

Analyse de la structure audio de la console 8 bits de Nintendo et des implications dans le langage sonore

Écrit par Diogo Andrei Benvenuti et Gabriel Duarte Barbosa
Version française corrigée par Martin Bussy-Pâris

Résumé

Le jeu vidéo est un phénomène audiovisuel. Puisqu'il s'agit d'un langage multimédia assimilé partout dans le monde, l'expérience sonore demande une investigation détaillée pour que l'on puisse expliquer ses particularités. Cet article se propose d'analyser une période spécifique de l'évolution du son comme facteur déterminant pour l'établissement de ce qu'est le jeu vidéo aujourd'hui en tant que média sonore: l'apparition de la console 8 bits de Nintendo. Pour que l'on puisse comprendre les possibilités et particularités du langage musical de la console, nous commencerons le travail par un bref rappel de son contexte historique, pour ensuite s'atteler à une étude de ses capacités techniques. Finalement, cette analyse va faire la lumière sur l'importance de la NES, du point de vue sonore, dans l'histoire des consoles, ainsi que sur les caractéristiques de son langage sonore et ses références aujourd'hui dans la culture « 8 bits ».

Mots clés: Nintendo, musique, audio, langage sonore, 8 bits.

Contact des auteurs:

diogoandrei@yahoo.com.br
presto.gabriel@gmail.com

1. Introduction

Pour analyser les aspects particuliers de la musique et du son de la console 8 bits de Nintendo, il faut situer son apparition dans l'histoire des jeux vidéo, parce qu'il faut comprendre non seulement ce que permettait techniquement la console au niveau du son, mais aussi les objectifs de Nintendo pour apporter des nouveautés à ce milieu.

Conçue et commercialisée dans un premier temps au Japon en 1983, sous le nom de Famicom¹, la console y connut un énorme succès commercial [Kent 2001, p. 279]. En revanche, elle n'atteignit les États-Unis qu'en 1985. Ce retard était dû au déclin du marché américain du jeu vidéo au début des années 1980 où la console n'attira pas le public comme on s'y attendait.

Au début de ladite décennie, l'entreprise la plus importante du marché des consoles, Atari, vit subitement la valeur de ses actions décroître de façon vertigineuse. Cette crise atteignit tout le marché vidéoludique [Kent 2001, p. 235], rendant le public et

les développeurs de plus en plus méfiants vis-à-vis des jeux vidéo. Comme principaux exemples de la mauvaise gestion d'Atari, Kent évoque la production de 12 millions de cartouches Pac-Man défectueuses et le développement bâclé du jeu "décevant" (dixit Kent), E.T., basé sur le film éponyme de Spielberg. La chute fut grande et, comme le souligne Kent, "ces nouvelles ont provoqué la panique".

À part cela, le début des années 1980 fut aussi marqué par l'apparition des premiers ordinateurs personnels, qui permettaient également de jouer et qui avaient, de plus, des performances supérieures et mieux exploitées que la console d'Atari. Le résultat fut que les consommateurs ainsi que les développeurs de jeux commencèrent à préférer les ordinateurs personnels aux jeux vidéo.

Il est extrêmement important de comprendre que les consoles, devenant progressivement inférieures aux ordinateurs personnels, avaient également de moindres capacités audio. Comme l'a bien remarqué Howard Phillips, de la filiale américaine de Nintendo, le son était véritablement un facteur décisif dans la réussite d'un jeu vidéo. Dans un article de l'époque intitulé « Les faits sur les jeux vidéo », Phillips remarque que l'un des problèmes des consoles était que leurs jeux « étaient limités par la faible qualité du support et qu'ils contenaient une variété restreinte d'effets sonores ». [Kent 2001, p. 349]. Par conséquent, Nintendo décida d'inclure des nouveautés à sa console dans l'optique de présenter des jeux différents, notamment au niveau du son, par rapport à ce qu'il se faisait à l'époque.

La première mesure prise par Nintendo pour s'attaquer au marché occidental fut de changer le nom de la console par NES² (Système de loisirs Nintendo), de manière à s'écarter un peu du concept de jeu vidéo, alors en crise aux États-Unis. L'objectif principal était de créer et de vendre un nouveau produit. À part les capacités du système lui-même, il est intéressant de se pencher sur la façon dont Nintendo mettait l'accent sur le terme « système de loisirs » en proposant beaucoup de périphériques pour sa plate-forme. Les premières apparitions de la NES étaient toujours accompagnées de R.O.B. (un petit robot qui interagissait avec quelques jeux) et une arme appelée Zapper. Le vice-président de la filiale américaine de Nintendo, Howard

¹ Diminutif de Family Computer, parfois abrégé FC.

² Désormais, ce travail va toujours faire allusion à la version américaine (NES), dans laquelle une grande partie de la recherche a été développée. Lorsqu'il le faudra, nous présenterons les différences existantes entre la Famicom et la NES.

Lincoln, met en évidence le nouveau concept de l'entreprise: « nous avons changé notre position. Nous vendions un jeu de robot et non pas un jeu vidéo » [Kent 2001, p. 288].

La structure audio de la console et l'usage qui en a été fait par la suite sont issus de ce même principe, avec l'intention de rendre la NES singulière et bien différente de sa prédécesseur l'Atari, tout en regagnant la confiance du marché nord-américain.

Il est important de remarquer que, avant l'apparition de Nintendo, les sons dans les jeux avaient des fonctions bien distinctes: attirer l'attention du joueur qui passent dans une salle d'arcade et sonoriser les actions du jeu de façon naturelle³ [Collins 2008, p. 9] - bien que les sons fussent simples et techniquement limités. En général, la musique était employée dans l'introduction et à la fin (game over) du jeu. Il y a quelques exemples rares de musique pendant une action victorieuse ou alors lorsque le jeu possédait un rythme marqué, comme dans le cas de Space Invaders (Taito, 1978). Sa bande sonore marque le mouvement des ennemis et donne un sens de tension pendant qu'ils s'approchent, ce qui donne, de façon délibérée ou non, le premier usage de musique continue extradiégétique (affirmation de Collins [2008, p. 9]).

