

Just Adding Value Using Voice Identification

Introduction [00:00:01] RTI international's justice practice area presents Justice Science.

Introduction [00:00:10] Welcome to Just Science, a podcast for justice professionals and anyone interested in learning more about forensic science, innovative technology, current research and actionable strategies to improve the criminal justice system. In episode four of our case study season, Just Science sat down with Josh Yonovitz, expert witness and forensic audio. To discuss the history of utilizing voice identification and investigations and the current state of forensic audio. In the 1960s, voice ID started being used in forensics, but the original voiceprint analysis used was proven to be inaccurate. Nowadays, forensic scientists use a methodology known as oral acoustic speaker identification, which is scientifically accepted but poorly understood. Listen along as Josh describes components of forensic audio, the software and training needed for forensic voice identification, and how forensic audio analysis has helped solve cases. This episode is funded by the National Institute of Justice's Forensic Technology Center of Excellence. Some content in this podcast may be considered sensitive and may evoke emotional responses or may not be appropriate for younger audiences. Here's your host, Jaclynn McKay.

Jaclynn McKay [00:01:17] Hello and welcome to Just Science. I'm your host, Jaclynn McKay, with the Forensic Technology Center of Excellence, a program of the National Institute of Justice. On today's episode, we will discuss forensic voice identification and the current state of the science here to guide us. And our discussion is forensic audio examiner Josh Yonovitz. Welcome, Josh, thank you for talking with us today.

Josh Yonovitz [00:01:40] Thank you so much. It's such a privilege to be here.

Jaclynn McKay [00:01:43] Josh, can you provide our audience with a little bit about your background and how you got involved with the world of forensic voice identification?

Josh Yonovitz [00:01:50] It actually started with my dad. My dad was a researcher at the University of Texas in the seventies, eighties and nineties, and sometimes attorneys and law enforcement departments would come to him for help with audio forensic and voice I.D. issues. And he got kind of a name for himself because he was involved in the Waco trial in Waco, Texas, with the Grant Davidians involved in the firearm identification based on the audio recordings from that trial. And he started his own forensic laboratory with a business partner of his named Herbert Cho. And in 2010, I got my bachelor's degree and I was going from you know job to job and hadn't really found something that I wanted to do yet. And he asked me to help him out at the lab and work for him. And it started part time and it was just sort of minor kind of scut work that slowly started to develop into more of an interest and more responsibilities. In 2013, we were retained on the George Zimmerman trial and that's when things really picked up and I took a much larger role in the laboratory. I was very involved in that trial. Then I started to become more involved in our research and publications and presentations, got my degree in forensic science from the University of Florida, my master's degree. And now my dad passed away a couple of years ago. So I have started my own forensic laboratory after closing his cases. It's called adept forensics, and I've been operating it for about a year now.

Jaclynn McKay [00:03:27] Nice. Could you tell our audience what exactly is forensic voice identification?

Josh Yonovitz [00:03:33] So forensic voice identification is sort of an umbrella term. It really includes three different areas. The first area is voice recognition, which is a automated process and mostly has to do with security and isn't really a forensic process. I'm just going to put that to the side. The second area is ear witness identification, which is exactly what it sounds like it is. If a person is witness to a crime and believes they can recognize the perpetrators voice and identify them in an ear witness lineup. Experts like myself are obviously not ear witnesses, but we are sometimes asked to determine the reliability of the statement of an ear witness. And lastly, the real meat and potatoes of the field is forensic speaker identification, which is when an expert examiner is given a recording in which the identity of the speaker is not known. And that's a forensic value. And then we determine the identity of the speaker by comparing it to known voice samples that we collect or generate.

Jaclynn McKay [00:04:39] With regards to forensic speaker identification. Would you mind elaborating on some of the strengths as well as limitations of this forensic discipline?

