Comprendre les fonctions ludiques du son dans les jeux vidéo : pour la formulation d'un cadre théorique de sémiotique multimodale

Sébastien Genvo, Nicole Pignier

▶ **To cite this version:**

Sébastien Genvo, Nicole Pignier. Comprendre les fonctions ludiques du son dans les jeux vidéo : pour la formulation d'un cadre théorique de sémiotique multimodale. Communication - Information, médias, théories, pratiques, 2011, 28 (2), 10.4000/communication.1845 . hal-02169909

HAL Id: hal-02169909

https://hal.univ-lorraine.fr/hal-02169909

Submitted on 1 Jul 2019

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.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Comprendre les fonctions ludiques du son dans les jeux vidéo : pour la formulation d'un cadre théorique de sémiotique multimodale

Sébastien Genvo[1]
Nicole Pignier[2]

Résumé : Malgré l'importance du son dans l'expérience des joueurs, cette dimension du jeu vidéo est encore absente des théories du domaine. Un bilan des quelques travaux sur le sujet nous permet tout d'abord de cerner les apports et limites des théories filmiques, narratives quant à la compréhension des effets de sens du son dans ces jeux. La vocation ludique du médium incite à proposer un cadre théorique complémentaire permettant de guider les travaux sur ce terrain de recherche.

Mots-clés : Sémiotique multimodale – Théories filmiques – Aire intermédiaire d'expérience – Jeu – Fiction – Son

Abstract : Despite of the importance of sound in the player experience, this aspect of video games is still underrepresented in game theories. Starting with an overview of the few existing works on the topic, we identify the contributions but also the limits of narrative and film theories for the understanding of game audio. Due to the playful aspect of the medium, we propose a complementary theoretical framework in order to guide works on this extensive field of research.

Keywords : Multimodal semiotics – Film theories – Intermediary area of experience – Play – Game – Fiction – Sound

[1] Université de Lorraine, Centre de recherche sur les médiations, sebastien.genvo@univ-lorraine.fr
[2] Université de Limoges, Centre de recherches sémiotiques, nicole.pignier@gmail.com

À l'inverse de ce que laisse augurer son appellation, le jeu « vidéo » n'est pas uniquement un objet visuel, mais aussi sonore, comme en atteste aujourd'hui le succès commercial conséquent de titres fondés sur une pratique musicale (*Guitar Hero*, *Singstar*,…), l'emploi de la bande son de jeux vidéo pour diffuser des morceaux de groupes et musiciens connus[3], les nombreux concerts donnés en hommage aux musiques de jeux de toutes époques[4], ou encore l'utilisation dans la musique électronique de sons produits par des consoles en vogue dans les années 80[5]. Ces quelques exemples montrent que la dimension sonore a des implications sur la pratique vidéoludique en tant que telle et qu'elle affecte aussi suffisamment les joueurs pour aller au-delà de ses fonctions narrative et descriptive. Pourtant, le son est encore le parent pauvre des théories naissantes du jeu vidéo, puisqu'il a quasiment été ignoré jusqu'à présent. Si l'on considère par exemple les quelques anthologies de références publiées sur le champ naissant des « *game studies* » (Raessens, Goldstein, 2005 ; Perron, Wolf, 2003, 2009 ; Salen, Zimmerman, 2006), aucun article n'est consacré au sujet. Au regard de l'importance non négligeable que recouvre le son dans l'expérience ludique des joueurs, ce manque théorique incite à se demander en quoi le son participe à l'engagement, à la sensation de « présence » dans l'univers du jeu. Comment le son, dans ses rapports aux autres modalités (visuelle, cinétique, gestuelle), se propose-t-il d'affecter, de déformer, de modifier le champ de perception du joueur et par-là même, de susciter des émotions ? Quel fondement sensible donne t'-il à la jouabilité de l'œuvre ? En somme, de quelles façons le son fait-il sens dans un cadre de jeu sur support numérique pour impliquer le joueur et lui permettre de jouer ?

Pour esquisser des pistes de réponse, il faut relever que quelques publications (presque toutes anglophones) ont posé récemment les premiers jalons quant à la spécificité du son vidéoludique et quant aux similitudes entretenues avec d'autres médias. À de nombreux égards, le son dans les jeux vidéo reprend plusieurs conventions établies par le cinéma, tout comme le 7ème art a par exemple emprunté à l'opéra, au théâtre et au cirque le modèle de la « fosse d'orchestre » pour les musiques extra-diégétiques (Jullier, 2002 : 14). Les deux domaines ont d'ailleurs

[3] Le jeu à succès *Grand Theft Auto 4* (Take Two, 2008) propose par exemple au joueur, qui incarne un gangster, une radio lui permettant de choisir sa propre musique lorsqu'il est en voiture. Chaque station de radio a sa propre thématique (Rock, Jazz, Hip-Hop,…) et permet d'écouter des artistes aussi variés que David Bowie, Phillip Glass ou encore Duke Ellington.

[4] Le 18 décembre 2008 a par exemple eu lieu au Palais des congrès de Paris la manifestation *Video game live*, concert symphonique consacré à la musique de jeux vidéo.

[5] Le festival Data Airlines à Marseille repose ainsi sur plusieurs jours de concerts animés par des groupes employant ce genre de matériel.

toujours fait preuve d'une dynamique d'échanges depuis près de trente ans, tant pour des éléments de contenus (franchises, sorties simultanées d'un film et de son jeu vidéo, etc.) que de formes (techniques de mise en scène de la narration, etc.)[6]. Ces liens entre l'objet vidéoludique et le cinéma ont incité une grande part des auteurs s'étant consacrés au sujet à revisiter les théories filmiques dédiées au son pour identifier les similitudes ou les différences. Plusieurs des textes portant sur le sujet ont par exemple avant tout abordé le rôle narratif du son dans des jeux vidéo d'aventure (Collins, 2007 ; Whalen, 2004).

Un bilan synthétique des travaux s'ancrant dans cette perspective de recherche nous permettra dans une première partie de cerner les apports mais aussi les limites des approches narratives quant à la compréhension des effets de sens véhiculés par le son dans les jeux vidéo. À ce titre, si les modes d'utilisation du son vidéoludique héritent d'une généalogie qui doit être prise en compte, la nature intrinsèquement participative de l'objet incite également à proposer un cadre théorique complémentaire pour tenir compte de ses spécificités. Afin de formuler celui-ci, nous interrogerons ce qui fait la singularité d'*une expérience ludique*. Nous verrons plus particulièrement que jouer implique de se sentir présent au sein de l'« aire intermédiaire d'expérience » que constitue chaque jeu. Cette caractérisation nous amènera à élargir notre propos, afin de clarifier certains fondamentaux quant à la conceptualisation du jeu vidéo en tant que médium[7] et en tant que forme particulière de jeu. Loin de constituer des digressions théoriques, ces précisions notionnelles permettront de mieux cerner dans une seconde partie le rôle ludique du son au sein des autres modalités sensibles (gestes, images), en convoquant une approche qui allie à la sémiotique de la présence (définie par Jacques Fontanille, 1999) une sémiotique multimodale (Pignier, Drouillat, 2004). Cette dernière propose précisément une appréhension des liens entre les fonctionnalités sonores, visuelles, gestuelles en intégrant une sémiotique du sensible. Il ne s'agira pas ici d'énumérer,

[6] Pour approfondir la thématique de l'« intermédialité » afférente au jeu vidéo et au cinéma, nous recommandons au lecteur l'article d'Alexis Blanchet (2008), « Cinéma et jeux vidéo : trente ans de liaisons » (comme le souligne Éric Méchoulan [2003 :10], l'intermédialité « observe qu'une œuvre ne fonctionne pas seulement dans ses dettes plus ou moins reconnues envers telles autres œuvres, ou dans la mobilisation de compétences discursives (au besoin usurpées), mais également dans le recours à des institutions qui en permettent l'efficacité et à des supports matériels qui en déterminent l'effectivité »).

[7] Comme le souligne Éric Méchoulan (2003 : 15-16), le médium est « ce qui permet les échanges dans une certaine communauté à la fois comme dispositif sensible (pierre, parchemin, papier, écran cathodique sont des supports médiatiques) et comme milieu dans lequel les échanges ont lieu ». L'emploi du mot médium pour qualifier le jeu vidéo rejoint pleinement cette définition dans le cadre de notre article, où il s'agit de considérer comment la modalité sonore dans le jeu vidéo concourt à un processus de médiation ludique (Genvo, 2009), en orchestration avec les autres modalités suscitées par l'objet.

avec une quelconque prétention d'exhaustivité, l'ensemble des effets sémiotiques du son présents dans les jeux vidéo, vu la diversité des types de jeux et la spécificité de chacun. L'objectif de notre approche est bien de proposer un cadre permettant de guider des études sémiotiques sur ce champ de recherches conséquent. Afin d'en montrer de possibles voies d'application, deux jeux (récents) serviront à l'exemplification de cette seconde partie[8].

Le premier, *Dead Space* (Electronic Arts, 2008), se déroule dans un univers de science-fiction horrifique. Comportant de nombreuses références à des films tels qu'*Alien* (Scott, 1979), *2001 Odyssée de l'espace* (Kubrick, 1968) ou *Event Horizon* (Anderson, 1998), le joueur incarne Isaac Clarke, un ingénieur envoyé sur un croiseur interstellaire victime d'une avarie, au bord duquel se trouvait justement sa petite amie. À son arrivée, le joueur va néanmoins découvrir de nombreux cadavres et la présence de créatures hostiles, les « nécromorphes »… La thématique de cette œuvre permet de pointer l'inscription du paysage sonore du jeu vidéo dans une généalogie médiatique tout en mettant l'accent sur l'impératif d'action qui est au cœur des processus de médiation du jeu vidéo. Comme l'indique Bernard Perron, chercheur spécialiste des jeux vidéo d'horreur,

> les notions de contrôle (ou la perte de celui-ci) et d'action sont au cœur de l'expérience du jeu vidéo. On comprend alors pourquoi la peur demeure l'émotion vidéoludique la plus exploitée. Elle est clairement orientée vers un objet et vers un but. C'est une émotion primaire et prototypique qui a une forte tendance à l'action (Perron, 2005 : 360).

