

Synthèse des principaux travaux de Kristine Jørgensen

- 2009. *A Comprehensive Study of Sound in Computer Games: How Audio Affects Player Action*. Lewiston : Mellen Press.
- 2008a. « Audio and Gameplay: An Analysis of PvP Battlegrounds in *World of Warcraft* ». *Gamestudies the international journal of computer game research*, vol. 8, n° 2 (décembre).
- 2008b. « Left in the Dark: Playing Computer Games with the Sound Turned Off ». Dans Collins, Karen (dir.). *From Pac-Mac to Pop Music: Interactive Audio in Games and New Media*. Burlington : Ashgate, p. 163-176.
- 2007. « On Transdiegetic Sounds in Computer Games ». Dans Bondebjerg, Ib (dir.). *Northern Lights Vol. 5: Digital Aesthetics and Communication*. Wilmington : Intellect Books, (2007), p. 105-118.
- 2006. « On the Functional Aspects of Computer Game Audio ». Dans les actes de colloque d'*Audio Mostly – A Conference on Sound in Games*, Piteå, 11-12 octobre 2006. Piteå : Interactive Institute, of the Sonic Studio, p. 48-52.

Les travaux ¹ de Kristine Jørgensen « cherchent à révéler la fonctionnalité du son dans les jeux vidéo, en prêtant une attention particulière à la relation qui se tisse entre les sons et les événements qui se produisent dans le jeu » (2009 : 3, ma traduction). L'objectif principal de Jørgensen est de répondre aux questions suivantes : en quoi et comment le son dans les jeux vidéo a-t-il une influence sur l'activité du joueur ? et en quoi l'information auditive fournie par le jeu permet-elle au joueur d'agir dans le monde virtuel ? Ses recherches visent aussi à déterminer comment le son contribue à « définir différents espaces d'action » (*ibid*). En d'autres mots, l'auteure cherche à montrer que le son épouse la double nature des œuvres vidéoludiques, celles-ci étant à la fois des systèmes informatiques réglés – manipulés à l'aide d'une interface – et des mondes fictionnels cohérents où s'entremêlent la représentation et la simulation (2006 ; 2008a ; 2008b ; 2009).

1. La majorité des articles signés par la chercheuse norvégienne ont paru alors qu'elle poursuivait des études doctorales. C'est le contenu de ces articles que l'on trouve, retravaillé et reformulé, dans la version publiée de sa thèse : *A Comprehensive Study of Sound in Computer Games : How Audio Affects Player Action*.

Les deux catégories de fonctions du son dans les jeux vidéo

Au moyen d'une analyse de données provenant d'une étude empirique sur les joueurs, puis à l'aide d'outils théoriques empruntés aux études sur le jeu (vidéo), aux études cinématographiques² et à l'écologie psychoacoustique³, Jørgensen distingue deux catégories de fonctions fondamentales, fonctions en rapport avec la double nature du jeu vidéo : le son a pour fonction d'assurer la convivialité⁴ (ou, plus généralement, la jouabilité) du système avec lequel interagit le joueur ; d'autre part, le son doit favoriser un sentiment de présence⁵ dans le monde du jeu. Ainsi, les sons fournissent de l'information pertinente au joueur tout en le faisant se sentir présent dans un environnement réaliste. Il est « important de noter que, dans un contexte de jeu, le monde virtuel et les fonctions liées à la convivialité ne sont pas distincts ; en effet, les sons qui semblent diégétiquement motivés ont aussi des fonctions de convivialité, et les sons qui semblent motivés par la convivialité sont ancrés dans l'environnement virtuel » (2009 : 82, ma traduction).

Cette double fonctionnalité représente d'ailleurs une spécificité du son vidéoludique qui s'observait déjà dans les premiers jeux vidéo commercialisés. Soit cette périodisation du développement vidéoludique que propose Jørgensen : les débuts (1930-1971) ; la commercialisation (1971-1984) ; l'établissement de la fonctionnalité (1984-1992) ; vers les temps modernes. La chercheuse y retrace l'histoire du son vidéoludique :

2. Pour Jørgensen, les théories sonores cinématographiques sont particulièrement utiles pour comprendre comment le son interagit avec l'environnement de jeu et comment il en vient à déterminer les différents espaces d'action qui composent cet environnement.

3. Partie de la psychologie cognitive, l'écologie acoustique s'intéresse à la perception auditive ; les travaux de recherche appliquée – sur les dispositifs de représentation sonore d'information et la « sonification » – s'intéressent plus particulièrement à la transmission d'information par le son dans telle ou telle situation bien réelle ou dans un environnement informatique. Selon Jørgensen, cette discipline est particulièrement utile pour comprendre l'apport du son à la convivialité du système.

4. Le terme doit être pris au sens qu'il a en informatique : la convivialité, c'est la facilité d'emploi d'un logiciel ou d'un matériel.