Selon Karen [2008, p. 12], jusqu'à la crise du début des années 1980, quoique les plateformes permettaient l'emploi de musique continue - avec des canaux audio distincts de ceux utilisés par les effets sonores - il était très fastidieux de la programmer. C'est là peut-être la principale raison qui explique le retard du milieu du jeu vidéo, en ce qui concerne la musique continue, par rapport aux autres médias audiovisuels tel que le cinéma, qui avaient déjà la musique extradiégétique comme pratique depuis leur naissance. La puce de son des consoles Atari, par exemple, avait un grave problème dans l'accordage des notes. Cela rendait la tâche du compositeur difficile, puisqu'il y avait un nombre extrêmement limité de notes relativement accordées avec lesquelles on pouvait travailler.

Le cas de Pitfall! 2 pour la console d'Atari est remarquable. En 1984, au faite de la crise, les développeurs de ce titre ont estimé que les capacités audio du système étaient très impressionnantes et que c'était un grand atout pour leur jeu [Forte et Garret 2008, p. 107-108]. Ainsi, pour optimiser encore davantage les performances sonores de la console, ils ont inclut puce dédiée au son dans la cartouche du jeu, permettant de programmer la musique d'une façon différente, avec des canaux distincts.

³ Il est intéressant de noter que le son était en fait un outil de marketing. Dans la publicité de Pong (Atari, 1972), on peut lire « son réaliste de balle ricochant et de raquette tapant », ainsi que dans celles de Jaws (Atari, 1975), « effets sonores réalistes - le plongeur crie lorsqu'il est attaqué par le requin ». Ces publicités, telles que beaucoup d'autres, sont accessibles sur <http://www.arcadeflyers.com/> (Consulté le 12 juillet 2010)

Nintendo, conscient de la phase dans laquelle se trouvait la technologie audio, connaissait également bien les tenants et les aboutissants du son dans le jeu vidéo. Ainsi, le son devenait une partie de son nouveau concept de produit : le système de loisirs. La compréhension de tous ces facteurs permit que, de fait, Nintendo fût capable de reconstruire seule le marché et de le conquérir avec ses nouveaux concepts [Collins 2008, p. 24].

2. Aspects techniques

Tout le son produit par la console 8 bits de Nintendo a quatre sources distinctes: l'unité audio de la console, les processeurs audio externes (trouvés dans quelques cartouches), le système de disquettes et le microphone du deuxième contrôleur (joystick) du Famicom. Le principal centre de traitement audio de la console est dans le CPU, la puce 2A03. C'est là où nous concentrons cette investigation. Dans la sous-section suivante, nous détaillerons le fonctionnement de ce générateur audio, après une brève description des deux autres sources audio.

Les développeurs de jeux pour le système 8 bits de Nintendo utilisaient des processeurs spéciaux pour étendre la capacité originelle de la console. Appelés MMC (Multi Memory Controllers ou Memory Management Controllers), il s'agissait de puces insérées dans le circuit même des cartouches. À part augmenter la capacité de calcul générale (utilisée en grande partie pour les graphismes), quelques-uns de ces processeurs offraient des canaux audio supplémentaires, tels que le MMC5 (deux canaux supplémentaires), le Konami VCR6 (trois canaux), le VCR7 (six canaux), et le NANCO106 (huit canaux)⁴. En revanche, il est important de noter que seule la version japonaise de la console avait accès aux canaux supplémentaires, parce que la NES ne possédait pas de récepteurs pour ces canaux dans son entrée de cartouches [Korth 2004]. Diskin affirme quand même que l'usage de ces puces « fut l'un des facteurs responsables par la longévité de la NES, permettant qu'elle surmonte ses limitations technologiques initiales » [2004, p. 27].

L'un des périphériques de la version japonaise était le système de disquettes Famicom Disk System (abrégé FDS). Capable de sauvegarder le contenu des jeux, ainsi que les progrès du joueur, dans quelques cas, le système s'intégrait à la console via un adaptateur, qui était connecté à son entrée de cartouches. Cet adaptateur possédait un circuit intégré personnalisé, le 2C33, qui, à part le fait de contrôler les fonctions du système de disquettes, ajoutait à la Nintendo 8 bits un canal audio supplémentaire, capable d'effectuer, de

⁴ Pour une liste complète de MMCs qui offrent des canaux audio supplémentaires, voir l'extrait « APU External Sound Channels » dans Korth [2004].

façon très simple, de la synthèse à tables d'ondes.⁵ Taylor [2004d] soutient que le canal supplémentaire « ajoute une nouvelle dimension à l'expérience de jeu » et commente avec enthousiasme la qualité audio du système : « le son conventionnel de la NES/Famicom peut presque être considéré comme rude lorsqu'il est comparé avec des jeux qui utilisent les avantages du canal de son supplémentaire mis à disposition via le FDS ». Comme l'affirme Collins [2008, p. 11], ce type de synthèse rendait possibles des résultats sonores beaucoup plus élégants et fut beaucoup plus utilisé à l'époque du 16 bits. Le système de disquettes de Nintendo, comme on l'a dit, est un cas très simplifié de synthèse à tables d'ondes - qui était cependant capable de résultats singuliers. En somme, le FDS rendait la console capable d'atteindre une nouvelle dimension sonore. D'un autre côté, le système souffrait de problèmes de stabilité et aussi de piratage.

Seules les versions les plus anciennes de la Famicom possédaient un microphone⁶, installé dans le deuxième contrôleur (avec un bouton de volume). Selon Korth [2004], le signal de ce microphone était envoyé simultanément vers deux chemins différents : (1) vers la sortie analogique de la console sans passer par aucune conversion, amplifiant ainsi simplement la voix du joueur dans le système audio auquel la console était connectée ; (2) vers un récepteur numérique interne, qui apparemment ne faisait que mesurer l'amplitude du signal pour le traitement de données des jeux qui employaient cette ressource. Il n'y a que peu de jeux qui utilisent ce microphone. *The Legend of Zelda*⁷ (Nintendo, 1986) est probablement le titre le plus célèbre qui bénéficie de cet élément. Même s'il s'agit d'un composant relativement rare, il est important de mettre l'accent sur son rôle en tant que générateur et (unique) récepteur audio directement lié au joueur.