[8] La constitution de notre méthodologie d'analyse pour ces deux jeux a été fondée sur les recommandations données par Bernard Perron et Mark J. P. Wolf (2009 : 10-13) au sujet des impératifs méthodologiques suscités par l'analyse du jeu vidéo. Ce dernier étant en effet un phénomène complexe, comportant de « multiples couches », il est nécessaire que le chercheur cerne précisément la question posée, et en même temps qu'il est la capacité de jouer le jeu, ceci avant même d'exercer ses compétences d'interprétation. Rejoignant ces remarques, nous avions en outre montré dans une autre recherche (Genvo, 2009) la nécessité d'adopter une posture réflexive après avoir expérimenté les modalités d'engagement ludique proposées, ce qui revient notamment à incorporer et à maîtriser l'interface logicielle (l'importance de cette expérimentation est aussi soulignée par Amato, 2003). Dans le cadre du présent article, l'analyse des deux jeux a avant tout été guidée par un souci d'exemplification des éléments théoriques présentés. L'un des deux auteurs a pratiqué pendant plusieurs dizaines d'heures chaque jeu, de façon à s'assurer de la bonne maîtrise de l'interface et de la réalisation des principaux objectifs. L'autre auteur a fait une observation de plusieurs cessions de jeu, selon des critères sémiotiques expliqués dans l'article, ce qui a garanti l'adoption d'une posture réflexive permettant de dégager les éléments signifiants pour l'exemplification. La méthodologie appliquée à l'analyse de ces deux jeux se veut donc avant tout qualitative et descriptive. Pour avoir davantage de détails sur les implications de la formulation d'une analyse descriptive dans les jeux vidéo, voir Genvo, 2009.

Le second jeu qui permettra d'illustrer notre cadre théorique est *Rez* (Sega, 2001). À la façon du film *Tron* (Disney, 1982) – auquel le titre fait référence[9] – le jeu se déroule au sein d'un système informatique, le joueur étant un virus à forme humaine qui doit nettoyer un réseau. Initialement intitulé *Project-K* en hommage au peintre russe Vassily Kandinsky, *Rez* est fondé sur la création répétée d'effets « synesthésiques »[10] ceci afin de renforcer la jouabilité de l'œuvre. Cette seconde étude de cas nous permettra d'éprouver l'opérationnalité de notre cadre d'analyse sur un jeu qui, tout en n'ayant pas de vocation narrative, fait un usage central du son dans la construction de sa dimension ludique. Le lecteur pourra trouver des vidéos présentant des extraits de ces deux jeux à l'adresse suivante : http://www.ludologique.com/publis/videos_sound_design.html

Des théories filmiques et narratologiques aux spécificités de l'objet vidéoludique

1. La notion de diégèse, un point d'ancrage récurrent

> Les jeux sont très différents d'autres formes de médias culturels, et pour de multiples façons l'utilisation d'anciennes formes de théories culturelles est inappropriée pour les jeux. Néanmoins, il y a des moments où les distinctions entre les diverses formes médiatiques – de même que les parallèles ou similitudes – éclairent certaines idées ou concepts intéressants, qui en certaines façons inscrivent les jeux dans une continuité avec les médias linéaires, et en d'autres façons les distinguent (Collins, 2008 : 5) [11].

La précédente citation est emblématique des positionnements épistémologiques des chercheurs à l'égard de théories forgées pour les autres médias ; les concepts issus de ces dernières sont repris afin de juger leur adéquation. À ce titre, le concept de

[9] Au sein d'un entretien, l'un des concepteurs principaux du jeu, Tetsuya Mizuguchi, donne l'explication suivante du titre : « Au début, effectivement, le jeu devait s'appeler Project-K. Et j'avais un copain anglais qui est venu essayer le jeu. Je lui ai demandé quel titre il donnerait, lui. Et il m'a répondu : "Rez", parce que cela fait référence à Tron. Quand les personnages du film meurent, ils sont "derez". Et quand ils apparaissent, ils sont "rez" ». Consulté en ligne le 11/03/09, http://www.playscope.fr/layout.php?&cnt=1267&page=3

[10] Nous entendons par « effets synesthésiques » l'association de sens différents qui trouvent une correspondance entre eux.

[11] Notre traduction de l'anglais : « Games are very different from other forms of cultural media, and in many ways the use of older forms of cultural theories is inappropriate for games. However, there are places where distinctions between various media forms- as well as parallels or corollaries- highlight some interesting ideas and concepts that in some ways make games a continuation of linear media, and in other ways distinguish the forms ».

diégèse est employé de façon récurrente pour introduire la question du son dans les jeux vidéo. La grande majorité des auteurs s'étant intéressés au sujet affirment en effet l'opérationnalité de la distinction, établie au cinéma, entre sons diégétiques et extra-diégétiques pour décrire les fonctions sémiotiques du son[12] vidéoludique (qu'il s'agisse de musiques ou de bruits[13]). Ce faisant, ils relèvent aussi la nécessité de prendre en considération la part de l'utilisateur dans les processus de génération sonore. Karen Collins (2007, 2008) avance ainsi que les sons de jeux peuvent généralement être classés selon les deux catégories précédentes, mais qu'à l'intérieur de chacune d'entre-elles, il serait nécessaire de distinguer les sons « non-dynamiques » et « dynamiques ». Il faut préciser que Collins définit les « sons dynamiques » comme :

- d'une part, les sons qui se déclenchent directement suite à une action du joueur. Elle les qualifie d' « interactifs » : « Si par exemple le joueur appuie sur un bouton, le personnage à l'écran agite son épée et fait "swoosh". Presser le bouton à nouveau causera la récurrence de ce son. Ce "swoosh" est un effet sonore interactif »[14] (Collins, 2007 : 264).
- d'autre part, les sons « adaptatifs » qui répondent à des modifications de l'environnement de jeu et qui peuvent relever en partie du joueur, mais sur lesquels celui-ci n'a pas de contrôle direct. Par exemple, dans *Super Mario Bros*. (Nintendo, 1985), lorsque le joueur met trop de temps à compléter un niveau, le tempo de la musique double.

Ces multiples distinctions permettent alors à Karen Collins d'aboutir à une grille de description du son dans les jeux vidéo :

---

[12] Dans le cadre de notre article le terme « son » renverra de façon générale à tout type de phénomène sonore, que nous préciserons quand nécessaire.

[13] Pour différencier les notions de « bruit » et de « musique », nous renvoyons le lecteur à la distinction faite par Pierre Schaeffer sur la double polarité de l'emploi des sons. Celui-ci relève en effet dans son ouvrage *Traité des objets musicaux* (1966) que « la question posée par l'objet sonore et les façons de l'écouter nous obligent à interroger sur ce que nous appelons le musical ». Dans cette perspective, Schaeffer relève que les sons ne cessent d'effectuer un « aller et retour du musical conventionnel au sonore encore sauvage ». En somme, ils oscillent entre « leur emploi en indices, tout tourné vers l'événement, leur emploi en signes, tout assujetti au code » (Schaeffer, 1966 : 349). Dans ce cadre, la musique, à l'inverse du bruit, est perçue comme un ensemble de sons structurés, « assujetti au code ». Néanmoins, Schaeffer souligne qu'une recherche musicale ne peut échapper à ce choix des polarités qu'en les assumant toutes deux, « l'essence du phénomène musical est peut-être dans cet écartèlement, dans cette ambivalence » (*Ibid.* : 350). Nous nous réservons pour une étude ultérieure une analyse sémiotique du bruit dans les jeux vidéo, qui mettrait en exergue précisément l'importance sémiotique du bruit parmi les autres catégories de sons.

[14] Notre traduction de l'anglais. Pour désigner le bruit de l'épée, l'auteur emploie le terme de « swooshing noise », qui, tout en étant une onomatopée, désigne un bruissement, un son rapide : « if for instance a player presses a button, the character on screen swings their sword and makes a "swooshing" noise. Pressing the button again will cause a recurrence of this sound. The "swoosh" is an interactive sound effect ».

- Concernant les *sons extra-diégétiques* : les sons non dynamiques correspondent par exemple aux musiques accompagnant une séquence « cinématique »[15] d'un jeu vidéo. Les sons adaptatifs correspondent entre autres à l'exemple précédent dans *Super Mario Bros*., tandis que les sons interactifs sont par exemple les sons générés par les menus de l'interface de jeu.
- Concernant les *sons diégétiques* : les sons non-dynamiques sont par exemple ceux qui peuplent l'environnement du personnage mais sur lesquels le joueur ne peut intervenir. Pour les sons adaptatifs, Karen Collins prend l'exemple du jeu *The Legend of Zelda : Ocarina of times* (Nintendo, 1998), où à l'aube il est possible d'entendre le chant du coq, le son des oiseaux durant la journée et les hurlements d'un loup la nuit. Enfin, les sons diégétiques interactifs correspondraient notamment à l'exemple précédent de l'épée agitée par un personnage.