5. Jørgensen emprunte ce terme à Alison McMahan (2003) dans son texte « Immersion, Engagement, and Presence : A Method for Analysing 3-D Video Games ». Du coup, Jørgensen associe l'efficacité de cet effet de présence au réalisme visuel et auditif de l'environnement de jeu, à sa « ressemblance » avec le « vrai monde » (2008b).

La première phase, celle de l'origine du son vidéoludique, est représentée par les premières machines de jeux électroniques⁶ et par des ébauches de jeux vidéo dans lesquels le son était rarement intégré en raison de limitations matérielles et d'une faible capacité de stockage. La seconde phase correspond à la période où les jeux vidéo sont devenus des produits commerciaux : des sons leur furent ajoutés pour attirer l'attention des passants vers ces jeux d'arcade installés dans les lieux publics. Dans une troisième phase, le son vidéoludique trouva son mode fonctionnel, conséquence d'une conception vidéoludique minutieuse et du progrès technologique. Cette période ouvrait la voie à un son qui à la fois fonctionnel et apte à créer un sentiment de présence dans l'environnement de jeu. La dernière phase concerne les plus récents progrès technologiques et fonctionnels, qui permettent une meilleure mise en œuvre du son vidéoludique (2009 : 14, ma traduction).

C'est donc à partir de *Pong* (Atari, 1972), mais plus particulièrement après la mise en marché de la console Nintendo Entertainment System (1984) que les concepteurs de jeux vidéo utilisèrent massivement le son à des fins de convivialité tout en cherchant à intégrer les fonctions de convivialité de manière naturelle dans le monde du jeu.

Système de signes, convivialité et monde virtuel

Dans les jeux vidéo, le son « fournit de l'information précise à propos d'un cadre ou d'une situation, et peut avoir une influence directe sur les actions et les événements qui se déroulent dans l'environnement [de jeu] » (2008b : 163, ma traduction). Le son vidéoludique se constitue ainsi comme un système de signes essentiel à la progression du joueur⁷ ; si bien que « l'usage du son à des fins de convivialité est appelé à être hautement informatif et communicatif,

6. Jørgensen fait ici allusion aux *pinball machines* qui dès les années trente faisaient entendre des sons artificiels, offrant ainsi une rétroaction autre que les simples bruits mécaniques normalement émis par l'appareil.

7. Certaines recherches menées par Jørgensen avaient d'ailleurs pour objectif de déterminer l'impact de l'absence du son sur l'activité des joueurs (voir 2008b et 2009). Selon les résultats d'une étude empirique menée auprès de quelques joueurs de *Hitman : Contracts* (Io Interactive, 2004) et de *Warcraft III* (Blizzard Entertainment, 2002), l'absence du son aurait pour effet de faire « décroître le sentiment de contrôle » (2008b : 167) du joueur en « faisant disparaître la sensation de spatialité » (*ibid* : 168) et en le privant d'information (portée de ses actions, portrait de la situation, identité de ses ennemis, etc.) relativement à tel ou tel événement. De plus, le retrait du son aurait pour effet de faire décroître radicalement le sentiment qu'a le joueur d'être présent dans le monde du jeu (*ibid* : 171).

et de fait les sons [agissent de manière] représentationnelle puisqu'ils renvoient à des situations particulières se produisant dans l'environnement [de jeu] » (2009 : 83, ma traduction). Entre le son et son référent intervient un processus associatif qui n'est pas étranger à la sémiotique peircienne⁸. Pour construire le sens, le joueur doit d'abord identifier le type de signal auquel il a affaire, puis il doit déterminer le type de relation qui se tisse entre le signe et son référent.

Basant sa réflexion sur des concepts empruntés aux recherches sur les dispositifs de représentation sonore d'information⁹, Jørgensen distingue deux types de signaux sonore : les icônes auditives¹⁰, « des sons caractéristiques pouvant être reconnus comme des sons correspondant à des événements du monde réel » (2009 : 84, ma traduction), et les carillons¹¹, « des bruits ou de la musique qui entretiennent une relation arbitraire ou symbolique avec leur référent » (*ibid.*, ma traduction). Utilisée de façon purement causale, la voix figure parmi les icônes auditives ; elle forme une catégorie à part si elle a une valeur sémantique (*ibid.*). Quant à la relation entre le signe et le référent, elle peut occuper l'une ou l'autre des trois positions remarquables que présente un continuum d'iconicité : la relation est dite iconique si « le signal correspond exactement au son qui serait produit par son équivalent naturel » ; elle est dite non arbitraire s'il existe « entre le signal et le référent une similarité physique, une proximité ou un contexte » ; elle est dite arbitraire si entre le signal et la référent la relation est construite » (2009 : 85, ma traduction).

