**Globally Speaking**

**Podcast 85 Transcript**

**Talking Transcription — Part 1**

**Featuring Jamie Hartz**

Michael I’m Michael Stevens and today on Globally Speaking, we’re starting our three-part series on transcription. This is an area where many of the projects that eventually translators work on starts. We have to get the ideas into text in order to get it localized. And while it’s at the very beginning of the process of everything we do, this is our first time covering it on the show.

We did a three-part series because we are starting [with] looking at the traditional methodologies of transcription, how people do it, and the work they do and the benefits that they bring to their clients. What we were led down is a path of where all of this is going and how is voice going to be used in our process in the future. So, we get to catch up with some really great technologists as well in this discussion.

So, I hope you guys will stick around for all three parts of the series. I’ll let our guest introduce herself now.

Jamie My name is Jamie Hartz. I’m a freelance translator. I do Spanish to English translation, mostly of commercial and legal documents and I’m interested in getting into the logistics translation field as well. And I do a lot of transcription work and that’s what we’re here to talk about today.

Michael This is our first time on the podcast talking about transcription. So, Jamie, what’s the basic definition of transcription?

Jamie Transcription is taking an audio or video file and converting it to text. So, creating a transcript of that audio and spoken language.

Michael Which is a very different skill than what they do in courtrooms.

Jamie It’s very different than things like subtitling or stenography or even closed-captioning. It’s a completely different field, a different practice.

Renato What type of audios and videos are you usually transcribing?

Jamie I get a lot of different types of transcription. Usually they’re audio files; I don’t get a whole lot of videos, so usually it’s like an MP3, WMA or WAV recording. A lot of them are law enforcement, public relations, like press conferences, and human resources also. So, a lot of the types of industries that come to the translation companies for translation work will also come to us for transcription work.

Renato Because you’re a translator and you do English-Spanish, are you getting more of your transcription work in Spanish or in English?

Jamie It’s actually a pretty good mix of both monolingual and bilingual. So, for law enforcement, a lot of the times I get both Spanish and English just because you’ll have, for example, a police officer who speaks only English and you’ve got a witness or a victim or a suspect who speaks only Spanish, and then you’ve got another person interpreting for them.

That type of situation will usually come to me in both Spanish and English and a lot of times, what they’re looking for is a transcription and then a translation of all of the Spanish text into English. But, for example, with human resources or with public relations, usually I’ll just get English audio recordings and they want basically a clean English transcript for them to use, either to give out to their journalists or for them to have on record in case someone ever, you know, brings up issues with whatever had been recorded, whatever interview it was. So, it’s a mix of both Spanish and English and a combination of the two.

Renato This is one of those areas of technology where automation has really had a huge advance and you have speech-to-text recognition. One of the first use cases tends to be transcription. But I think that you just described one of those cases where technology might not be very friendly: it’s the fact that you’re dealing with bilingual content.

While you as a human have no trouble switching from one language to the other, maybe technology would take longer. Have you encountered any opportunities where your service is favored over the automated services that are available on the market?

Jamie Yeah, that’s a great question and it’s something I’ve thought about ever since I started transcribing because, you know, I have this phone that I keep in my pocket that I press a button and I say, “Hey Google, navigate me somewhere,” or, you know, I ask it to look something up, and it knows what I’m saying. So, in theory, shouldn’t transcription be able to be outdone by a machine? You’ve definitely hit on one major aspect of why that’s not the case, and it’s the fact that humans can code-switch. Humans can discern one language from another much better than I think a machine can.

But another is the same story we’re seeing with translation, where we have machines that can do translation and people ask, “Can’t they just replace the humans?” And the answer is yes, but humans can do it better, and a lot of times, if you start with a machine transcription, what you get is very garbled, what you get is kind of nonsense, and it has to be cleaned up by a human anyway. And so, it’s almost better in that case to just clean it up by redoing it in the first place.

I think this is a funny example of how voice-to-text doesn’t always work right. So, I have an Android phone and in my contact list, I have my parents’ landline listed as *mi casa—*my house. And I used to be able to say, “Hey Google, call *mi casa*,” when I wanted to call my parents’ house, and it would dial them, and I wouldn’t have to touch the phone and it would be hands-free and it worked fine. That was voice-to-text that was very functional.

