

## Entre synchronisation et action : son et violence dans les jeux vidéo d'horreur

Bernard Perron et Guillaume Roux-Girard

Ce n'est pas d'hier que l'on reproche au jeu vidéo de faire la promotion de la violence. On se souviendra qu'au début des années 90, le Congrès américain avait dénoncé l'utilisation de la violence comme outil de commercialisation des jeux vidéo, s'inquiétant plus particulièrement de son impact sur les enfants. Pour le sénateur Joseph Lieberman, si le film-jeu *Night Trap* (Digital Pictures, 1992) mettait en scène des agressions trop réalistes contre des jeunes filles légèrement vêtues, le jeu de combat *Mortal Kombat* (Midway, 1993), avec ses personnages numérisés, était quant à lui d'une violence extrême. Pire encore, le jeu récompensait l'usage de la violence en permettant au joueur de tuer ou, même, de décapiter ses adversaires (voir Kent 2001 : 467-468). De telles idées préconçues perdurent encore aujourd'hui dans le discours d'une caste conservatrice de non-joueurs. Et force est d'admettre qu'alimenté par des images tridimensionnelles en temps réel de plus en plus photoréalistes, un tel jugement leur paraît justifié. Pourtant, il est tout aussi possible d'envisager autrement cette violence vidéoludique. Comme le formule l'éditorial d'un dossier du magazine français *PSM3*, « le jeu vidéo tire son essence de la violence sous-jacente qui régit notre société. Les meurtres, les exactions, la guerre, la violence ordinaire, le jeu vidéo est à la fois : un exutoire et un reflet décalé des travers de notre monde. Alors pourquoi s'en offusquer ? » (Belic et Tavish 2010 : 42). Comme au cinéma et à la télévision, les images violentes sont intégrées à l'esthétique ou au style de production.

Alors qu'on se tourne vers ces images pour pourfendre ces médias, on oublie trop souvent ce qui donne un caractère insoutenable aux diverses représentations numériques photoréalistes

de la violence, à savoir le son<sup>33</sup>. Nous ne sommes plus au temps du cinéma muet. Tel que l'a si justement montré Michel Chion, le son a davantage une fonction de valeur ajoutée sur l'image qu'un effet direct sur le spectateur ou sur le joueur. Pour Chion (1990 : 8-9), la valeur ajoutée désigne

la valeur expressive et informative dont un son enrichit une image donnée, jusqu'à donner à croire, dans l'impression immédiate qu'on en a ou le souvenir qu'on en garde, que cette information ou cette expression se dégage « naturellement » de ce qu'on voit et est déjà contenue dans l'image seule. Et jusqu'à procurer l'impression, éminemment injuste, que le son est inutile, et qu'il redouble un sens qu'en réalité il amène et crée, soit de toutes pièces, soit par sa différence même d'avec ce qu'on voit.

Même en admettant que leur réception a dû être différente à l'époque à cause de leur caractère singulier, les seules images de *Mortal Kombat* ne sont pas intenable au point de rendre les combats difficiles à regarder. Sans le choc des coups et les plaintes des rivaux qui ajoutent au sang qui gicle, la réaction du sénateur Lieberman n'aurait probablement pas été la même. Une telle observation s'applique également à un moment d'anthologie dans l'histoire récente du jeu vidéo d'horreur. En effet, c'est une chose de voir Leon, le personnage joueur de *Resident Evil 4* (Capcom, 2005), se faire trancher la tête par un maniaque armé d'une tronçonneuse, et une autre d'entendre le rugissement horrifiant de ladite scie, le dernier souffle de Leon et le giclement du sang lorsque sa tête se détache de son corps.

Parce que, d'une part, l'horreur demeure un genre très critiqué pour la gratuité de sa violence souvent très graphique et parce que, d'autre part, elle reste foncièrement sonore – les effets sonores étant à l'origine de la terreur, de la peur ou du dégoût que souhaite ressentir le joueur<sup>34</sup> –, ce texte souhaite analyser l'impact du son sur la mise en

---

<sup>33</sup> Pour ne citer qu'un seul exemple horrifiant : que l'image soit de très mauvaise qualité ou non, la décapitation du jeune homme d'affaires américain Nicholas Evan Berg par des terroristes islamistes en Iraq en 2004 ne serait pas aussi traumatisante sans les cris et les sons qu'on y entend.