Selon Jørgensen, les jeux vidéo utilisent principalement les carillons et les icônes auditives, signaux qui joueront différents rôles selon le genre

8. Jørgensen s'inspire de la taxonomie proposée par Peter Keller et Catherine Stevens, elle-même dérivée de la sémiotique peircienne : séparant en deux niveaux l'identification du signe et le processus associatif par lequel l'auditeur comprend à quel objet ou événement le son renvoie, cette taxonomie est tributaire de la théorie triadique qui met en relation le signe, l'objet et l'interprétant.

9. Traduction de *auditory displays*.

10. Traduction de *auditory icon* ; il s'agit d'un son enregistré à partir d'une source présente dans notre réalité donnée.

11. Traduction de *earcon*, terme inventé en 1985 par D. A. Sumikawa, dans son article « Guidelines for the Integration of Audio Cues into Computer User Interfaces », pour désigner l'équivalent auditif d'une icône visuelle. Ce jeu de mots repose sur la sonorité de *icon*, à laquelle peut correspondre la graphie *eye-con*, mot lui-même transformé en *ear-con*. Le terme a ensuite été associé plus étroitement aux sons synthétisés d'un environnement informatique. « Carillon » est le nom français donné à ce type d'occurrence sonore, le jeu de mots n'ayant par survécu à la traduction.

vidéoludique. Consistant essentiellement en sons abstraits, artificiels ou musicaux, les carillons « sont typiquement utilisés en connexion avec la musique extradiégétique ou des sons liés à l'interface graphique de l'utilisateur » (2009 : 87, ma traduction). Si un processus d'apprentissage est nécessaire pour construire la relation entre le signe et le référent, une fois cette relation établie et stable, le carillon peut glisser dans la catégorie des icônes auditives¹². De leur côté, puisqu'elles entretiennent un lien plus serré avec les sons naturels et qu'elles sont reconnues plus intuitivement par le joueur, les icônes auditives « sont habituellement utilisées comme des sons diégétiques » (*ibid*). C'est d'ailleurs la force des icônes auditives¹³, qui semblent se dégager naturellement du monde (sentiment de présence) du jeu tout en ayant une forte valeur informative (convivialité), alors que les carillons ne semblent pas émaner naturellement de l'environnement. De façon générale,

le concept d'icône auditive permet d'expliquer comment les sons qui semblent émaner naturellement de l'univers du jeu ont aussi une forte valeur informative, alors que le concept de carillon permet de comprendre pourquoi la musique de jeu vidéo et les sons artificiels peuvent avoir du sens sans perturber [l'expérience] du [joueur] (2006 : 49, ma traduction).

Cela dit, il arrive qu'une icône auditive ait « avec un certain objet une relation qu'intuitivement on dirait naturelle, mais, s'ils sont écoutés plus attentivement, on découvre que ces sons ne sont pas adaptés à la situation particulière qu'ils dépeignent » (2009 : 90, ma traduction). C'est à ce moment que les sons seront identifiés comme des icônes auditives à relation non arbitraire. Jørgensen décrit ces sons comme étant « moins transparents » (2006 : 49), puisqu'ils ne correspondent pas exactement au son qui serait produit par cet objet dans la réalité. Ces types de signaux sonores sont, par exemple, régulièrement associés à l'inventaire du joueur, puisqu'il est lui-même représenté d'une manière plus abstraite en connexion avec le monde diégétique. L'auteure note que « [d]ans *World of Warcraft*, un son générique sera associé à chaque item relié à l'inventaire, et ce, peu importe si l'item est récolté, mis de côté ou simplement déplacé entre différents emplacements de l'inventaire »

12. Pour expliquer ce phénomène, Jørgensen donne l'exemple du son produit par une sonnette de porte. Ce son est synthétisé, mais, une fois associé par convention à l'événement « quelqu'un sonne à la porte », il sera réinterprété comme l'équivalent sonore « naturel » de cet événement.

13. Il est intéressant de noter que, pour Jørgensen, les sons d'ambiance d'un jeu ne peuvent pas être considérés comme des icônes auditives. S'ils contribuent activement au sentiment de présence, ils ne sont pas liés à la convivialité.

(2009 : 91, ma traduction). Situés quelque part au milieu du continuum d'iconicité, ces sons conservent, grâce à une association métaphorique, une certaine connexion avec l'objet ou l'événement qu'ils représentent tout en favorisant l'identification de cet objet ou de cette situation.

