

## Episode 28: Transcription // Translation

*Run time: 30:44*

AH: Allison Hitt (co-executive producer, host)  
JR: Jana Rosinski (producer, host)  
SZ: Sean Zdenek (special guest, Texas Tech University)  
BB: Brenda Brueggemann (special guest, University of Louisville)  
SH: Steven Hammer (special guest, Saint Joseph's University)  
CV: Crystal VanKooten (special guest, Oakland University)  
BK: Ben Kuebrich (co-executive producer)

*Cue music: "Crunk in the Trunk" by Podington Bear*

AH: You're listening to *This Rhetorical Life*, a podcast dedicated to the practice, pedagogy, and public circulation of rhetoric in our lives.

JR: Now it's recording?

AH: Yeah cuz now it's counting. Um...[*digital reverb*] Oh, no, it's doing something weird. That ch-ch-ch-

JR: Test.

AH: Okay now it's gone.

JR: Test, test.

AH: That sounds fine.

JR: Does the sound sound balanced for both of us?

AH: Um...

JR: Greetings, listeners. This is Jana—

AH: —and I'm Allison.

For our 25th episode, we talked with other academic podcasters about the process and practice of podcasting and about sound production more generally. So for this episode, we were interested in thinking again about what it means to produce sound, to manipulate it through editing, to use it to craft logical and ethical and emotional arguments, to translate that meaning into words through transcribing, captioning, or asking students to critically reflect on their rhetorical aural choices.

JR: We asked, what does sound look like? How do you bring in the felt, embodied nature of verbal and non-verbal communication into written text? What is gained and lost when you translate across modes? What does it mean for transcribing and captioning to be accessible and rhetorical practices?

We recorded interviews with Sean Zdenek, Brenda Brueggemann, Steven Hammer, and Crystal VanKooten at this year's Conference on College Composition and Communication. In some clips, you may hear wind on the water, academic conversations echoing up from the lobby of the convention center, people clapping at the end of presentations. How do those ambient sounds complement or interfere with the content of the interviews? What are the implications of removing all the background sound from these clips, to make the audio as "polished" as possible? How do non-speech sounds contribute to how we make meaning?

The interviews featured in this episode begin to answer some of these questions. Sean and Brenda talk about making meaning with transcription and captioning specifically, and Steven and Crystal talk more broadly about making and manipulating meaning through sound.

*Cue transition music: "Human Transition" by Podington Bear*

AH: Sean Zdenek is Associate Professor of Technical Communication and Rhetoric at Texas Tech University and writes about the rhetoric of closed captioning on his website, [accessiblerhetoric.com](http://seanzdenek.com) (<http://seanzdenek.com>). His upcoming book, *Reading Sounds: Closed-Captioned Media and Popular Culture*, analyzes how closed captions make meaning, which you can check out at <http://readingsounds.net>.

We were interested in talking to Sean about what it means for captioning to be both an accessible and rhetorical practice and what choices a captioner must make when deciding what content gets captioned or left out. Sean offered insights into the difficulties of negotiating meaning with non-speech sounds, how English closed captioning sometimes involves translation across languages, and how captioning choices should follow traditional rhetorical considerations of purpose and audience.

*Cue transition music: "Human Transition" by Podington Bear*

*Cue music: "Alien Language" by Podington Bear*

SZ: Well you know I was hoping that there would be some background sounds—ambient sounds—which are really hard to caption. In my talk yesterday, I made the point, which I think is really easy to demonstrate, that meaning only becomes anchored or secured in context. So this sound here that I can see...I mean, I can see the guy spraying the deck on the pier because we're right here at the water. But when I get home, if I don't remember that, I won't know what that is. It might sound more like steam or something. I know we're on the water, so maybe I'll mistake it for a boat engine or something. Just air or

wind. So I don't even know if I can caption...or transcribe...I don't know if I can even transcribe the non-speech sounds coming back home not being able to see what they are.

