

Project Bar-B-Q 2007 Workgroup Report

Game Producer's Guide to Audio

Participants:

Scott Snyder, Edge of Reality
Chris Grigg, Beatnik
Simon Ashby, AudioKinetic
Jim Rippie, Invisible Industries
Linda Law, The Fat Man

Problem Statement:

There is a serious problem today with game audio and it is NOT production quality.

Most designers and producers do not understand the extent to which audio can be used to enhance their product quality, partially because they do not have a game audio language or style book that they can use when designing their games. This results in games that do not have audio integrated into their game design, engine design, budget, or production plan/milestones. Because such audio is only an overlay and not an integral aspect of gameplay, overall game quality suffers.

Solution:

This workgroup wanted to articulate game audio concepts and guidelines for the development process that game designers should use in early stages of product development in order to fully integrate music and sound into the creative design and project plan for the game.

Game Development Process Audio Guidelines:

Get the audio team on board as early as possible and keep them on board. This section looks at each phase of game development, starting with conception all the way through graduation from college, and recommends appropriate audio action items for each stage.

1. Conceptual design
 - a. Study audio design concepts and terminology in order to be prepared to include audio aspects in game design and project planning.
 - b. Bring in an audio artist and an audio programmer to be a part of the design team:
 - i. Especially if it's going to be an audio based game such as Guitar Hero.
 - ii. Especially if it's going to be a game for the visually impaired.
 - iii. Do it anyway.
 - c. Collect example music that you like for your design concept. This music is called a temp track.
2. Early formal design (may be concurrent with or after Project Planning)
 - a. Consult with audio artist regarding use of audio as a game design element.
 - b. Keep the game audio design concepts in mind as you flesh out the design.
 - c. Experiment with musical style, where possible. Do it early using the temp track. Iterations in musical style in subsequent project phases are expensive.

3. Project planning
 - a. Consult with audio artist and audio programmer when developing the project schedule! This is very important. Make sure there is protected audio development and implementation time.
 - b. When possible, make sure all parts of your audio design are reflected in specific, line item, audio tasks in the schedule and budget. Include localization (other language versions)!
 - c. Remember to allocate sufficient memory for the audio. If there will be memory limitations, discuss and negotiate with the audio team early in the process.
 - d. Make sure audio assets are included in manifests, bills of materials, etc.
4. Prototyping with continued design development and refinement
 - a. Put sound in as soon as possible. Use temp sounds/temp track until your game sounds arrive.
 - b. Listen and refine.
 - c. In the early stages, ignore relative volume levels. Mixing comes later!
 - d. Include game's audio director in design reviews.
 - e. Include game's audio programmer in design reviews.
 - f. Early audio reviews should be for style and feel, not specific details like "more French horn".
5. Production phase – almost all of the design is cast in silicon
 - a. Avoid audio design changes.
 - b. Avoid memory allocation reductions.
 - c. This is the time the sound department is cranking, and is constrained by the budget you set previously. Prioritize any new requests you have for them!
 - d. Finalize timing on as many cinematics as possible.
 - e. Make sure the audio team tells the QA team what to be listening for. Only the audio team knows what specific things QA needs to look for at any particular point in the development timeline.
6. From Alpha to ship
 - a. This is the time for mixing. Now you can care about relative volume levels. Louder is not better.
 - b. Silence is an option in dealing with audio problems at this stage.
 - c. Prioritize your requested audio changes. You may not get them all made prior to your ship date.
 - d. Make sure there are NO temp sounds left in your game!
 - e. Emergency measure: If you need to alleviate repetition in commonly seen parts of the game, you can repurpose music from obscure parts of the game.
 - f. Emergency measure: If necessary, use music that you have the rights to from previous games of the same style.
 - g. Emergency measure: If development of cinematics is running late in the schedule, time the cinematics to the music you already have.
7. Postmortem
 - a. Invite sound team representative(s)

Audio Design Concepts:

These are the building blocks of game audio. Your game will use some or all of the following as determined by the nature of the design, the resources available, and the audio team's plan.