<sup>34</sup> Avouons aussi que, par opposition à une relation au réel plus « transparente », l'horreur introduit une certaine distanciation face à la violence par l'entremise de son

scène de la violence dans les jeux vidéo d'horreur. Et puisque l'une des manières de définir la violence repose sur la nature des comportements ou des agressions physiques, il s'agit aussi d'étudier comment l'interactivité au cœur de l'expérience transforme la relation du joueur avec le son, relation qui s'appuie ainsi tout autant, sinon plus, sur la nature actionnelle du jeu que sur la dimension spectatorielle de l'œuvre audiovisuelle. Violence et action forment de toute évidence les éléments complémentaires d'un même ensemble au sein de notre analyse.

La panique morale au sujet de la violence vidéoludique est fort simple à expliquer. Alors que le jeu vidéo peut proposer des récits audiovisuels dérangeants, il ne se limite pas seulement à les représenter. Dès lors où celui-ci simule un monde cauchemardesque dans lequel il faut agir, la jouabilité nécessite que le joueur participe au déroulement narratif. Suivant la définition de Frasca (2003 : 223) : « Simulation does not simply retain the – generally audiovisual – characteristics of the object but it also includes a model of its behaviours. This model reacts to certain stimuli (input data, pushing buttons, joystick movements), according to a set of conditions ». En d'autres termes, le fait d'agir reste au centre de l'expérience vidéoludique. Pour arriver à la fin de l'histoire, le joueur doit continuellement prendre des décisions, faire des choix, remplir des objectifs, atteindre des buts, surmonter des obstacles, vaincre des ennemis, etc. Dans cette optique, il doit tirer parti du répertoire de commandes que lui offre le jeu, principalement basé sur l'attaque et la défense dans les jeux vidéo d'horreur. Les diverses entrées (*inputs*) extradiégétiques – ce que le joueur est physiquement en train d'effectuer à l'aide de la manette de jeu – sont définies par Andreas Gregersen et Torben Grodal comme des actions primitives (« actions-P »). Suivant les auteurs, il faut faire une distinction

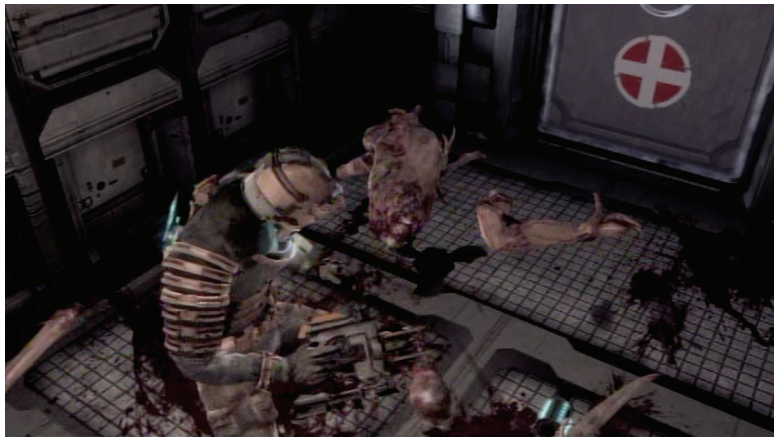
between primitive actions meaning actual body movements and on the other hand actions in the wider sense: moving the index finger (to pull a trigger) is a primitive action, whereas

---

approche plus fantastique. Analyser *Resident Evil* (Capcom, 1996) ou *Silent Hill* (Konami, 1999) ne soulève notamment pas les mêmes questions qu'une étude des controversés *Manhunt 1* et *2* (développés par Roskstar en 2003 et 2007) dans lesquels les personnages joueurs sont amenés à tuer de façon gratuite et sanguinaire des êtres humains (criminels ou non).

discharging a firearm is an action. [...] [A] primitive action (P-action) is thus defined as merely a movement of the body. A given P-action may be part of many different actions, and an action may be constituted by different P-actions (Gregersen et Grodal 2009 : 70).