So I'm really interested in non-speech. I'm drawing on like Armory Schaffery has a book, *The Soundscape*, from the late 1970s in which he says things like, "The same sound has two different meaning. Context matters." And I think we can apply this directly to closed captioning. We tend to assume that captioning is objective. It's just copying down. We tend to privilege speech sounds, and there's just something about speech that sort of makes it seem easier to transcribe. It's straightforward and objective, but it's so much more complex than that—especially when you add in non-speech sounds, especially when you consider that everybody has a different way of speaking. But when you're captioning and there are reading speed guidelines that need to be taken into consideration, when do...when do you note someone's manner of speaking in the captions? There are a whole bunch of considerations I think that need to be made when you're trying to...when you're trying to channel or represent the intentions of the content creator.

*Cue transition music: "Human Transition" by Podington Bear*

SZ: We tend to think of English closed captioning as requiring a really good working knowledge of English, but closed captioning—even English closed captioning—does involve some understanding that there are other languages, and the captioner may need to hunt down some information about those languages. I've chatted with Stephanie Kerschbaum about this a little bit because she's collected some examples, which I cite in my book, and I have some other examples of really simple words in other languages. Non-English. Like "hola."

I've got an example from...I can't think of the name. It's the movie with Brad Pitt—the baseball movie with Brad Pitt and Jonah Hill? Brad Pitt in that movie does not know Spanish, but they bring on a Spanish-speaking baseball player, and the baseball player says, "Hola, señor." Something like that. And Brad Pitt says "uh uh hola." But it's captioned as, "Both speak Spanish," something like that. So you get the impression that Brad Pitt in this movie speaks Spanish, but we're supposed to understand that he doesn't speak Spanish. They want to bring in this guy who doesn't speak English but he's great at baseball, and Brad Pitt kind of mumbles through "hola." So all the captioner needed to do was "hola" or maybe even "h- h- hola" in order to convey that meaning. This happens all the time where they just...they just sort of summarize or describe the other language without really trying to convey that meaning by transcribing those other, non-English words.

*Cue transition music: "Human Transition" by Podington Bear*

SZ: This is why a rhetorical perspective here is so valuable. I mean, image description is also driven by the purpose of the image in a particular piece, just as captioning is, just as transcribing is. And you could go into great detail describing my California accent or whatever or your accent or the pitch or the volume or the way that we're pronouncing

certain words or my...my stammering here or repetition. But I don't know if that's...it really depends on the purpose. I cite an article in my book by Ronald Macauley, a linguist, who I don't know if he uses the term rhetoric, but he's really offering a rhetorical view of transcription because a linguist could go into incredible detail, but you know their transcription in their article needs to be driven by the purpose of the piece, what they're...what they're interested in. Same with captioning.

There are some sounds that probably *should not be captioned*. And this is kind of a controversial point because, well, I've been accused of being kind of paternalistic. Like I get to decide what access that people who can't hear have. But it's really a matter of the kind of access that writing provides. Just a real quick example is a dog barking in an establishing shot, so you see the outside of a suburban home, and you hear dogs and crickets. One captioner shared with me how repeatedly captioning that dog could lead readers and viewers to mistakenly assume that the dog is playing some significant role in the piece. As soon as you start going wild and over-captioning all these ambient sounds, it can lead us in a direction that might lead us astray in thinking that some of these non-speech sounds might be more important. So I think captioning, as with image description, needs to be driven by the rhetorical purpose, maybe genre too, the context, the audience's needs.

The meaning is different. So we talk about equal access, but that's just a name. Writing and sound offer different affordances, and the meaning acquired through captioning is not going to be the same.

*Cue transition music: "Human Transition" by Podington Bear*

AH: Sean and Brenda presented together at CCCC on a panel titled "Caption That! Critical, Creative, and Contextual Encounters with Closed Captioning in Multimodal Composition," so we were also interested in talking with Brenda about captioning as a rhetorical practice. Brenda Brueggemann is Director of Composition and Professor of English at the University of Louisville and has published multiple books and articles in Deaf/Disability Studies, and she spoke with us about the accessibility of captioning, the rhetorical and pedagogical affordances of captioning, and the difficulties of translating meaning across modes.