1. Types of sound:
 - a. Music
 - b. Dialog (VO)
 - i. Dialog in cinematics (story relevant)
 - ii. In-game voice (can be story relevant but is not necessarily so)
 - c. Sound effects (sfx)
 - i. Hard sound effects
 1. Gameplay effects
 2. UI effects
 - ii. Backgrounds/Ambience
 - d. Voice chat (players talking to one another)
2. Output formats
 - a. Mono
 - b. Stereo
 - c. 5.1 surround

- d. 7.1 surround
 - e. Virtual speakers using 3D positioning algorithms
 - f. PC HD Audio
3. Audio data source
 - a. Hard disk streaming
 - b. Optical disk streaming (CD, DVD)
 - c. Memory
 4. Parameterized sound: a soundscape that reacts to game parameters
 - a. Scaling
 - b. Materials/Surface types
 - c. Music intensity levels (driven by things such as number of enemies/NPC's)
 - d. Music track muting
 5. 3D positioning. Sounds are associated with objects in the world and are heard by the player in appropriate 3D space. (Example: enemy sneaking up from behind. Counter-examples: danger music, narrator VO)
 6. Environmental reverb (see glossary entry)
 7. Real time DSP (e.g. Doppler, obstruction/occlusion, low pass filter)
 8. Repetition. Design to avoid annoying levels of repetition. Design to avoid annoying levels of repetition.
 9. Silence is not bad. Use it.
 10. Audio that enhances and supports gameplay (not detracting and distracting).
 - a. Sound effects tell the player what happened.
 - b. Music tells the player how they feel about what happened.
 11. Music can also be used in foreshadowing, leitmotiv, etc.
 12. Continuous sounds, such as background music, are assembled from smaller parts to achieve soundtrack goals, such as repetition avoidance and tracking game states.
 13. Transitions
 14. Mixing
 - a. Volume of all elements must be balanced.
 - b. Prioritization of sounds/music
 - i. Sounds important to gameplay must be heard.
 - ii. Soundscape shouldn't seem cluttered.

Game Audio Glossary:

audio engine: A software layer that manages audio playback via inputs such as audio files, playback parameters/variables and playback scripts.

channel: Typically refers to the available number of simultaneous sound or instrument tracks. For example, old 8-bit games often had 3-channel sound, which meant three simultaneous tonal sounds or voices could be used (they also often had a fourth noise channel). Today hundreds of simultaneous channels are available in most game consoles, although many portable game players remain very limited in the number of channels.

compression: 1. Also known as dynamic range compression or DRC, is a process whereby the dynamic range of an audio signal is reduced. Limiting is a type of compression with a higher ratio of reduction. 2. File size reduction, as in MP3. A loss of audio fidelity usually results.

cue: A named sound event that the game signals to the audio Engine. The audio engine responds with a corresponding, predefined action that was designed by the audio artist. A cue can furnish any combination of operations that the audio engine is capable of performing. Examples of cue actions are playing sounds, loading media, and setting variables

DSP: Digital Signal Processing/Processor: refers to the processing of a signal (sound) digitally, including using filters and effects.

environmental reverb (I3DL2): Audio processing that conveys a sense of the space where the listener is located. I3DL2 is a guidelines document published by the Interactive Audio Special Interest Group that defines requirements for minimum system features and functions needed for an audio renderer providing, among other things, environmental reverb capability.

sound event: A sound event is not an audio file. A sound event contains all the information needed to appropriately play back an audio file or combination of audio files. See cue.

falloff: In the real world, a given sound from a given sound source is perceived as louder when it's closer to the listener, and quieter when it's farther away. Eventually it's too far away to be heard at all. In game audio, 'falloff' is the manner in which loudness decreases with distance, usually described as a distance vs. attenuation curve. See also min/max distance.

file: A sound file. The name of the sound file is not the same as what the sound is called in the program. For example, the program may refer to something called gunshot_01 and the file might be MightyKaboom3a.wav.

format: A sound file format. Examples are .WAV, .MP3, .OGG, and .MID.

hook: A call in code (a stub) that initiates a cue/sound event.

listener: The point in the world where the sound is heard.

loop: The playback of an audio file, or series of audio files, repetitively such that when the end point is reached, playback continues immediately from the beginning until a command is issued to stop the loop. An audio file that is intended to be used in this manner is sometimes referred to as a loop. Often a looped sound will sound unnatural when it stops unless the stopping event also triggers a release sound or fades before it stops.