Cette utilisation non arbitraire du son va en quelque sorte de pair avec la vision qu'a Jørgensen de la conception sonore vidéoludique. En ce qui a trait à la création d'un sentiment de présence dans le monde numérique, elle souligne que le réalisme n'est pas le but visé par les concepteurs, lequel préfèrent parler de crédibilité. Bien évidemment, un son enregistré ne peut jamais être la reproduction parfaite de son équivalent naturel. En effet, tout enregistrement implique une part de subjectivité, certains aspects de l'objet sonore ayant été filtrés par la prise de son. C'est donc le rendu¹⁴ du son qui importe, dans les jeux vidéo. Jørgensen va même plus loin :

C'est la valeur fonctionnelle du son qui est importante et, en ce sens, ce n'est pas tant qu'un son semble crédible ou adapté pour un environnement qui est en cause, mais bien qu'il semble adapté à ce qu'un objet précis peut faire ou à ce qu'un événement particulier signifie pour le joueur (2009 : 61, ma traduction)¹⁵.

Selon Jørgensen, il est alors peut-être plus pertinent de parler d'un rendu fonctionnel¹⁶, la convivialité du système prenant toujours le dessus sur le réalisme de la diégèse.

14. Jørgensen parle plutôt de *perceptual fidelity*, terme qu'elle emprunte à Lars Nyre. Cependant, le concept de « rendu », que j'emprunte à mon tour à Michel Chion, me semble être une traduction plus juste de l'idée de Nyre que l'expression « fidélité perceptuelle ».

15. Ceci ne signifie pas qu'il faille négliger le potentiel d'affect des sons. C'est simplement que pour accéder à l'expérience esthétique d'un jeu (voir Arsenault, 2011), qui passe en grande partie par le sentiment conféré par le pouvoir d'action du joueur (Murray, 1997), il faut tout d'abord maîtriser les mécaniques du jeu et obtenir les informations qui permettent de réagir adéquatement aux situations qu'il propose.

16. Ma traduction de *functional fidelity*. À l'appui de son affirmation, Jørgensen donne l'exemple des monstres qui émettent des sons au moment d'attaquer le joueur. Au regard de la réalité, ces sons semblent effectivement étranges puisqu'il est fort peu probable qu'un prédateur s'annonce ainsi avant d'attaquer. Par exemple un lion demeurera silencieux jusqu'à ce qu'il soit prêt à enfoncer les griffes dans la chair de sa proie. Or, dans les jeux vidéo, ces sons offrent une information indispensable à la prise de décision du joueur, d'où l'idée d'un rendu fonctionnel.

Implémentation¹⁷ et types de fonctions sonores vidéoludique

En ce qui concerne les fonctions de convivialité, Jørgensen souligne que les sons peuvent être implémentés de deux manières différentes : ils peuvent avoir une valeur proactive ou réactive. Les sons à valeur proactive servent d'annonce, de préavis ou d'alerte : les sons peuvent représenter différents degrés d'urgence, les annonces ne demandant pas une attention immédiate de la part du joueur, alors que les sons de préavis et d'alerte demandent au joueur de prendre immédiatement, face à la menace qu'ils incarnent, les décisions qui s'imposent (2006 : 50). Les sons à valeur réactive servent de réponse – sous forme de confirmation ou de rejet –, ils offrent une rétroaction sur les actions du joueur (2009 : 62). Les fonctions de convivialité sont probablement les plus communes et « ont une relation directe avec les actions qui se produisent dans les jeux » (2009 : 158, ma traduction). Ces fonctions seront assurées par des icônes auditives, des icônes auditives à relation non arbitraire ou des carillons, selon le genre vidéoludique et l'environnement de jeu.

Les jeux vidéo emploient aussi les sons pour assurer des fonctions liées au contrôle ou à l'orientation (2006, 2009). Selon certains joueurs interviewés par Jørgensen, acheminer une même information par différents canaux sensoriels (vue et ouïe) facilite l'enregistrement et la compréhension de certains messages. De plus, « le son contribue à faciliter la gestion du jeu [...] en offrant de l'information qui serait difficile à fournir par le visuel seul » (2006, 51, ma traduction). Ce canal supplémentaire facilite aussi l'orientation du joueur dans le jeu. De manière générale, les fonctions d'orientation « informent [le joueur] d'une présence¹⁸ ou de la localisation relative des objets et des événements » (2006, p. 50, ma traduction). En permettant au joueur de localiser approximativement leur source dans l'environnement, les sons étendent sa perception au-delà de ce qui est perceptible visuellement. Le joueur peut ainsi repérer des ennemis avant même de les voir, ce qui lui procure un avantage considérable.

¹⁷ « Implémentation » doit être compris au sens informatique du terme (dérivé de l'anglais *implementation*). Par exemple, certains sons sont implémentés dans le code du jeu de manière à être déclenchés en réponse aux actions du joueur (réactifs) ou pour illustrer les changements dans l'état du jeu (proactifs).

¹⁸. À ne pas confondre avec le sentiment de présence que peut ressentir le joueur.

