Rachid Finge (00:01):

Hi, and welcome back to the Made by Google Podcast. I'm your host Rachid Finge, and today's episode is all about sound. If you think of it, many of Google's products make sound, whether it's your pixel phone, your Nest Mini, your Nest hub, or your pixel buds. Of course, recently, Pixel Buds Pro got something new, something magical called Spatial Audio. It's basically like a movie theatre in your ears. And today we have someone with us that helped to bring that magic to life. So please join me in welcoming Lu Silverstein from the Pixel team. Lu, thank you for joining. Please tell us a little bit about your role at Google and how you ended up here.

Lu Silverstein (00:51):

Yeah, and thanks for having me on.

Rachid Finge (00:54):

Sure

Lu Silverstein (00:54):

Well, there's how I came here, why I came here. I came here from previous I was working at a streaming TV company. Probably a lot of people know it, Roku, and it was a fun company to work at, but we sort of weren't large enough to do much research and development. And so I was really interested in doing a little bit more of bringing research and development innovation into the product space. And so, you know, Google came, came to me and I was really excited to be able to have that opportunity moved over.

Rachid Finge (01:26):

Today's guest is all about how Pixel sounds. Lu Silverstein is a senior product manager in the Pixel team. He has a passion for making hardware and software melt together, and that's exactly what was needed to make spatial audio work. Lu loves a challenge, so making movies sound like you're in a real theatre while all you're doing is wearing a pair of Pixel Buds Pro is right up his alley. I hope you'll enjoy our conversation. All right. And now you work on spatial audio. Maybe just for starters, explain to us what that is before I ask you, you know if you have like any special connection with the topic.

Lu Silverstein (02:07):

Yeah. You know, my role here is as what's called a product manager for Pixel audio experiences. And so I cover everything that has to do with pixel sound, which means anything that plays, it could be media, like music or movies, right? It could be ringtones, or notifications. It also has anything to do with the microphone, anything that comes in.

Rachid Finge (02:28):

Oh, and is it only like the technical part, or are you also involved in, you know, maybe I hear a certain ringtone? Were you involved in creating that ringtone or having people on the team create those, for example?

Lu Silverstein (02:38):

Yeah, it's all of the above. So in the product management role, it's around working on anything from the sort of business aspects to the design, to the development, all the way through to making sure that it's launched and out there and doing well.

Rachid Finge (02:54):

Perfect. Well, I have you to thank them because I love the alarm sound on my pixel, pixel phone. So I guess that's from your team when I wake up every morning. So, so thank you for that. But that's a clear description of your job as a senior product manager. And then today's episode is specifically about spatial audio, which sounds very fancy and it actually is, but can you explain how you would define what that is?

Lu Silverstein (03:20):

Yeah, sure. I think the best way to think about it, it's sort of the simple way is that it's, it's like surround sound for headphones. And it's sort of what the title or the name kind of implies spatial is that it creates the sound in space so it gives it more depth kind of like in the real world, like if you're standing outside and there is a, you know, a bird chirping in one location and a train goes by in another, those are in space and they're a certain distance from you in the location. So spatial audio, we take a lot of what happens in the real life and we actually simulate that using algorithms to make it sound like those sounds are actually in those locations in space. But you've got your headphones on.

Rachid Finge (04:00):

Yeah. Because you got to apply some sort of magic, right? Because I know a lot of people, they spend maybe thousands and thousands of dollars putting multiple speakers in their living room so they can have this home cinema surround set. Yet we sort of create the same thing using just two earpieces. So could you tell us a little bit about how you do that? Because it seems impossible to me

Lu Silverstein (04:28):

There is a bit of magic, but it's all math magic in the end. You have to sort of recreate how we actually hear sounds, so the way we hear sounds is through, it starts through a sound actually coming in and reflecting off the outside ear that you have and then going to the inside ear and then your brain processes it. But there are a lot of things happening there because you have two ears, right? And so when a sound actually comes toward your head, it actually reaches your ears at different times, the left and the right. And so we, it actually bounces around differently. This sounds a little crazy, but actually transfers through your head from one ear to the other, some of the sounds

Rachid Finge (05:10):

Oh wow. It does.