Josh Yonovitz [00:04:47] Well, the strengths are obviously these days everyone has a voice recorder in their pocket. Crimes have never been recorded with so much frequency as they are now. So the need to determine the identity of a perpetrator or to assign correct liability in a civil case has never been higher than it is now. In terms of weaknesses. It can be difficult to collect the necessary evidence to compare to voice samples. For instance, people are not always willing to provide a voice sample which might be later used to incriminate them as part of an examination. There's also a natural tendency for people to disguise their speech. Even people who aren't guilty, studies have shown, will try to disguise their speech when giving known voice samples that experts like myself might use to compare to unknown samples. It's just a human tendency people have.

Jaclynn McKay [00:05:40] That's really interesting. So the forensic voice identification used today is much different than what was used back in the 1960s when it first came online. Can you describe what the previous technology was and why this was inaccurate?

Josh Yonovitz [00:05:57] Yeah, it started with someone you might have heard of named Alexander Graham Bell. Not many people know this. His wife was deaf from the age of five, profoundly deaf, and he couldn't stand sign language. And it was really his mission in life to try to forcefully integrate deaf people into hearing society. He established Bell Labs in New Jersey with the goal of making language visual so that deaf people wouldn't have a reason to keep using sign language. In the 1960s, an engineer at Bell Labs named Lawrence Casta. He thought he found a forensic application for this visual speech technology. At the time, they called them voice prints. We call them audio spectrograms today, but it's basically the same thing. And he thought that if you take two voice prints and line them up together, same way you do with fingerprints, you could figure out if they're from the same person or not. And he believed that every person had an intrinsic voiceprint characteristic that was unique to that person. And he loved to talk about voice prints and fingerprints as being equally reliable, which was very appealing to the courts and forensic scientists at that time. He also wanted to prove that his method was easy to do and that anyone can do it. And this was 1961, 1962, which was an awful, terrible time to be alive. And he he tried to prove that in the worst way. He got a group of 18 year old women and he put them in a room together and he taught them how to read voice prints. And then he made them do comparisons. And then he said, look, I taught young women how to do this. Anyone can do it. And he reported a 3% error rate from this experiment, which was extremely good. And then he started to tour the country. He started to go all over the country from trial to trial on his own dime, testifying in trials, trying to use his method of

voiceprint analysis and voice ID cases. Eventually, the Court of Military Appeals in 1967 passed a Frye challenge for him, and it was considered an acceptable method at that point. He got a couple of followers, a couple of people, someone named Oscar Toshi, who was a professor at the University of Michigan, and Ernest Nash, who was a Michigan detective. They were best friends and they traveled around together, going from trial to trial, pushing this voiceprint analysis method exactly the same way. Lawrence Kersten And that's when some speech scientists started to notice what was happening. Real speech scientists, people named Empress Flosser and Baumbach were the the Big three, and they did their own repeat studies of Lawrence Kirsty's experiment. But the error rate they got wasn't 3%. The error rate they got was ranging between 40 and 60%.

Jaclynn McKay [00:08:54] Wow.

Josh Yonovitz [00:08:54] Yeah. When they introduced disguised speech, which is very prevalent, it's when a person tries to hide their voice or make their voice sound like someone else's. The error rate jumped up to about 70 to 77%. So they started chasing these voiceprint analysis experts around the country, kind of like in the election of 1860. Abraham Lincoln would chase Breckinridge around the United States at every campaign stop. Lincoln would turn it into a debate. That's what these speech scientists did. Every time one of Oscar, Toshi or Ernest Nash would testify, one of these speech scientists would show up and they would argue with him. And sometimes they would be successful and sometimes they wouldn't. And this went on for more than a decade. But then eventually a wave of successful appeals came in, provable false convictions of people who were convicted using voiceprint analysis. To this point, Toshi and Nash kind of faded into the background like they were never there. And that's when the Innocence Project started up in 1990 and 1991, and DNA evidence was getting started. And that wave of provable false convictions got a lot bigger. And that was the end of voiceprint analysis. It is not an acceptable science anymore. Nowadays we use what's called the oral acoustic method of speaker identification, which was developed by a University of Florida professor named Harry Hollin, and it has completely replaced voiceprint analysis. It is a approach that combines acoustics with linguistics and phonology, which is a study of speech sounds.