On comprendra que c'est la traduction des actions-P dans l'univers diégétique qui pose problème. Qu'il s'agisse de contourner un obstacle ou de trucider un monstre, c'est bien le joueur qui est responsable de cette conduite. Qui plus est, une seule action-P peut donner lieu à une action diégétique sanglante. Par exemple, une fois que les monstres sont assommés dans *Silent Hill: Homecoming* (Double Helix, 2008), le joueur peut leur assener un coup de grâce dévastateur en pressant simplement sur la touche ornée d'un symbole carré de sa manette de PlayStation 3. Dans *Dead Space* (EA Redwood Shores, 2008), c'est l'utilisation de la touche « R2 » qui permet d'écrabouiller d'un furieux coup de pied les anciens membres de l'équipage maintenant transformés en monstrueux morts-vivants (des nécromorphes) [figure 1].



**Figure 1 :**  
**Isaac, le personnage-joueur de *Dead Space*, écrase la tête d'un nécromorphe à l'aide de ses bottes spatiales.**

Dans *ObsCure: The Aftermath* (Hydravision, 2007) pour la console Nintendo Wii, l'action-P du joueur avec la télécommande Wii

(Wiimote) imite cette fois le mouvement du personnage qui frappe entre autres les monstres avec un bâton de baseball. Aussi violents et dangereux que puissent être les monstres rencontrés dans l'univers horrifique, ces derniers rappellent que, pour survivre au cauchemar, le personnage joueur doit lui aussi être impétueux.

L'impact des actions-P repose sur le rapport de causalité entre cette action et le résultat à l'image et au son. La question du synchronisme demeure au cœur de cette relation de cause à effet<sup>35</sup>. Comme l'explique Michel Chion (2003 : 430), un point de synchronisation est « [...] un moment plus saillant de rencontre synchrone entre un moment sonore et un moment visuel concomitant [...] ». Dans un contexte vidéoludique, la simultanéité d'un bouton pressé et de l'image de la flamme jaillissant d'une arme accompagnée du « bang » produit par le coup de feu génère un premier niveau de synchronisation. Celui-ci crée un lien entre le fait d'appuyer sur ledit bouton et le coup de feu tiré sur l'axe diégétique du jeu. L'action-P entraîne ainsi, dirait Chion, l'« expression privilégiée de l'instantanéité » ou, plus exactement dans le présent contexte, de l'interactivité. Car le lien de causalité établit entre les sons et les images permet d'installer un second niveau de synchronisation plus important qui a trait à un événement ou à une action, ici la réaction du monstre touché, blessé ou tué. Le joueur contrôle cette action. Il agit dans le domaine de l'immédiat. Surtout que le joueur n'est pas amené à s'apitoyer sur le sort des monstres à fuir ou tuer puisqu'ils sont à la fois nombreux et l'incarnation du Mal.

La distinction entre les deux niveaux de synchronisme est le signe d'une importante différence entre la valeur d'un point de synchronisation dans un jeu vidéo et celle présente dans un film. Au cinéma, tout en participant aux enjeux narratifs, ces points de synchronisation ont une grande importance puisqu'ils font partie d'un événement « esthétisé ». Par exemple, dans *The Descent* (Neil Marshall, 2005), le piolet que plante accidentellement l'une des spéléologues dans la gorge de sa camarade en croyant se défendre de l'attaque d'un *cranler* possède une force émotive du fait que le spectateur est entièrement soumis à la soudaineté de l'action, à la vie autonome et à l'indépendance du geste posé. Cela est d'autant plus vrai qu'il est hors

---

<sup>35</sup> Notons qu'une large part des réflexions théoriques reliées aux sons ainsi qu'à ses causes et effets est tirée du mémoire de maîtrise de Guillaume Roux-Girard (2010).

