MARK: I would like to welcome Dimitri from Google to talk about Live Transcribe.

DIMITRI: OK. Can you put here the slides? Thank you, Mark. Thank you, Mark, for the introduction. You saw Live Transcribe when Mark spoke but now you see Euphonia. Euphonia understands me speech, Live Transcribe understands other people's speech. It is what we are talking about. This is what I'll be talking about as many examples, how you can communicate with people, we have Euphonia that allows me to speak.

But let me tell a little bit of history. 30 years ago, I moved from academia to a 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 was years away. Five years passed, I thought would another five years ... it lasted 25 years. Then I moved to Google.

The last five years Google gave me a lot of critical applications for speech technology, please show me the next slide. So, I was talking about this Live Transcribe and Euphonia and something grew around them. Live Captions like captions ... but I have to tell 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 you the idea how Euphonia understands me. When other people speak you watch this, it gives you an idea how Live Transcribe understands me but let me first show you a video about Live Transcribe. Please start the next slide. Please play the video.

[Video]

I completely lost hearing when I was 11 years old. Technology provided me hearing. Hello.

>> Hello, how are you?

DIMITRI: I'm fine, how are you?

>>: I'm fine, what can I get you today?

DIMITRI: I would like tea. Live Transcribe is more an application than dragon and android. It provides incredible speech ...

>>: Dimitri worked with the speech team to develop the technology and Google has made strides but the app sits ... I saw an opportunity for something useful and we dug into that. E evolved naturally.

>>: Did I mention I'm having a famous party this weekend?

>>: This was something born out of Dimitri's needs. It was a matter of hooking the wires together. Gallaudet University is the university for the deaf and hard of hearing and for the world. We have partnered with them to see how well they have used Live Transcribe.

>>: I started using it the first time with my two sons. Both are hearing, thank you can sign a little but not a lot. So, a use Live Transcribe with them.

>>: Why do they say go?

>>: Live Transcribe is changing the way we communicate with hearing people.

>>: How are you doing in your maths club?

>>: Good.

>>: We started with everyday examples and problems ...

>>: Information is power. 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. Let me take it out. So, we developed at Google also Euphonia. Euphonia is different. This is Live Transcribe and Euphonia. Euphonia is a model. It is located in small Androids. Live requires model on clouds. It also understands other people too. So, if I get stuck in Iowa, I can use Live Transcribe to discuss with people how to get out. Because no Wi‑Fi usually, but I can get Euphonia and it understands my voice and it understands other people's voices. So, you will hear more right now about Euphonia.

Euphonia is very good. But if you look at most of people, it will not understand you.

>> No‑one has collected large data sets whose speech is hard for people to understand.

>> People who have multiple sclerosis, who are deaf, who have had a stroke, who stutter.

>> They're not used in training the speech recognition models. The aim is to record things and then have it recognise things that you say that aren't in the training set. Dimitry recorded 15,000 phrases. It wasn't obvious that this was going to work. He just sat there and he kept recording.

DIMITRI: 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 recogniser to work for Dimitry and 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 voice utterances, from sounds alone you can communicate. But there might be other ways of communicating.

>> Most people would use an off-screen keyboard and type each letter with their eyes and it makes communication slow.

>> Steve might crack a joke about something that happened a few minutes ago. The idea is to make a machine that Steve can train to understand his visual expressions. To be able to laugh, to be able to cheer, to be able to boo. Things that seem maybe superfluous but actually are so core to being human.

>> I still think this is only the tip of the iceberg, we are not scratching the surface yet. If we can get people to work with other people to build something that really works for everyone.

DIMITRI: To understand and be understood is absolutely unbelievable!

I can now plan to give a mathematical lecture, because my background is mathematics. I can tell you now algebraic geometry. Commutative Moufang loop. I plan to give a lecture in Paris to my mathematical friends, with my friends who will speak all of these wonderful words. It is absolutely unbelievable! It is a transformative moment for society. Parents of deaf children now have completely new opportunities. Now, their children could be fully integrated in society in new ways. You know, it is very difficult for speech therapists to teach a deaf person to speak. It takes many, many years of practice. Sometimes with no results. I know someone, a deaf person, she told me she spent many years of speech therapies. She came to shop, she asked for beer, and for beer, she asked to buy beer. She could not get this; they did not understand what to buy. But now she feels different. It is difficult to teach people to speak well and it is difficult to understand the person, but it is much easier to teach persons that machine understands that person. The way that people can have more motivation start to speak better and to be integrated at school and in many societies, where they can inevitably get CART services or some interpreters. There are many, many thousand users of Live Transcribe now. They have stories. They come to hospitals. The hospitals forget to hire for them a sign interpreter, so they try Live Transcribe. They are talking to doctors and the doctors are amazed, someone talking to the doctors.

Okay. Now let's see what the next slide is. Euphonia, it is one method to make speech and to understand non‑standard speech. But we have another research method, a very new one we have started recently. It is called Parrotron. Parrotron maps directly audio non‑standard audio into another audio that could be standard speech. And it trained neural networks to make this standard map, to make MP from standard speech from non‑standard speech to standard. Euphonia is different. Euphonia is speech recognition that helps many models, many systems. It has acoustic systems. It has language and audio systems. So, let's play the next view that shows how Parrotron works.

DIMITRI: Hey Google, tell me how many [inaudible] is in one litre?

>> I'm not sure how to help with that yet.

DIMITRI: How many yards in a metre?

>> Sorry, I'm not sure how to help with that.

DIMITRI: How far is San Jose from Sunnyvale?

>> Southern university and A&M college is 645 miles away from Missouri...

DIMITRI: What is the definition of rhythm?

>> Here's the definition of rain: Moisture condensed from the atmosphere that falls in separate drops.

DIMITRI: What is the definition of rhythm?

>> Hey Google, what is the definition of rhythm?

>> Here's the definition of rhythm, a strong regular repeated pattern of movement and sounds. [Laughter]

DIMITRI: Let me show you what happens next, so live caption. Let me choose a random video. If we switch to my phone. If we switch to the Live Transcribe phone, please. I'm choosing a random video. I did not prepare it in advance. So, let me see what I find.

DIMITRI: You can see the captions. You can now for any video... let's go back to the slide. So, you can now play it for any video and to have the transcription run in Pixel 4. Soon it will be available in Pixel 3. It is fantastic. Because a lot of videos do not have captions. Now, you just run this and you can see the captions. The next slide. We also have remarkable Slide Caption. I love you. Please run Slide Caption and you can see what I'm talking about. The next slide. You can change the font; how big the slides are. The next, the next. The next slide. Oh, now, when I'm using transcription, I kind of need to look at people. I need this and I want to see people when I'm talking. We are developing now as a research project a tactile wearable device. It maps audio into tactile information. You can feel your environment. You can feel sounds. It represents several sounds. So that I think you can start to lip‑read people and you can feel sounds and then you have more interactions and then for example, you look at a person.

[SUBTITLED FILM]

DIMITRI: The next slide, please. I'm finished. Thank you for your attention to this critical application that helped to change people's lives. I don't know if I have time for questions. You can ask me a question on Live Transcribe.

MARK: I haven't looked at Slido. I was transfixed, sorry. Does anyone have a question? I have one. What is the next thing?

DIMITRI: I need Live Transcribe. Just a moment.

MARK: What is the next thing you want to develop?

DIMITRI: This tactile device, this wearable, the tactile device I'm talking about, it is the next project that I'm focusing on.

MARK: What timescale have you given yourself? Another five years!

DIMITRI: I hope it will be faster. Google is very fast!

MARK: Fantastic. Thank you so much. Thank you. [Applause]