Cette classification semble toutefois très perméable, un son non-dynamique, comme le chant des oiseaux dans un environnement donné, pouvant finalement se révéler « adaptatif », voire interactif si l'on considère que le joueur en déplaçant son personnage dans une autre zone de jeu va peut-être faire changer l'ambiance sonore. Karen Collins précise à ce titre que les degrés d' « activités dynamiques » dans un jeu sont « fluides », pouvant rendre une entreprise de classification ardue. Dans cette optique, la grille précédente est davantage à comprendre comme un outil descriptif pouvant aider à comprendre les différentes fonctions signifiantes des sons rencontrés par le joueur.

2. Son dans les jeux vidéo et spatialité

Karen Collins, comme d'autres auteurs (Whalen, 2004 ; Stockburger, 2005 ; Arsenault, 2008), insiste de façon importante sur le rôle sémiotique du son comme aide pour évoluer dans l'espace de jeu. Cette fonction peut prendre plusieurs formes, Stockburger et Collins relèvent notamment la récurrence des procédés de visualisation de sons acousmatiques dans les jeux vidéo, en prenant tous deux

[15] Ce terme renvoie à une séquence dont le déroulement est prédéfini et dans laquelle le joueur ne modifier le contenu. Un jeu comme *Metal Gear Solid 4* (Konami, 2008) est par exemple connu pour faire un usage répété de ce genre de séquences.

comme référent théorique les écrits de Michel Chion sur l'audio-vision et sur le rôle du son au cinéma. Selon ce dernier, le terme acousmatique signifie :

> Que l'on entend sans voir la cause originaire du son […]. Dans un film, un son peut accomplir dès ses premières apparitions deux sortes de trajet : soit il est d'emblée visualisé, et ensuite acousmatisé, soit il est pour commencer acousmatique, et n'est visualisé qu'après (Chion, 2000 : 63-64).

Dans ce cadre, les films ont la possibilité de préserver longtemps la cause d'un son avant de la révéler, l'exemple le plus caractéristique étant la voix de la mère de Norman Bates dans *Psycho* (Paramount, 1960), que l'on entend sans qu'il nous soit donné de la voir. L'effet de surprise réside alors dans la découverte de la véritable cause de cette voix lors du dénouement finale… Du fait du contrôle du joueur, la visualisation des sons dans les jeux vidéo est bien plus aisée et le mystère est rarement entretenu aussi longtemps que dans le film d'Hitchcock. Ainsi, selon Axel Stockburger (2005), la fonction acousmatique dynamique contrôlée par l'utilisateur est l'une des fonctions de spatialisation les plus importantes dans les jeux vidéo. Comme nous l'avons indiqué, Karen Collins rejoint ce point de vue :

> L'action d'anticipation est une partie critique de la réussite dans de nombreux jeux, particulièrement dans les jeux d'aventure et d'action. Notamment, les sons acousmatiques – c'est à dire les sons sans origine visuelle définie – peuvent nous inciter à regarder dans la direction d'un son (Collins, 2007 : 270)[16].

Si les réflexions précédentes peuvent s'appliquer particulièrement aux bruits, Zach Whalen et Dominic Arsenault relèvent également un rôle similaire de la musique comme « aide à la perception de la diégèse spatiale » (Whalen, 2004). Dans de nombreux jeux, des leitmotivs musicaux sont notamment employés pour caractériser les lieux, ceux-ci pouvant répondre aux exigences d'orientation de l'action du joueur. Zach Whalen montre notamment comment les musiques de *Super Mario Bros*. permettent, selon les environnements, d'indiquer les zones de sécurité et de danger. Mais les deux auteurs précédents indiquent aussi que la musique permet de

[16] Notre traduction de l'anglais : « Anticipating action is a critical part of being successful in many games, particularly adventure and action games. Notably, acousmatic sound— that is, sound with no clear origin visually— may inspire us to look to the direction of a sound ».

renforcer la crédibilité de l'univers fictionnel et de la diégèse. Dominic Arsenault relève ainsi que dans les années 80,

> « l'expressivité sonore dont disposaient les compositeurs de musique de jeux vidéo était mise activement à contribution dans la construction d'un espace. Le graphisme rudimentaire de l'époque ne permettait pas les subtiles nuances nécessaires pour différencier l'atmosphère de deux niveaux : souvent, c'est à la musique qu'incombait cette tâche » (Arsenault, 2008).

Dans cette perspective, Zach Whalen propose de discerner deux grandes fonctions sémiotiques complémentaires du son vidéoludique. D'une part il s'agit de donner une plus grande crédibilité à l'univers diégétique afin de renforcer l' « immersion » du joueur dans la fiction ; d'autre part, il s'agit d' « engager » le joueur dans une certaine dynamique d'action vis-à-vis des données visuelles qui lui sont transmises.

3. Limites de la notion de diégèse dans la compréhension des phénomènes ludiques

Les réflexions précédentes nécessitent selon nous de revenir sur l'emploi du concept de diégèse fait par la plupart des auteurs comme socle de catégorisation du son vidéoludique. En effet, comme nous l'avons souligné en introduction, le corpus qui a composé une grande partie des analyses effectuées dans les articles précédemment cités reposait sur des jeux ayant une composante narrative, même minime (comme ce peut être le cas dans les aventures de Mario). Il est utile de rappeler que, selon Roger Odin, le concept de diégèse a été introduit dans le vocabulaire de la filmologie par Etienne Souriau en 1951 pour dénommer tout ce qui appartient, dans l'intelligibilité, à l'histoire racontée, au monde supposé ou présupposé par la fiction du film (Odin, 2000). Si cette définition peut s'appliquer à des jeux comportant effectivement une histoire, elle est plus difficilement applicable à d'autres types de jeux tels que le célèbre *Tetris*. Où se situerait dans ce cas la diégèse, et donc les sons diégétiques, en l'absence de tout personnage et d'histoire racontée ? On pourrait, pour répondre à cette question, rejoindre la définition qu'apporte Gérard Genette sur le terme. Dans sa perspective, la diégèse ne serait pas tant l'histoire mais l'univers où elle advient (1982 : 13). Il s'agit en somme de s'abstraire du

narratif. En conclusion de son article, Zach Whalen concède toutefois la difficulté de qualifier dans les jeux vidéo de nombreux sons selon la distinction diégétique / extra-diégétique, ce que remarque également Karen Collins (2008 : 125) qui souligne une grande fluidité des sons de jeux vidéo entre diégèse et extra-diégèse. À ce titre, ces deux notions ne sont pas précisément définies par Collins ou Whalen, ce qui peut parfois amener ces auteurs à prêter au concept une signification bien éloignée de son acception d'origine sans toutefois revenir sur les conséquences épistémologiques de cette modification. Karen Collins indique par exemple que « jouer à un jeu vidéo implique à la fois une activité diégétique et extra-diégétique : le joueur a une interaction consciente avec l'interface (la diégèse), de même qu'une réponse corporelle à l'expérience et à l'environnement de jeu (l'extradiégèse) »[17] (Collins, 2008 : 3). Comme on peut le voir, la qualification du terme est ici fort éloignée des définitions avancées par Souriau ou même Genette.

On constate donc une pluralité d'acceptions perméables prêtées à ce concept dans le cadre particulier du jeu vidéo, pouvant mener à des confusions. Il nous semble par conséquent nécessaire de trouver un concept complémentaire et approprié pour penser plus spécifiquement, d'un côté, le lien entre l'espace de fiction généré par un jeu et la réalité « ordinaire » du joueur et, de l'autre, les fonctions que peut recouvrir le son entre ces deux pôles. Pour ce faire, il nous semble adapté d'employer le concept *d'aire intermédiaire d'expérience* forgé par D.W. Winnicott dans son ouvrage *Jeu et réalité, l'espace potentiel* (1975).

4. Aborder le lien entre jeu et fiction par la notion d'aire intermédiaire d'expérience

Pour présenter ce concept, il faut en premier lieu revenir sur le lien qui existe entre jeu et fiction. De nombreux auteurs indiquent en effet que le joueur, lorsqu'il est en action ludique, adopte une *attitude mentale* qui est de l'ordre du faire « comme si », pour reprendre une expression de Roger Caillois, « faire comme si » il faisait quelque chose de différent. Pour cette raison, cet auteur considère le jeu comme une activité fictive », « accompagnée d'une conscience spécifique de réalité seconde ou de franche irréalité par rapport à la vie courante » (Caillois, 1956 : 43). Si le joueur

[17] Notre traduction de l'anglais : « Playing a video game involves both diegetic and extradiegetic activity: the player has a conscious interaction with the interface (the diegetic), as well as a corporeal response to the gaming environment and experience (extradiegetic) ».

d'échecs s'applique à déplacer minutieusement chacune de ses pièces en envisageant le poids que chacune de ses décisions prendra dans la suite du jeu, son geste ne saurait par exemple se résumer à cette seule action pour qu'il puisse effectivement y avoir jeu. Comme le disent les amateurs de jeu d'échecs, il ne s'agit pas uniquement de « pousser du bois ». Pour entrer véritablement dans le jeu, le joueur doit faire « comme si » le plateau et les pièces qui se présentent à lui ne se résumaient pas uniquement à ce qu'ils sont ; il doit entrer dans un univers de sens singulier qui va guider ses actes, *régit à la fois par des règles de la vie courante et par des règles spécifiques qui n'y ont pas habituellement cours*.