Jaclynn McKay [00:10:37] That was beautiful, really eloquently put together. So, Josh, you said that oral acoustics Speaker identification is now the current methodology. Would you mind just elaborating more on that for us?

Josh Yonovitz [00:10:49] Sure. The idea of this system is that you break down speech to all of its component pieces and look at them individually. So whereas Laurence Kersten was only interested in the frequency that someone's spoken and the visual representation of that, we look at how people articulate their speech sounds, how they articulate their consonants, how they articulate their vowels, which is very important because it's very hard to disguise your vowels as opposed to your consonants. We look at a mode of characteristics. For instance, we consider those speaking rate and we look at melodic pattern. The best way to understand the difference between speaking rate or prosody and melodic pattern is to think about William Shatner. He can talk quickly, he can talk slowly, but he has a very distinctive pattern to his speech. William Shatner always talks like this. We also look at linguistic observations. These are how words are formed in different parts of the body. We look at of voices of resonance, which is determined by resonance chambers in your chest, in your mouth, in your sinuses, and in your cranium, which is very important because you do have some control over the other residents chambers. If anyone ever told you to speak from the diaphragm, that's what they're referring to. But you don't really have any control over the one in your head. Your nasality, which is the amount of

airflow that goes through your sinuses with more nasality. You sound like this vocal fry, which is the amount of air that passes through your larynx. The less air that pushes through the air your larynx, the more you sound like this vocal fry. And we also look at acoustic criteria like fundamental frequency.

Jaclynn McKay [00:12:41] Mind blown. So you said that it's harder to disguise your vowels. Can you give us an example of what that might look like?

Josh Yonovitz [00:12:49] Yeah. Well, the idea of vowels, the reason the vowels are different than consonants is that consonants are speech sounds where your airflow is either blocked off or constricted. For instance, if you're making a P sound, as in pat for those at home, go ahead and say Pat and pay attention to what your what your lips are doing. You're closing your lips just to the end of that sound. But then think about the R in pat. You're just opening your airway and air is just going right out. You have sort of less manual control over it. So someone who is trying to disguise their voice or impersonate someone else has less control over how they articulate their vowels than their consonants or some vowels.

Jaclynn McKay [00:13:32] So I'm curious. In the different regions of the United States and all across the world, we have various accents. Are the changes in the accents focused mostly on the way people pronounce the consonants as opposed to the vowels? Or is that just a completely different concept?

Josh Yonovitz [00:13:48] It depends. It is a dialect, and accents happen more in the brain. And vowels can be formed differently. In consonants can be formed differently. For instance, Canadians. Have you ever heard of the Canadian rise of the Canadian lift? It's the. When speaking, it confuses Americans because we think they're asking us questions. They'll say, I went to Tim Hortons and tried to get a cup of coffee. And then when they're done with their statement, when they're ready for you to say something, then then it drops and they'll say, But Tim Hortons was closed. Now you can talk. So that's more of a behavioral. And it is certainly something that we do look at.

Jaclynn McKay [00:14:29] It's really interesting. So with all these things that you're looking at when it comes to forensic voice identification, can you discuss some of the tools that you need in order to do this and who is actually qualified or how you become qualified to actually do this type of analysis?

Josh Yonovitz [00:14:46] It's good news. It's really good news and it's terrible news. The good news is that you actually don't need much equipment wise to get going. In fact, I could make a very functional computer system that you can do top tier of voice ID work on with only free software. And the reason for that is that the software tools that we use comes from the field of speech language pathology. Because the methods of voice ID are very similar to the methods that speech language pathologists use to diagnose speech disorders. And that software is available for free. My favorite is a piece of software from Denmark called perhaps that is really excellent software and it's completely free. The downside, the terrible news is it takes quite a lot to be qualified to do forensic speaker identification. That's because of people like Lawrence Casta. The field has kind of a terrible legacy. So now we really protect that legacy very jealously. So in order to be qualified, you really need to have some kind of higher degree in a related field. You need to do coursework in acoustics, linguistics, understanding of phonology, and you need to have been supervised. The IAI requires that a person has been supervised for at least five years, performed at least 100 comparisons and 40 different cases under supervision,

which can take quite a while to build up that kind of profile. So that's the good and the bad of it.