Lu Silverstein (05:11):

Yeah. Your head actually, it's kind of like, you know, you can hear sounds through a wall, while there are sounds that go through your head. And so we recreate all of that using these algorithms to sort of fake your brain into thinking that that sound is coming from the outside and actually has reflected off of your ears on the outside ear and on the inner ear. And then we play it back like that. And that's why we have a special algorithm that Google has developed that is called the Spatializing algorithm, which runs on the phone. And when you listen to a movie it has to have a soundtrack with a certain number of channels. We take all of those, it's kind of like you hear about 5.1 surround sound. We take all of that 5.1 sound from a movie and we then process it on the phone and we play it back through your earbuds to sound like it's coming actually from the outside world with all those things that I mentioned earlier that we take into consideration and we have special models that we've trained to do that.

Rachid Finge (06:08):

Right. So if I go back to my home cinema example, so 5.1, if I'm not mistaken, that means someone bought a sort of five speakers to put in every corner of the room, one in the middle on above the TV perhaps, and then the. one is the subwoofer. So if there's a bird coming from behind in the movie, then I can hear that from behind because I've actually placed a speaker behind me, but then with the pixel buds in the spatial, you use ulcerative math to trick my brain into thinking that bird is from behind, even if it's not.

Lu Silverstein (06:42):

Yeah, that's right. That's exactly it.

Rachid Finge (06:44):

That's still crazy, but it's all math. So that's really cool. So now Lu, when I'm watching that home cinema movie, I can easily, I guess, you know, move my head from left to right or turn my neck a little bit and of course, the sound will change since the positions of my ears are changing. But how does that work then when you have your pixel buds in? Because as I move my head, I move the pixel buds equally, right?

Lu Silverstein (07:10):

Yeah. So now we're gonna peel the onion back a little bit. There are two essentially different types of spatial audio that are available on Pixel. And what you're asking about is a version where, and this is sort of the full-blown immersive experience where the Pixel Buds pro has these little sensors in them called accelerometers that measure the movement of your head. The two different types of spatial audio are what we call dynamic, which tracks the movement of your head. And then we have another one called Static, which is sort of a subset, I would call it a little bit simpler version of, of spatial audio. So let's start with actually the dynamic one since you asked about that and then I'll explain how the static one works, but the dynamic one, it takes those little sensors in the Pixel Buds Pro and it knows the position of your head when you're moving left and right. And so in the real world, let's keep that bird example, let's say it's a hummingbird that's actually floating there right in front of you. When you move your head left and right, the hummingbird sound stays where it is in space. So we have to recreate all of that and, and we have to know when you're moving your head left and right to literally sort of do the negative, then keep the bird in the space where it is. And so that's what the Pixel Buds Pro can actually do with these little accelerometers. A lot of headphones don't have those accelerometers it's obviously an additional cost and there are other things you have to do from a software perspective. Most all of your standard headphones don't have those accelerometers. So they have, what is that simpler version, which is called Static Spatial Audio. And so what happens is, let's say again the bird is floating in one spot, if you don't have those little accelerometers and we don't know where your head is when you move left and right, it actually gets stuck to your head and moves left and right. Not so much like the real world. So it's a little bit less immersive and lifelike, but it does still have the distance from you to the bird. That's the spatial part. And so you get all, if you don't move your head very much, you get a pretty good experience and even in the static version, but the Pixel Buds Pro gives you that extra kind of oomph with a more realistic immersive experience.

Rachid Finge (09:25):

I guess from a product management perspective, that must make it extra cool to work at Google because, you know, you can work on all the algorithms, but you only get to the dynamic spatial audio if you can work with the hardware team and tell them, Hey, I need those sensors in the earbuds, otherwise

we cannot make this happen. And it's maybe one of the few places where you can create that software and hardware combination.