Les sons permettent aussi au joueur d'identifier des objets ou des événements, et de leur accorder une « valeur » relative. En conséquence, les fonctions d'identification « sont reliées au son en tant que système de reconnaissance » (2009 : 158, ma traduction)¹⁹. Cette fonction n'est pas présente dans les jeux « uniquement sous la forme d'icônes auditives automatiquement reconnues, mais aussi sous la forme de carillons [qui doivent passer par un processus d'apprentissage] avant d'être reconnus comme appartenant à une source spécifique » (2006 : 51, ma traduction). Jørgensen donne ici l'exemple du leitmotiv musical. Le son permet aussi d'accorder une certaine « valeur » à une source. Prenant le cas du jeu de stratégie en temps réel *Warcraft III : Reign of Chaos* (Blizzard Entertainment, 2002), l'auteure conclut que « plus une entité est puissante, plus son [signal sonore] de reconnaissance sera distinct » (*ibid*, ma traduction)²⁰. Ainsi, « le son devient un système d'outils qui offre de l'information à propos des mécaniques de jeu » (2009 : 159, ma traduction).

Quant à elles, « [b]ien que fonctionnant de manière plus subtile, les fonctions d'atmosphère du son peuvent tout de même être considérées comme centrales [dans l'expérience du jeu] » (2006 : 50, ma traduction). À cet égard, tout comme dans les films, la musique joue un grand rôle, mais les sons de l'environnement et des objets, ainsi que les dialogues, contribuent aussi à l'ambiance d'un jeu ou d'une situation. Ainsi, « [g]lobalement, le paysage sonore contribue à un sentiment de présence ou même à l'immersion dans un jeu en créant l'illusion que le monde du jeu est un espace réel » (2006 : 50, ma traduction). L'auteure insiste aussi sur le fait que les sons d'ambiance peuvent « influencer le comportement du joueur, par exemple en le faisant agir plus prudemment, de manière à éviter les ennemis dangereux et les situations déplaisantes » (*ibid*, ma traduction)²¹. Cette influence peut être plus directe

19. Dans une écoute causale, les sons servent toujours à glaner de l'information sur leur source.

20. Dans son exemple, Jørgensen oppose les sons produits par les ouvriers Orcs qui, par leurs signaux de reconnaissance – « work, work », « ready to work » et « be happy too » –, semblent « humbles et obéissants », alors que les signaux vocaux des *Warchiefs*, l'entité la plus puissante du jeu – « I have an axe to grind », « for my ancestors », et « an excellent plan » – « mettent l'accent sur leur agressivité, leur honneur et leurs perspectives stratégiques » (2006 : 51, ma traduction). Cet écart est d'ailleurs perceptible à même la morphologie du son, la voix des *Warchiefs* étant beaucoup plus grave que celle des ouvriers et leurs pas, beaucoup plus lourds (*ibid*).

21. Il s'agit ici d'un autre exemple de l'influence réciproque entre le pouvoir affectif du son et ses qualités informatives (information par l'affect, affect par l'information).

encore puisque, « lorsque la musique est utilisée avec une fonction de réponse ou d'urgence, elle aura aussi des propriétés [liées à] l'atmosphère » (*ibid*, ma traduction). C'est là d'ailleurs le meilleur exemple qui soit d'une situation où la convivialité du son s'intègre à un sentiment de présence dans le monde du jeu.

Fonction transdiégétique du son

Comme le spécifie Jørgensen, une autre des techniques²² « [q]ui favorise la fusion du système de jeu avec le monde virtuel est de brouiller la frontière existant entre l'environnement de l'avatar²³ [...] et celui du joueur [...] » (2008a, ma traduction). Cette technique consiste le plus souvent à fournir de l'information sur le système « par des canaux inhérents à l'environnement de jeu » (*ibid*, ma traduction). Le brouillage est le résultat de l'amalgame de deux niveaux de communication : communication entre le système du jeu et le joueur ; communication entre le monde du jeu et l'avatar (2008a). En effet, en plus de l'espace diégétique, un espace conceptuel s'installe entre le joueur et le système de manière à laisser un individu, qui se trouve à l'extérieur de l'environnement de jeu, transférer de l'information vers le monde du jeu. Devant cette réalité interactive, Jørgensen souligne que, dans un contexte vidéoludique, la division naturelle entre le diégétique et l'extradiégétique, telle qu'elle se trouve définie par les études narratives et cinématographiques, semble nettement se complexifier, conséquence directe de cette union des niveaux de communications. Par conséquent, elle choisira d'utiliser un nouveau terme, qui met l'accent sur les aspects fonctionnels spécifiques du son dans les jeux vidéo, expliquant que le son échappe à l'opposition communément admise, en études cinématographiques, entre le son diégétique (un son qui a pour origine le monde filmique) et le son extradiégétique (un son qui découle d'une source externe) (2007 : 105, ma traduction).

22. La première technique étant de favoriser l'utilisation de certains types de signaux sonores, nommément les icônes auditives, plus que d'autres.