Jaclynn McKay [00:16:15] Josh, you mentioned that forensic voice identification is an umbrella term and there are some subset disciplines underneath of that. Would you mind giving us a case example of what ear witness identification would look like?

Josh Yonovitz [00:16:30] I would love to. I worked on a case a few years ago that involved a defendant who was a member of a neo-Nazi group. It was an organization that made the news a few years ago because they were swatting black churches across the American South. For anyone that doesn't know, swatting is when you call 911 and report a dangerous situation at a place with the hopes that the police will harass that place and potentially there may be an overreaction or a misunderstanding that causes people to be injured or killed, which has happened. It's very dangerous. This neo-Nazi group also targeted a university, the university where the defendant attended as a student. The defendant was using an app on their phone that would disguise their phone number when they called Campus 911, which they did. And they initially said that there were pipe bombs hidden around campus. And then they later said that because they were depressed, they were going to start shooting people on campus. The 911 operator was a student who answered the phone. And by asking questions like, What major are you in? You sound smart. Are you an engineering student? What's your favorite classes? Do you have a lot of friends? Do you have a girlfriend? She encouraged this person to stay on the phone with her for a full 45 minutes before he hung up, and then he felt compelled to call back and stay on the phone for another 20 minutes. A few days later, the defendant accidentally dialed Campus 911 on his phone without using that app, which disguised his number. And this was probably the most unlucky person I'd ever encountered because he got the same 911 operator and he said, Oh, I'm so sorry, I accidentally called you and hung up. And she recognized his voice. And it wasn't just that he got the same operator. He had no way of knowing this. But she wasn't just a student. She was a graduate student in speech language pathology in her second year. So she had advanced academic training in linguistics and acoustics and phonetics, the same things that we look for in an expert in voice I.D. One of the major criteria for a successful ear witness identification is familiarity with the person. And she talked to him for over an hour. The amount of time between hearing the person and making the identification is very important. And he only waited a few days before he accidentally dialed that number. So it was a perfect ear witness identification. In this situation, we were hired by the defense to determine if her identification was reliable. We concluded that it was. And we did our own comparative analysis, our own speaker identification, and did determine the defendant was, in fact, the caller.

Jaclynn McKay [00:19:32] Wow. That's amazing. It reminds me of the scene in My Cousin Vinny where the individual gets on the stand and she's supposed to be an expert in automobiles.

Josh Yonovitz [00:19:43] And that's exactly.

Jaclynn McKay [00:19:44] The right person. Yes. That's the exact type of person that you want to answer that phone call. So let's kind of pivot to Speaker ID. Do you have a case example of that?

Josh Yonovitz [00:19:54] I do a very difficult one. This was a civil litigation, and this also had to do with disguised speech, which makes it more difficult in this situation. A very

wealthy elderly man had created a living will through his bank. And then he passed away and his two living relatives were his second wife and his son. His adult son. His second wife was a step mother to the son. The stepmother believed that before the man's death, his son had called the bank, impersonating him with his Social Security number and bank account information and pretended to be him impersonating his voice and had the will altered just before his death. This was a remarkably difficult situation because fathers and sons have a tendency to sound very similar to each other. But this was a situation where we were fortunate because in order to get to a representative, when you call a bank, you have to give the automated service quite a few numbers. Social Security numbers and bank numbers. And even though the son was unwilling to provide voice samples to us, those numbers were perfect for comparative analysis because numbers have a high vowel to consonant ratio, and it's very hard to disguise your voice when you're using vowels. So when he's saying three, four, eight, nine, it was really all we need to make a yes or no determination. Is this the elderly man who would shortly pass away? And in this case it was not. And because his son was not willing to provide samples, we could not say it's the son, but we could certainly say it's not the father. It's definitely a different person called and made the request of changes to the will that greatly benefited the son and left the stepmother with nothing.