*Cue transition music: "Human Transition" by Podington Bear*

BB: The first question that you asked is what does it mean for transcribing or captioning to be an accessible practice, and I guess a lot of what I've been thinking about lately is the expansion of the practice beyond accessibility just for deaf and hard of hearing people. It's an accessible practice for a wider range of audiences who might need multiple channels. Non-native speakers. But it's also an accessible practice in that when you create a transcript and a caption with something multimedia and you load it back up, you've now created a search engine. You've created text for the web crawlers to actually find discourse.

That's how it works with TED. TED takes most of its major talks, and they transcribe them and have a transcript—an interactive transcript—so you can find and start the talk at any place by simply pushing a word in the transcript. Then you can take the transcript and do the transcript in 26 different languages right now. Captioning creates that template for that to happen.

In our field even though composition and communication get hinged together, we have always talked about speech and writing as two different things, and they are and they will be, but I think captioning has this uncanny ability to merge those two back again to turn speech into writing and vice versa.

On the other hand, I think that...I don't ever believe it's absolutely necessary that you can actually make everything completely, fully accessible. For example, talk radio or sometimes an audio documentary is just meant to be that way, and there's no way you're going to be able to transcribe or reproduce all the nuances of that voice. And I don't think that's any more necessary than describing every work of art on the planet in an audio-visual description for blind people. I mean, that's why we have art of all different kinds—or different modes.

*Cue transition music: "Human Transition" by Podington Bear*

BB: Last year at Society for Disability Studies, I delivered a talk. I had been working on a blog—an educational blog about the Nazi's Aktion T4 program. I've made two trips to Germany and studied it, and a blog and writing about different artifacts is the best way I found to do it. So I had an opening video that briefly encapsulates the mention of the Nazi's T4 program, and the video uses still images and has my voice reading the text that I wrote about it. People say it's haunting. It's both because of the material, but it's also because I read it in that way. It's 3:45. Then I created an audio-visual description track to describe the images that are coming behind my voice and reading, and that audio-visual description track took ten minutes. Most people who have tried to listen to it say they can't even bear it because the description of the images...you know, it's one thing to look at them but another thing to hear the voice. Again, I did it out of accessibility, but ironically even my good friend at Berkeley who's blind said, "That was almost more than I could take!"

I think that people don't realize that to caption something...to caption something or even to audio-visually describe it, you have to imagine in some ways who the most likely audience is for this material. It makes a difference whether you're captioning this for a 65-year-old man or a 10-year-old girl. I mean obviously—level of language. The other things, what Sean Zdenek was talking about...how would you want to caption the music and imagining who the audience might be for it? You're never going to know that entirely, but I think there's more thought about that. That it's about the subject matter—fitting the subject matter, fitting the audience addressed and audience invoked. You have to imagine both of those. And then also thinking about yourself. You're there really interpreting.

*Cue transition music: “Human Transition” by Podington Bear*

AH: Brenda’s response reinforces the argument that practices such as captioning, transcribing, and audio-visually describing are very much tied to rhetorical considerations of audience. For the rest of our time together, Brenda reflected on a question that she commonly is asked by students— “how do you hear music?”—and talked about what it means to experience sound as someone with partial hearing and how we can turn those questions into teaching moments about the many different ways we all experience and translate different sounds into different meanings.

*Cue transition music: “Human Transition” by Podington Bear*

BB: So I have some hearing. You know, I can hear voices, but discrimination gets really hard for me. And it’s so cool like on Audacity to see the sound waves because then it just fills in the rest of my imagination. It helps me catch a lot of it.