23. Ce terme, de manière générale désigne l'incarnation numérique, dans un monde virtuel, de l'individu qui prend part au jeu. Si cet individu interagit avec un personnage déterminé, ayant en quelque sorte une personnalité propre, on parlera de personnage-joueur.

Le nouvel espace conceptuel est un espace « transdiégétique »²⁴. Certains sons provenant de l'extérieur de la diégèse peuvent avoir un impact sur les événements qui se déroulent à l'intérieur de celle-ci « parce que le spectateur se change en un joueur qui ne fait pas qu'entendre le[s] son[s] extradiégétique[s], mais qui utilise aussi l'information qu'ils contiennent pour influencer sur ce qui se produit dans le monde du jeu » (2007 : 106, ma traduction). Le phénomène inverse serait aussi observable. Suivant son analyse sonore du jeu de stratégie en temps réel *Warcraft III : Reign of Chaos*, Jørgensen affirme que certains sons « ont des sources diégétiques, mais ne semblent pas s'adresser [à une entité présente dans] le monde du jeu. Ces sons semblent plutôt communiquer directement avec le joueur qui est situé dans le monde réel » (2007, p. 112). Conséquemment, Jørgensen divise le transdiégétique en deux fonctions qui représentent les trajets effectués par le son à travers l'espace conceptuel décrit ci-dessus. Les sons ayant une fonction transdiégétique externe représentent ainsi « des sons qui, à proprement parler, doivent être étiquetés extradiégétiques, mais qui semblent communiquer avec des personnages ou des dispositifs internes à la diégèse » (2007 : 112, ma traduction). À l'opposé, les sons transdiégétiques internes « ont des sources diégétiques, mais [...] semblent communiquer avec le joueur qui se situe dans l'espace réel »²⁵. Ce brouillage de la division entre le diégétique et l'extradiégétique, qui fait appel aux aptitudes

24. Jørgensen a d'abord basé sa réflexion sur la définition de la diégèse qu'offrent David Bordwell et Kristin Thompson, puis sur la distinction d'Edward Branigan entre diégétique, extradiégétique et extrafictionnel. Dans un texte tout récent (2011), elle laisse entendre que les concepts liés à la diégèse sont inaptes à décrire efficacement les espaces de communication du jeu vidéo, remettant d'ailleurs en question son propre travail sur le transdiégétique. Comme la réflexion menée dans ce texte découle d'une définition inexacte de la diégèse et que s'y affirme parfois une rhétorique fallacieuse, nous avons choisi d'écarter cette publication pour l'instant, quitte à y revenir éventuellement pour en faire la critique.

25. Cette fonction demanderait à être critiquée davantage puisque Jørgensen la développe à partir d'un postulat erroné. Elle affirme que l'absence d'avatar qui caractérise les jeux de stratégie en temps réel signifie une absence de représentation du joueur sur l'axe diégétique, les entités s'adressant directement au joueur lorsqu'elles acquiescent à ses ordres. Or, ces entités chevaleresques s'adressent plutôt à l'instance incarnée du joueur, qu'elles nomment Milord. Le fait est plus saillant dans le précédent jeu de Blizzard, *Starcraft* (1998). Ayant la possibilité de contrôler une race extraterrestre, les Zergs, le joueur se fera appeler Cerebrate par les entités, un Cerabrate ayant la forme d'une masse de chair visqueuse sans rapport avec l'apparence du joueur confortablement assis dans son salon. Ainsi, bien qu'« invisible » à l'écran, on peut tout de même déduire que le joueur est bel et bien « incarné » dans la diégèse.

de métacommunication²⁶ du joueur, témoigne une fois de plus de la double réalité des jeux vidéo (système et diégèse) tout en démontrant qu'une compréhension aigüe du contexte vidéoludique dans son ensemble est nécessaire à la progression du joueur dans le jeu.

Contexte et affordance

Pour que le joueur puisse interpréter correctement l'information que lui fournit le son, il ne doit pas seulement comprendre la relation qui se tisse entre celui-ci et l'événement qui le déclenche. « À cet égard, contexte est un mot clé et je soutiens qu'un son spécifique ne peut pas être compris isolément, mais que la situation dans laquelle il est entendu détermine toujours l'interprétation du contenu informatif d'un signal sonore » (2008a, ma traduction). En ce sens, « le son vidéoludique est intégré à la compréhension globale de ce que jouer au jeu signifie et, de manière à bien saisir ce qu'un signal auditif indique, le joueur doit comprendre la situation de jeu [...] dans son ensemble » (*ibid*, ma traduction).

