



The Transom Review
Volume X/Issue Y

Ben Shapiro

(Edited by Sydney Lewis)

- [Intro from Jay Allison](#)
- [Ben Shapiro's Manifesto - Joining the A/V Club: Storytelling with Image and Sound](#)
- [About Ben Shapiro](#)
- [Ben Shapiro in Talk](#)



Intro from Jay Allison

These days, newspaper reporters carry microphones and radio producers carry cameras. Everybodys scrambling to get on the multimedia bandwagon. We can debate whether this convergence is value-added or watered-down, but its certainly a trend. Radio producers, accustomed to rendering stories in time and having some mastery of technical tools, are well-positioned to move on this turf. But what do images add or take away? And how do you get the damn software to work anyway?

Radio and video maker Ben Shapiro offers a thoughtful guide, complete with plenty of examples and tips in our new Transom Manifesto. In the coming weeks, well feature some sound-driven stories with images to help us ponder this new media era in which radio is once again about to become obsolete.

Joining the A/V Club: Storytelling with Image and Sound

Film Storytelling vs. Radio Storytelling...an Example

Here are two versions of the same piece that I produced some years back, in collaboration with Joe Richman, called *The Starting Five*. We did all the interviews on camera, cut the video, then pulled the audio off the master videotapes for the radio piece. Notice that the story is told with a series of chapters or scenes, only in the radio piece scenes are set-up by the narrator and in the film version the images tell the tale without narration.

[Listen to the radio](#) first, then [go to the videotape](#).



I kind of love that first shot of the video piece where the old players are awkwardly standing on the court in their tuxedos. This is an example, I think, of how images are useful to tell story. This shot helps set up a theme of the piece: these players come from a time and world so removed that they're uncomfortable even visiting the NBA court of today, let alone playing on it. We "read" their slightly hunched posture, their discomfort, the speed at which they are honored and then dispatched. The shot that follows, of the guy walking through the parking lot to the restaurant, returns them to their own world, that of the retired basketball players. They return to their context, the past and their memories. Without any voice-over we've already managed to set up the prolonged flashback that follows. Their Florida retirement clothes, their gestures and relationship to their environment these are described by the image, and they are just as much a part of the story, its themes and structure, as the simple facts of the situation.

What are facts after all? Documentaries are about how characters, with their own inner worlds, live through history. The "facts" are merely a frame which allows us to explore these issues of theme and character. Even if we are producing a "news" piece, we cannot ignore that as journalists we deal with human beings in particular historical circumstances. The expressive possibilities of sound and image allow us to tell that story.

Films Are Complicated, Get Used to It

An editor I know says "knowing how to use Final Cut or Avid and saying you can edit is like knowing Word and thinking you're a novelist". Harsh, but point taken. Crafting films is an art unto itself. Know that if you are coming from other fields, radio or print, many of your skills will be useful, but many formerly useful habits may just lead you down blind alleys. People spend years learning how to edit well, and even then spend months and months cutting a film. There are rules of thumb and techniques to cutting, but each film is a unique process to find specific solutions.

Creating a film or slideshow is like writing, with a syntax of sound and image instead of words. Like a well-crafted sentence, each image has a particular purpose and place in a sequence of images. Like sound cuts in a radio piece, each image should follow logically from the previous shot and set up the next. The whole thing is multi-dimensional, its complexity and possibilities make editing both exciting and daunting.

It doesn't have to be this way. Television documentary pieces often are cut in standard, repeatable structures. Pieces are pre-scripted, and images are shot and cut based on the script. But these aren't the kinds of films that interest me. Films have much more depth and impact if they are approached more openly and find their shape in the edit. Even "hard-hitting" current-affairs documentaries can be poetic and musical--look at "Darwin's Nightmare" or "Iraq in Fragments." I don't see why slideshows should be any different. All the usual filmmaking challenges apply to still image sequences: choosing the right images, finding a clear and economical way to tell the story, etc.

For some filmmakers, each film is a journey to answer a question or come to some understanding. As one filmmaker (I cannot recall who) said "I don't know how the film or even the story will turn out. Why make a film at all if I already know the answer?"