La notion d'aire intermédiaire d'expérience semble alors tout à fait appropriée pour qualifier ce lieu où se déroule le jeu : « Cette aire où l'on joue n'est pas la réalité psychique interne. Elle est en dehors de l'individu, mais elle n'appartient pas non plus au monde extérieur » (Winnicott, 1975 : 67). Pour reprendre la terminologie employée par cet auteur, le jeu ne prend place ni au sein de la réalité « extérieure », « du dehors », dont la perception peut être « partagée » et qui est de l'ordre du « non-moi », ni au sein de la seule réalité « intérieure », qui est de l'ordre du « moi » (elle ne saurait être uniquement considérée comme une construction mentale fictionnelle, comme un fantasme). La réalité dans laquelle se réalisera l'action du joueur prend place dans le cadre d'une aire qui se situe entre ces deux pôles, où le joueur va transposer les choses du monde au sein duquel il vit dans un ordre nouveau. Pour reprendre une expression employée par Jean-Marie Schaeffer, on peut avancer que le joueur adopte un comportement « biplanaire », ce qui est aussi avancé par Jesper Juul ou encore par Jacques Henriot. Ce dernier souligne que le joueur vit constamment sur deux plans, présent à ce qu'il fait mais également animé par un dessein qui porte ses actes au-delà du seul instant présent et de l'immédiateté de ses actions. Le joueur ne peut donc jamais totalement s'abstraire de la « réalité ordinaire ». En ceci, jouer, ce n'est pas rêver. Le jeu n'est donc pas tant une activité « fictive », comme le suggère Roger Caillois, que « fictionnelle ». Dans ce cadre, il est possible d'avancer que des jeux tels que *Tetris*, *Super Mario Bros*. comportent bien une dimension fictionnelle lorsqu'ils sont actualisés : comme l'indiquent Bernard Perron et Dominic Arsenault, jouer à un jeu vidéo ne revient pas à croire que l'on interagit avec une machine pour avoir un effet sur le monde réel (Perron, Arsenault, 2009 : 112). Mais en revanche, *Tetris*, à l'inverse de *Super Mario Bros*., ne demande pas à l'individu d'identifier une diégèse (au sens d'univers dans

lequel va prendre place une *histoire*), comme c'est le cas pour le second. Tous deux, en tant que jeux, constituent en revanche des aires intermédiaires d'expérience.

5. Les spécificités du jeu sur support numérique

Si le concept d'aire intermédiaire d'expérience s'applique à tout jeu, il faut néanmoins prendre en compte certaines spécificités du médium numérique en partant d'un constat simple mais essentiel pour comprendre les particularités ludiques du jeu vidéo : dans le cadre du jeu sur support numérique, le joueur va prendre connaissance des règles lors de l'actualisation même du système, alors que généralement dans un jeu traditionnel les règles sont données au joueur avant qu'il effectue son action de jeu. En d'autres termes, s'il n'est pas possible de jouer aux échecs sans connaître préalablement les règles précédant l'action, un logiciel informatique d'échecs peut permettre à l'utilisateur de jouer sans en connaître par avance les règles. La notion spécifiquement vidéoludique de *gameplay* renvoie à ce processus de découverte de la structure de jeu – compris comme un système de règles (en anglais *game*) – dans l'action ludique (*play*). Pour reprendre une réflexion de Vincent Mabillot (2001) au sujet des dispositifs de « médiations interactives » – une « partie du contenu est indéfinie. On ne se contente pas d'activer la matérialisation d'un contenu, mais bien de participer à son énonciation signifiante »[18]. Dans le cas des jeux vidéo, au-delà des logiques d'exemplification technologique visuelle ou sonore, c'est avant tout cette dynamique de co-énonciation signifiante propre à chaque jeu, en somme son *gameplay*, qui va susciter le (dé)plaisir du joueur[19], son engagement, son immersion, bref, sa sensation de « présence » au sein de l'aire intermédiaire d'expérience.

Comme le souligne Xavier Rétaux (2003), alors que certains chercheurs tels Rheingold (1993) et Cadoz (1994), pensent que c'est la capacité d'un système à simuler la réalité et à la reproduire fidèlement qui favorise l'immersion, à savoir le sentiment d'être pris dans le jeu, d'autres tels Marsh, Wright et Smith (2001) montrent que le sentiment d'être présent au sein d'un monde fictionnel est un construit. Dans cette thèse, l'utilisateur n'a pas besoin, pour se sentir dans le jeu, de systèmes qui simulent fidèlement le quotidien. Ce sont les automatismes qu'il se crée

[18] <http://grm.uqam.ca/activites/cmo2001/mabillot.html>, consulté le 27/03/06.

[19] Il est important de préciser que ce terme est employé de façon récurrente par les joueurs et professionnels pour qualifier ce qui fait la qualité d'un bon jeu indépendamment de ses caractéristiques techniques.

au fur et à mesure qu'il acquiert une bonne maîtrise des règles du jeu qui tiennent lieu de « naturel » et qui expriment, tout en l'amplifiant, le sentiment d'être présent au jeu et d'être absent du réel. Pour parvenir à ce sentiment de présence, il faut, d'après ces auteurs, que l'utilisateur soit maintenu dans le flux d'attention. Une étude menée par Xavier Rétaux montre que c'est avant tout cette seconde hypothèse qui serait à retenir dans le cas des jeux vidéo, ce qui rejoint notre remarque au sujet de l'importance de la qualité du *gameplay* pour le joueur.

Dans une enquête menée auprès de joueurs du jeu *Quake 3* (Id Software, 1999) - reconnu à l'époque de sa sortie pour faire preuve de graphismes particulièrement photo-réalistes vis-à-vis de la concurrence - Xavier Rétaux constate en effet que l'illusion physique de non médiation (qui est l'apanage de systèmes cherchant la reproduction fidèle du réel) n'a qu'une vie éphémère chez le joueur dans la création de la sensation de présence au sein d'un univers virtuel, même si elle concourt à l'impliquer dans un premier temps.

> Les joueurs nous apprennent que le jeu est surtout une activité construite. Les jeux vidéo sont conçus pour être maîtrisés. Ils ont chacun leurs règles, leurs voies vers la victoire et leurs contraintes techniques. Chaque jeu est un espace de liberté dans lequel le joueur s'exprime, un monde à part créé par un contrat ludique (Duflo, 1997). Le sujet va construire la transparence de son activité […] non pour masquer la réalité mais pour être efficace dans le jeu (*op. cit*.). Cette construction constitue le jeu et permet de se sentir présent en son sein (Rétaux, 2003 : 80).

Ce qui permet notamment à Xavier Rétaux d'arriver à ces conclusions c'est que les joueurs expérimentés de *Quake 3* préfèrent entre autres sacrifier la qualité des graphismes si cela leur permet d'être plus efficaces dans l'aire de jeu. En effet, dans ce jeu, il est possible de diminuer la qualité du rendu visuel, qui nécessite d'importantes ressources de calcul machine, afin que la vitesse d'affichage des images soit plus élevée, ce qui permet conjointement au joueur d'être plus réactif à l'ennemi. Mais, plutôt que de considérer que la dimension sensible du système de jeu serait secondaire, cette dernière réflexion montre au contraire l'importance de considérer l'adéquation de celle-ci vis-à-vis du *gameplay*. Dans ce cadre, comment le son, dans son interaction à l'image et au geste, concourt-il à impliquer l'individu dans une expérience typiquement ludique ? En quoi le son vient-il fonder la dimension

sensible de la jouabilité de l'œuvre ? Autrement dit, de quelle façon construit-il un lien entre l'ailleurs de l'espace fictionnel postulé par le jeu et la réalité extérieure du joueur de sorte à rendre ce dernier présent au sein d'une aire intermédiaire d'expérience ?

Une sémiotique multimodale du sensible appliquée au jeu vidéo

1. Le jeu vidéo comme acte de co-énonciation

En sémiotique, plusieurs travaux portent sur la théorie de la présence, regroupés sous le terme de sémiotique du sensible. Le concept de *présence*, chez Jacques Fontanille[20], désigne une expérience sensible, perceptive et affective. La présence suppose un acte d'énonciation qui permet de se sentir présent au monde et de rendre le monde présent à soi. Énoncer, dit Jacques Fontanille, c'est se rendre présent quelque chose à l'aide du langage. Cette assertion se fonde sur l'approche phénoménologique que Merleau-Ponty adopte pour analyser l'acte de perception[21] : *« Percevoir, c'est se rendre présent quelque chose à l'aide du corps »* (cité par Fontanille, 1998 : 92). Si chez Xavier Rétaux le concept de présence ne prend pas en compte les fondements phénoménologiques de l'approche sémiotique fontanillienne, il pointe cependant l'idée que le sentiment de présence vécu par le joueur se construit au fur et à mesure de l'action ludique. De notre côté, nous reprenons l'approche méthodologique de J. Fontanille pour analyser le champ de présence parce que, nous le verrons, elle semble fortement appropriée pour qualifier l'expérience sensible et perceptive du joueur au fil du jeu. Pour autant, nous préférons au postulat phénoménologique[22] de la fusion du sujet au monde un

[20] Jacques Fontanille préside la chaire de sémiotique à Paris depuis de nombreuses années. Ses travaux, dans la lignée de l'école du structuraliste Greimas, ont d'abord beaucoup apporté à la sémiotique narrative et à la sémiotique des passions. Depuis 1997, le chercheur a largement contribué à l'émergence d'une sémiotique du sensible et, plus récemment, il a développé une sémiotique des pratiques culturelles.

[21] En outre, Jacques Fontanille emprunte à la phénoménologie d'Husserl le concept d'*intentionnalité* qui repose sur l'imperfection de la présence. Le sujet va s'efforcer, au cours de l'acte d'énonciation, de retrouver une plénitude existentielle, une sorte de fusion avec le monde, un sentiment de coïncidence entre l'apparence des choses et ce qui émane de l'être même. C'est dans cette quête de sens qu'il va chercher à viser et à saisir les choses, au-delà de leur apparence, dans leur apparaître. Pour plus de précision à ce sujet voir Fontanille (1999 : 229-230).

