

## Editing to Sound

by Chris & Trish Meyer, Crish Design

### A primer on spotting hit points in audio, plus aligning video to these points.

Nothing can bring down good video like bad audio, and nothing can help lift bad video like good audio. Learning how to synchronize your video edits and effect keyframes to audio will result in a sum that is greater than its parts. In this article, we're going to give a basic primer on spotting hit points in audio tracks, plus give a few pointers on aligning video and keyframes to these points.

#### Spotting Audio

When we are animating or editing visuals to sound, the most interesting points in the audio tend to be the loudest ones: the moment a door slams, lightning cracks, a drum is hit, or a guitar is thrashed. By looking for these peaks – taller points in the audio waveform, going in either the upward or downward direction – we have a tremendous head start in finding the more interesting audio events, which we can then use as a starting point for visual edits and effect keyframes. Strong drum beats produce these peaks, as do syllables in words. Areas with no peaks or other visible waveform indicate pauses between words and sentences. We often refer to these interesting audio events as hit points.

Music has patterns in it which makes it easier to guess where useful edit points might be. Most music has a central timing pulse known as its tempo or rhythm. The elementary division of this pulse is called a

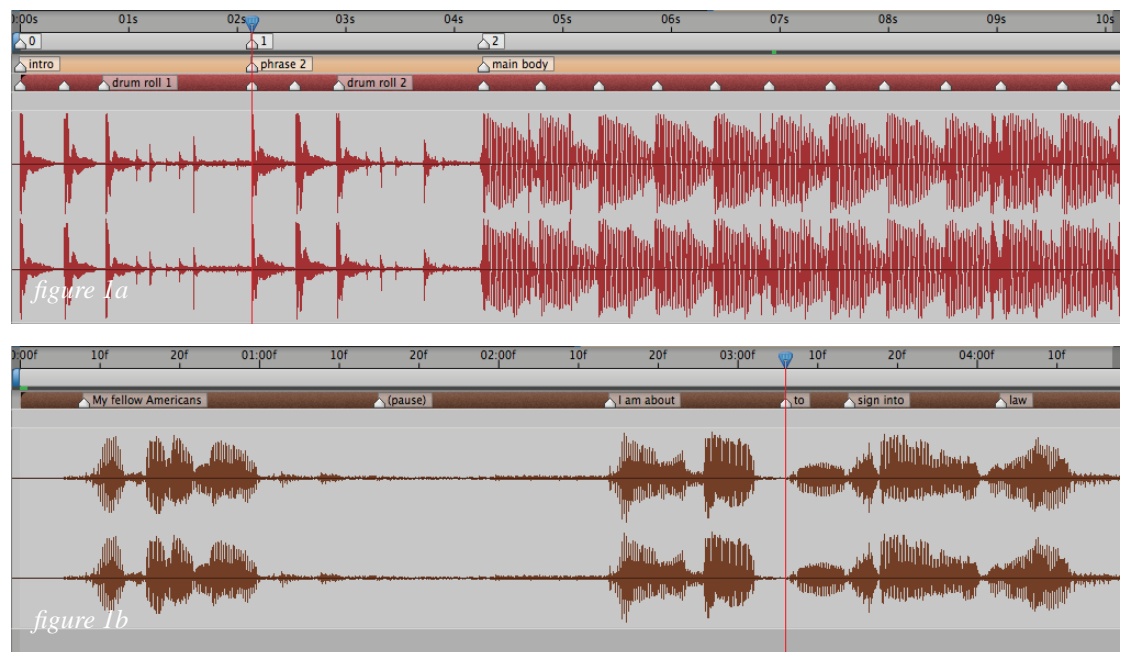
beat. When you're dancing, or just tapping your fingers or toes along to a piece of music, you're usually following these beats. Music is then structured around simple phrases of these beats, known as bars or measures. A large amount of music has four beats in a measure, although there are exceptions: For example, a waltz has three beats to a measure. The first beat in each measure – the one that you instinctively want to tap loudest on, and which often marks the beginning of a repeat of a rhythmic pattern – is referred to as the downbeat.

A basic approach to spotting music is to spot the downbeats, and mark them as potential hit points. Listen to the music and practice tapping along until you are confident you have a feel for its pulse. Counting out loud – ONE, two, three, four – often helps reinforce this, including identifying the downbeats. Then place markers on these downbeats in your editing or compositing software. Most software has a special key – such as an asterisk or the letter M – that allows you to do this in real time while previewing the music track.

When attempting to place markers in real time, chances are you will place them slightly late. The reasons for this include your personal reaction time, delays in the software recognizing your taps, or even the time it takes for sound to travel from the speakers to your ears! To correct your timing, zoom in on the audio waveform and slide these markers in time to align with the nearest spike in the waveform display which indicates where a drum hit or note landed.