En affirmant cela, Jørgensen prend comme point de départ l'approche cognitive du cinéma de Birger Langkjaer : dans un contexte audiovisuel, la meilleure façon « d'expliquer la représentation est d'outrepasser la supposition qu'il y a une relation représentationnelle entre le son et l'image. Il faut plutôt regarder ce à quoi ils réfèrent ensemble [...], nommément une situation telle qu'un événement ou une activité » (2009 : 69, ma traduction). Pour Jørgensen, cette perspective orientée vers la situation gagnera à être mise en relation avec le concept d'*affordance*, tel que l'a défini James Gibson : « n'importe quel objet dont on fait la rencontre présentera des manières d'interagir avec lui » (2009 : 70, ma traduction). Pour reprendre l'exemple de Jørgensen, « l'image et le son d'un monstre offriront conjointement de l'information à propos d'une situation que *l'on peut fuir* ou que *l'on peut combattre* » (*ibid*, ma traduction). C'est donc le potentiel de dangerosité de la situation qui est de première importance et pas seulement la présence du monstre. C'est ensuite en regard de ses connaissances contextuelles (déterminées par la connaissance que le joueur a du

26. La métacommunication correspond « à notre habileté à comprendre simultanément plusieurs cadres de référence et explique pourquoi nous n'avons aucun mal à interpréter la relation qui existe entre le diégétique, l'interface et l'extradiégétique dans les jeux vidéo » (2007 : 115, ma traduction).

jeu et du genre, ainsi que du contexte particulier) appliquées à la situation diégétique que le joueur pourra déterminer l'action ou les actions appropriées.

Comme Jørgensen le souligne, « [é]couter est une activité cognitive complexe [et] un joueur doit souvent donner un sens à des situations où [cohabitent] plusieurs sons simultanés » (2008a, ma traduction). Il devient alors crucial de classer les sons efficacement, en fonction du degré d'urgence associé à la situation qu'ils dépeignent. Du coup, la signification des sons individuels devient secondaire par rapport à la compréhension de la situation globale. C'est d'ailleurs ainsi qu'un même son, en fonction du contexte dans lequel il est entendu, peut prendre une signification différente. C'est aussi pourquoi, dans « Audio and Gameplay : An Analysis of PvP Battlegrounds in *World of Warcraft* », Jørgensen s'attaque à la question des sources sonores²⁷ en mentionnant que, puisque la progression du joueur dépend plus de la situation (d'écoute) que de la nature de l'objet sonore²⁸ lui-même, il devient impératif d'identifier le générateur du son plutôt que sa source.

Les générateurs sonores

Pour l'auteure, « le générateur d'un son n'est pas équivalent à la source du son. Alors que la source est l'objet qui produit physiquement (ou virtuellement) le son, le générateur est la cause de l'événement qui produit le son » (*ibid*, ma traduction). Pour reprendre son exemple, lorsqu'un son d'impact est entendu au moment où un avatar est frappé par un ennemi, celui-ci sera considéré comme la source du son, mais l'ennemi en sera le générateur. Cette distinction permet au joueur de mieux juger du niveau d'urgence d'un son puisqu'il dépend directement de l'événement dont il découle. Dans son étude, Jørgensen distingue cinq grandes catégories de générateurs²⁹ : les sons générés

27. Jørgensen donne suite à un article intitulé « Sound and Sources in *Sacred* », publié en 2004, dont cette synthèse ne rend pas compte pour des raisons de concision, mais aussi de cohérence. En effet, la majorité des intuitions de l'auteure ont été conceptualisées plus à fond dans ses articles subséquents.

28. À ne pas confondre avec l'objet sonore de Pierre Schaeffer. Il s'agit plutôt de la relation entre le son et le référent.

29. Nous croyons important de souligner que les définitions que Jørgensen propose des générateurs sonores découlent de l'analyse du seul *World of Warcraft*, qui appartient à un genre précis, le jeu de rôle en ligne, qui se joue en mode multijoueur. On peut supposer que, dans un autre contexte vidéoludique, les définitions de ces générateurs ainsi que les fonctions qui s'y rattachent varieraient. De plus, le texte de Jørgensen porte uniquement sur les fonctions de

par le joueur, les sons produits par les alliés et les ennemis, les sons engendrés par l'environnement de jeu et enfin, ceux propagés par le système de jeu.

« Un son généré par le joueur est causé par l'action de celui-ci. Si un joueur n'est pas activement impliqué dans la production du son, il n'est pas le générateur de ce son » (2008a, ma traduction). Selon Jørgensen, la principale fonction informative d'un son généré par le joueur est de fournir des renseignements liés à la convivialité, ou, plus précisément, de transmettre une réponse [du système] puisqu'il est [audible] immédiatement après l'action du joueur (2008a, ma traduction). Bien que les fonctions de convivialité demeurent toujours prédominantes, les sons générés par le joueur fournissent aussi de l'information sur l'espace, parfois sur la temporalité, ainsi que sur l'état du personnage-joueur (*ibid*). Lorsqu'ils sont diégétiques, ces sons sont généralement incarnés par des icônes auditives ou des signaux conventionnels, ce qui favorise leur intégration au monde du jeu. Selon Jørgensen, certains sons générés par les joueurs peuvent avoir un caractère transdiégétique. Les sons reliés à l'inventaire sont dits transdiégétiques externes, alors que les réponses de l'avatar aux commandes du joueur, qui ne peuvent être entendues par les autres joueurs, sont considérées comme transdiégétiques internes. Ces sons peuvent offrir de l'information spatiotemporelle ; le plus souvent ils servent des fonctions de confirmation ou de rejet. Selon la définition qu'offre Jørgensen, tous les sons extradiégétiques du jeu sont générés par les joueurs et concernent les différents menus du jeu.