MIDI: MIDI is a technology that represents music in digital form, but it is not like other digital music technologies such as MP3 and CDs. A MIDI file is not a digitized sound file, it is a message file. The messages contain individual instructions for playing each individual note of each individual instrument. MIDI encodes musical functions, which includes the start of a note, its pitch, length, volume and musical attributes, such as vibrato.

min/max distance (as it applies to falloff): When a listener is within the minimum distance of a sound source, it is heard at full volume and the volume is not automatically adjusted for distance. Between the min distance and the max distance, the falloff curve is used to automatically adjust the volume for distance. When the distance between a listener and a sound source is greater than the maximum distance, it is not heard at all (fully attenuated).

mix: *verb* To combine individual sound elements in an appropriate way by controlling their volume, panning, reverb, EQ, and other effects. *noun* The resulting audio playback experience.

Nyquist Frequency: The highest frequency that can be represented in a digital signal of a specified sampling frequency. It is equal to one-half of the sampling rate. For example, audio CDs have a sampling frequency of 44100 Hz. The Nyquist frequency is therefore 22050 Hz, which is an upper bound on the highest frequency the data can unambiguously represent. To avoid aliasing, the Nyquist frequency must be strictly greater than the maximum frequency component within the signal.

occlusion: Muffling a sound because an object comes between the sound source and the listener. Amount and character of the muffling can depend on the material properties of the blocking object, and on how completely the sound is blocked (i.e. less blocked when the source is behind an edge of the blocking object, vs. more blocked when behind its center).

Redbook audio: The standard audio for CD's. It is a 16-bit, 44.1k stereo uncompressed format.

release sound: A transition sound to provide a graceful exit from a loop. Examples: A bell and a ricochet have long releases. A short piano note has a short mechanical sounding release as the hammer comes back in contact with the strings.

sample: 1. A measurement of amplitude. A sample contains the information of the amplitude value of a waveform measured over a period of time. 2. A collection of sound files and definitions that are used to make up a single virtual instrument (i.e. violin).

sample rate (also known as sample frequency): The number of times the original sound is sampled (measured) per second. A CD quality sample rate of 44.1 KHz means that 44100 samples per second were recorded. If the sample rate is too low, a distortion known as aliasing will occur, and will be audible when the sample is converted back to analogue by a digital to analogue converter. Analogue to digital converters will typically have an anti-aliasing filter which removes harmonics above the highest frequency that the sample rate can accommodate.

script: A simple text file created by the audio artist, for the purpose of controlling the behavior (including adaptive response) of a cue/sound event.

stem: A mix that does not contain a complete set of the audio elements but does have appropriate volume, reverb, pan, etc. applied to the elements that it does have. For example, a .WAV file that contains only mixed drums.

streaming: A technique for transferring data such that it can be processed as a steady and continuous stream. For example, in online applications with streaming, the client browser or plug-in can start playing the data before the entire file has been transmitted. For streaming to work, the client side receiving the data must be able to collect the data and send it as a steady stream to the application that is processing the data and converting it to sound or pictures. This means that if the streaming client receives the data more quickly than required, it needs to save the excess data in a buffer. If the data doesn't come quickly enough, however, the presentation of the data will not be smooth.

track: Depending on the media type, an audio file may have multiple, individually selectable or controllable tracks intended to play in parallel. For example, an audio file or file image may have multiple channels, each of which may be individually muted, faded, or processed with DSP.

trigger: An event that signals the beginning of a sound or series of sounds

voice: In referring to digital audio, voice is used to describe an instrument or other type of sound, rather than specifically a vocal part. A music keyboard, for instance, may be pre-programmed with 64 *voices*, or instrument sounds, which will include (typically) piano, strings, guitar voices, and so on. Someone speaking is called VO (voice over) not voice.