Building Meaning by Combining Images

We media-makers are in the business of providing the key bits of information that allow the audience to create the story. Here's one approach I've suggested to students about how to think about delivering the key details of a scene to the audience. When you enter a room for the first time, how do you experience the new surroundings? You take a

moment to orient yourself by looking around, registering the pertinent details of a place and a situation. We as filmmakers make that choice for the viewer, since we sequence what they will see and hear. If we pick the right details, and present them compellingly and clearly with the fewest extraneous bits, then the audience will do the work of piecing together a coherent world of situations, characters, events. They will literally "make sense" of the story fragments we provide. In radio we do this with sounds and voice, in films or slideshows we add the image.

When we juxtapose two images, a third meaning is created. In his famous experiment, Russian filmmaker Kuleshov took a single image of an expressionless face and cut it against shots of a crying baby, the desert, a plate of food. When he showed these to audiences they described the face as having expressions of, respectively, concern, thirst and hunger. The syntax of films or slideshows involves the dynamic interactions between images.

Film stories are told in sequences of images. One might say that the most useful images for audio-visual storytelling are those that have both specificity and forward momentum. Like a good piece of writing, a series of strong images lead to the image that follows, like a bridge built of many pieces, one leading into the other.

Trust Your Images, and Know Them Well

Trust your images to deliver the story. Images are a wonderfully strong and economical storytelling tool. A sequence of a few shots can immediately set mood, introduce characters and a setting. But they have to be the right images.

Be deadly honest with yourself about every image in your piece. Here's an example of what not to do. Years ago (all my examples of failures begin this way), I was hired to create a promotional video for a company that produced high-end artsy corporate events. I shot documentary footage of an awards dinner, celebs and donors in a carefully lit theater. In the middle of the piece, I put a shot of a steam table with food, fancy stuff, with chefs and tuxedoed waiters hovering around. But the client objected, and rightfully so. Sure it was a shot of food at a fancy event, but when you looked at, what you literally saw was dark unspecified food mass in a plain steel steam tray.

So when you're building a piece look, *really look* at what is exactly in the frame--what will the viewer literally see? Is that an image of an angry mob, or of a meandering group of bystanders? Don't talk yourself into hoping the image tells something that it doesn't. Each image should express itself unambiguously, and if you're not sure, the audience won't be either.



photo by Gregory Crewdson

Privilege the Images over Words

...Or, put another way, if words and image are giving the same information, dump the words.

So what is the place of narration or voice in an audio-visual piece? How do we figure out what narration there should be, and what information can be conveyed solely by images? The answer, like so much about audio-visual storytelling, is through trial and error.

Here's an example from a television documentary I made about the photographer Gregory Crewdson. I was trying to show the amazingly elaborate process he goes through to make a photograph, so on the morning of a shoot I did an interview where I had him list the events of the day, everything from testing the gas jets used to simulate a house burning, to working with the actors. Later in the cutting room, I made a narrated montage of the day. I began by assembling a string of the useful parts of Crewdson's interview. Then I laid on top of this audio a sequence of shots of the events that he was describing, adjusting timing here and there as necessary to make it all fit. Pretty literal "explaining" filmmaking, but I wanted viewers to know the process step by step.

When I showed the cut to a friend, the problem was obvious: too much talk. It was frustrating to watch the images with a voice chatting continuously, describing actions that we could see perfectly well on the screen. In fact, the voice-over made it hard to focus on what we were seeing—one could listen or watch, but not both. To fix matters, I radically trimmed back the voice, leaving only the bare essentials that I felt were needed to guide the viewer through the images, and to maintain the feeling that Crewdson was narrating the sequence. Then, I adjusted the timing of the shots.

In a nutshell, I first used the audio to get him to say all I wanted, but when I added the images I found I didn't need all the words because the images said it better.

The original rough cut of the sequence is lost to history, but here is [the final version](#).

Timing of Voice and Image



Another thing to note about this bit of film is the timing between voice and image. Notice that when a shot hits, especially in a new setting, there is a beat before the voice starts. This gives the viewer a chance to register the image before having to pay attention to the voiceover. Again, like a person entering a room, the viewer needs a moment to orient themselves—otherwise they *literally won't hear* what is being said to them. Often a relatively small adjustment of timing will fix the problem,

sliding the voice or picture one way or another.

Here is **another excerpt** from near the end of that same film.

One thing I like about that bit is that it suggests that for Crewdson, the moment of creating an photograph takes on some of the quality of mystery and fascination that he tries to imbue in the image itself.

It's More than Radio with Pictures

The strongest moments of filmic or slideshow storytelling integrate the audio and visual elements to create a larger point. Rather than simply adding-on information, the sound and image work synergistically.