du contrôle du spectateur ; le choc du coup entre la gorge et le piolet est en même temps un choc perceptif. Dans un jeu vidéo d'horreur, les points de synchronisation peuvent aussi avoir cette valeur lorsqu'ils sont utilisés en concordance avec un événement qui n'est pas relié aux actions-P du joueur. Dès qu'elle est enclenchée, la courte cinématique qui montre la décapitation de Leon dans *Resident Evil 4* ou celle d'Isaac dans *Dead Space* par les excroissances tranchantes du *hunter* participe activement à la représentation de la violence faite au personnage joueur. De même, un monstre qui frappe par surprise le joueur alors qu'il explore un lieu sombre entraîne ce type d'effet. Cependant, durant les séquences de jouabilité, les points de synchronisation qui coïncident avec les actions-P du joueur prennent moins une valeur esthétique qu'une valeur pragmatique dans la mesure où ils sont les produits d'une volonté en acte. Ces points de synchronisation contribuent grandement au sentiment insufflé par le potentiel d'action (*agency*) et participent à l'effet de présence du joueur dans l'univers du jeu.

Autrement dit, ces points saillants de synchronisation deviennent la preuve concrète de l'influence du joueur à l'intérieur du monde cauchemardesque dans lequel il est plongé. C'est pour cette raison qu'il est moins pertinent d'examiner la source sonore comme nous le ferions au cinéma, que de porter une attention particulière à ce que Kristine Jørgenson (2008 : s.p.) a appelé des générateurs sonores : « Notice that the *generator* of a sound is not the same as the *source* of a sound. While the source is the object that physically (or virtually) produces the sound ; the generator is what causes the event that produces the sound ». Le générateur relève donc bien de l'événement et non de l'émission du son. Si un cri de douleur est de toute évidence émis par celui qui reçoit un coup, le générateur de ce son demeure la personne qui a asséné ledit coup. Pour mieux étudier les sons vidéoludiques, Jørgensen les regroupe selon cinq grandes catégories. Nous allons nous concentrer sur deux d'entre elles, à savoir les sons générés par l'action du joueur ainsi que ceux produits par l'action des ennemis<sup>36</sup>. Conséquemment, dans *Condemned: Criminal Origins* (Monolith Productions, 2005), les nombreux combats corps à corps – avec par exemple un « 2x4 », une barre de renfort, une masse ou une

---

<sup>36</sup> Les autres étant les sons produits par les alliés, les sons engendrés par l'environnement de jeu et finalement, ceux propagés par le système de jeu.

hache à incendie – alternent des sons générés par le joueur incarnant l'agent Ethan Thomas (vu à la première personne) et par les nombreux détraqués qui rôdent dans les lieux d'investigation [figure 2]. Le son produit par l'impact du coup d'une barre de renfort sur la tête de l'un de ces frénétiques est généré par l'action-P du joueur sur la gâchette droite (RT) de la manette de la *Xbox 360*, action transposée par un coup sec du personnage joueur suivi par les plaintes de l'ennemi frappé. Inversement, associées aux éclats rouges à l'écran, les réactions douloureuses de l'agent Thomas sont causées par les assauts qu'il subit. Ici, c'est l'intelligence artificielle qui est porteuse de l'action. Victime de violence, le personnage joueur doit alors devenir lui aussi sanguinaire pour survivre. Autant cette nécessité cautionne les actes du joueur, autant elle entraîne une expression gratuite de comportements considérés par plusieurs détracteurs des jeux vidéo comme malsains.



**Figure 2:**  
Ethan Thomas, le personnage-joueur de *Condemned : Criminal Origins*, se défend avec une masse de l'attaque d'un détraqué.