Concernant la reproduction, indépendante de la source génératrice de son, la console est monophonique et possède deux sorties analogiques audio. L'une de ces sorties va vers un modulateur RF (radiofréquence, destiné aux téléviseurs), où les signaux audio et vidéo sortent par le même connecteur. L'autre est dédiée, via un connecteur RCA, à n'importe quel système audio pouvant recevoir ce type de signal. Tout le son (excepté le signal provenant du microphone, traité dans le paragraphe précédent), avant de sortir de la console, passait par un convertisseur analogique-numérique standard, de 4 bits [Taylor 2004c], ce qui donne 16

niveaux de tension⁸. Suivant le modèle et la version, la console peut n'avoir qu'une ou deux de sortie.

2.1 L'unité de traitement audio de la console

Presque tout le traitement audio est concentré sur l'unité centrale de la console, celle-ci n'ayant pas d'unité dédiée⁹. Cette unité centrale, le « coeur de la console » [Taylor 2004a], est basé sur le microprocesseur 8 bits produit par MOS Technology (EEUU), le 6502¹⁰, sorti originellement en 1975. Pendant le processus de développement de la console de salon, Nintendo demanda à MOS Technology de personnaliser le 6502 pour qu'il puisse servir aux buts du projet. La puce gagna ainsi de nouvelles capacités. Parmi celles-ci, le microprocesseur serait alors apte à générer et produire des sons¹¹. La nouvelle unité personnalisée fut nommée 2A03¹² et était produite par Ricoh, au Japon [Diskin 2004, p. 7].

Le 2A03 est catégorisé comme une pseudo unité de traitement audio (pAPU¹³, de l'anglais « pseudo Audio Processing Unit ») car il ne traite pas que du son, mais aussi d'autres informations. Cette unité de la console est composé de cinq canaux audio. Brad Taylor affirme que comme le système est « une collection de plusieurs unités indépendantes, qui opèrent toujours en parallèle » [2004b], les cinq canaux audio travaillent aussi de manière autonome. Alors, les altérations provoquées dans l'un des canaux n'interfèrent pas avec les autres.

Les deux premiers canaux¹⁴ sont des générateurs d'impulsions, capables de moduler leurs rapports cycliques¹⁵ en 12,5%, 25%, 50% (ce qui engendre une

⁸ Probablement en niveau nominal -10dBV, conformément aux standards d'appareils audio domestiques.

⁹ Exceptées les puces externes dans les cartouches qui, comme expliqué, ne font pas partie de l'univers de cette recherche.

¹⁰ Lié aussi à l'histoire d'autres consoles, tels que l'Atari 2600, et aussi de quelques ordinateurs, comme le Commodore 64.

¹¹ Pour une liste détaillée des fonctions intégrées au 2A03, voir Brad Taylor [2004a].

¹² À cause des deux systèmes différents de télévision, la puce 2A03 (1,78 MHz) servait le standard NTSC et sa version pour le PAL était le 2A07 (1,77 MHz) [Korth 2004]. Par question de praticité et des sources de recherche, ce travail se référera toujours au 2A03. De possibles différences entre les versions NTSC et PAL ne seront pas prises en compte ici. Taylor affirme que le canal d'échantillons possède la même sonorité dans les deux systèmes [2004a].

¹³ Des auteurs tels que Diskin et Taylor utilisent le terme pAPU, mais également APU, apparemment par question de brièveté.

¹⁴ Pour une description complète du fonctionnement de chaque canal, voir Brad Taylor.

¹⁵ Rapport entre la partie positive (lorsque le haut-parleur avance) et négative (lorsque le haut-parleur recule) de l'impulsion. Avec la valeur de 50%, les parties positive et négative ont la même durée, ce qui engendre une onde carrée. Dans le cas de 12,5%, un quart de l'impulsion est positif et le reste négatif. Avec 75%, trois quarts de l'impulsion est positif et le reste négatif. De telles variations créent de différentes sonorités et des variations de cette onde sonore.

⁵ Technique de synthèse dans laquelle la création du son est basée sur des ondes sonores pré-traitées. Cette synthèse possède une sonorité très typique. La synthèse à tables d'ondes tend à générer des sons beaucoup plus réalistes (lorsque le but est de reproduire synthétiquement un instrument réel).

⁶ Nous avons constaté lors de nos recherches qu'il y a peu d'informations à propos de ce composant de Nintendo, Martin Korth étant notre unique source d'information.

⁷ La version américaine du titre n'adapta pas le jeu à l'absence du microphone, probablement parce qu'il n'était utilisé qu'en peu de moments et ne gênait pas la progression du joueur.

onde carrée) et 75%. Les fréquences générées sont dans l'intervalle de 54,6Hz jusqu'à 12,4kHz. Ils possèdent un contrôle de volume de 4 bits (un total de 16 variations d'intensité) et un contrôle de portamento¹⁶ (pitch bending) qui comporte toute la tessiture entre 54Hz et 28kHz.

Un troisième canal, générateur d'ondes triangulaires, est capable de produire des fréquences entre 27Hz et 56Hz avec une définition de 4 bits et un volume fixe.

Le quatrième canal est un générateur de bruit blanc qui crée aléatoirement des fréquences entre 29Hz et 447kHz. De même que les canaux d'impulsion, il possède 16 niveaux de volume.

Le cinquième canal agit simplement comme un joueur d'échantillons. Puisqu'il utilise la technique de conversion connue sous le nom de modulation delta¹⁷, ce canal est appelé en général DMC (Delta Modulation Channel). Il utilise des fréquences d'échantillonnage pré-programmées qui varient entre 4,2kHz et 33,5kHz et deux méthodes d'exécution : DMA (Direct Memory Access) et le standard Pulse Code Modulation (PCM) [Collins 2008, p. 25 ; Segulja et Dai 2008, p. 2 ; Taylor 2004a].