Jaclynn McKay [00:21:49] Wow. So when you ask for voice samples, what sort of things are you having people say in order to capture these samples?

Josh Yonovitz [00:21:58] Yeah, that's a really great question. The first step is when making what we call known exemplars, which is known samples of speech where the identity of the speaker is known. The first step is to just strike up a casual conversation with them. You know, was it hard to get here? You know, a good movie is what you do for dinner last night, things like that. It helps establish a baseline and it helps with disguised speech detection later because they might not have enacted a strategy yet. The next step is we build a linguistic inventory for the person. There are some statements that will have them read. There are about 3 to 5 paragraphs long. My favorite is a linguistic passage called the Rainbow Passage, which is from a children's book about colors, kind of introduces children's to color theory, but we use it because it has every speech sound in the English language. And so by reading it, we can build a full linguistic inventory of how this person speaks, and then we will ask them to repeat portions from the unknown recording word for word. Being very careful not to read it out loud to them, because people do have a tendency to repeat things the way they last heard them. Certainly no coaching. It's just words on a piece of paper that they read in a fairly monotone voice and will make them repeat that 3 to 5 times.

Jaclynn McKay [00:23:19] So you said you like to use that passage because it has all of the phonetic sounds in the English language. So what other sounds exist in other languages that you might have to pivot from using this passage?

Josh Yonovitz [00:23:31] There are so many the the best representation of that that everyone knows is the rolled R that's in Spanish. Mandarin has a collection of phonemes or speech sounds which are not in the English language, which I will not try to do because I do not speak Mandarin. But fortunately, you don't need to speak a language to perform a voice identification of it because of this linguistic inventory that we built. But it is important to find a linguistic passage that does have all those speech sounds when working with other languages.

Jaclynn McKay [00:24:01] You mentioned that the ear witness had advanced training in some of the disciplines that forensic audio examiners use. Could you describe a little bit more in detail about how you evaluate the reliability of ear witness identifications?

Josh Yonovitz [00:24:17] Yeah, absolutely. It's very similar to eyewitness identifications with a few differences. One, people tend to remember voices in a traumatic situation much better than they remember faces, which is very, very beneficial for criminal investigators. And secondly, when someone is unsure about ear witness identification, they have a tendency to not guess like eyewitness identifications do. Eyewitness identifiers often feel a need to try to be helpful and try to make their memory change to fit the people that are in front of them. And there's less of a tendency for ear witnesses to do that. The reason for that isn't very well understood yet, but the research is very conclusive about it, which is very helpful for criminal investigators. The downside of ear witness identifiers is that it's much harder for them to describe a person's voice because we're not very good at describing people's voices. That's why there aren't voice sketch artists. There isn't a person that can replicate a voice based on someone's description of it. People are also very good at remembering voices over long periods of time. The shorter the span of time between the identification and the ear witnessing of it, the more reliable. But still people are able to remember voices from years and years ago and make identifications. I had a case once involving a fast food chain that's known for breads and soups, and someone robbed a store at gunpoint and then left. And then the person at the counter said, Hey, I remembered that person even though they were wearing a mask. I remember their voice. They did a shift here once, three years ago, and I trained them. And very remarkably, it was an accurate ear witness identification, which led to direct evidence that convicted the person.

Jaclynn McKay [00:26:10] Wow. That's fascinating. So based on the study you were talking about how people remember voices better than faces during traumatic experiences. I know that with the way the nose links up to the brain, smells can trigger memories. So is remembering voices versus faces have anything to do with how the eyes and the ears link up to the brain?

Josh Yonovitz [00:26:39] Yeah, absolutely. It's about the chemical process that stores memories in different parts of your brain and the process isn't the same for each part. So your visual processing center stores information in a slightly different chemical way than in your auditory and language processing center. And that's really the most important part, is that there's two parts of your brain that form your auditory and language functions, but they both store speech information, which is one of the reasons why trauma, which does have often a profound impact on people's ability to recollect faces, has less of an impact on people's ability to recollect voices. It can also be harder to retrieve that information deliberately, though, which is fine because in an ear witness identification, ideally you would hear the perpetrator's voice again and then something would click and you would remember it and you would remember it with certainty.

