception et la mise en œuvre de tâches sur le web social ». *Médias sociaux et apprentissage des langues : (r)évolution?* Alsic , 15 (1).
- Schmoll, L. 2011, « Usages éducatifs des jeux en ligne : l'exemple de l'apprentissage des langues ». *Revue des sciences sociales*, n° 45.

¹¹ Le « *flow* » est l'état mental qu'une personne atteint lorsqu'elle se trouve complètement engagée dans ce qu'elle fait et qui se mesure notamment par la réticence qu'elle montre à s'arrêter.

² Le dispositif est constitué d'une TV stéréoscopique 3-D, d'une paire de lunettes de stéréovision, d'une capture de mouvement à 6 degrés de liberté, d'un micro-casque, d'un smartphone et de deux PC (un premier pour l'affichage graphique, et un second pour la reconnaissance de la parole).