Bien que le 2A03 travaille avec cinq canaux audio indépendants, la puce ne possède que deux sorties¹⁸ (identifiées comme broches 1 et 2 du microprocesseur). Les deux premiers canaux audio sont mis ensemble à la broche 1 et les trois autres le sont à la broche 2 [Taylor 2004a]. De là, ils suivent vers les sorties de la console, où ils sont mélangés de nouveau pour la sortie monophonique.

3. Musique et effets sonores

Étant donnés les aspects techniques et les ressources offertes par le 2A03 prises en considération, passons maintenant à l'analyse, du point de vue pratique, des possibilités offertes au « musicien ». Pour ce travail, nous considérerons le musicien, ou compositeur, comme l'entité qui simplement crée et met en oeuvre le son (musique et effets sonores) d'un jeu donné - indépendamment du fait que ce musicien ait

¹⁶ Brad Taylor [2004c] note que la seule différence connue entre ces deux canaux d'impulsion est dans l'un des aspects de programmation liés au portamento. Une grande partie de la littérature traite les deux canaux comme étant identiques. Le même sera fait dans ce travail.

¹⁷ Méthode de conversion analogique-numérique et numérique-analogique simple, dont le principal but est de transmettre l'information, sans se soucier beaucoup de la qualité du son. Pour une description complète de la méthode, voir Steele, R., 1975, Delta Modulation Systems. London: Pentech Press.

¹⁸ Il y a un diagramme de blocs qui montre ces sorties (broches) du 2A03, disponible sur <http://www.neshq.com/hardgen/schem4.jpg> (Consulté le 20 juillet 2010).

programmé le son ou non. Nous ferons comme cela, car le plus souvent, comme l'observe Collins [2008, p. 35-36], le programmeur lui-même créait la musique, les effets sonores et organisait tous les éléments audio d'un jeu.

3.1 Échelles et constructions harmoniques

Pour travailler des gammes musicales, dans le concept occidental de gamme comme organisation de hauteurs définies, nous avons besoin de notes musicales. Si l'on analyse les canaux de la Nintendo 8 bits, on peut spéculer que, excepté le canal de bruits, tous peuvent générer des fréquences de hauteur définie. En théorie, le canal d'échantillons peut exécuter des échantillons de sons avec une hauteur définie, par exemple, quelques notes pré-enregistrées d'un piano. Cependant, en pratique, à cause des limitations de la mémoire, le DMC était généralement utilisé pour des sons de hauteur non-définie (percussion) ou effets sonores (SFX¹⁹), comme nous le verrons plus tard. Donc, nous passons directement à l'analyse des canaux d'impulsion et des ondes triangulaires.

Les canaux d'impulsion, comme vu avant, sont capables de produire un large spectre de fréquences. En prenant comme base le la⁴²⁰ (440Hz) dans cette analyse, ces canaux couvrent toutes les notes entre la1 (55Hz) et fa9 (11,1kHz). Ainsi, les canaux d'impulsion couvrent à peu près huit octaves. Collins [2008, p. 25] corrobore cette observation.

Les ondes triangulaires de la console commencent une octave au-dessous par rapport à celles d'impulsion (27Hz)²¹, la0 (zéro). Théoriquement, elles vont jusqu'à 56kHz. Non seulement cette valeur dépasse largement la limite supérieure de l'audition humaine (20kHz), mais, en plus, elle est au-delà du seuil de perception de hauteur. En pratique, comme la perception humaine de la note (pitch, hauteur définie) décroît aux hautes fréquences, nous prenons comme limite de discernement de notes la valeur 12kHz - des valeurs au-dessus de celle-ci sont rarement trouvées dans la pratique musicale [Olazabal 1954, p. 21]. Ainsi marquons-nous la limite musicale supérieure du canal d'ondes triangulaires en fa9 (11,1kHz).

Les deux canaux sont capables de générer une gamme chromatique tempérée. Cela signifie que le plus petit intervalle entre les notes sera toujours le même, ce qui permet à la console de Nintendo de générer une division de fréquences respectant le système occidental

¹⁹ Forme réduite de « sound effects », en anglais.

²⁰ Dans la notation musicale scientifique, les notes sont accompagnées d'un numéro pour désigner ses respectives octaves et ainsi distinguer des notes avec un même nom mais des hauteurs différentes. Un la 2 est plus grave qu'un la 3 - il est une octave au-dessous, par exemple.

²¹ Il serait difficile de trouver une télévision capable de générer une fréquence si basse, mais cela va au-delà de la portée de ce travail.

d'organisation de hauteurs. On peut établir, de manière très simple, qu'ainsi la NES a à sa portée toutes les notes qu'un piano parfaitement accordé serait capable de produire. De cette gamme, le musicien peut extraire les notes qui lui sont nécessaires dans la composition. De même, puisque les canaux du 2A03 travaillent de façon indépendante, le système est capable de travailler la simultanéité des notes, respectant les normes harmoniques (conception verticale de composition) sans avoir de problème d'accord. En plus des possibilités harmoniques, il y a aussi des possibilités de polyphonie (mélodies simultanées). Ainsi par la nécessité de travailler la juxtaposition des lignes mélodiques. Dans le cas de Nintendo, il existe trois lignes mélodiques (voix) possibles, étant donnés les canaux. A propos de la division des voix, Collins discourt que « les trois canaux générateurs de ton étaient typiquement utilisés de manière très conventionnelle, avec un canal de mélodie, un autre pour l'accompagnement, et le dernier pour la basse (canaux de bruits et DMC pour la percussion). Les deux canaux d'impulsion servaient habituellement pour les accords ou la mélodie, le canal triangulaire jouant l'accompagnement à la basse » [2008, p. 25].