Quant à eux, « [l]es sons générés par les ennemis et les alliés sont tous diégétiques [...] et sont produits de manière externe à la perspective du joueur, puisqu'ils sont détachés de ses actions et émergent de l'environnement de jeu » (2008b, ma traduction). Dans une situation de jeu, « il est crucial de déterminer si un son est généré par un ennemi ou un allié, puisque [chaque situation] aura des conséquences différentes pour le joueur » (*ibid*, ma traduction). Comme Jørgensen l'explique, les sons produits par les alliés et les ennemis peuvent parfois être difficiles à distinguer (*ibid*)³⁰. C'est pourquoi il devient essentiel pour le joueur d'évaluer la situation entière, dépassant du même coup la simple

convivialité rattachées aux différents générateurs, mettant de côté, pour l'occasion, un important pan de sa théorie sur le son vidéoludique.

30. Ceci vaut particulièrement pour *World of Warcraft*, car tous les sons émis par un type d'avatar, par exemple les voleurs, sont tirés du même fichier, et ce, qu'ils soient alliés ou ennemis.

identification des sons. Ce sont les possibles implications associées à un son qui priment alors.

Les sons générés par les ennemis et les alliés « transportent de l'information spatiale [orientation et identification] en relation avec une présence relative et fournissent des indications sur des changements d'état de l'avatar » (*ibid*, ma traduction). Cependant, en termes de convivialité, les deux catégories de sons demandent à être distinguées fonctionnellement (*ibid*). Les sons générés par les ennemis doivent généralement être traités avec un sentiment d'urgence plus élevé et requièrent, comme le mentionne Jørgensen, une attitude proactive de la part du joueur. En fonction du contexte, les sons des ennemis peuvent aussi être évalués comme de simples annonces qui témoignent de différents états du jeu. Selon l'auteure, il en va de même pour les sons des alliés qui ne font qu'offrir de l'information, par exemple sur leur présence dans l'environnement, sans susciter de réaction proactive à leur égard³¹.

Les sons générés par l'environnement sont décrits par Jørgensen comme étant non dynamiques et n'ayant donc pas d'importance en ce qui a trait aux actions posées par le joueur. Elle affirme qu'une fois que le joueur aura reconnu un son appartenant à l'environnement, il le relèguera parmi les préoccupations secondaires. C'est que « les sources de ces sons ne sont ni interactives, ni produites par les joueurs, leur inclusion ne servant aucune fonction opérationnelle sinon celle d'identifier [des objets] et d'accentuer un sentiment de présence dans le monde du jeu » (2008a, ma traduction).

Enfin, les sons générés par le système de jeu sont, de l'aveu de Jørgensen, de loin ceux qui portent le plus à l'interprétation. L'auteure les définit comme des sons « générés par le système, afin de fournir l'information qu'un avatar ne peut produire par lui-même, et qui transportent les renseignements directement liés aux règles du jeu ainsi qu'à propos de l'état de jeu et de l'avatar » (2008a, ma traduction)³². Leur fonction principale est

31. Puisque dans bon nombre de jeux la survie des alliés devient aussi importante que celle du personnage-joueur, il est possible de penser que, dans un contexte autre que *World of Warcraft*, plusieurs sons alliés pourraient atteindre un certain niveau d'urgence.

32. Dans son étude sur les séances de jeu joueur-contre-joueur, dans *World of Warcraft*, Jørgensen ne prend pas le temps d'approfondir ce qu'elle entend par « sons générés par l'environnement de jeu » puisque ces derniers jouent un rôle secondaire dans la démonstration de son hypothèse.

d'offrir au joueur de l'information sur une modification touchant l'avatar ou l'état de jeu. La majorité des sons générés par le système seraient transdiégétiques, mais pourraient aussi être diégétiques quand l'avatar produit des sons qui ne sont pas le résultat d'une action faite par le joueur. Comme l'explique Jørgensen, « le fait que le son vidéoludique reflète ainsi les mécaniques de jeu souligne le fait que le son offre de l'information importante pour la jouabilité et qu'il conditionne le comportement du joueur dans le jeu » (2008a, ma traduction).

Guillaume Roux-Girard
Janvier 2012