Lu Silverstein (09:48):

Yeah, that's right. So I probably should explain that there are a number of pieces in the chain of and it's exciting working here because we have to actually tie all those things together. If, if you just start from the earbuds you then have to make those work with the phone, but then the phone has to also work with the content in the movies. So we have to work with the content providers as well. This project was really complex in a lot of ways and that we had to work with all the different parties involved, including the operating system on the phone. So there were many, many different groups that we had involved to pull those pieces together to make that simple experience for the user where they just tap play on a movie and they have their pixel buds pro in and it just works.

Lu Silverstein (10:30):

There was actually a lot of work in the background that had happened to tie those not only tie the groups together, like from an organizational standpoint, coordination, but technically we have to thread the needle of all the content. So the way to sort of a user to think about this is that when, they find a movie they want and they see that it says either Dolby on it or 5.1 and they click play that what happens at that point is the movie itself has a video track and an audio track, right? We take the audio track, we see that it's that 5.1 that you mentioned earlier. So there's, there's actually six channels, there's five plus one, and we have to then decode because it's compressed to go over the network. Decode it and uncompress it and then we have to specialize it and this is all being done on the phone and then we have to send it typically over Bluetooth. We have to then send it over Bluetooth to the headset and then, then playback, and then we have to figure out where the person's head position is. So there's a lot of communication going on back and forth between the phone and the earbuds at that point.

Rachid Finge (11:37):

Now, I can't possibly ask you, you know, how you trick our ears because that's just well, rocket science I guess. But I am wondering, you know, there are so many things we do with AI at Google. What's the role of AI in creating that dynamic spatial audio? Is there a role for AI to play in there?

Lu Silverstein (11:56):

We kind of delineate between AI and algorithms that are more fixed and the difference is how you, how you train or create those algorithms. AI is more generative where it's you train it and then it can kind of go off and do things on, it's on its own right. This is more of, we, we program it, but it's, it's pretty fixed like set of algorithms that are in there. And so this is more of a spatializer algorithm but it still does a lot of work at the time when the movie is playing. It still has to decode a lot of those things. Like I was saying, it has to specialize them and put them in real-time in these different locations.

Rachid Finge (12:38):

Okay. So we have the accelerometers in Pixel Buds Pro, what else is there hardware-wise to make sure that spatial audio, you know, sounds as intended?

Lu Silverstein (12:48):

So there are a few things to make it sound like it's in space and it's actually a lot of it's done in the software side of things with on the, on the earbuds. So what we do in software is we actually recreate

the room and the room that, that you're in. And we do a lot of assumption of it, but you can imagine like if you're in a stadium and you yell, right, the sound is very different than if you're in a very, very small, let's say room. There's a lot of reflections in a small room. And so when you speak, it'll reflect off the left wall, the right wall. And so we recreate a lot of that and the reverberation that happens to make it sound like the person is in a specific location and typically in a movie there are different scenes, right? There are scenes that are inside, there are scenes that are outside and so that helps re recreate it through these algorithms that we've built in the spatializer. The other thing that we do is we actually model essentially your head and how it acts with three-dimensional sounds. On the headphones there's sort of a nerdy term we call head-related transfer function or HRTF, which is also built into the algorithm to make it sound more like 3D sound.

Rachid Finge (14:00):
All right. So you already mentioned 5.1 movies. What kind of content other than that might work with spatial audio? If I have my Pixel phone and Pixel BuDS Pro?

Lu Silverstein (14:09):
Yeah. So any kind of five-out-one content that is played back should work. So for example, games if a game has 5.1, it would work with that. We focus primarily on the movie experience though, because that's where you get the most immersion from that spatial audio effect. And we've done a lot of studies around it to see how people really do find they can identify sounds even better with spatial audio, you know, specific sounds where they're located and they feel more immersed in it, but there, there is some other content that it would work on. We are working to expand that to additional content types in the incoming years.

Rachid Finge (14:48):
I was just also wondering when it comes to spatial audio, you mentioned those sensors in the earbuds, but our help center page also says that you need to keep your phone steady. Like why is that?