À partir d'une analyse des ondes employées dans le processus de composition, on peut conjecturer sur la raison de cet usage des canaux de la Nintendo 8 bits. Si l'on part du concept de Toch, « on pourrait dire que la mélodie consiste en la succession de sons de hauteurs distinctes, par opposition à leur audition simultanée » [1985, p. 12]. La nature des ondes générées par les canaux d'impulsion permet d'atteindre la qualité recherchée pour la composition musicale. D'abord, si le cycle de l'impulsion est à 50%, nous aurons une onde carrée. Avec une valeur efficace (RMS²²) de facteur 1 (plus grande que celle de l'onde triangulaire, comme on le verra plus tard), une plus grande extension d'harmoniques²³ (impaires) et en plus un contrôle de volume (offert par l'APU), ce canal offre un meilleur contrôle et une plus grande versatilité au compositeur. Si le cycle de l'impulsion ne fournit pas d'onde carrée, le RMS ne sera pas de 1, mais encore proche de celui d'une onde triangulaire. Il faut se rappeler que, comme

²² RMS (Root Mean Square), valeur efficace d'amplitude. En somme, plus la valeur RMS est grande, plus le volume réel d'un signal audio l'est. Des ondes possédant la même tension ne seront pas forcément perçues avec le même volume. La nature de l'onde (carrée, triangulaire, etc.) utilise cette amplitude pour donner une valeur efficace d'utilisation. En résumé, une onde triangulaire avec amplitude de 1V sera perçue avec un volume plus petit que celui d'une onde carrée de même tension, à cause du fait que la valeur efficace (RMS) de l'onde carrée soit plus grande que celle de l'onde triangulaire. Pour une description technique complète, voir Ballou, G. 2008. Handbook for Sound Engineers. Massachusetts: Focal Press.

²³ Les ondes carrées contiennent, en théorie, tous les harmoniques impaires. Les ondes triangulaires, quant à elles, possèdent également des harmoniques impaires, mais avec une moindre extension (les harmoniques perdent rapidement leur énergie lorsqu'elles s'éloignent de la fréquence fondamentale).

les canaux d'impulsion offrent encore un contrôle de portamento, les compositeurs peuvent créer plus d'expressivité et d'ornementations, en caractérisant la mélodie avec plus de variations. Le canal triangulaire, sans contrôle de volume ni de portamento, avec un RMS d'environ 0,57 et une moindre extension d'harmoniques, peut cependant atteindre des notes une octave en-dessous que les ondes d'impulsion et est ainsi adapté à la création d'une ligne de basse.

En revanche, de telles questions techniques, bien qu'elles dictent la pratique de la composition, n'excluent pas la possibilité de travailler les voix et les canaux d'autres façons. Quoique la plupart des jeux utilisent des configurations d'arrangement comme celle-là, différents cas existent. Dans l'une des musiques de Metroid (Nintendo, 1987), le compositeur Hirozaku Tanaka utilise le canal triangulaire pour la voix principale et l'un des canaux d'impulsion pour la basse [Collins 2008, p. 26]. Nous verrons plus tard que ces canaux sont aussi utilisés pour les SFX.

3.2 Percussions, effets sonores et échantillons

Comme le canal de bruits ne produit pas de notes, il servait généralement au compositeur comme instrument de percussion, ainsi qu'à la création de SFX [Karen 2008, p. 26]. Il est important de noter que ce canal n'est pas un simple générateur de bruit blanc (tout le spectre entre 20Hz et 20kHz), car il permet également de filtrer ce bruit blanc. Comme Taylor [2004a] le note, le canal est capable de créer « plusieurs simulations d'effets sonores ». Dans *The Legend of Zelda*, le canal est chargé d'exécuter les cymbales. Dans *Mega Man 6* (Capcom, 1993), en revanche, le canal est chargé de toutes les percussions (dans quelques niveaux) et exécute les lignes de grosse caisse, de caisse claire et les cymbales. C'est un bon exemple pour noter que le canal n'émet pas que du bruit blanc, mais peut atténuer des régions du spectre pour créer des bruits accentués dans les régions les plus graves ou les plus aigües. Employé comme SFX dans *Metal Gear* (Ultra Games, 1988), le canal y est responsable des sons d'attaque du personnage-joueur, alors que dans *Super Mario Bros 3* (Nintendo, 1988), il sert à sonoriser les blocs se brisant.

L'unité d'échantillons, DMC, exécute dans la plupart des cas de courts échantillons et partage des fonctions semblables à celles du canal de bruits²⁴. Il était fort commun que ce canal exécute des échantillons de SFX (*Legend of Zelda*, sons de coups d'épée), ainsi que de percussions, comme dans *Metal Gear* (Ultra Games, 1988) où il joue la grosse caisse. En ce qui concerne la percussion, ces deux canaux travaillent généralement

²⁴ L'unique exception trouvée (également relevée par Collins [2008, p. 26]), est le jeu *Journey of Silius* (Sunsoft, 1990), où le canal d'échantillons exécute les samples de basse et le canal d'ondes triangulaires les lignes de batterie.

ensemble. Dans Metal Gear, la grosse caisse est dans le DMC, mais la caisse claire et les cymbales sont dans le canal de bruits. Dans Teenage Mutant Ninja Turtles 3 (Konami, 1991), le DMC exécute la grosse caisse et la caisse claire, et le canal de bruits a uniquement en charge les cymbales.

Puisqu'un échantillon peut garder l'information de n'importe quel son, la Nintendo 8 bits était alors théoriquement capable d'émettre n'importe quel type de son. C'est pourquoi elle put apporter des voix aux jeux, une pratique très commune dans les jeux d'arcade via les speech chips, des composants dédiés uniquement à l'exécution d'échantillons de voix [Collins 2008, p. 15, 19]. Comme la résolution (en bits) et le taux d'échantillonnage (en Hertz) sont relativement bas, les voix possèdent une coloration typique, nettement métallisée et généralement distordue. Quelques exemples de jeux avec des échantillons de voix : Blades of Steel (Konami, 1988), The Adventures of Bayou Billy (Konami, 1989), Gauntlet II (Tengen 1990), S.C.A.T. (Natsume, 1991) et High Speed (Rare, 1991), ce dernier ayant à peu près cinq secondes d'échantillons, soit une grosse quantité de voix au regard de la pratique habituelle de l'époque.