[22] En phénoménologie puis en sémiotique, plusieurs travaux ont été consacrés à la théorie de la présence. Chez Husserl en particulier, la présence est une question de qualité, de degré avec lequel un sujet apprécie dans son univers de conscience un corrélat noématique. Ce dernier, défini comme l'objet intentionnel de la pensée, ne peut en aucun cas se réduire à l'actuel, au réel, à l'existant. En effet, le sentiment de présence ne nécessite pas un objet perçu pour avoir lieu. Portant la marque temporelle de la rétension comme dans le cas du souvenir ou de la protension comme dans le cas de l'imagination, la présence chez Husserl ne peut pas se confondre avec l'acte de perception et elle comporte en outre de fortes implications métaphysiques. Le jeu vidéo constituant une activité perceptive à part entière, l'approche conceptuelle husserlienne ne nous

postulat qui serait plus modestement énonciatif. Ainsi, nous observons le jeu comme un acte d'énonciation auquel le joueur est invité à participer, non pas pour retrouver l'apparaître des choses et un état fusionnel avec le monde, y compris le monde fictif proposé, mais pour vivre une continuité de sens entre le monde réel et le monde ludique. Comme nous l'avons souligné en présentant le concept de *gameplay*, tout en jouant, dans le cas des jeux vidéo, le joueur co-énonce[23], mentalement et par la gestuelle. Ce faisant, il cherche à donner sens à son activité en établissant une continuité de sens entre les mondes réel et fictionnel qui l'entourent au sein du jeu, défini comme une aire intermédiaire d'expérience. Pour le dire en termes sémiotiques, le réel et la fiction sont homogénéisés autour de l'opérateur de l'acte de co-énonciation[24], à savoir un corps sentant et percevant.

Cette construction du sens suppose que le joueur effectue, en même temps mais à un autre niveau que celui des actions, des prises de position sensible. Son corps sentant s'instaure alors comme zone de référence réagissant à la présence qui l'environne. Pour décrire les effets de présence vécus par le sujet sensible, Jacques Fontanille propose, en l'empruntant au linguiste Emile Benvéniste, le concept de *champ positionnel*. Le corps sentant prend en effet position dans un espace environnant – un champ- dans lequel il se vit comme présent au monde et dans lequel les choses et les êtres se rendent présents à lui. Le champ de présence du sujet sensible a les propriétés élémentaires[25] suivantes (*ibid.* : 97-98) :

- Le *centre de référence* qui est institué par le corps sensible lui-même dans la mesure où c'est ce dernier qui ressent et perçoit ce qui advient autour de lui ;

---

permet pas de comprendre le sentiment de présence au jeu vécu par le joueur. Greimas, de son côté, dans son *Dictionnaire raisonné de la théorie du langage*, exclut l'approche ontologique que la phénoménologie fait de la présence pour lui substituer une approche structuraliste : la présence est un concept opérationnel qui permet de déterminer, par une opposition catégorielle présence/absence, l'existence sémiotique d'un terme. Par nature constituée de pôles opposés et se focalisant uniquement sur une approche descriptive du langage mis en discours, l'approche structuraliste ne nous permet pas non plus de cerner les variations du sentiment de présence du joueur. En revanche, chez Jacques Fontanille, le concept de présence a peu à peu basculé du structuralisme vers une sémiotique tensive, apte à interroger l'expérience sensible et perceptive du sujet de discours.

[23] Le concept de *co-énonciateur* a été introduit par le linguiste A. Culioli, pour rendre sa place à la part d'activité mentale, émotionnelle revenant au destinataire d'un énoncé. Le linguiste Dominique Maingueneau, a lui aussi repris ce concept qui s'est étendu en sciences du langage avant d'être repris en sciences de l'information – communication. A fortiori, la participation gestuelle du joueur permet d'intensifier la co-énonciation, corporelle, et plus seulement mentale, que le joueur accomplit lors de son activité.

[24] Nous entendons par *énonciation* l'acte de prédication (en linguistique, prédiquer c'est dire quelque chose à propos de l'objet dont on parle) multimodal. Les sujets d'énonciation et de co-énonciation sont à la fois des sujets sensibles de la perception et des sujets discursifs de la prédication.

[25] Le champ du discours est le domaine spatio-temporel que l'instance de discours se donne en prenant position en vue d'une énonciation (Fontanille, 2004 : 98).

- Les *horizons du champ* qui délimitent le domaine de la présence du sujet au monde et du monde au sujet. Le sujet est sensible en effet à ce qui advient dans un espace-temps qui lui est relatif et au-delà duquel il ne se sent plus présent. Au-delà des horizons du champ qui lui est propre, il n'a plus de prise sur le monde et le monde n'a plus de prise sur lui ;
- La *profondeur du champ*, qui met en relation le centre et les horizons, profondeur dite *progressive* quand le corps sensible identifie peu à peu ce qui advient à l'horizon, profondeur *régressive* quand ce qui advient à l'horizon prend prise sur le sujet sensible à son insu ;
- Les *degrés d'intensité* et de quantité propres à cette profondeur. La profondeur est perçue comme une tension entre le centre et les horizons, tension qui dépend des variations d'intensité et de quantité ou étendue. Quand apparaît une très forte intensité, non plus au centre du champ mais à son horizon, se forme alors un autre champ positionnel, concurrent du premier, c'est-à-dire le champ de l'altérité. Le corps sentant perçoit alors émotionnellement la présence d'un autre corps sentant.

2. Quelques spécificités du champ de présence du joueur

- Les centres de référence

Dans les jeux vidéo, le joueur co-énonce à l'aide d'un avatar. Il y a alors une sorte de dédoublement des centres de référence qui permet tantôt au joueur de percevoir ce qui se passe dans le champ positionnel de l'avatar, comme si les deux centres de référence n'en faisaient plus qu'un, tantôt de s'instituer comme centre de perception distinct de celui qu'occupe l'avatar (le joueur perçoit alors autre chose que ce que l'avatar est censé percevoir[26]). Dans le premier cas, il s'agit par exemple des jeux dits « en vue subjective » (usuellement dénommés *First Person Shooters*), tels que *Doom* (Id Software, 1993), où les points de perception visuel et auditif correspondent au positionnement d'une instance de la fiction (en terme filmologique, il s'agirait d'une ocularisation et d'une auricularisation internes). Dans le second cas, les points

[26] Les personnages des jeux vidéo sont des algorithmes tout comme les personnages des livres sont des êtres de papier et il serait possible de rétorquer que, dans tous les cas, ils ne perçoivent rien. Pour autant, l'expérience fictionnelle suppose que les co-énonciateurs conçoivent les personnages comme des êtres doués de perception.

de vue et d'écoute sont extérieurs à un personnage ou à une entité, le joueur peut donc percevoir de l'extérieur l'instance qui le représente dans le monde du jeu. Ce dédoublement des champs positionnels permet au joueur, à la fois ou tour à tour, d'être, de vivre dedans et dehors la peau de l'avatar. Les choses se complexifient encore lorsque cette instance extérieure de perception (la « caméra ») est manipulable par le joueur. C'est notamment le cas dans un jeu de notre corpus : dans *Dead Space*, la caméra, qui est aussi le point de perception sonore, se trouve derrière l'avatar, le suivant constamment dans ses déplacements, un peu à la façon du jeu *Tomb Raider* (Eidos, 1996). Néanmoins, tout en pouvant faire agir son avatar (déplacement, tir, etc.), le joueur a aussi la possibilité de déplacer la caméra autour de celui-ci, il peut donc par exemple visualiser l'avatar de face pour percevoir ce qui passe dans son dos. La caméra n'est pas ici une entité qui prend corps dans la fiction, elle n'est pas non plus présente dans le monde extérieur du joueur, bien que manipulable via son contrôleur de jeu. En somme, trois corps percevant s'instaurent comme zone de référence, celui de l'avatar, celui de la « caméra » (qui permet la représentation audio-visuelle du monde fictionnel), et celui du joueur. Ce dernier doit orchestrer au mieux ces trois champs positionnels et, sous réserve d'une intégration parfaite des dispositifs de commande[27], se voit doté d'une *hyperperception*. Autrement dit, le joueur peut glisser d'un champ positionnel à l'autre, pour augmenter, décupler, ses capacités perceptives.

- Les horizons de champ

L'horizon du champ de l'avatar ainsi que celui de la caméra dépendent de celui du joueur et réciproquement. En fonction de la position spatiale qu'il va attribuer à son avatar, le joueur oriente son champ de perception dans tel ou tel lieu. L'horizon de champ de la caméra, quant à lui, dépasse ce qui est visualisé à l'écran par l'intermédiaire du son. La caméra peut discerner tel ou tel bruit (en tant qu'indice d'événements) dont la source n'est pas visualisée. Le joueur, orientant alors celle-là vers la source sonore, est amené à redéfinir sans cesse la périphérie de son propre champ de perception ; il est en ce sens doté d'*hypermobilité*. Effectivement, le fait de

[27] Sur Xbox 360 par exemple, via la manette de jeu, le joueur doit contrôler avec un *joystick* dépendant de son pouce gauche le déplacement de l'avatar, tandis qu'un *joystick* activé par le pouce droit permet de déplacer la caméra. Si la plupart des joueurs occasionnels intègrent rapidement cet impératif de commande, il faut noter que cela requiert un certain apprentissage pour le néophyte.

pouvoir glisser d'un champ de perception (celui du joueur) à un autre (celui de l'avatar, celui de la caméra), incite le joueur à faire bouger les limites de son propre champ de perception. Cette mobilité perceptive continue ou hypermobilité participe à la complexité de l'expérience co-énonciative dans les jeux vidéo.