Aussi efficace que soit la mise en scène de la violence dans *Condemned : Criminal Origins*, les jeux vidéo d'horreur – associés ou non au *survival horror*, genre vidéoludique qui lui est spécifiquement lié – se jouent essentiellement à la troisième personne parce que l'on voit son personnage à l'écran (voir Perron 2009). Le joueur peut certes être effrayé ou envahi de panique dans un jeu à la première personne, mais

il est certainement davantage paralysé par l'horreur lorsqu'il voit le corps de son personnage violenté, mutilé et battu à mort par des monstres. De nouveau, la dimension sonore tient un rôle fondamental dans l'équation. Tout comme le son a besoin de l'image, cette dernière a besoin de la valeur ajoutée du son pour être efficace. Chion (1990 : 23) l'a bien perçu lorsqu'il écrit que le « son transformé par l'image qu'il influence reprojette finalement sur celle-ci le produit de leurs influences mutuelles ».

Au-delà de la synchronisation, les sons doivent avoir un bon rendu pour produire le résultat escompté. Encore une fois, suivant Chion (2003 : 431), il est nécessaire que « le son [soit] reconnu par [le joueur] comme vrai, efficace et adapté, non s'il reproduit le son que fait dans le réel le même type de situation ou de cause, mais s'il rend les sensations, pas spécifiquement sonores [...] associées à cette cause ou à la circonstance évoquée dans la scène [...] ». Le traitement appliqué à un son en temps réel peut lui aussi contribuer à la qualité de ce rendu. Dans *Dead Space* ou même dans le jeu de tir sur rail *Dead Space : Extraction* (Visceral Games, 2009), où le joueur peut utiliser le *Wii Zapper* simulant une arme à feu, le son gluant qui accompagne l'impact d'un tir de pistolet au plasma faisant exploser les entrailles d'un nécromorphe en une masse difforme et cramoisie aide le joueur à percevoir cette éviscération comme réelle alors qu'il assiste plutôt à une simple translation de polygones à l'écran. La relation causale, que ces effets contribuent à établir entre les sons, les images, les événements du jeu et les actions du joueur, remplit un rôle important dans l'impact de cette décharge. L'effet d'un son tient aussi à la nature du phénomène. Comparativement à l'image,

la valeur figurative, narrative, d'un bruit est très faible, très floue. Un même son peut donc, selon le contexte dramatique et visuel, venir raconter des choses très diverses, puisque pour le spectateur de cinéma, plus que le réalisme acoustique c'est le critère de *synchronisme* avant tout, et secondairement de vraisemblance globale (une vraisemblance qui est affaire non de vérité, mais de convention), qui l'amène à coller un son sur un événement ou un phénomène. Le *même son* pourra donc venir bruiser de façon convaincante dans une comédie la pastèque qui s'écrase, et dans un film de guerre le crâne réduit



en bouillie. Ce même bruit sera ici jubilatoire, et là insoutenable (Chion 1990 : 23).

C'est ce bruit combiné avec un son rappelant le sifflement d'un serpent, qui est utilisé dans *Resident Evil 5* (Capcom, 2009) lorsque la tête d'un *cephalo* explose pour laisser place au tentacule épineux du parasite. Comme la plupart des jeux vidéo d'horreur, *Resident Evil 5* pullule de ce genre d'événements. C'est alors le macro-contexte générique de l'horreur qui prévaut. Nul besoin de se poser trop de questions devant une telle micro-situation. À chaque fois, le joueur doit éliminer les monstres ou, si possible, les fuir pour continuer sa quête.<sup>37</sup> Le traitement des éléments sonores entendus repose sur les schémas cognitifs du joueur qui orientent sa prise de décision tout en servant de perspective d'écoute. Pour Jørgensen (2008 : s.p.), « *Context is a keyword in this respect, and [...] a specific sound cannot be comprehended in isolation, but that the situation in which it is heard always decides the interpretation of the informative content of the sound signal* ». Cette idée, qui vient de toute évidence compléter les propos de Chion au sujet de la valeur ajoutée, permet de mieux saisir comment s'opère cette relation entre l'image, le son et la violence. Car au bout du compte, c'est la mise en scène globale de l'effet audiovisuel qui donne naissance à la représentation des actes violents perpétrés par le joueur ou ceux dont son personnage est victime. Le contexte général et situationnel unifie les sons aux images et leur donne leur signification.