Il est important de comprendre que le canal de bruits, autant que le DMC, n'étaient pas attribués qu'à un seul type de son ou fonction dans un jeu (tels que, en fait, n'importe quel autre canal de la Nintendo). Chacun de ces canaux pouvait exécuter, par exemple, une partie des percussions et être en même temps responsable d'un SFX. Puisque les générateurs du 2A03 ne pouvaient exécuter plus d'un son par canal, il fallait choisir à un moment donné entre la percussion et l'effet, les deux ne pouvant fonctionner en même temps. Dans des situations où les percussions, par exemple, exécutent la musique, et où, au même moment, une action du joueur déclenche un SFX, une décision devait être prise par les programmeurs et les musiciens : quel son fallait-il choisir ? Lequel est prioritaire dans l'esthétique du jeu ? De telles questions amènent à un concept de hiérarchie de sons dans les canaux audio, que nous discuterons par la suite.

4. Système de hiérarchie

En somme, le 2A03 peut exécuter cinq sons simultanés et indépendants, qui sont alors transformés en un signal analogique monophonique. En revanche, cela ne signifie pas forcément qu'un jeu ne peut avoir que cinq éléments sonores dans un niveau donné, par exemple. Il est normal que chaque canal soit responsable de plusieurs éléments sonores. Si, à un moment donné, on demande à un canal d'exécuter ses deux éléments sonores, il n'en exécutera qu'un, en respectant une hiérarchie d'importance programmée dans le jeu. Pour illustrer cette hiérarchie, on peut analyser Super Mario Bros 3.

Dans le premier niveau du jeu, nous avons pu

identifier au moins 14²⁵ éléments sonores. Il y a des canaux avec peu d'éléments, comme le canal de bruits, qui exécute les cymbales et le SFX de dérapage du joueur. En revanche le premier canal d'impulsion est responsable d'une grande quantité de sons: mélodie (musique), SFX saut du joueur²⁶, SFX coup de tête, SFX saut sur un ennemi, SFX coup de pied sur une carapace, SFX carapace ricochant, SFX joueur atteint, parmi d'autres. L'importance des éléments sonores peut être observée sur la figure 1. Les éléments dans les bas niveaux sont rendus silencieux par ceux qui sont au sommet de la pyramide, soit une hiérarchie progressive plus on monte dans le diagramme.

Il est intéressant de noter que, d'abord, le canal exécute de nombreuses tâches en même temps et qu'en plus, il possède un système de hiérarchie assez complexe. La mélodie a la moindre priorité, avec six éléments pouvant la couvrir. Son placement à la base de la pyramide, en revanche, garantit qu'il n'y ait pas besoin d'action du personnage du jeu pour qu'elle soit exécutée. Donc, elle est hors du contrôle du joueur, qui peut dans le pire des cas la couvrir, mais ne peut arrêter son exécution. On garantit ainsi une continuité musicale le long de l'expérience de jeu, comme l'observe Collins [2008, p. 12, 15]. Il s'agit d'une préoccupation typique de l'époque, vu qu'à la fin des années 1970, nombre de développeurs cherchait à atteindre cette exécution continue. A l'opposé sur la pyramide, au sommet de la hiérarchie, on trouve le signal sonore du joueur atteint par un ennemi. Le SFX en question est simplement une alerte pour le joueur (l'informant qu'il a fait une erreur dans ses mouvements), et, comme toute alerte, l'intérêt d'informer le joueur de cette situation est une priorité. Beaucoup d'autres SFX liés à des actions du monde du jeu - provoquées par le joueur même - se trouvent entre ces deux éléments sonores. En revanche, le canal de bruits, lui, n'est pas si chargé.

Un exemple très intéressant est le jeu Ninja Gaiden 3 (Tecmo, 1991). Le canal de bruits n'exécute que les cymbales de l'arrangement musicale et le SFX lorsque l'ennemi est atteint, alors que le DMC exécute les autres parties de la batterie (grosse caisse et caisse claire), et aussi un échantillon de voix, actionnée par le coup d'épée du joueur. Le SFX est dans un niveau hiérarchique supérieur à celui de la percussion dans le canal DMC, donc cette dernière est interrompue par l'assaut du joueur. Cependant, le canal de bruits est aussi affecté par l'action, dans une structure conditionnelle : si l'ennemi est atteint par l'attaque du joueur, le SFX correspondant est exécuté, sinon, les cymbales sont interrompues par un silence. Nous pensons qu'il s'agit d'une solution pour résoudre un

²⁵ Il est vrai qu'il y a beaucoup d'autres éléments à part ceux-ci, mais par raison de simplicité nous nous sommes concentrés sur 14 éléments - ce qui suffit pour exemplifier et pour illustrer le système de hiérarchie.

²⁶ Ces descriptions sont simplement illustratives, liées aux actions, aux phénomènes et aux faits du monde du jeu.

problème de masquage : un son est couvert, masqué, par un deuxième, à cause d'aspects divers de ses composantes spectrales [Menezes 2003, p. 89]. Alors, pour que le SFX d'attaque (un échantillon de voix) soit mis en avant et ait de la présence, le canal de bruits doit être muté²⁷ (Voir la figure 2). Il est intéressant de constater, à l'oreille, la cohérence du mixage de tous les canaux, alors que c'était rarement, à l'époque, une problématique pour les développeurs de la plupart des jeux, dans lesquels les SFX et la musique se couvraient l'un l'autre de façon très anarchique [Collins 2008, p. 26].

À la base de la hiérarchie il y a un élément de percussion, les cymbales, qui n'est couvert que par le bruit des blocs se brisant. Au niveau du son, il en résulte que, alors que la mélodie du canal d'impulsions est constamment interrompue, les cymbales restent constantes la plupart du temps : cela donne le rythme au joueur.

²⁷ C'est un procédé très semblable aux techniques de mixage (dans l'ingénierie audio) qui emploient une ressource connue sous le nom de side-chain, via des compresseurs de dynamique : un canal subit une perte de volume lorsqu'un signal audio, dans un autre canal, dépasse une amplitude donnée. Si l'on fait un parallèle entre ce qui se produit dans le jeu en question et une situation de mixage, le compresseur de dynamique serait du style « ducker », puisque le canal subit une perte totale de volume - il est muté.

