

Transcript for Sound by AI: AI-Driven Sound in Education
2024 ETUG Spring Workshop: Digital Literacy Today – Day 2, May 10, 2024
Educational Technology Users Group
Presenter: John Born

FACILITATOR:

Our next speaker is John Born, who is the manager of educational media at Simon Fraser University. And John is gonna talk about Sound By AI: AI-Driven Sound in Education. So take it away John.

JOHN BORN:

Thank you so much. Thank you. It's important for me to acknowledge being a guest today on the unceded traditional Coast Salish lands of Squamish Tsleil-Waututh and Musqueam Nations. It's important for me to do that. So welcome. It's great to be here at ETUG and be a participant and be a presenter. It's exciting to me that it's now been 30 years of ETUG. I was trying to think about my history of ETUG and I presented in 2013 on sound and then 2014 with Ken Jeffrey on transmedia and then again in 2017, again on transmedia on some of the stuff that we were doing at SFU. So it's been seven years since presenting again. So it was good to, it's, it's good to be back up here. So just like my past sound presentation, the core of this one will be on sound and its power in the teaching and learning environment. The AI tools we'll be looking at are just that, tools. Although some people don't like referring to AI as a tool, but I'll be using that term. So as someone who manages a team of designers at SFU and somebody who is a sound designer personally outside of, in the professional environment, just being a little bit more of an artist. It's important to me to be having these types of discussions about AI looking closely at what's happening out there. So I was talking to Brian, my colleague, yesterday and he raised a good point and, and I actually just put it in my introduction too and it's, it either has the potential to add creativity or lessen creativity. And but I believe that that will come down to the way that we approach some of these things. So my work with research, like close research began September 2023, not that long ago, when I took a communications class here at Harbour Centre called Social constructive Construction of Communication Technologies, where we started to look at AI and how it was impacting the social aspects of communication. So today, I'll be using a mix of different types of information on sound. So hopefully there'll be something that resonates with everybody: the pedagogy side, the technology side, and hopefully that will resonate with, something will resonate with everyone.

So before we start a couple of things. So a YouTube channel I watch called Bryce, Bruce and Stephanie's Comics always starts with a disclaimer, "For information purposes only, do your own research." And so the presentation focuses on my investigation of generating sound with AI. So my kind of notebook that I've decided, hey, how can I turn this into a presentation that I could share with my colleagues here at ETUG. And so, I'll be looking at some of the more popular tools. And, and talking a little about, showing some of them, and having some examples of them. And also, so, for me, it's important to create and use your own personal content. For me personally, when I input and work with these tools. For me personally, I like to

make sure that I'm putting in my own content for examples and things like that, rather than taking other people's examples and putting into some of the AI. But again, it just depends on the situation and, and how you're collaborating with people. The irony of all this is the information that you're inputting of your own personal work. And then AI sometimes gets a handle on it. You may have to subscribe or you're kind of you have to look at some of the licenses and copyright once AI gets a hold of your own personal information and then spits it back to you. Sometimes you end up having to subscribe, use money in order to get some of these things back. And also it's important to understand the AI policies and copyright issues at your institution.

Again, you know, we're closely looking at some of the things using generative AI and AI tools at SFU. Also my presentation, everything I talked about today is on my blog. and it's [Bornthinkers.wordpress.com](https://bornthinkers.wordpress.com) and it just, it'll give you all the links. I'll put my slides up there and all the research articles that I looked at today so I can. And another thing too, kind of one of my philosophies if you want to know my philosophy on AI is kind of despite the, the optimistic title of my presentation on the, on the website is, you know, one of the things that I really believe about AI in terms of media in is it's useful as an idea generation tool and a placeholder sometimes for using your own original designs. So there's a research article out there called "AI as a Mediator Between Composers, Sound Designers and Creative Media Producers by Sebastian Löbbers. The link is on my website there. So AI-driven music research can find valuable application in aiding human-to-human ideation. For example, as a translator for non-verbal descriptions of sound and music. However, AI tools may have to be adjustable to user needs and project requirements to avoid unpredictable or inadequate behaviours. So, depending on what your field is, it may not be media, a media generation. You know, sometimes if somebody who is not, they don't have those types of tool sets of sketching or are working, sometimes generating something out of AI and then having a conversation with a designer can be an extremely useful tool and I've seen that happen at my centre. Another example or another quote is "Artificial intelligence to facilitate communication between stakeholders and accelerate the consensus-building process." So I believe that using AI can sometimes be an effective way to discuss ideas about generating original content.