The filmmaker Chris Marker makes essay films that utilize a constant interplay between images and a voice track that often comments on the images. His films are poetic, stuffed with ideas, brilliantly edited. Even at their most dense, from a viewing standpoint they remain comprehensible and clear. His film *La Jetee* is made up entirely of still images. Below is a link to the beginning of his film *Sans Soleil*. Watch the rhythm of the voiceover, where it starts and stops in relation to the cut points of the image, and also how the phrasing of spoken ideas works alongside sequences of shots, like that of the people sleeping on the ferry.

The very first images and sentences are a remarkable example of simple but powerful combining of sound and image. Here is the sequence in script form:

BLACK

Narration: The first image he told me about was of three children on a road in Iceland, in 1965.

Image of three kids walking on hill, backlit by sun.

Cut to BLACK

Narration: He said that for him it was the image of happiness and also that he had tried several times to link it to other images...

CUT to Aircraft Carrier and Plane

Narration: ...but it never worked. He wrote me...

Cut to BLACK

Narration: ...one day I'll have to put it all alone at the beginning of a film with a

long piece of black leader; if they don't see happiness in the picture, at least they'll see the black.

This sequence does several things. As we know, beginnings are very important, and this one immediately draws us into the film. It also sets up themes which play out in the rest of the piece: memory and the past, how images relate to our experience of time and space. It also establishes the form of the narration, which consists entirely of letters written to the narrator by an unknown filmmaker. Two shots and three sentences—it's a model for economic audio/visual storytelling that works because it makes creative use of the relationships between voice, image, and the rhythms between the two.

Screening is a Must

Maybe because films are inherently complicated and hard to predict in their effects, screenings are a key part of the filmmaking process. It's like playing a radio piece for others: you suddenly see it through their eyes, the parts that work well run smoothly, the parts that are awkward or just wrong make you sweat. Something you thought vaguely amusing gets a laugh. When it works, it seems a bit miraculous.

After you screen, always ask your test viewers some questions. Did specific key story points come across clearly? How did people feel about the situations and the characters? Were they ever bored or confused? You may get a bunch of different answers from different viewers, but if everyone tells you the same thing is a problem, they're right.

An editor I know has this rule about screenings: "pay attention to peoples' problems with a cut, but not to their ideas of how to fix them." In other words, people can tell you what isn't working, but it's mostly up to you to find out why and fix it.

Why Make Slideshows?

I kind of wonder, as more and more print and radio folks pursue web-movies and slideshows, is this about healthy cross-media experimentation, or fear of getting left behind? What exactly is so creepy about the words "value added"? Are radio people in 2007 really concerned they'll be out of the loop if they don't jump on this television bandwagon? Didn't that ship sail about 50 years ago? We as producers need to choose where to best spend our limited energies and resources.

But it is true, new production and distribution models are at hand. Maybe *filmmaking* is changing, so that it might have more of those qualities that attracted many of us to radio in the first place: widespread accessibility of cheap tools, a potential audience that is diverse geographically and in many other ways, a place where historically marginalized voices can find expression, a site for individual creative flexibility.

These attributes make radio and audio-visual storytelling rich with possibility--they don't ensure success. The work we admire grows from individuals' abilities and willingness to refine their craft, and develop their own storytelling language.

Tools and A Few Slideshow Notes

Tools

The main software tools for making audio-visual pieces fall into two camps. First are the

video editing applications like AVID Xpress, Final Cut, or iMovie. Second are the more web related applications like Flash, and those designed especially for slideshows like SoundSlides.

Final Cut is remarkably flexible. Want an image to rotate while it pans to three different points on the image while it shifts colors? No problem! You can build pretty much any effect you can imagine by combining multiple effects and specifying each of their actions over time. I find AVID applications more user-friendly, and their media management more reliable, though they lack some nifty editing features of Final Cut.

iMovie may be good enough for most simple "slideshow" type pieces. It lacks the flexibility of AVID and Final Cut but it's a good starting-off tool.

For any of these video applications, you'll need to export the movie into a web-streamable file. Video compression for web streaming is a whole topic unto itself, with different codecs, sizes, and data rates. On a Mac this process is relatively simple, you can export a piece via QuickTime and use one of the "streaming for web" settings. The result is a movie that is compressed and small enough to stream. Try a setting and then view the exported file.

One thing to keep in mind when compressing video for streaming, fast moves and pans tend to make the image go soft, jitter, and just generally fall apart.