People always want to know...when they think about deafness, the first thing they think about is the lack of music. So they say, “Oh how sad, you don’t hear music.” And I say, “No! I hear music,” and they want to know “how do you hear music?” And I say, “Well how do you hear music?” Having people describe. I do an exercise with my students when they ask that. I play a piece of music and ask three or four different students to describe it back to me. And of course you don’t get the same description. Different people go after different kinds of things. That sound had a different *look* or *feel* to them. Some people will describe it really emotionally. Some people will go after it more technically—you know, real technical. Some people will describe certain instruments they heard. We map into it.

The other part of what does sound look like. I have a good friend and colleague who’s a very famous deaf storyteller, and he has a line in an essay he wrote—one of my favorite lines of all time—he says, “sound has a way of bouncing off visual cues.” When someone’s cell phone rings in class, I don’t hear the cell phone, but I know immediately a cell phone has rang because it’s registered all over their bodies, and they’re eyes are darting and they’re doing that things where [*laughs*], they’re trying to pretend. And sound does that. People say well, like a siren coming. I can tell a half-mile back because the traffic just changes. There’s this pause, there’s this shift. Yeah.

*Cue transition music: “Human Transition” by Podington Bear*

AH: While Brenda and Sean addressed transcribing and captioning—the processes of translating audio and video into text, how these are accessible and rhetorical practices, and what it means to make meaning through practices that we typically think of accommodating—Steven and Crystal address the processes of translation that occur when we produce, manipulate, and make meaning with sound.

JR: Steven Hammer is Assistant Professor of Communication and Digital Media at Saint Joseph’s University. With his work in sound production and dirty new media, we were

interested in his thoughts about how different tools, technologies, and interfaces impact how we record sound and produce new media and whether or not there is a rhetoric of cleanliness in sonic texts in the field. He also talked with us about the affordances and limitations in the production and manipulation of sound and of translating and creating meaning through different media.

*Cue transition music: "Human Transition" by Podington Bear*

SH: You know, we think about this in writing and video. We think about this with sound that tools are tools are tools—like objective things. They really, really shape of course the way that we have to gather information, present data, publish it, disperse it. All of those. I think at the core of what I'm thinking about is how our technologies and tools and then everything else like this room...which then you have to start thinking about architecture and especially modernist architecture, which was really meant to blockade the public and the private and noise and signal and everything like that. So that's where maybe the work of Rickert and ambient rhetoric comes in, right? That there are all of these factors in terms of the way that we produce. It's not, "I have an idea," and I turn it into this thing. It's all of these things from the file format to the bugs of the recorder we're using to construction sounds in the background. I think it's really just taking an approach and saying, there are all these other things that we're playing with, not on.

*Cue transition music: "Human Transition" by Podington Bear*

SH: There's a project called The Ghost in the Mp3. So we all know that when we compress any file, there are artifacts left over. This is why when we're streaming on Netflix or we're watching something and we get these moments of breakage. What we're really seeing there are moments of compression, right? Because of course in order to distribute this media very quickly, we need to compress it so much. Sometimes those ghosts of compression peak out and show us for a second that, *Oh I'm not watching a film* as in moving film. *I'm watching an ultra-compressed mp4 streamed over the web*. So when we compress something into an mp3, certain frequencies have to be excluded. Now they've of course chosen which frequencies and which parts of sound don't make it to minimize the change from a lossless audio file to an mp3. But you can also when you convert to an mp3, you can save all the stuff that's been discarded and listen to it. If you were downloading music in the early days...I'm always thinking of like Kazaa or something like that, and how shitty it sounded, like [*high-pitched hissing sound effects*]. You always knew if you were listening to something that was from a CD that you put on your computer or if it was some crappy version of a song. To me, that's really beautiful, and those are beautiful moments because they are these moments where you realize what you're really interacting with. That *this isn't a good jam*. This is a thing that was compressed: downloaded, uploaded, and played on a specific device.