So with that said, this is my first ever presentation with AI art. So it's an experiment for me. All the AI art here that I'm using in my presentation is from Copilot or Canva, which I built this in. And one of the things about using AI-generated content is, it's for me, it's similar to using Microsoft Office clip art in a sense. Expect duplications and non-originality in your work. And I've heard more than one designer call AI art the new presentation clip art. So I wanted to feel what it's like to leave my office with a presentation with AI art. How comfortable did I feel once I finally got up here to show you guys art that's been generated in AI. So, oh, sorry, before moving on with that quote, The quote comes from a research article called "Theoretical Foundations for Sound Use in Multimedia Interaction : Instruction to Enhance Learning by M. J. Bishop and Ward Kates from Maryland University. It discusses how sound can support the learning process, especially when it comes to acquisition, processing, and retrieval of new information in a variety of ways. So definitely an article worth investigating. I was really

appreciative of Luke's presentation yesterday when you talked about accessibility. Because I've got research articles on my blog where I'm really interested in how sound can become an accessibility tool for the learner. And how maybe music can aid a learner in order to be able to focus, Talk a little bit about ADHD, ADD so, you know, music therapy bolsters the attention and focus. Research articles have shown and I know that background ambient music certainly helps me focus.

So AI sound design tools are software applications or platforms that utilize artificial intelligence techniques to assist in the creation of, manipulation, or analysis of sound. These tools leverage machine learning by algorithms and other AI methods to automate or enhance various aspects of the sound design process. And a note that I made here is that AI is not always generating quick sound content. Sometimes it's automating a conversion process in your software. Like in one of the digital audio workstations I use is called Apple Logic Pro. And what it could be doing in there, it could be doing, it could be mastering the file, it could be doing optimal stereo width, it could be doing EQ doing parameters within it. And one of the things that I see is especially with mastering tools or some of the tools like Adobe Premiere Audition. In the future there may be an integration of AI within those types of programs that you may not even have an opportunity to opt out of, which is that for me, I think there absolutely has to be that choice. In fact, there's a new term out there. We have DAW for digital audio workstation, but we now have GAW which is the generative audio workstation and this is a DAW with the integration of AI tools. Even the hugely popular Audacity now uses OpenVINO AI suite of tools useful for spoken word, audio, and music. This includes noise suppression and a transcription option available within the tool. They are offered by Intel right now. currently only running on PC.

So I want to talk a little bit about some of the popular tools that I've been looking at. The audio tools that they usually generate content or modify audio content like I mentioned, doing sample generation manipulation. They can be used for audio mixing and mastering sound effects, which we'll take a look at in a second, voice synthesis and manipulation. And I've got a guessing game on SoundCloud where I'm gonna play a couple of AI sound tools and get you guys an opportunity to kind of guess what's been generated from text to audio.

So if I go here, I'm gonna put you guys to work for a second. Hopefully this works we'll see for both audiences online and in the room. So this is an AI-generated audio sample from text and your job is to guess what this is. Doesn't seem to be sharing. So let me just try one more thing here.

[AI audio of animal sound]

Any guesses on what that is? What was that? So the actual text prompt was cow mooing. So I played that for my wife and it just kind of scared her and she didn't want it. She didn't want anything to do with it. Of course, of course, absolutely. Yell out for anything.

It was a cow mooing. So the actual text I put in was cow mooing, and I'll show you the tool that I put that into as well. That's really bad. So here's some other, so that was from Meta's Audiobox tool. And this one is another one from my edit.

[AI audio of animal sound.]

That's a cow mooing. And here's the last cow mooing.