Jaclynn McKay [00:27:38] That's really interesting for those listening. If they want to start paying more attention to the voices that they're hearing and trying to describe them. What are some things that they should take note of or what would be some things that would be pertinent to an ear witness identification when you're listening to voices?

Josh Yonovitz [00:27:59] I would suggest that they think if there were any words that the person mispronounced, if they had a dialect or an accent, how quickly they were speaking on what words they stopped, did they maybe stop and start a sentence easily to remember

details like that is what I would suggest. And if you're speaking to an ear witness, those would be great questions to ask them before doing an ear witness lineup, which is what it's really called.

Jaclynn McKay [00:28:28] Interesting. So kind of along those same lines, for those who have never heard of forensic voice identification prior to this episode, like myself, what are some key takeaways that you would like them to leave here knowing?

Josh Yonovitz [00:28:42] Forensic speaker identification especially is an area that has a real terrible legacy to it because of the actions of those people that I discussed at the beginning of of this meeting. But it's has it's very reliable now. It has very good track record. It reliably passes Frye and Daubert challenges. It meets those acceptability standards, the confidence interval. There's a lot of studies of the confidence interval, and it's reliably placed at about 0.96, which is very good. The error rates are extremely low. It is a reliable method now, but what's critical is finding an examiner who is actually qualified to perform the function. Unfortunately, this is not a method where knowing the methodology and having the right tools is enough. You do have to also have a formal education in the topic and had some supervised specialized training to perform it, which is very important. But if you don't meet those requirements but you still need to include Speaker ID in an investigation, there is a path for you. You can reach out to local universities and put together what's called a listening panel. This is something that law enforcement agencies do and it's considered to be reliable. Usually you would reach out to linguistics departments or speech language pathology departments. You're probably going to have to fork over some cash. Be willing to pay some grad students 50 or 60 bucks for an hour of their time and explain the task. And there's a specific way that you need to prepare the materials for them. But you can ask them to make a determination.

Jaclynn McKay [00:30:23] If investigators wanted to start incorporating this type of analysis in their investigation. Where would they begin?

Josh Yonovitz [00:30:29] Well, reaching out to a local university that has a program in speech language pathology or linguistics is a great way to build a listening panel, which is a good approach. But the best way is to reach out to someone who actually is a practitioner of forensic speaker identification. We're happy to talk to you. We can talk to you about collecting materials, making exemplar recordings, the right process and methodology to do it. If you're a research it yourself kind of person. Dr. Harry Hollin wrote a couple of books on the subject. Forensic voice identification is sort of the Holy Bible of forensic speech ID, so I think that's a great place to start.

Jaclynn McKay [00:31:09] What are your hopes for the future of forensic voice identification and are there any other areas that still need improvements?

Josh Yonovitz [00:31:16] Yeah. So forensic science in a lot of ways is a natural science. And one of the most interesting competitions in nature is between bats and moths. Bats, everyone knows, developed the ability to echolocate, to navigate through caves and to find their prey, which are usually moths. Moths develops the ability to make clicking sounds at the same frequency that bats use to echolocate, which disperses the echolocation, pulses and hides the moths just the same way that submarines can sometimes hide from sonar. So bats over millions of years evolved to modulate the frequency that they send those echo locating pulses out so those clicks can't disperse them. So then over millions of years, moths developed these incredible attachments to their wings and tails that disperse those sound waves and scatter them in random directions. And the back and forth just