- Les profondeurs de champ

Doté de plusieurs centres de références, démultiplié dans plusieurs champs positionnels en mobilité continue, le joueur s'invite du coup à une expérience sensible inédite, à une sorte d'*hyperprésence*. Cela, dans la mesure où il vit des expériences sensibles et perceptives multiples, en variation continue. En effet, la pluralité des champs qu'il occupe directement ou indirectement est en tension permanente entre *profondeur progressive,* quand un des centres de référence perçoit peu à peu le monde dans son champ et arrive à l'identifier, *profondeur régressive* quand quelque chose, à l'horizon d'un des champs, semble se rapprocher sans se faire connaître. Le joueur a alors conscience que quelque chose advient de dangereux, par exemple quand il s'agit d'un cri, mais n'arrivant ni à identifier ni à localiser ce « quelque chose » avec précision, il a le sentiment que son champ de perception se fait envahir par celui d'un sujet ennemi. Tel peut être l'effet émotionnel du son de nature acousmatique. Dans *Dead Space*, l'apparition d'un ou plusieurs monstres dans les alentours de l'avatar se fait fréquemment de façon abrupte. Précisément, le son de la créature précéde bien souvent sa visualisation. Un son récurrent peut être par exemple celui d'une grille d'aération qui tombe et par laquelle arrive un ennemi. La « musique », composée essentiellement de violons et de percussions, vient redoubler la présence de cette menace puisque la proximité d'un ennemi entraîne une « cacophonie dissonante »[28]. Ceci renforce le sentiment d'étrangeté pour le joueur ; la musique se rapproche davantage d'une texture sonore qui s'additionne aux cris et vociférations des monstres. Cette complexité sonore brouille la limite entre ce qui appartient ou non au monde postulé par la fiction (comme l'indique Pierre Schaeffer [1966], le son ne cesse ici d'osciller entre un

[28] C'est par le terme de « *blaring dissonant cacophony* » que le superviseur de la conception sonore de *Dead Space* qualifie la forme que prend à certains moments la « musique » du jeu. Voir http://www.originalsoundversion.com/?p=693, consulté le 11/05/09. Pour être plus spécifique dans la qualification de l'environnement musical, on peut rapprocher celui-ci des compositions de György Ligeti, dont les œuvres *Lontano* et *Lux Aeterna* ont notamment été employées dans *Shining* et *2001 Odyssée de l'espace* de Stanley Kubrick (de nombreuses références étant faites à ce dernier film dans *Dead Space*, l'histoire nous met sur la piste d'un mystérieux monolithe…).

emploi comme indice d'événements et un emploi respectant davantage les conventions et codifications musicales). Il est parfois difficile de faire la distinction entre ce qui relève de sons provenant d'un monstre, du vaisseau spatial (craquement de la coque, etc.) et les nappes sonores de la musique, d'autant que ces différents aspects sont liés. Il s'agit de sons dynamiques, pour reprendre le terme de Karen Collins.

Dans le même temps, le son, qui témoigne d'une présence menaçante et envahissante, prend aussi pour le joueur une fonction informationnelle, d'avertissement. Par exemple, plus le nombre d'adversaires est important, plus la musique gagne en intensité. De même, le déplacement de la caméra entraîne le changement de la localisation des bruits émanant des « nécromorphes » sur les baffles du joueur. L'instance de perception sonore n'est pas l'avatar mais la caméra, ce qui démontre l'importance du procédé de visualisation des sons acousmatiques dans ce genre de jeux[29]. Si le monstre se trouve à gauche de la caméra, les bruits qu'il émet seront localisés sur l'enceinte correspondante. À cet effet, les développeurs du jeu préconisent l'utilisation d'un système à cinq enceintes[30], afin de pouvoir localiser plus facilement l'adversaire en fonction du son qu'il émet.

Tournant sa caméra vers la source sonore de certains bruits, ou partant à la recherche du monstre avec l'avatar, le joueur va pouvoir identifier, mesurer, comprendre le degré de dangerosité et agir en conséquence une fois la source sonore visualisée. Comme nous l'avons déjà souligné, à la différence des films d'horreur ou des écrits fantastiques, le jeu vidéo ne laisse jamais le joueur vivre de façon durable l'expérience de la peur, du vertige, de l'angoisse, qui laissent le corps sentant dans l'incapacité à dire, à catégoriser, à mesurer. En effet, s'il est en position de rejet ou de fuite, le joueur doit très vite restaurer le sens de l'action, en retournant à une profondeur de champ progressive, dans laquelle il va pouvoir identifier ce qui se passe à l'horizon et ne pas rester prisonnier d'une menace sans forme qui l'anéantirait dans un puissant sentiment de présence étrangère et une obsession destructrice. L'hyperprésence, notamment dans les jeux vidéo qui activent la passion de la peur, se caractérise de la sorte par des expériences très intenses, ponctuelles, itératives, toujours en tension entre sentiment d'étrangeté et de contrôle perceptifs.

---

[29] Etienne Armand Amato (2006) relève à ce titre, en faisant référence à Michel Chion, la nécessité de faire dans les jeux vidéo, tout comme au cinéma, la différence entre point de vue et point d'écoute, ce dernier pouvant être rattaché selon les jeux aussi bien à l'avatar, à la caméra, voire au pointeur de la souris.

[30] Soit trois enceintes à l'avant (à gauche, à droite et au centre) et deux enceintes à l'arrière (à gauche et à droite).

3. Les actants[31] de la présence

Selon que le corps sentant a prise sur le monde grâce à la fonction d'un son d'avertissement par exemple ou qu'il se sent sous l'emprise des choses à cause de textures sonores suspectes et indéfinissables, il n'a pas le même statut actantiel, à savoir la même force, le même rôle. Dans le premier cas, il se sent *source* de perception, dans le second cas, il se sent *cible* de quelque chose qui le vise, le traque. Pour comprendre précisément les fonctions du son dans les jeux vidéo et ses effets de sens pour le joueur, il est fort utile d'identifier les actants à l'œuvre dans le champ de présence du sujet sensible (Fontanille, 1998 : 99), à savoir :

- une *source*[32] dotée d'une intentionnalité minimale, qui oriente sa perception vers quelque chose présent dans son champ, avec plus ou moins d'intensité. Cette source de perception accomplit un acte de *visée* au sens où elle cherche à percevoir, elle vise quelque chose. Dans les jeux vidéo, la source du champ de présence doit viser plusieurs cibles, d'où l'intense nécessité de concentration : tandis que le joueur cherche à saisir ce qui se passe dans son champ, il doit aussi chercher à saisir ce qui se passe dans le champ de son avatar, voire de la caméra ;
- une *cible* qui est perçue, identifiée par le corps sentant. Jacques Fontanille explique que, quand le corps sentant reconnaît sa cible comme origine d'une intensité et dotée d'une intentionnalité, il est source de la visée mais cible de l'intensité : il est visé par une autre intentionnalité, source de perception. Dans les jeux vidéo qui activent la peur, quand le son acousmatique sert à faire ressentir au joueur une présence informe, étrangère, tout en invitant à en découvrir l'origine, il fait osciller le joueur entre cible d'un son non identifié (et par là-même inquiétant) et source d'une perception stratégique. Le geste du joueur, initié par le son, va alors permettre de saisir et comprendre l'autre

[31] Les *actants* sont des forces et des rôles nécessaires à l'accomplissement d'un procès. On distingue, selon J. Fontanille, les *actants narratifs* proposés par le schéma narratif structuraliste, tels le sujet, l'objet, le destinateur, le destinataire, l'adjuvant et l'opposant, des *actant positionnels* qui sont définis par leur place dans le champ positionnel du discours ou champ de perception. Pour en savoir plus, voir Fontanille (1998 : 139-140).

[32] Afin d'éviter toute ambiguïté, il faut préciser que nous ne faisons pas référence ici à la source fictionnelle du son mais bien à la source de perception (par exemple le joueur).

menaçant, pouvoir évaluer sa force, sa stratégie et ses intentions, pouvoir adapter ses propres intentions.

- un *contrôle*, qui gère la relation entre la source et la cible, affaiblissant ou renforçant les effets de présence sensible dans le champ du corps sentant. L'avatar du jeu peut servir de contrôle, permettant au joueur d'aller sentir et d'évoluer dans le monde du jeu, mais la caméra peut elle aussi jouer ce rôle. Si le joueur est maladroit, celle-ci se transforme en *obstacle*.

À l'aide de cette méthodologie, nous cernons ainsi la spécificité de la présence dans les jeux vidéo. Elle se définit comme une expérience d'hyperperception, d'hypermobilité, d'hyperprésence. Pour notre part, considérant la présence dans le jeu vidéo comme une expérience ressentie par le joueur, de continuité entre son monde physique et le monde du jeu, il nous faut désormais préciser, à l'aide de la théorie sémiotique de la présence exposée ci-dessus, comment la dimension sonore, en orchestration avec les autres modalités (visuelle, gestuelle, cinétique) contribue par ses diverses fonctions à accompagner les effets de présence chez le joueur et les sous-tendre.