Dans les jeux vidéo d'horreur, et dans tous les autres jeux associés à la violence (jeux de combats ou jeux de guerre), le contexte de jouabilité contraint le joueur à user de ses compétences sensori-motrices et cognitives afin de survivre et de triompher. Comme nous avons voulu le montrer, le son prédispose de manière considérable le joueur à créer et à répondre à la violence simulée par le jeu. Car, pour se référer une dernière fois aux lumineuses affirmations de Michel Chion (1998 : 56),

la nature d'un bruit en soi, au cinéma [et dans le jeu vidéo], est presque indifférente. Il y a bien sûr quelques tabous,

---

<sup>37</sup> Il y a en fait un schéma d'événement principal dans les jeux vidéo d'horreur : « faire face au monstre ». Ce schéma se traduit dans la diégèse par trois actions du joueur : se figer sur place, combattre ou fuir (voir Perron 2006).

notamment sur le plan des convenances (bruits organiques, dont s'est amusé quelquefois le cinéma sonore débutant), mais l'important est que le son n'a pas de sens et de valeur fixe ; il est essentiellement ambigu, et, plutôt qu'un lieu de fixation (que son essence temporelle lui interdit d'être), il est le lieu d'un *glissement*.

En définitive, et peut-être plus pour le non-joueur témoin d'une volonté en acte dans un univers fictionnel vidéoludique, le son entraîne bel et bien un *glissement* : *glissement* de l'image polygonale vers une image « réelle », *glissement* des actions-P vers des actions diégétiques et *glissement* d'un événement stylisé vers un événement qui ferait oublier la présence du média par celui qui regarde et écoute<sup>38</sup>. Bref, le son introduit du « jeu » dans l'expérience de la violence.

## RÉFÉRENCES BIBLIOGRAPHIQUES

- BELIC, Vinz et Hub Tavish (2010). « Violence et jeu vidéo : génération hardcore », *PSM3*, no 32, janvier, p. 42.
- BOLTER, Jay David et Richard Grusin (1999). *Remediation : Understanding New Media*, Cambridge, MIT Press.
- CHION, Michel (2003). *Un art sonore, le cinéma. Histoire, esthétique, poétique*, Paris, Cahiers du cinéma.
- CHION, Michel (1998). « Trois sons », *Vertigo, L'infilmable*, No 3, p. 55-56.
- CHION, Michel (1990). *L'Audio-vision*, Paris, Nathan.
- FRASCA, Gonzalo (2003). « Simulation versus Narrative. Introduction to Ludology », dans Mark J.P. Wolf et Bernard Perron (dir.), *The Video Game Theory Reader*, New York, Routledge, p. 221-235.
- GREGERSEN, Andreas et Torben Grodal (2009). « Embodiment and Interface », dans Bernard Perron et Mark J.P. Wolf (dir.), *The Video Game Theory Reader 2*, New York, Routledge, p. 65-83.

---

<sup>38</sup> Il s'agit là de la fameuse notion d'« immediacy » de Jay David Bolter, Richard Grusin (voir 1999).

- JØRGENSEN, Kristine (2008). « Audio and Gameplay : An Analysis of PvP Battlegrounds in *World of Warcraft* », *Gamestudies*, Vol. 8, No 2, décembre, disponible en ligne : <http://gamestudies.org/0802/articles/jorgensen>.
- KENT, Steven L. (2001). *The Ultimate History of Video Games*, New York, Three Rivers Press.
- PERRON, Bernard (2009). « The Survival Horror: The Extended Body Genre », dans Bernard Perron (dir.), *Horror Video Games : Essays on the Fusion of Fear and Play*, Jefferson, McFarland, p. 121-143.
- PERRON, Bernard (2006). « The Heuristic Circle of Gameplay : the Case of Survival Horror », dans Manthos Santorinereos (dir.), *Medi@terra 2006 Conference Proceedings*, Athènes, Fournos, p. 62-71.
- ROUX-GIRARD, Guillaume (2010). *L'écoute de la peur : une étude du son dans les jeux vidéo d'horreur*, non publié, Université de Montréal.