goes on and on and on. Every time bats develop better detection moths develop better evasion, detection, evasion, detection, evasion. And that's the same situation that voice ID practitioners are in now with people who practice disguised speech, hiding their voice, or more insidiously imitating the voice of someone else, disguised speech and voice. It has been an issue since the the Lindbergh baby kidnaping where the kidnaper shouted ransom demands to Charles Lindbergh over a park and in the middle of the night doing a terrible Scottish accent. And Charles Lindbergh was still able to identify his voice in a courtroom saying that was the guy that's the guy that kidnaped my baby. But now we're at a technological precipice where deepfakes and A.I. technology give people the ability to type a sentence and have it be read in someone else's voice. And it will take a credible amount of research to keep up with and combat these very rapidly advancing technologies. It's evasion and detection again, just like in nature. Right now, it's very easy to detect deep fakes and it's very easy to detect A.I. speech, but it won't stay that way much longer. And there's very little research being put into the subject. So I think that the future in this field is in our ability to respond to these advancements. I think that this area isn't just important legally and forensically, but it also affects our entire sense of reality. If we don't have certainty that a person's familiar voice is actually them or not.

Jaclynn McKay [00:34:14] Josh, I know you gave a presentation here at the IAI conference on Forensic Voice ID. Were there any questions that came up from the audience members that you wanted to discuss on this podcast?

Josh Yonovitz [00:34:25] Yeah, there was there was I received a question at the end of my presentation that was very good and it had to do with the length of voice samples. They wanted to know if if you could do a speaker ID based on a small voice samples, maybe one sentence or just a few words. And the answer is it depends. There used to be an organization called the ABRE, the American Board of Recorded Evidence Examiners, but they abbreviated to ABRE and they don't exist anymore. But they in the nineties developed what they called the golden rule, which is that you need at least 16 seconds of speech in order form speaker ID. And this was the issue in the George Zimmerman trial. That's what made that trial forensically significant for the field. Because in that instance, for those that don't know, George Zimmerman instigated a situation which caused a violent altercation with the victim, Trayvon Martin. Neighbors, people who are across the street from the fight called 911 and threw their phone through the closed window. And across the street, you could hear someone screaming for help. And it became very relevant to the trial, which one of the two were screaming for help. And there were four experts in that trial, including us. And all four were saying different things because of the 16 second issue. There was only about 4 seconds of speech in that recording, not 16. It was our position that in this situation you can do a voice identification because there's only two possible people they could be, and their voice profiles were very different from each other. It was the opinion of another expert. I won't say their name, but it was their position that identification could not be done. They ultimately failed their Frye challenge and were determined to be a not credible expert witness. And our Frye challenge was successful. And so the methodology changed. As long as you have a limited number of people for whom could be the speaker and as long as their voice characteristics are significantly distinct from each other, you can do a voice comparison using a small number of words.

Jaclynn McKay [00:36:53] Before we close, do you have any final thoughts for our listeners?

Josh Yonovitz [00:36:56] Yeah. Forensic Speaker Identification is just going to become more and more relevant as time goes on. More and more of our lives are recorded and

there's a need for more and more certainty as to who's speaking. Especially in legal and forensic issues right now, there is no university that offers a program in forensic voice identification. There's no certification board that has certifications for forensic voice identification. Right now, there are very strict requirements for it, but someone might not know what those are until they're in front of a judge and at the mercy of a Frye or a Daubert challenge. So one of my hopes for the IAI going forward and for the fields in general is for a return to a formalized education in forensic speaker identification to give space for the next generation of forensic voice examiners.

Jaclynn McKay [00:37:58] Josh, thank you for your time today. It has been a pleasure discussing this topic with you.

Josh Yonovitz [00:38:03] Thank you so much. This has been a wonderful experience.

Jaclynn McKay [00:38:06] If you enjoyed today's episode, be sure to like and follow just science on your platform of choice. For more information on today's topic and resources in the forensics field, visit [Forensic COE dot org](http://ForensicCOE.org). I'm Jaclynn McKay and this has been another episode of Just Science.

Introduction [00:38:26] Next week, Justin sits down with Dr. Justin Shore and Timothy Primrose to discuss cases that highlight the use of forensic technology. Opinions or points of views expressed in this podcast represent a consensus of the authors and do not necessarily represent the official position or policies of its funding.