The phone’s getting smarter and smarter and it’s working less and less because now, as I think, you know, AI has come in and started to try to learn what I want better, you know, I would say, “Okay Google, call *mi casa*,” and it would say, “Okay, calling Michigan casa.”

Michael Ah, so it’s the abbreviation for Michigan.

Jamie Right. It was still calling the right phone number, but it was responding with something it thought was smarter, and it was really more stupid. And then even more recently, I would say, “Okay Google, call *mi casa*,” and it would say, “Okay, you want me to call you ‘casa’.” Because I said, “Call me ‘casa’”; ‘call me *this*’. And so, they’re getting smarter and smarter and they’re understanding us less and less.

Renato My son has played some jokes with me and he changed the phone to call me Darth Vader. So, I would say, “Hey Siri,” and, “Yes, Darth Vader,” or something like that.

[Laughter]

Jamie Great.

Michael So it’s sort of an advancement you’re seeing in the technology in front of our eyes as yes, it’s gotten smarter, but it’s complicated things more, so there’ll need to be another leap before it’s as useable as we would like it.

Renato Do you use these technologies to accelerate your productivity? Amazon, web-based, transcription software will do stuff in a fraction of the time that a human would do in 13 languages. Do you use technologies available in the market commercially to accelerate your process? Is post-editing of transcription an activity that you’re developing?

Jamie That’s quite a good question. Yes, I use it, but only to the extent that it makes me more efficient. There are certain softwares that will, you know, you drop in the audio file and it will spit out for you a transcription, but in order for that to actually be more efficient, it has to be a single-speaker audio file that’s only in English, and he’s speaking really clearly and there’s no background noise and there’s not interruptions in the audio.

Whereas for these law enforcement transcriptions where there’s three people talking at the same time and they are speaking in two different languages, that would never be more efficient, I would say. At least at the point that the technology is now, because you need to be able to separate out who is speaking at what time. You need to be able to separate out what language they’re speaking, discern what is going on in the background. One of the things that I’ll transcribe kind of often is DUI, ‘interviews.’ Someone gets pulled over…

Renato You’ve got the slur [laughing].

Michael Yeah, does that mean you’re fluent in drunk?

Jamie I am—Spanish drunk, no less. So, yeah, the cop is wearing a body camera, the suspect is, you know, pretty audibly drunk and they are on the side of the road, usually with cars flying by and people walking by, and there’s a million different sounds in the background. And so, to be able to distinguish a voice that’s right next to me, you know, right next to the body camera from a voice that is a hundred feet away is something that I think a machine would really struggle to do. But I can do it because I can hear how loud they are and, you know, what distance they would be from the camera or from the recording device.

Michael There are a lot of things right out of the gate that you’re able to process more quickly than a machine, and whether a machine can be trained on that is still to be determined. You mentioned recently a couple instances where you’ve been able to hear about the results or impact of the work you’ve done. Are you able to share those stories with our audience?

Jamie One of the most meaningful projects I think I’ve ever done was about two years ago. I received an audio file to transcribe, bilingual audio file. There was a bilingual detective interviewing an alleged victim.

He was actually a child victim of sexual assault and he was very young. I was able to listen to how the detective had such skill at going back and forth between the two languages, recognizing that this child was uncomfortable, making him feel more comfortable, establishing a basis for him to be credible in court and for the DA.

So, the bilingual detective was interviewing this child about his experience and what had happened to him and I, you know, was, like, emotionally moved by it because it was basically like “Law and Order SVU” but real life. So, it’s, you know, maybe an hour-long interview and it took me I don’t even know how long to transcribe and translate, and then I turn it in, and you don’t really hear anything. You don’t expect to hear anything, right?

But just a few weeks ago, I got another project, a translation to do from the same client, and realized pretty quickly that it was related to that same case, because that kind of thing really sticks in your mind—you know, I was not going to forget those names and that situation—and it was a letter from the mother of that victim to the court. She was thanking the court for everything that they and the district attorney and the police had done for her child and how that situation had changed their lives. It had brought her so much peace to know that justice had come to their family.

And basically, this criminal had been convicted and put in jail. That was really rewarding for me to (a) be able to see some closure for such an alarming situation and also (b) to be able to see that I had played a role in it. Which is something that you don’t really expect to do as a translator sitting in your home office. You don’t feel like you impact anyone, and so, to see that I had played a very small role in that situation, in bringing someone to justice, was a really cool situation.