*Cue transition music: "Human Transition" by Podington Bear*

SH: Everything is so polished, right? I think our culture of technology...and this is why people in glitch and dirty new media pick on Apple so much. They have that slogan for a

long time of “it just works.” If you look at their design, it’s meant to be sort of invisible, to try to remove the appearance of interface altogether. I can’t even take my battery out of my MacBook anymore. They’re really locking everything down in the spirit of beauty, simplicity, and functionality, but what they’re removing is our attention to their design and how it shapes how we consume and produce and think. It really shapes us, but it’s being increasingly obscured from us. Largely, what glitch and dirty new media folks are doing is going to those sites in culture and in technoculture that are doing the most to conceal their own production and disrupt those moments.

And these are things that I do with my students, too. So I teach digital media production, and for instance when they’re working with digital images they think of them as photographs. But they’re not. They’re 0s and 1s. So I have them open up a jpeg with a text editor, and they’re like, “Oh! When my computer, even when I double-click on it, my Mac says this should be interpreted this way. It should be interpreted with Preview in a photograph form.” But you can open it with Audacity, and you can put reverb on a picture, and it will mess it up. And it really signals to students that the way we are taught to and in a lot of ways forced to interact with different file types—like a jpeg or an mp3—there are other ways to do it. Practices like that for my students gets them to think about the jpeg as something other than a picture.

*Cue transition music: “Human Transition” by Podington Bear*

JR: Steven also shared thoughts about what this work means for the field in terms of how we teach multimedia and multimodal writing, specifically in terms of digital media production and encouraging students to reflect critically on the choices they made by translating into words what they tried to do with their aural texts.

*Cue transition music: “Human Transition” by Podington Bear*

SH: If we want to talk a little bit more about the field and in terms of how we think about and how we teach multimedia writing/multimodal writing. In my panel yesterday, I was talking about remix and my problems with remix. So remix started as a really intensely material thing. Cutting up audio was taking magnetic tape and understanding the material and the affordances and limitations of the material and then altering that material. And now so often when we think about remix, it’s an important I think economic and social movement, but in terms of actually doing it and teaching it, it’s so often like, [*boring voice*] “Drag an mp3 into this really nice interface that allows you to use the scissors to cut.” So it’s kind of like this drag and drop magic that undercuts really the origins of remix itself, so it becomes like an app that discourages us from thinking about what we’re doing and the material that we’re working with.

So much of what we do isn’t that *pure* intentionality that we like to think about and that to a large extent we *teach*. So one thing I really find valuable to try to get around that is to encourage them to talk about their work in a way that isn’t just like, “Oh it’s art, don’t you get it?” Saying, “It’s okay to make sense of this stuff after you did it.” I had some students recently...I was doing a sound production class, and a lot of them did these



really great projects, and I had them write a 500-word essay explaining their choices and their decisions and all of that. A lot of them were like, “We tried to do this. We tried to do this.” And I was like, “Actually you’re drawing from so much more, and it doesn’t matter that you didn’t know it at the time.” That’s a really awesome moment to teach! So here’s where I can teach you about that, and here’s how you can talk about your work, even though it’s after the fact. I think it’s important to open that door.

*Cue transition music: “Human Transition” by Podington Bear*

JR: At the end of his interview, Steven emphasized translating meaning from aural texts to reflective textual essays and how the struggles that students sometimes fail to articulate can become teachable moments. In the last interview, Crystal VanKooten, Assistant Professor of Rhetoric and Writing at Oakland University, echoes some of the pedagogical considerations of teaching students to make meaning with sound.

*Cue transition music: “Human Transition” by Podington Bear*

CV: Hi podcast listeners! This is Crystal VanKooten, and today I’m going to talk with you about a project that I’ve been working on that’s called “Music Goes to Work in Multimodal Video Assignments.”

JR: With her work in audio-visual composition, we were interested in hearing her thoughts about the rhetorical considerations that students make when they layer sounds and work with sound. Crystal shared with us that she became interested in the rhetoric of sound and music in 2009 in response to Cindy Selfe’s article calling compositionists to pay more attention to sound in order to open up “more persuasive and more communicative possibilities to our students.”