Lu Silverstein (15:01):
Ah, so when we do the tracking of positions, if the phone moves around a lot, we assume that we may need to lock the head position so that you don't, you don't feel disoriented. And so it helps to keep the phone. There are sensors, the same kind of sensors on the phone. There's Acceleron that measure movement.

Rachid Finge (15:23):
Those usually make sure that when you tilt the phone right from portrait to landscape, the screen rotates, I guess, right?

Lu Silverstein (15:30):
Yeah, exactly. Yep. And so we want to make sure that you have the best experience and that if you're moving around a lot, you're also not having orientation issues with the head bud earbuds.

Rachid Finge (15:41):
Right. Okay. So that's basically to, to temporarily maybe stop the de-tricking of the ear in order to make sure that, you know, I don't get disoriented or maybe a little bit nausea or something like that.

Lu Silverstein (15:52):

Yeah, exactly.

Rachid Finge (15:53):

Okay, well thanks for that. I'm really happy you thought about that. So what's next for spatial audio on Pixel or maybe even more widely when it comes, to our products?

Lu Silverstein (16:04):

You know this is the fun part about working here is that we can imagine a lot of things that could be done with it. These are not definitive plans that we have, but I can tell you just from a personal perspective, I'm excited about the possibilities and what these things are like in the real world, when you think about sounds and space it helps you get oriented. So you can think of things like what we're doing right now, even a podcast, right? You could hear people in different locations, especially if you have multiple people or even things like newscasts where there's like a panel of people. You can tell where people are. Some people do a lot of conference calls and so you can imagine hearing people in different locations, it feels much more realistic and you can kind of identifying people in who's speaking. So there's a lot of different directions we, we can go, you know, we're working like I said, on figuring out where to expand next. We just launched spatial audio at the beginning of this year, and we're really waiting here, like from users and from our customers, what kind of feedback they have and what really they enjoy and what they're, they're interested in as well.

Rachid Finge (17:12):

All right. Sounds like stay tuned and give feedback. So that will will help you and your team decide where to go next with spatial audio. Wonderful. Yeah. So we always close every episode of the Made by Google podcast with a top tip from our guest. For our listeners, Lu, give us a tip of something that our listeners and users of Pixel devices can do to get the most out of spatial audio.

Lu Silverstein (17:39):

Yeah, this is a good one. So one of the things that are complicated with spatial audio, I think for our customers is sometimes to know whether you have, I mentioned all those different pieces of the chain that you have to have. You have to have the right piece of content. You have to have the phone working, right? You have to have earbuds. A simple way to just check that you're getting all of that. And if you have your Pixel Buds Pro in and you have your Pixel, let's say, you know, Six, Six Pro, Seven, Seven Pro, and you want to know if you're getting spatial audio, the dynamic kind, when people are talking in dialogue on the movie, you can move your head left and right and see if their voices stay centred even though you're moving your head back and forth. That's a really quick, simple way to check that you're getting the spatial audio and that it's working properly. It's sort of like the analogy, you know when you go into like a retail store where TVs are, and you can see the difference between the TVs on the wall there, but when you get home, you really don't know what your TV looks like compared to the other ones. This is of course quick check, is everything working? Do I have a difference between spatial audio on and spatial audio off?

Rachid Finge (18:43):

That's a good one. And well, I have no plans tonight, so definitely going to try that and listen to some content in 5.1 with my Pixel Buds Pro. Lu, thank you so much for giving us an explainer on what spatial audio is and looking forward to seeing what you come up with next.

Lu Silverstein (19:00):

Yeah, thank you. Thanks for having me on.

Rachid Finge (19:03):

Oh, I really enjoyed that. Thank you so much, Lu, for your time, and thanks for explaining the magic behind Spatial Audio. Now. If you don't have them yet, maybe this is a great time to visit the Google Store and get yourself a pair of Pixel Buds Pro. SCR is the ultimate companion to your pixel phone. I'm already looking forward to the next Made by Google Podcast episode. If you haven't yet, make sure you subscribe so you won't miss it. Thank you so much for listening and hope to catch you on the next one. Take care. Talk to you later.