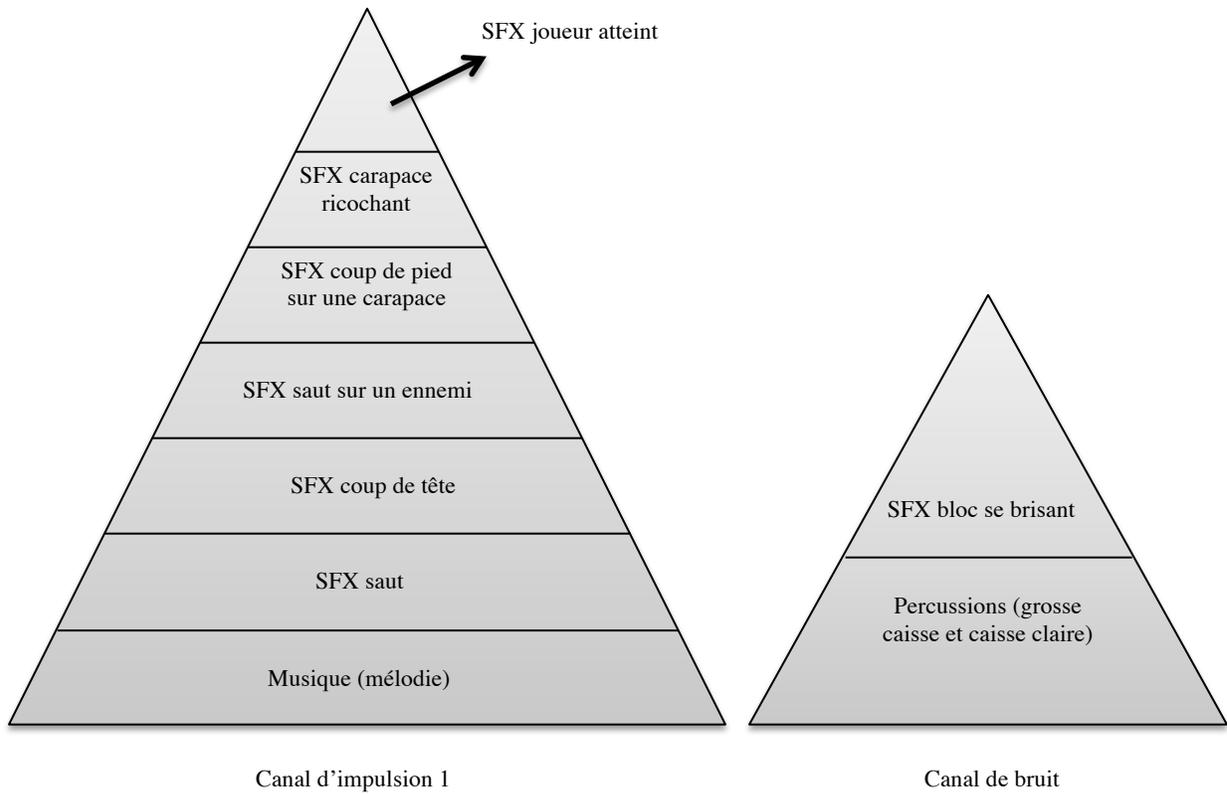


Figure 1: Hiérarchie de deux canaux audio du jeu Super Mario Bros 3.

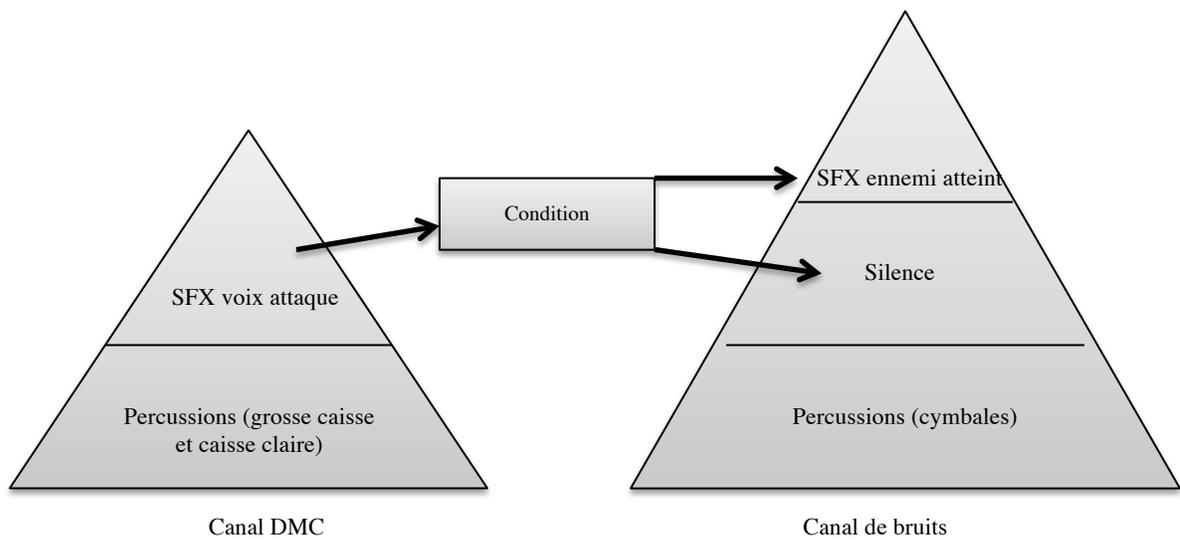


Figure 2 : Hiérarchie et interaction de deux canaux audio du jeu Ninja Gaiden 3.

Pour éviter la cacophonie entre les canaux, bien que ceux-ci soient indépendants (en tant que générateurs de son), la programmation peut créer une hiérarchie entre eux (figure 3). Dans cette perspective, l'importance des

éléments d'un canal conduit les éléments des autres canaux - ce qui crée un nouveau degré de complexité dans le système audio et affecte la sonorité finale du jeu.