*Cue transition music: “Human Transition” by Podington Bear*

CV: I started really paying attention more...*trying* to pay more attention to sound in the ways that students were using it and talking about it in my classes like first-year composition and upper-level writing classes as well. What I noticed as I paid more attention was that sometimes students didn’t have the words that they needed to describe their choices with sound and with music in the classroom, and they struggled to articulate rationales for why they chose the sounds and the music that they choose. So this all led to me wanting to do a more systematic inquiry you could say into the ways that sound is rhetorical and that music can be rhetorical and the ways that students in classes like first-year composition can leverage that rhetorical power of music.

*Cue transition music: “Human Transition” by Podington Bear*

JR: Crystal created a study and interviewed students from four first-year composition courses who created video compositions for an assignment. She wanted to listen to students talk about how they were using music in their compositions and what they were using it for, as well as pay attention to the language they used to talk about music. Based on these

interviews, Crystal shared some of her findings in terms of helping students create meaning and translate across modes by collaboratively creating a sonic vocabulary and developing a culture of rhetorical listening in the classroom.

*Cue transition music: “Human Transition” by Podington Bear*

CV: Sometimes words weren't sufficient for students when they were trying to describe their musical choices, so they used a lot of vague descriptor words I guess you could say. Or if the words weren't vague, I wasn't sure what quite they meant with them. So they threw words like “tone” and “feel” and “fit” and “flow.” And these are words that students in first-year composition use to describe their prose as well, and you have to interrogate, “what do you mean by that?” But they would use “the music just flowed” or “I felt it” [or] “the music fit.” And I would have to say, “What do you mean by that? How did it fit?” And sometimes students could answer, and sometimes they struggled a little bit to answer.

So one thing I've been thinking about is the pedagogical opportunity that we have to craft what I would call a sonic vocabulary that students can use to really piece out their rhetorical, musical choices and think about new choices that they could make, potential choices that they didn't think about because now they have *tools*—linguistic tools—to describe what those moves are and what they sound like. However, I do think that such a vocabulary can and should be developed by composers and students themselves, so I don't think it's pre-packaged vocabulary that we can just sort of hand them and say, “Here, start saying dynamics and start saying tempo.” You know, maybe the terms “dynamics” and “tempo” will be useful for students, but maybe not. And maybe they want to come up with their own terms for why a fast moment or a slow moment would be more appropriate.

Sometimes words are not adequate to describe music. And I think we all know that from listening to music—that music is extra-discursive at times. It's felt in your body. It's what Steph Ceraso calls “a multimodal event” where you're feeling it physically and mentally feeling it. And it is logical and emotional and bodily. And sometimes that can't be fully described in words, and I think *that* is part of the reason where in the study students would be mimicking music with their voices to me to try to get me to understand what a particular sequence in the song would sound like. So they would go, “de-de-de-de-de” to describe a part instead of in words. Or they would bounce in their seat and move their hands to try to show the pacing of the music--whether it was fast or slow. And so these examples are evidence that the rhetoric and composition of music is extra-discursive.

*Cue music: “Old Skin” by Podington Bear*

AH: Special thanks to Sean Zdenek, Brenda Brueggemann, Steven Hammer, and Crystal VanKooten for taking the time to share ideas about making and manipulating meaning with sound, how practices like editing and captioning are rhetorical, and the affordances and limitations of translating sound into other modes. Full-text transcripts for each

episode are available at [thisrhetoricallife.syr.edu](http://thisrhetoricallife.syr.edu). Thank you all for spending time with us today--whether listening or reading along.

*Cue music: "Old Skin" by Podington Bear*

BK: *This Rhetorical Life* is brought to you by graduate students in the Composition and Cultural Rhetoric program at Syracuse University. Executive producers of *This Rhetorical Life* are Ben Kuebrich and Allison Hitt with additional production from Karriann Soto, Tamara Issak, and Jana Rosinski.

JR: [white noise]