- Thank you Mark for introduction. You saw Live Transcribe when Mark spoke, but now you see Euphonia Euphonia understands my speech. Live Transcribe understands other people's speech. This is what I'll be talking about. As many examples how you can communicate with people. We have now Live Transcribe that allows me to understand other people and we have Euphonia that allows other people to understand me, even I have strong, Russian accent. But let me tell a little bit history. 30 years ago, I moved from Academia to speech technology team. My dream was to develop speech recognition systems that would help people who do not hear to communicate. I thought the solution for this problem five years away. Five years passed, I thought would more another five years ... And this lasted for 25 years, then I moved to Google. The last five years, Google developed a lot of critical applications for speech technology. Please show me next slide. So, I'll be talking about this Live Transcribe Euphonia and something that grew around them. Live captions like captions ... but let me ask you something, you have a lot of information now. You have information from many streams, so when I speak, what is this? Because this gives the idea how Euphonia understands me. When other people speak, you watch this. This will give you ideas how Live Transcribe understand me, but let first show video about Live Transcribe. Please start the next slide. I completely lost hearing when I was only a year old. Technology provide tremendous opportunities to hear. I research science speech recognition technology I could fulfill my dream for whole life. Use speech recognition to communicate with people. Hello

- [Woman] Hello, how are you?

- I'm fine, how are you?

- I'm doing good. What can I get you today?

- I would like tea. Live Transcribe provide communication ... It provide transcription of speech for deaf and hard of hearing.

- Dimitri worked with a speech team to develop that technology and Google has made strides in this in the past couple of years, but the app itself is what I've been working on. I'm not an android engineer. I saw an opportunity for something useful and dug into that and kind of evolved naturally.

- [Man] Did I already mention to you I'm having one of my famous chili parties again this weekend?

- This was something that was born out of Dimitris' needs. All the technology existed already, it was kind of just a matter of hooking the wires together. Gallaudet University is the primary university for the deaf and hard of hearing in the world, really. We have been partnering with Adam to see how people are going to use the Live Transcribe.

- [Narrator] I started using it and the first time was with my two sons. Both of them are hearing and they can sign a little bit, but not a lot. So, I use Live Transcribe with them.

- [Child] Why do they say go?

- [Narrator] Live Transcribe is changing the way we communicate with hearing people.

- [Dr. Obiedat] How about your math class?

- [Child] I'm doing good.

- [Chet] We still have to review every day. Here you can see an example of multiplication problems.

- [Narrator] Information is power. And Live Transcribe gives us deaf people an opportunity to be more empowered.

- I want to show you another way to use Live Transcribe. I have this small phone. I have a band and a magnet. Now my hands are free. It is very convenient. I can eat and talk to people. It has Android, I can hear every word. You can buy it cheap on Ebay for $50! Small magnet is very strong. I can wave, it does not fall. ... Let me take it out. So, we developed at Google also Euphonia. Euphonia is different. This is Live Transcribe and Euphonia is a model. It is located in small Android. Live Transcribe requires model on clouds. And it also Euphonia understands other people too. So if I get stuck in Iowa, I can not use Live Transcribe to discuss with people how to get out because no Wi-Fi usually. But I can get an Euphonia and understand my voice and it understand other peoples voices. So, right now you will hear more about Euphonia. Euphonia is very good. It has speech recognition. But if you do not sound like most people, it will not understand you.

- No one's ever collected large data sets of people whose speech is hard for others to understand.

- People who have Multiple Sclerosis, who are deaf, who had stroke, who stutter.

- They're not used in training the speech recognition models The game is to record things

- What's the temperature today?

- And then how to recognize things that you say that aren't in the training side.

- Algebraic Geometry

- Dimitri recorded 15,000 phrases. It wasn't obvious that this was gonna work. He just sat there and he kept recording.

- We need to make all voice interactive devices be able to understand any person who speaks to them.

- You can see that it's possible to make a speech recognizer to work for Dimitri. It should be possible to make it work for many people. Even people who can't speak because they've lost the ability to speak. The work that Shanqing has done on, you know, voice utterances, from sounds alone, you can communicate. But there might be other ways of communicating.

- Most people with ALS end up using an on screen keyboard and having to type each individual letter with their eyes. For me, communicating is slow.

- [Irene] Steve might crack a joke and it's related to something that happened a few minutes ago. The idea is to create a tool so that Steve can train motion learning models himself to understand his facial expressions. To be able to laugh, to be able to cheer, to be able to boo. Things that seem meaningless, but actually are so core to being human.