[AI audio of cow sounds]

okay. That's painful. That sounds like a sheep bawing. All right. So here, here's another AI example. You got to guess what this one is.

[AI audio of animal sound.]

Anybody have any guesses on that one? That was a wolf howling. Yeah. The actual input was wolf howling at the moon. I don't know if "at the moon" had any effect on it versus wolf. co early composition by man. And then I'll play two more. So this one here.

[AI audio of machine sounds]

I'll play another one. These are two different AI generated SFX generators and I'll play the other one. Supposed to be the same thing.

[AI audio of machine sounds]

Any guesses on that. That's a coffee grinder. Yeah. The actual coffee grinder grinding coffee beans. So okay, that's kind of funny. And the tool that I actually used for this. It's like, okay, you show me that. But where, what if I wanted to do the same thing is a tool called Audiobox.

And I've got that link there too and I went down and tried on demos. Winter sound effects. Oh And put cow mooing in there. Try it again here. Generate. Thank you for sharing the link in there too. Yeah.

[AI audio of animal sounds.]

Let me get one more. Okay. okay. So how do you think some of these tools are generated?

Well, a lot of it has to do with. So a lot of it has to do. Sorry, go back here. So some of these SFX are generated by a data set of existing sound effects, AI algorithms. They learn and to analyze parts of the characters of different sounds via machine learning techniques. So using this input, the AI models are able to generate new sound effects that closely resemble the training data that responded to training data that they have and respond to relevant prompts or

descriptions. So how some of these are generated is based on synthesis algorithms as well too that I've been looking closely at. I use a tool called iZotope Iris 2, which is a sample-based synthesizer that combines sampler and modular synth and spectral filter. And inside there you're able to look at additives, subtractive, and granular synthesis. And AI tools right now are being able to use these and manipulate these same types of tools to use sine waves, use square waves, use saw waves, use triangle waves and make the modifications for these particular types of sounds. And it's not just text based. It can be based on the parameters that you're using in the digital audio workstation that you're using. And it could be Pro Tools, it could be Audacity if you have those plugins. It could be Logic Pro that starts to manipulate and use those types of things. And the difference between a synthesizer and a sampler would be like a sampler would be kind of a natural sound that's, that's, that could potentially be recorded then inputted in and then, then it would be manipulated versus the synthesizer where from the ground up, the sound would be synthesized using an oscillator. And using one of the waveforms that I talked about.

So moving on to something completely different. Talking a little bit about sonic pedagogy. So, you know, those are some of the tools, those are some of the reasons of how that's being done. And you know, I want to talk a little bit about, about why and some of the contexts for, for potentially using these tools but also just using sound in general. So my personal definition of sonic pedagogy is the intentional use of sound in terms of content creation activities and assessments in order to achieve the desired learning outcomes. In my first presentation back in 2013, for me, it was important to look at sound and I talked about the instructional design approach from William Horton who talked about using absorbed do and connect activities in, in order to approach and work with the development of different types of activities and how sound was used. And I still use that today. So, the quote here is a beautiful quote, "Sound shapes and is shaped by pedagogical practice. Sound permeates and resounds in pedagogical context and interactions in ways that are cultural material and relational." For me, just extremely passionate about sound. Oh I kind of mute here, extremely passionate about sound. And sometimes I think it's really underutilized in the learning context when you go to start to develop curriculum. And that quote is from "Resounding Pedagogy" in a themed issue on Critical Communication Pedagogies on for in sound. So I had a Zoom call recently with an instructor from UBC. I'm at SFU but Kar On was nice to hook me up with an instructor who teaches sound and music and doing some sessional work at UBC, and he asked me, he said, "John as far as sound design and AI, you've got the presentation coming up. What is, what do you see as one of the biggest benefits of using sound to generate AI?" And it resonates with, with what Luke mentioned is accessibility, and for me, accessibility factors for language multimodal learning. And that text-to-speech technology enables students to take their lessons in preview in multiple languages, allowing students to hear in their native spoken language if they want to with dubbing. And it's definitely one of my favourite uses of sound. And again, I'll show you one of my favourite tools for TTS text to speech. So again, someone mentioned as far as sonic pedagogies go, looking at Meyer's multimedia learning, the 12 principles and the, the modality principle in terms of not only having just text and also video, or excuse me, images, but using it in, in relationship to having speech along with visual information, not that you can't

use text, but it's important to have if you're gonna be using speech, it's good to connect it with images. The personalization principle, which is getting better in AI I think, and that's humans learning best from informal conversational voices rather than a formal voice. And then again, the voice principle where it's better to be learning some of these instructional activities that you use to be using from a human voice rather than a text, speech generator. And I think in that aspect, we're getting better with the technology.