Renato That’s very interesting because you had the opportunity to go full circle and understand exactly the consequences and the impact of your work. I’ve heard stories of interpreters that are helping victims in dangerous situations, 911 situations, and the call drops, and they don’t know what it is, they don’t know what happened and they get hung up on.

Michael It’s left completely unresolved.

Renato Yeah.

Michael Yeah.

Jamie Yeah, that’s one of the biggest challenges.

Michael That’s a rather impactful story and situation to be in, but then some of your work is also focused on helping businesses be more successful. So, for some of your clients who want to be in RFPs that may not be in the native language of their company, that’s work you do as well?

Jamie That’s right, yeah, that particular situation was a translation project where some agency from a Spanish-speaking country’s government put out an RFP. They wanted to build some kind of infrastructure, and an American company wanted to submit their information for the RFP. So, obviously they needed to know what all the specifications were, what the requirements were, how to apply, who to get in contact with.

So, they had me and a team of translators translate the RFP, and again, like, you don’t expect to hear anything back, and then a year later, I get a project to translate and it’s the award letter saying you have officially gotten this contract, we’d love to work with you. Which is really cool because, again, I get to see them being successful and maybe this is more work for me in the future.

Michael Yeah, absolutely! And how disappointing would it be if they wouldn’t have been able to read the award letter? They’re just like, “Oh man, we’ll never know.”

Jamie Or, they just dropped it into Google Translate and were like, “Oh, we’re not sure what this says, let’s just ignore it,” yeah.

Michael Exactly, exactly.

Renato I’m curious, what automated transcription technologies do you use and why?

Jamie Yeah, yeah, let me talk more about that. There’s a whole range of different technologies. I think it’s transcribe.com offers an option for transcribers to basically upload an audio file and it spits out a transcription. I’m sure there are more nuanced versions of that, and for the situations I’m working in, I would never want to upload something online and have it be spit out into, you know, the cyberspace.

So, I don’t personally use that, but one of the things I do use is Dragon Naturally Speaking, is a voice-to-text for dictating. I’ve used that on occasion. There’s also some transcription software that people will use such as Express Scribe or InkScribe, and those are software that basically allow you to stop and start audio more quickly than you would be able to, like, for instance, with your mouse. You’re not got going to switch from your keyboard to your mouse as quickly as you need to.

I use Express Scribe which helps me to drop the audio file into that software. It reads it and it will allow me to use a foot pedal that I have, which is another great tool for transcribers to start and stop the audio. So, while my fingers are moving on my keyboard, they don’t have to be distracted by stopping and starting audio. That’s done by my feet. I can use multiple appendages to start and stop audio, I guess that’s great.

Renato In the beginning of my career, I did a lot of transcriptions but I’m an old person, so it was with cassette tapes. I actually had some of those roll tapes that I used, and it was very hard. But one of the key things that I had with the pedal that you connected to your cassette player, was that one of the key features of that pedal was that it rewinded the tape five seconds.

You hit your foot on the pedal and it plays, and then you take it off, it stops, right? It would rewind five seconds because you always lose something. I don’t know if it does the same when you’re doing it with software electronically.

Jamie Yeah, that’s funny, it works the same way. Yeah, I think it jumps back three seconds.

Renato Three seconds, there you go. People don’t realize how hard it is, because sometimes you will spend several minutes trying to understand one word. And it’s because of the accent, it’s because of the background noise, it’s because of two people speaking at the same time. And there is one word that kills you and you don’t know what that word is, and it might be significant to the context.

So, I empathize! I know how hard that work is, and I was asking you about the technology because we transcribe our podcast and we use a human transcriber who has been working with us from the beginning and she’s fantastic. Say hi to Val!

But I also use in my work, in interviews that I’m not going to publish, it’s not content that is going to be shared--there are tools like Temi, which is fantastic. You upload a file, it’s one hour of content, and ten minutes later, you have the whole hour transcribed, sent back to you, for, I don’t know, ten cents a minute or something like that. But the quality—it’s just for information only. You cannot use it for publishing content and for using it professionally.

I believe that machine translation is the same thing. I mean, machine translation used by a professional translator has a completely different impact on the end result than the one used by the general public who doesn’t know the quality of that machine translation.

Thank you so much for this.

Jamie You’re very welcome. This was fun!

**End of conversation**