- I still think this is only the tip of the iceburg. We're not even scratching the surface yet of what is possible. If we can get speech recognizers to work small numbers of people who learn lessons, which we can then combine to build something that really works for everyone.

- [Dimitri] To understand and be understood, it's absolutely unbelievable. And I can now plan to give mathematical lectures. My background is mathematics. So I can tell now algebraic geometry. Commutative moufang loop. I plan to give lecture in Paris to my mathematical friends, who can finally speak all these wonderful words. It's absolutely unbelievable. It's transformative moment for society. Parents of deaf children now have completely new opportunity. Now their children could be fully integrated in society in new way. You know, it's very difficult for speech therapists to teach a deaf person to speak. It take many, many years of practice. Sometimes with no results. I know someone deaf person, he told me he spent many years in speech therapies. He came to shop, he asked for beer. For beer, he asked to buy beer. And, he could not get this. They did not understand what to buy. But now, a situation is different. It's difficult to teach persons to speak well so others understand that person. But it's much easier to teach persons that machine understands that person. This way people can have more motivation start to speak better and to be integrated at school in many societies. Where they can inevitably get CART services, or some interpreters. There are many, many, thousand users of Live Transcribe now that are stories they come to hospitals and hospitals forget to hire for them sign interpreter so Live Transcribe and they're talking to doctors. Doctors are amazed. Someone told me when doctors said the whole hospital can't wait to see what wonderful transcribes. Okay, now lets see what the next slide is. Oh. Euphonia it is one method to make speech and to understand nonstandard speech. But we have another research method, very new that we started recently. It's called Parrotron. Parrotron maps directly one audio nonstandard audio into another audio. That could be standard speech and it trained neural network to make this standard map to make maps from nonstandard speech to standard speech. Euphonia is different. Euphonia is speech recognition that have many models, many systems. It has acoustic systems. It has language model system. So let's play next video that show how Parrotron work.

- [Dimitri On The Video] Hey google, how many ounces in one litre?

- [Google] Sorry, I'm not sure how to help with that yet.

- How many yards in a meter?

- [Goggle] Sorry, I'm not sure how to help with that.

- How far is San Jose from Sunnyvale?

- [Google] Southern University and A&M college is 645 miles away from Missouri.

- What is definition of rhythm?

- Here's the definition of rain: moisture condensed from the atmosphere that falls visibly in separate drops.

- This one.

- What is the definition of rhythm?

- [Parrotron] Hey google, what is the definition of rhythm?

- [Google] Here's the definition of rhythm: a strong, regular, repeated pattern of movement of sound.

- And let me show now what is next for live caption. Okay, I make demo for this. Let me choose random video. We switch to my phone. We switch to Live Transcribe phone please. I am choosing random video. I did not prepare it in advance. So let me see what I find. You can see the captions. You can now for any video... lets got back to slide. So, you can now play any video to have it's transcription run in pixel 4. And soon it will be available in pixel 3. It is fantastic because a lot of videos do not have captions at all. Now just run these and you see captions. Next slide. Also have remarkable slide caption. I love you. Please run caption so I can see what they are talking. And next slide. You can change font, how big are slides. Next. Next. Now, when I'm using transcription, I can not look at people. I need but I want to see people when I'm talking. So, we are developing now a research project tactile wearable device. It maps audio into tactile information. You can feel environment, you can feel sound. It represents several cast of sounds. That I think of this you can start to lipread people and fish sounds so you have more interaction than when you look for using on persons then on this. Please run video for this.

- [Man On Video] This video shows our concept for a tactor device to enable deaf and hard of hearing users to feel sound. Our concept is a device with multiple vibrating tactors worn on the forearm, or some place out of the way. Each tactor responds to a different vowel. Okay, now for a demo. Here I am using the interface. The hexagons visualize what the tactors are doing. Beet, bat, boot, aloha, catch you later, catch you later.

- Next slide please. Okay, so it's finished. We're finished. Thank you for your attention to this critical application that helps to change people lives. I don't know if I have time for questions, if not, I just go. Otherwise if I have time for questions, you could ask me on Live Transcribe.

- [Mark] I haven't looked at the slide. I was transfixed, sorry. Does anybody have a question? Anybody who wants to know one thing? I have one. What's the next thing you want--

- Please turn on Live Transcribe. Just a moment.

- What's the next thing you want to develop?

- This tactor device. I want to develop this tactor device that I'm talking about. It is the next project that I'm focusing on.

- And what time scale have you given yourself? Is it another five years?

- I hope it will be faster. Google is very fast.

- [Mark] Fantastic. Thank you so much.