4. La modalité sonore et ses effets de présence dans les jeux vidéo

Dans les jeux vidéo, les sensations et les impressions intimes vécues par le joueur sont intensifiées ou abaissées par l'orchestration du son avec le geste, en corrélation avec l'image visuelle. Dans le cas du son acousmatique, le joueur est tenté de saisir l'origine sonore, de l'évaluer, et d'agir en conséquence, reprenant ainsi le contrôle de son champ de perception. En revanche, dans le cas où c'est le geste du joueur qui compose les figures visuelles et sonores (comme c'est notamment le cas pour les sons désignés d'interactifs par Karen Collins), la multiplication des gestes fait advenir les figures visuo-sonores. Le joueur, bien que devant rester maître de sa composition, se sent happé par une harmonie parfaite entre son champ positionnel et les champs sonore, visuel. En tension entre une position de source de perception, de contrôle permettant l'ajustement entre les corps en interaction, cible de l'intensité visuo-sonore, le joueur est hyperprésent au jeu et le monde du jeu hyperprésent au joueur. Une interaction harmonieuse que le joueur vit profondément dans certains jeux, de façon à provoquer un état de transe, comme si ses gestes provoquaient

réciproquement la même intensité, la même itération dans les autres corps du son et de l'image. C'est le cas de *Rez*, où l'avatar se déplace automatiquement dans un long travelling avant, au sein d'une multitude de figures géométriques aux apparences de vaisseaux spatiaux. Dans ce jeu, le joueur n'a que peu de contrôle sur la caméra. Son rôle est essentiellement de déplacer un réticule de visée qui lui permettra, lorsqu'il appuie sur le bouton du *joystick*, de sélectionner un ou plusieurs ennemis (jusqu'à 8 simultanément), puis de faire partir, en relâchant le bouton, autant de rayons lasers qui provoqueront l'explosion de ses adversaires dans d'immenses tâches colorées. Chacune de ces actions provoque différents sons de percussions d'une hauteur aigüe, qu'il s'agisse par exemple de la sélection de l'adversaire ou de son explosion (ces sons ne cherchent en rien à être figuratifs, ils se rapprochent davantage de sons de cymbales ou de claquements de mains). Chaque action de visée et de tirs du joueur provoque de la sorte une rythmique qui trouve un équivalent visuel.

Ces séquences sonores discontinues s'orchestrent avec une musique d'ambiance continue. Cette dernière, s'apparentant très clairement à de la « techno », est principalement composée de sons basse fréquence au rythme marqué. Plus le joueur progresse dans la partie, plus le tempo de la musique s'accélère. Conjointement, les ennemis deviennent eux-aussi de plus en plus nombreux, la rythmique sonore créée par le joueur s'intensifiant de fait. On assiste alors à une multiplication des « points de synchronisation », pour reprendre un terme de Michel Chion (2000 : 52), soit « dans une chaîne audio-visuelle, un moment saillant de rencontre synchrone entre un moment sonore et un moment visuel », ce moment s'accompagnant également du geste dans le cas de *Rez*. Plus le joueur vise ses cibles avec justesse, plus le son s'intensifie et plus les composants de l'image se multiplient. Le succès de la visée permet de même la poursuite de la partie, puisque les ennemis qui ne sont pas détruits tirent sur l'avatar, qui finira par être éliminé si le joueur n'est pas assez habile. Le joueur est de la sorte happé dans un crescendo visuel, sonore et gestuel. Son geste entraîne des corps visuo-sonores qui entraînent un nouveau geste dans une sorte d'expérience jouissive, l'intensité et la discontinuité des séquences sonores prenant le joueur aux tripes et le poussant à faire durer cette danse des corps en interaction. Il faut à ce titre noter que ce sont des sons aigüs qui sont majoritairement générés par le joueur, or « un son richement doté en fréquences aigües créera une perception plus en alerte » (*Ibid.* : 17). En somme, plus

le joueur réussit, plus le son l'incite à être concentré sur l'action de jeu. On peut d'ailleurs noter également dans *Dead Space* que les cris des « nécromorphes » sont particulièrement aigües, ce type de fréquences permettant par ailleurs de localiser plus facilement l'origine d'un son, à l'inverse des sons basses fréquences qui envahissent tout l'espace.

Ces quelques orchestrations multimodales révèlent une fonction fondamentale du son dans les jeux vidéo ; former l'espace-temps de l'expérience ludique, cette « aire intermédiaire » dont nous parle Winnicott. En orchestration avec l'image et la plupart du temps le geste du joueur, le son forme un domaine spatio-temporel singulier que le joueur peut prendre en main, vivre dans un maintenant intense au sens propre du terme de main-tenant, c'est-à-dire tenant dans sa main. Cela, d'autant plus que l'installation sonore recommandée dans de nombreux jeux actuels est circulaire : haut-parleurs devant et derrière le joueur viennent former les limites physiques, matérielles, du champ de présence du joueur. Le son sert à créer une continuité entre espace fictif, espace physique et plus encore, entre champs de perception. Il intensifie le sentiment de présence du monde du jeu dans le monde physique du joueur et vice-versa. Modifiant le champ de perception du joueur, ou bien l'aidant à le contrôler, il permet de renouveler sans cesse la qualité des instants vécus et d'attribuer une dynamique aux micro-événements perceptifs qui adviennent dans son champ de présence.

5. Un cadre pour appréhender les fonctions des sons dans les jeux vidéo

Si les classifications actuelles des fonctions des sons dans les jeux vidéo permettent d'en clarifier la spécificité médiatique, elles ne révèlent cependant que peu de choses sur ses capacités à faire vivre au joueur une expérience de présence sensible du joueur au jeu. Nous proposons ci-dessous un cadre général dans lequel les différents statuts du son (acousmatique, interactionnel, dynamique, adaptatif…) dans les jeux pourraient être regroupés selon leur aptitude à main-tenir le joueur dans son « aire d'expérience ».

En orchestration avec les modalités visuelle et/ou gestuelle, le son peut :

- permettre au joueur de maîtriser son champ de perception, celui de l'avatar, voire celui de la caméra. En ce sens, le son sert de contrôle entre la source de perception (le joueur) et sa cible. C'est le cas des sons codifiés liés aux règles de fonctionnement du jeu, par exemple. C'est aussi le cas quand le son indique le résultat de l'action du joueur. Etablissant une distance mesurée entre le joueur et ce qu'il vise à l'horizon de son champ de perception, le son permet de « régler » les positions stratégiques du joueur. Il a alors une fonction de *présentation* au sens de mise devant soi, à distance, des règles des événements dans un espace-temps propice à la compréhension, à l'assimilation du fonctionnement du jeu et des interfaces matérielles ;
- permettre au joueur de s'imaginer le monde du jeu, de le figurer comme vraisemblable, comme intensément palpable. C'est le cas, entre autres, des sons qui illustrent les visuels, laissant le joueur percevoir les cris d'animaux, la pluie qui tombe, le chant d'oiseaux, le vent,… Le son a alors une fonction de *présentification* au sens qu'Herman Parret (2006 : 22-24) donne à ce terme ; donner à voir, préciser les formes des choses, leurs figures, laisser appréhender par l'imagination. C'est le cas, entre autres, des sons figuratifs ou iconiques[33], qui renvoient au main-tenant fictif. Le son invite alors le joueur à la fois à bien saisir ce qui advient dans son champ de perception et à se laisser immerger quelque peu dans la fiction qui, par sa vraisemblance, happe le joueur dans un « comme si on y était » ;
- permettre au joueur de vivre des sensations intenses, en amont de la perception des formes, de vivre des émotions intimes en sentant une présence l'envahir : une matière, un timbre, un rythme dont il n'identifie pas encore la source ou qui l'excède, venant envahir le champ de perception du joueur. C'est le cas de certains sons acousmatiques provoquant la peur ou de certains sons interactionnels qui fondent un corps à corps entre le corps du joueur et d'autres corps sonores ou visuo-sonores appartenant au jeu. La transe, la jouissance, relèvent alors la capacité du son à créer des effets de *présence sensible*.[34] Dans les jeux non figuratifs, plus abstraits, le son revêt

[33] Ces termes synonymes dans notre propos signifient renvoient aux capacités des sons à imiter des figures du monde naturel. Dans le parcours génératif du sens proposé par Greimas, le niveau figuratif du discours est celui où se construisent les champs lexicaux, les isotopies qui renvoient à un espace-temps et à des objets, personnages, éléments naturels aptes à produire un monde à part entière.

[34] Nous reprenons d'Herman Parret la distinction entre présentification et présence (2006 : 22-23).

en prédominance cette fonction. Ce ne sont pas les formes ou pas seulement les formes qui créent un effet d'immersion, mais c'est l'ajustement entre le corps du joueur et les corps visuo-sonores.

Bien entendu, présentation, présentification et présence sont trois grandes fonctions du son appelées à s'orchestrer au fil du jeu de manière variable, l'une prenant le pas sur les autres et vice versa. Dans *Rez*, le son a au départ une fonction évidente de présentation en ce qu'il permet au joueur de s'assurer de sa bonne maîtrise des règles du jeu, les sons successifs liés aux explosions de plusieurs ennemis lui indiquant la réussite de son action de façon bien plus précise que le visuel, qui est vite saturé de nouveaux adversaires ou de tâches colorées résultant de la destruction des adversaires. Sur ce point, le jeu vidéo s'appuie sur une propriété de la chaîne audio-visuelle déjà relevée par Michel Chion pour le cinéma :

> Pourquoi par exemple, les mouvements visuels rapides qu'accumulent les films de Kung-fu ou d'effets spéciaux ne créent-ils pas une impression confuse ? C'est parce qu'ils sont aidés et "pointés" par des ponctuations sonores rapides (sifflements, cris, chocs et tintements) qui marquent perceptivement certains moments et impriment dans la mémoire une trace audio-visuelle forte (Chion, 2000 : 14).