It can be easy to fill up a layer with markers. Often we will create a second dummy track that runs parallel with the audio track just to hold more markers. We will place simple markers for beats on the main track, and use the extra track for additional text comments such as “vocals start here.” After you have placed markers, you can hide the waveform display (since it can take awhile to draw in some

Figure 1: When spotting music (a), we will place a marker at the start of each measure or even each beat. We often add a second layer (the orange track here) to hold additional markers and comments that describe the sections of the music, as well as master markers along the timeline (the numbered markers here) to quickly navigate between these sections. Hit points in speech are usually not as distinct (b) as the spikes in music, but are important to note nonetheless.



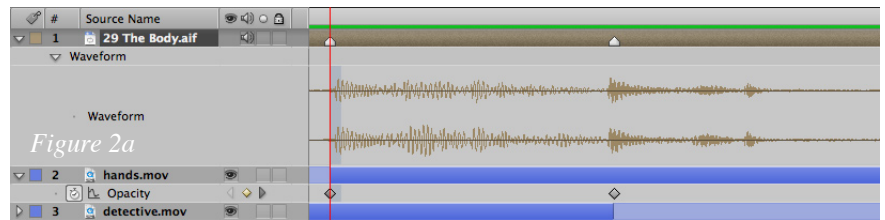
# Tips & Tricks

More available at [artbeats.com](http://artbeats.com)

**ARTBEATS**  
FOOTAGE YOU CAN USE



Figure 2: This fade up (a) begins at the start of a low-pitched drum roll (the first marker on the top track) and ends at the start of a following high-pitched piano phrase (the second marker). This makes it appear as if the drum roll is motivating the transition, and the piano is announcing the arrival of the new scene. Note that at the very first keyframe of the transition (a), you're not seeing any of the new clip yet, so it's a good idea to place this initial keyframe just before a hit point. That way, when you hear the hit, you should be seeing part of the next image. Clips: BE123 from the Artbeats Business Executives collection and VNC122 from the Incarcerated (V-Line) collection.



programs), and animate based on the position of the markers. Once the music and script layers have markers, you can copy and paste these layers into other compositions, timelines, or sequences for easy animation decisions.

## Placing the Cut

Whenever possible, we like to start our visual work with the music and look at this to decide where edit and action points in our video and animations will be. At the most simplistic, we will look at the downbeats of the music as our first cuts at where edits should go. If we are building motion graphics animations such as opening titles, we look at the finer resolution of the beats inside the measures of music as first choices to place keyframes for times when flying logos hit their mark, text pops on, elements fall into place, etc.

Beats should not be thought of as just points in time: You can think of the length of a beat as a good duration for events that happen over time. For example, the length of one beat is a good time over which to stretch a transition. One or two beats, or even a whole measure, are good lengths to fade or zoom up or down elements such as text, and also to stretch special effects such as particle system outbursts or lighting sweeps. Typically, these transitions or effects will start or end on a downbeat; occasionally, straddling the downbeat also works. The main point is to think of downbeats as your emotional or dynamic pivot points.

Whether your events start or end with the downbeat depends on the psychological interaction you are trying to evoke between the music and the visuals: Is the music motivating the visuals, or are the visuals motivating the music? For example, if there is a build-up in the energy of the resolving in a big cymbal crash, you might start a transition or effect on that crash, and make it last the duration of the crash. This

will make it seem that the music caused the visuals to change. On the other hand, if you are approaching a change in the music – such as the start of the guitar solo – you might have the transition, move, or effect end when this new musical section starts. This will make it seem the visuals built up the tension, and caused the music to change in reaction. Either way, the interaction between the music and video will increase their overall impact.

Going back to the micro level, tension games can also be played with exactly when you “hit” the individual beats. Sound events happen more quickly and with finer resolution than video or film frames. As a result, you might find an audio peak occurs somewhere between two frames. Which side should you fall on? If you hit before the peak, you will increase the tension: It will seem the video is impatiently rushing the audio. If you hit after the peak, it will relax the tension, because it will seem the video is reacting to the audio. While nudging keyframes, remember that a keyframe often indicates the point in time just before an event begins: For example, a keyframe set to 0% opacity means you won't start to see the effects of a fade up until the next frame, when opacity is greater than 0%.

In general, don't be afraid to run lots of test previews, trying out tweaks in your timing to see how the end result feels.

## Zooming Out

Of course, the other part of this magic formula is spotting interesting hit points (visual cues) in your video to line up with your audio hit points. We assume you're already familiar with this practice, as it is something you do every time you edit video.

However, having to live in the dual world of hi-def widescreen and standard-def 4:3 screen aspect ratios has brought a new challenge when timing audio to video. Quite often, the left and right sides of the widescreen version are just chopped off to create the 4:3 version – this technique is known as center cut. The implication is, if action is moving left or right, the time they pop on or off screen is different for the widescreen and 4:3 versions.