I'll show you some AI-generated text I use from samples from AI

[Audio of male AI voice]

The civilization of the galaxy spread across 200 million worlds. The black void of space swarmed with the ships of the empire. But far off the trade routes almost forgotten, lay the dying planet earth. A backwater and galactic culture. A place of exile, a pebble in the sky.

[Audio of female AI voice]

The civilization of the galaxy spread across 200 million worlds. The black void of space swarmed with the ships of the empire. But far off the trade routes almost forgotten lay the dying planet earth. A backwater in galactic culture. A place of exile, a pebble.

[New recoding]

Tonight we find Nick and Nora at a wedding being held in a fashionable hotel suite in Manhattan. The groom, Arthur Hornsby is an old friend of Nick's. The bride, Jane is breathtakingly lovely. A fact which does not escape our hero.

JOHN:

Thanks, John. Yes, the first one was generated from an AI program called 11 Labs and that was based on text that I inputted into the program on the web page. Amazing inflection. Yeah, I think you know, to some degree, it's getting better and better the technology. So let me go back quickly tonight.

So in practice creative ways to use this. So not just AI generated but also just different types of activities that you'd be interested in cooperating with sound. And also I made a note here because it was really important to me to mention that adding activities that allow students to use AI tools in their lessons. It doesn't have to be an instructor creating content with AI, it can be a student lesson or assignment that integrates AI in a useful and intentional way, which for me is important. We have a kind of a community of practice at SFU where I, I'm new to the practice. I just came in but we recently had an instructor that came in and talked about how he was integrating AI into his classroom. And for him and a lot of people out here, for him, it was really kind of about assessment and how it's going to be used. But it came down to kind of transparency and the different activities that we're gonna be using AI. For me, I think it's maybe specific on what the content of your course might be. You know, what type of different AI that you may be using and for what different type of purposes. So for me, language learning using text-to-speech tools. ChatGTP has different types of conversion examples for pronunciation,

dialects, intonation. And I've got some samples and I put those samples on my website, but I'm also gonna go in here and use them here. John prompting AI tool: I am a beginner learning English. Let's have a conversation. Please ask questions and correct me every time I make a mistake and give me the opportunity to try the sentence again.

AI VOICE: Of course, let's start. How are you feeling today?

JOHN: I'm feeling pretty good.

AI Voice: Great, remember to use very before good to make it "I'm feeling pretty good." Let's try that again. How are you feeling today?

JOHN: I'm feeling very good,

AI Voice: Good job. Now, what did you do yesterday?

JOHN:

Yesterday, I cooked dinner for my wife. AI Voice: Nice. Just a small tweak. Use "my" before "wife" to show possession. So it's I cook dinner for my wife. You say that again?

JOHN: I don't want to say it again. There's a different one here, which is called the Vocabulary Quiz. So let's give that one a shot.

JOHN prompting AI: Let's try a vocabulary quiz. You say an Italian word and I'll try to guess its meaning in English. Let's start with some basic words. AI Voice: Sure. Let's start with ciao. What do you think it means in English? Ciao means hello and goodbye? AI Voice: Ciao in Italian means both hello and goodbye in English. Well done. Now, let's try another word. Grazia. What do you think it means?

JOHN: It means thank you.

AI Voice: That's correct. Grazia means thank you in English. You're doing great with these vocabulary words. Would you like to continue with more words?