Par la suite, le rythme de l'orchestration multimodale (orchestration du geste avec le son et l'image) va doubler la fonction de présentation du son d'une fonction de présence, en allant crescendo. En effet, cependant que le corps sensible progresse vers une situation de maîtrise dans son champ de présence, il est sous l'emprise des événements visuo-sonores. Sous l'effet du rythme, le son, en s'orchestrant de façon harmonieuse avec l'ensemble des autres modalités visuelle et gestuelle, vient constituer et intensifier l'immersion.

Conclusion

En clôture de son article sur l'aspect participatif et « non-linéaire » du son dans le jeu vidéo, Karen Collins pose la question de son impact culturel sur les modes actuels de réception de la musique : « Est-ce que la réception du son non-linéaire des jeux

correspond à d'autres types de son non-linéaires contemporains, ou est-ce que l'expérience de jeu change ou influence la réception de la non-linéarité ? » (Collins, 2007 : 293 - 294)[35]. À ce titre, Karen Collins avance qu'une réflexion sur le domaine permet d'ouvrir de nouvelles pistes aux théories du son dans les médiums audio-visuels en général :

> Est-ce que le son des jeux a permis aux boucles [sonores] de devenir de plus en plus acceptables dans la musique populaire, en changeant la façon dont elles sont perçues et écoutées ? […] Tout comme MTV a aidé l'influence de nouveaux styles dans la réalisation de films, nous devons à présent considérer les jeux vidéo comme une influence culturelle majeure, avec des implications à la fois dans les pratiques et les théories de la musique populaire et dans les théories culturelles en général[36] (*Ibid*. : 294).

Pour apporter nos propres éléments de réflexion quant à ces perspectives et pour ouvrir sur un cadre plus large d'application, nous pouvons avancer que notre analyse rejoint certaines conclusions émises par Hervé Zénouda sur les ruptures et les continuités qu'apporte le numérique dans le rapport qui se crée entre les images et les sons. Cet auteur remarque une logique historique de cette relation (à travers la peinture, le cinéma et le cinéma expérimental) qui peut être résumée comme une dynamique de convergence allant de la correspondance à la « fusion » dans le numérique. Cette fusion serait double pour les hypermédias. Elle s'exprimerait d'une part « au niveau des différentes structures langagières d'organisation de la substance audiovisuelle (méta-data, méta-langages, unités sémiotiques diverses…) qui s'imbriquent les unes les autres en différentes couches » (Zénouda, 2005). D'autre part, il y aurait une fusion entre ces « couches sensibles » et la logique du système informatique, la « couche computationnelle ».

> L'addition de ces deux logiques de fusion modifie en profondeur la notion même de média pour faire émerger un nouvel objet audiovisuel programmable, multimodal et

[35] Notre traduction de l'anglais : « Does the reception of non-linear games audio correspond with other contemporary non-linear audio, or does the experience of gaming change or influence the reception of non-linearity? ».
[36] Notre traduction de l'anglais : « Has games audio helped looping to become more acceptable in popular music, by changing the way it is received and listened to? [...] Just as MTV helped to influence new styles in film editing, we need to now consider video games as a major cultural influence with implications on both practice and theory in popular music and cultural theory in general ».

pluri-média qui n'est plus vraiment au niveau des modalités une image, un son, ni même au niveau média une vidéo, mais un objet audiovisuel composite, structuré et complexe. Ce changement de statut a des répercussions à la fois techniques et esthétiques et concerne tout autant la production que la réception (*Ibid*., 2005).

Si l'auteur ne considère pas dans son article le cas des jeux vidéo, il nous semble justement que ce médium est aujourd'hui celui qui met au cœur même de ses modalités d'expression esthétique cette double logique de fusion. Le « *gameplay* » engage le joueur à éprouver des émotions à travers la découverte sensible d'un système, dans un processus de co-énonciation. En ceci, le jeu vidéo peut constituer pour la recherche un « paradigme » pour les arts et médias du XXIème siècle (Natkin, 2004). Dans cette perspective, le son ne peut y être considéré de façon indépendante, il faut prendre en compte la façon dont il s'orchestre avec l'image, le geste et les règles du jeu, le cadre théorique que nous avons formulé permettant de répondre à cet impératif.

Références bibliographiques

Amato E.A. (2003), « Coopération dans un jeu vidéo de rôle persistant », Caen, Communication lors du colloque IHM 03.

Amato E.A. (2006), « Reformulation du corps humain par le jeu vidéo : la posture vidéoludique », dans Genvo S. (Dir.), *Le game design de jeux vidéo, approches de l'expression vidéoludique*, L'Harmattan, Paris : 300-323.

Arsenault D. (2008), « Paysages 8-bit. Musicalité et spatialité dans le jeu vidéo des années 1985-1990 », *Inter, art actuel*, 98.

Blanchet A. (2008), « Cinéma et jeux vidéo : trente ans de liaisons », *Médiamorphoses*, 22 : 33 - 38.

Cadoz C. (1994), *Les réalités virtuelles*, Flammarion, Paris.

Caillois R. (1958, 1967), *Les jeux et les hommes*, Gallimard, Paris.

Chion (2000), *L'audio-vision. Son et image au cinéma*, Armand Colin, Paris.

Collins K. (2008), *Game sound*, MIT Press, Cambridge, MA.

Collins K. (2007), « An Introduction to the Participatory and Non-Linear Aspects of Video Game Audio », dans Hawkins S., Richardson J. (Dir.), *Essays on Sound and Vision*, Helsinki University Press, Helsinki, 263 - 298, disponible en ligne : http://www.gamessound.com/interactive.pdf.

Duflo C. (1997), *Le jeu. De Pascal à Schiller*, Presses universitaires de France, Paris.

Fontanille J. (1998)*, Sémiotique du discours*, Presses Universitaires de Limoges, Limoges.

Fontanille J. (1999),*Sémiotique et littérature*, collection « Formes sémiotiques », PUF, Paris.

Fontanille J. (2004), *Soma et séma, figures du corps*, Maisonneuve et Larose, Paris.

Genette G. (1982), *Palimpsestes. La littérature au second degré*, Éd. du Seuil, Paris.

Genvo S. (2009), *Le jeu à son ère numérique. Comprendre et analyser les jeux vidéo*, L'Harmattan, Paris.

Greimas A.J., Courtès J. (1993), *Dictionnaire raisonné de la théorie du langage*, coll. « Hachette supérieur », Hachette, Paris

Jullier L. (2002), *L'analyse de séquences*, Nathan, Paris.

Juul J., 2005, *Half-Real,* MIT Press, Cambridge.

Mabillot V. (2001), « Les dimensions proxémiques recomposées de la communication interactive », *La communication médiatisée par ordinateur : un carrefour des problématiques*, Sherbrooke, disponible en ligne, <http://grm.uqam.ca/activites/cmo2001/mabillot.html>.

Marsh T., Wright P., Smith S. (2001), « Evaluation for the design experience: Modeling Breackdown of Interaction and Illusion », *Journal of cyberpsychology and behavior*, 4(2).

Natkin S. (2004), *Jeux vidéo et médias du XXIe siècle*, Vuibert Éd., Paris.

Odin R. (2000), *De la fiction*, De Boeck Université, Bruxelles.

Parret H. (2006), *Epiphanies de la présence,* Presses Universitaires de Limoges, Limoges.

Perron B., Wolf M.J.P. (dir.) (2003), *The video game theory reader*, Routledge, New York.

Perron B., Wolf M.J.P. (dir.) (2009), *The video game theory reader 2*, Routledge, New York.

Perron B. (2005), « Jeu vidéo et émotions », dans Genvo S. (Dir.), *Le game design de jeux vidéo. Approches de l'expression vidéoludiques*, L'Harmattan, Paris : 347 - 366.

Perron B., Arsenault D. (2009), « In the frame of the magic cycle », dans Perron B., Wolf M.J.P. (dir.), *The video game theory reader 2*, Routledge, New York : 109 - 132.

Pignier N., Drouillat B. (2004), *Penser le webdesign, modèles sémiotiques pour les projets multimédias*, Paris, L'Harmattan.

Raessens Joost, Jeffrey Goldstein (Dir.) (2005), *Handbook of computer game studies*, MIT Press, Cambridge, MA.

Rétaux X. (2003), « Présence dans l'environnement : théories et applications aux jeux vidéo », dans Roustan M. (Dir.), *La pratique du jeu vidéo, réalité ou virtualité,* L'Harmattan, Paris : 69-82.

Rheingold H. (1993), *La réalité virtuelle*, Dunod, Paris.

Salen K., Zimmerman E. (Dir.) (2006), *The game design Reader*, MIT Press, Cambridge, MA.

Schaeffer P. (1966), *Traité des objets musicaux*, Éd. du Seuil, Paris.

Stockburger A. (2003), « The game environment from an auditive perspective », *Digital Games Research Conference*, 4-6 Novembre, University of Utrecht, Pays-Bas, disponible en ligne : http://www.audiogames.net/pics/upload/gameenvironment.htm

Whalen Z. (2004), « Play Along: An Approach to Videogame Music », *Game Studies: The International Journal of Computer Game Research*, 4(1), disponible en ligne : http://www.gamestudies.org/0401/whalen

Winnicott D.W. (1971, 1975), *Jeu et réalité, l'espace potentiel*, Gallimard, Paris.

Zénouda H. (2005), « Images et sons dans les hypermédias. De la correspondance à la fusion », *H2PTM'05*, Hermes, Paris, disponible en ligne : http://zenouda.free.fr/texts/Hypermedia.pdf