A Few Slideshow Notes

Don't underestimate the sheer number and quality of images you will need. Think 2-5 seconds per (count it, five seconds is a long time), maybe 15-25 images/minute--and those are the selected images that you choose to best tell the story or illustrate your point. So pre-production planning to gather the right images, and enough of them, is essential.

Moves such as pans, zooms, tilts, are a valuable editing tool for stills sequences or slideshows. Moves make images feel less static and allow you to hold them longer. Moves also have a beginning, middle, and end, so cut points have to be adjusted in response to that "action". They allow you to "direct" the still, zooming in on a specific detail or figure, or revealing information progressively by panning across to another element of a photo. They can also easily be overused, too many and the piece gets "swoopy". Use a variety of directions and rhythms of zooms and pans.

In general moves should be, like anything else in your piece, motivated by the needs of your story.

Zak - Jun 21, 2007 - #5

I think Alan Berliner is amazing and brilliant for a number of reasons. Because his movies are autobiographical, and because he explores big ideas (identity, family, sleep) through the metaphorical lens of his own complex mind, the form of his movies take on an endearing complexity as well. His stuff grabs me because there's good reason for his films' elaborateness. It's not heavy handed just for the sake of being heavy handed. The form and content in his films seem happily married.

Elizabeth White - Jun 24, 2007 - #6

Hello Ben Shapiro,

I really appreciate your comment about allowing a piece's form to emerge as you tinker - trial and error - with the material.

Your point that "films have much more depth and impact if they are approached more openly and find their shape in the edit" seems applicable to all kinds of creative production.

Given that non-perspective attitude, I found your notes about the number of images slideshows require a bit surprising. I've recently fallen in love with the medium, and part of what I like so much about it is the way it can encourage a deeper, more sustained kind of looking.

For example, I was in New Orleans in 2006 with my microphone. I teamed up with a portrait photographer working with a 4X5 camera. We made daily trips to a FEMA trailer park. Diederik photographed people and I talked to them.

The slideshow ended up being a series of 1 - 3 minute monologues, each set to just one photo: the portrait of the speaker. 7 photos in about 14 minutes.

I believe - hope - that the effect of this is that the viewer looks at the portrait for much longer than he might if he were looking at the series in a magazine ... and that the photographs are good enough that they reward sustained attention. I hope the viewer might say things like: 'hearing the voice deepened my initial reading of the subject's expression' or 'i didn't even really notice the fence in the background until the subject talked about being caged in'.

Occasionally I'd intersperse a text slide with the words the person was saying, so you'd read and listen at the same time. I found this technique allowed me to ensure the viewer's attention was on the audio when a really great line came along. Like: "And we get plenty to eat, could be more nutritious and all. We get a lot of pork and cheese. You don't need all that. Too much swine - pork, pork." People laughed at that, which was a nice relief.

Anyway, I guess my point is that I recognize there will certainly be times when a slideshow is best served with a quick succession of images. But I also think audio-slideshows have an amazing potential to use sound to stay the viewer's attention - to create time to ponder, notice things, and let meaning unfold.

Ben Shapiro - Jun 25, 2007 - #7

Elizabeth, thanks for the remarks. It sounds like you found an interesting solution for your piece, the longer single images, and interspersed text. I like the idea of playing with duration and attention that way. Is it viewable on-line? Maybe you could post a link for us to see.

My guesstimate of 2-6 secs/image is meant to suggest the surprisingly high number of images one typically might need to build a piece, and the value of thinking strategically. It can be useful to have a variety of images so that if you find a certain approach works, you have a few images to play with in the edit. Sort of like colors on a palette—you won't use them all but it's useful to have a variety at the ready to explore options. I always learn a lot from that.

About Alan Berliner's films...I share your appreciation, Zak. One thing I admire about them is that they are experimental while being totally engrossing and fun to watch. He

understands how to build a story. And he writes really well.

In the context of our topic of slideshow/films and radio, it's worth noting how inventively Berliner uses archival sound and images. He never uses archival material merely as straightforward illustration of an historic era or event. That's part of it, but the archival is also there for us to mine for clues as to characters' memory and inner experience, or as chapter breaks or links between topics, or to make jokes.

There's a theory that the first thing a film or radio piece has to do is explain to the viewer/listener how it's going to work. So at the beginning of "Nobody's Business" Berliner and his father engaging in a combative conversation (mostly about whether he'll get his father to tell his story), and we see archival footage of a boxing match and hear a bell ring. It's both funny and resonant. Immediately we get it: this film is going to use archival in this particular way.