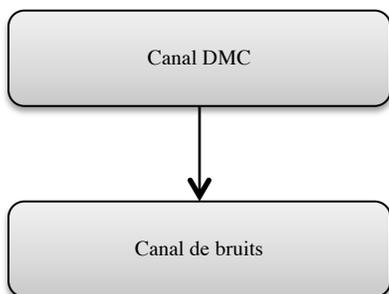


Figure 3 : Hiérarchie entre deux canaux du jeu Ninja Gaiden 3.

5. Conclusion

Le contexte dans lequel la NES parut sur le marché nord-américain montre que, en cherchant un nouvel attrait commercial pour le jeu vidéo, Nintendo trouva dans le son et la musique un champ à exploiter et à développer. Cette préoccupation, ainsi que les nouvelles capacités sonores de la console furent l'un des facteurs qui garantirent la réussite de la console.

Ce nouveau positionnement plaça, par la suite, le langage sonore comme une partie essentielle de l'expérience vidéoludique. De plus, cela contribua à l'établissement d'un nouveau langage musical propre au monde des jeux vidéo.

Un ensemble de facteurs historiques et techniques permit que la bande sonore des jeux devienne polyphonique, et dans la plupart des cas tonale ou modale²⁸. Tout cela fut rendu possible par le fait que, pour la première fois dans une console de salon, le système audio était capable d'offrir des canaux indépendants et capables de générer une gamme chromatique tempérée. Si, comme l'affirme Olazabal [1954, p. 77], Johann Sebastian Bach fut celui qui consacra le tempérament dans l'histoire de la musique occidentale (avec son ensemble d'oeuvres « le clavier bien tempéré »), on peut affirmer que ce fut la NES qui consacra le tempérament dans les consoles de salon, même si sa prédécesseur, l'Atari 2600, était déjà capable d'utiliser la gamme tempérée. Avec le tempérament, les éléments typiques de la musique occidentale, tels que la mélodie, l'accompagnement, ou la ligne de basse, pouvaient alors être exécutés en harmonie - avec des possibilités de modulations (changements de tonalité).

L'insertion d'un canal d'échantillons permit aussi à la console de salon d'exécuter des sons qui étaient

²⁸ Tous les jeux analysés pendant cette recherche, ainsi que tous ceux trouvés dans la bibliographie, sont de nature tonale ou modale. Aucun exemple de musique atonale n'a été trouvé, mais on ne rejette pas la possibilité de qu'il en existe, vu l'importance du catalogue de jeux de la console.

auparavant typiques des jeux d'arcade. Ainsi, Nintendo fit le premier pas vers le (future) rapprochement technologique entre les deux univers (le monde intérieur, domestique, domaine de la console, et le monde extérieur, domaine des jeux d'arcade). Ce canal d'échantillons permit également que les percussions de la bande sonore atteignent un plus haut niveau de complexité.

La musique et les effets sonores partageaient leur exécution entre les canaux de la console, en obligeant le musicien à pondérer les priorités d'exécution. Il est évident que cette hiérarchie n'est pas qu'une simple pratique de la programmation, mais aussi une question d'esthétique sonore. Ainsi, les limites du système sont déterminants dans son langage sonore.

Tout ceci permet de comprendre que le langage sonore développé avec la NES fut primordial dans l'établissement d'un nouveau style de jeu vidéo dans les années 1980, qu'évoquent des manifestations sonores comme la « culture 8 bits ». De telles manifestations ont tendance à se référer directement à cette époque, dont Nintendo est l'acteur principal. Encore aujourd'hui, des oeuvres sont réalisées en référence à cette esthétique, indépendamment du style de jeux ou de la plate-forme. Nous pourrions donner comme exemple les logiciels (émulateurs) qui reproduisent le son de la console, comme le Famitracker, et l'instrument virtuel mis à disposition par le groupe japonais YMCK, lui-même un bon exemple de la culture internationale du 8 bits.

Bibliographie

- COLLINS, K., 2008. *Game Sound: an introduction to the history, theory, and practice of video game music and sound design*. Cambridge, Massachusetts; London, England: MIT Press.
- DISKIN, P., 2004. Nintendo Entertainment System Documentation. <http://nesdev.parodius.com/NESDoc.pdf> [Consulté le 27 juin, 2010].
- FORTE, A. AND GARRET, M. 25 anos no cipó: Pitfall! completa um quarto de século. *Electronic Gaming Monthly*. 2008 Mai; 63(1): 106-108.
- KENT, S. L., 2001. *The ultimate history of video games: from Pong to Pokemon – the story behind the craze that touched our lives and change the world*. New York: Three Rivers Press.
- KORTH, M., 2004. Everything about NES and Famicom. <http://nocash.emubase.de/everynes.htm> [Consulté le 3 juillet, 2010].
- MENEZES, F., 2003. *A acústica musical em palavras e sons*. Cotia, São Paulo: Ateliê Editorial.
- OLAZABAL, T., 1954. *Acústica musical y Organología*. Buenos Aires: Ricordi Americana.

SEGULIA, C. AND DAI, B., 2008. FPGA Implementation of the NES Audio Processing Unit. University of Toronto.
<http://www.eecg.toronto.edu/~pc/courses/432/2008/projects/nes-audio.pdf> [Consulté le 3 juillet, 2010].

TAYLOR, B., 2004a. 2A03 technical reference.
<http://nesdev.parodius.com/2A03%20technical%20reference.txt> [Consulté le 18 juin, 2010].

TAYLOR, B., 2004b. NES APU Sound Hardware Reference.
http://www.slack.net/~ant/nes-emu/apu_ref.txt [Accessed 18 June, 2010].

TAYLOR, B., 2004c. NES Sound Channel Guide.
<http://web.textfiles.com/games/nessound.txt> [Consulté le 18 juin, 2010].

TAYLOR, B., 2004d. The Famicom Disk System technical reference.
<http://nesdev.parodius.com/FDS%20technical%20reference.txt> [Consulté le 9 août, 2010].

TOCH, E., 1985. *La Melodía*. Calabria, Espanha: Editorial Labor.