JOHN: I've got two more. Which one is basically direct translation. How do you say "I would like a cup of coffee" in Spanish? And another one which is the conversational practice, which let's practise conversational English together by practising a real-life scenario, like ordering food at a restaurant. you initiate the conversation using only English and I'll respond in order to have that type of conversation. Another one that is really exciting to me is collaborative music projects. So students generate music that might best fit their assessments. They may have a video or animation piece. they defend their selection on why this would fit to that piece. Another one is sound analysis. Selecting a piece of music or an audio recording and have students analyze its elements, rhythm, melody, harmony, dynamics, timbre. Encourage students to listen and

discuss how these elements contribute to the overall meaning, mood, and effectiveness of the piece. Also to ask yourself if you're using AI and you decide that you're gonna use this in your course. What would be the process of recording these sounds yourself in an original context to have those types of conversations? And again, if you're working with AI sound and describing how it could be manipulated to be suitable for a particular project. I've been working with sound for probably 30 years. In 1991 I got a four-track recorder. It's called the Tascam 424. So, I mean, in just in like literally in two years, three years, I've even thought about having AI as a tool to kind of generate that sound. So I'm very much a somebody that, you know, really appreciates the craft that goes along, the artistry that goes along with, you know, doing these types of recordings. But AI is out there. Again, people are asking about it. And they will be asking more about it. I find you might be finding this at your teaching centres.

Our help desk, we're finding when people have questions about AI tools, it's usually about plagiarism or it's usually about, you know, how I can incorporate it in terms of essay writing, but you might not be seeing a lot of questions about multimedia generation. You know, we've never had a question that came to the help desk in our centre that said, you know, I'm gonna, I'm about to use a, you know, an image from an, AI and add it into my course. So, you know, those, those questions are becoming more and more and as somebody that manages a design team, when I have these conversations with my team, you know, what is it like to discuss AI with a faculty member or someone else that decides, okay, maybe now I'll use AI and I'm not gonna use the educational media team at my centre because I have these different types of options. So, you know, there's those competing aspects. So it's something that we really have to, we really have to think about and be on top of because you want to share the benefits of. If anybody that's using AI generates media content, it be a video, sound. You know, you want them to use it thoughtfully and intentional, you want them to think about copyright, where this content may be coming from, the source that they've got it from. But I'm also truly. I truly believe that any type of sound you use almost needs to have that, that human element that is going to be able to look at the sensitivities of a design and then be able to customize it to that particular type of media. And those are some of the things that I want my team and other people to look at closely. I'm gonna keep an eye on time here. Actually, this is my last slide. So we're, we're getting close here. Sound storytelling. I was gonna, there's generic radio scripts that are copyright free and you can enter that content into the into programs like TTSMaker.

Again, for me, that's my own blog, that's my own writing. And I put that in there. And for me, if I wanted to take my content, be able to slow it down with permission using content from, from other people who are interested in using and adding their content to the collective AI you know, consciousness out there. But again, for me, this was, this is really interesting to have the ability to say "Go in here." I mean, this is not new technology and, and if you're interested in this, you've probably seen a lot of this. But if I want to go to Japanese 0620, convert that to speech, I'll need a new verification code 4106. Convert that to speech.

[AI audio of person speaking Japanese]

JOHN:

So the interesting thing is you can get in here, you have the ability to go in and manipulate the audio, the voice speed, the volume. You can do some pitch adjustments. And I mean, even back in the day, like, I was really interested in using a program called Max MSP to create a program that was like this. So I'm not, just because it's using AI technology, you know, I'm not against it. I like TTSMaker because you own all the content that is going up on, in, like once you put it in, you own the content and keep it up there for...

FACILITATOR:

So you need to wrap up. Yes. Yes. Yeah, absolutely. And I'm gonna go back to my last slide.
[Inaudible person talking]

JOHN:

Okay, go and see what he's doing. Yeah, absolutely.

So in order to get some of the questions, there's my contact information and I'd love to hear about what you might be doing with media. I can hear about it now if there's any time for questions, but also that's my contact information if any of you want to reach out.

FACILITATOR: Ok, thanks so much, John. We should probably head straight to lunch if that's okay. Thank you, John.