I wonder if archival could be used like that in radio? Some pieces use archival tape to set mood or establish a character's inner state, or for sonic punctuation, or for jokes. In the Sonic Memorial Special a recurring music bed was actually a live concert recorded at the World Trade Center--identified as such once during the hour-long piece--it works through of its mood and rhythm as much as because its "archival".

But could a radio piece use archival as Berliner does when, say, he cuts to a crowd of news photographers from the 1960s when a character is talking about their imagined success?

Joe Richman - Jun 27, 2007 - #10

Ben - this is a great manifesto.

As I was reading, I was thinking about the value of getting feedback on your work from friends or people you trust. You mentioned screenings, but you also told the stories of two old pieces that changed after you showed them to friends.

I think the scariest thing is to be stuck in the feedback loop of your own brain (in my case, it seems to happen a lot.) Getting another set of ears/eyes/brain cells to react to your work is just so damn important. It even makes you hear your story differently just knowing someone else is listening.

I mention this because I wanted to give you a shout out. Everyone should find a friend who can act as an informal editor. For every story I do, Ben Shapiro is that guy.

I'm looking forward to reading more.

Jackson - Jul 5, 2007 - #11

Ben -- I did as you asked back in #1. I listened to the radio version of the Starting Five and then watched the video. Both are terrific, but they are not telling exactly the same story to my mind. The video seems more nuanced -- when the story turns to the anti-Semitism segment in the radio version, it feels a touch hefty, not exactly heavy-handed, but there is nothing left for the listener to connect in his or her own brain.

The same segment in the video -- it feels somehow lighter. True, it feels like there's more music in the video (or more different music), but there's also fewer words -- and I'm not even taking into account the whole narration thing.

As a radio guy, I found it disconcerting to encounter the faces after hearing the voices. I

don't mean that in a bad way, and it may have been a result in part of having listened to the story before watching it. But I found my attention turning more to the image than to the word.

It calls to mind the early Puritans back in England, stripping out all the imagery from the old Catholic and Anglo-Catholic churches -- the statuary, the pictorial windows, the icons.

John Milton in his unusual stage piece *Comus* tells his audience never to trust what they *see* -- the Papist fondness for imagery merely feeds superstition, he suggests, while the Puritans are concerned with the pursuit of "right reason".

In other words, only believe in the word, what you can actually hear (or read).
Appearances can be deceptive.

Do you feel the differences between the video and the audio of the starting five?

Ben Shapiro - Jul 8, 2007 - #12

I think its possible that the Starting Five video piece may seem more nuanced because we figured it out as a video piece in the cutting room first. Of course that was back in the day when Joe narrated his pieces, and I could see this working well as a narrationless piece, full of recollection. It's funny that you bring up the anti-semitism angle because there was discussion about how much weight to give it in the tv story. At one point we were going to drop the line where the guy said: "at that point I didn't want to play any more". If anything, we felt the radio piece gave less emphasis to that end of things. That section speeds by pretty quickly in both versions, maybe the radio narration gives it more weight.

If you're interested in the religious/cultural issues attached to sound and "the word", see Jonathan Sterne's "The Audible Past: Cultural Origins of Sound Reproduction". Among many other topics, he relates how sound reproduction was regarded by Victorians as death-defying, because your voice could speak after you were gone—hence the original Nipper the dog "his master's voice" RCA logo. Supposedly in the original painting of the image, Nipper was sitting on his master's coffin while listening, perplexed, to his voice on the gramophone!

I just watched the FINAL SALE, posted as an accompaniment to our discussion, and it is full of engaging images and tape. There's a nice rhythm to the relationship between the images and the sound, for example the image of cart and the squeaky wheels—not too literal but clearly connected. It established a pattern with the sound preceding the image, so we hear the sound, wonder what it is, and then the answer is revealed/clarified in the image. The pieces of black keep our eyes fresh to new images, and also to return the emphasis to the sound. I might have liked a bit where there was just ambient sound and images but no specific talk. Periods of talk/non-talk could be another variable to play with.

Watching this piece also brings to mind the issue of context and presentation. With this piece, I sat down and focused on it in a way that I probably wouldn't if I was listening to the radio, in part because I deliberately requested the piece. I wonder if there could/should be a col

Page printed from:

http://www.transom.org/guests/review/200706_ben_shapiro/