# Sound Scene Classification in Hearing Aids
Recorded September 10th, 2019

Presenter: Donald Hayes
AudiologyOnline.com Course #33830
Partner: Unitron

- [Donald] The presentation today is about Sound Scene and Spatial Speech Classification in Hearing Aids. So we have the learner outcomes here, are all listed, as far as what you should be able to expect out of the course. However, you've probably already seen this when you signed up for the course, so let me go to the beginning. So to start with, I just wanna begin with the question of what do you think a modern hearing aid really is? And I know as a clinician, and I was a clinician, I worked in a clinic for 11 years in a busy teaching hospital, when we look at hearing aids we tend to view what goes on in there the way we were trained by our professors at the universities. You've got a mic, you've got an amplifier, you've got a receiver, you've got a vent. There's a very few specific parts. However, modern hearing aids really don't actually work that way anymore. A hearing aid is really now much more of a system. You can't really look at things in isolation. Even though we show them to you in the fitting software like the different elements of an automatic program, the speech enhancement, the noise reduction, the direction mics, and you can set them all differently and you can use them like some sort of a menu in a Chinese restaurant and you take one from column A and one from column B and go like that.

In reality, that's not how hearing aids are designed, and it's not how hearing aids actually work. Hearing aids work on the basis of a series of components that have been designed to operate together as a system, and that system is based on a philosophy. So what we're gonna do today is to talk a little bit about the philosophy, but more so about how that impacts things like classification and then later on spatial detection. One of the things that I think most of us as clinicians don't ever stop to think about is the underlying architecture at the bottom of the hearing aid system. We all know that hearing aids can classify signals into speech in noise, speech in quiet, music, noise only, but did you ever stop to think about how that's all done and whether or not some hearing aids are better at it than others? We do think about these things. And so let me just give you an overview of how the whole system comes together in terms of environmental classification or really any component, any larger component of a hearing aid system. Everything begins with the philosophy of the developers.

So Unitron has a philosophy. Phonak has a philosophy. Oticon, Sivantos, Starkey, everybody has their own philosophy not just about what's important, but how the things that are built into the hearing aid underneath all work together to achieve a goal. The consequence of that philosophy is how microphones change from omni to directional to split directional, when speech enhancement or noise induction or noise reduction may engage or disengage, or even how strong the different features are within the processing. When it comes to speaking about classification, there's an additional thought that you need to bear in mind, and that is when you go in and you fit a hearing aid, you look at the individual destinations that a classifier can go to. Conversation in a crowd, conversation in noise, music. And when someone has a complaint about how speech sounds in noise, you will go in and you'll meddle around with the various parameters to try and optimize that fitting for them. And odds are you will do a very good job of improving the values that exist in that particular destination of the automatic program.

However, to quote the VP of R&D here, the best possible signal processing is only as good as the accuracy of the classifier. And the quality of the signal processing decisions that are driven by the classifier results. What he meant by that is if you had absolutely the very best conversation in noise settings for this individual, and whenever that person goes into a loud noisy environment, the hearing aid thinks they're listening to music and switches automatically to the music program, then all your work is undone. If the hearing aid can't classify accurately, and precisely and then direct its components to operate the way you want it to in the correct listening environment, all your work is wasted. So we looked at that and we asked the question, how well does our classifier work? So this whole talk began with our own benchmarking. Why did we start with benchmarking? And the answer comes back in Unitron's case, to Log It All.

Log It All is a feature that we have that basically data logging tells you what the hearing aid is doing in any given situation. What Log It All tells you is how did we acoustically classify that situation? Did we think of that situation as a, did we think of that situation as a conversation in noise? Or conversation in a crowd? And how much time have you spent in each situation? But then the question arises, well how accurate is Log It All? If your classifier is inaccurate, then your Log It All results will be inaccurate. And that's a valid question.

So in answering that question we did a few different things. One of the first things we did was we did a Global Listening Environment Study. This took place a few years ago, maybe 2016 or so. And what we did was we went out to 10 different countries. Everything from Australia to the USA. We went to clinics where people were fitted with our hearing aids, and all we did was we got from them the records of the fittings for the number of people that you see. So Australia, there were 134 fittings. The Netherlands there was 141. The USA, there was 208. At multiple different clinics around each country. So this led us to a total of 1055 records. Everybody who was fitted and we used, oops, sorry, whose results we used from those countries had worn the hearing aids for a minimum of two weeks each, and we looked at the data logging results to make sure that they wore the hearing aids for a minimum of 10 hours or so a day. Anybody in fact with an average of less than six hours per day was excluded from the later study. And what we tried to do was we tried to find out if you looked at all the different aspects of the people who were wearing the fittings or wearing the hearing aids, what were the listening environments that they were in?

So, one of the first things we looked at was age. Because as a clinician, when somebody who's 55 years old walks through the door and says I want a pair of hearing aids, versus somebody who's 85 years old and walks through the door and says I want a pair of hearing aids, in my head I automatically assume that the guy who's 85 is in very different acoustic scenes from the guy who's 55. And that's true. They are in different acoustic scenes. But are they in similar acoustic environments? That is the question. So to answer that question, if we just look at people who are in the age range from 50 to 60 or the age range from 70 to 79, what you're looking at here is two graphs. The graph on the left is younger people who were in the age range from 50 to 59. Each line represents one person. If you look along the bottom axis of the left line you'll see conversation in quiet. Conversation in a small group, conversation in a crowd, conversation in noise, quiet listening, noise, and music. These are the seven listening environments in Log It All. For each person, we looked at the average amount of time they spent in each one of those seven listening environments. And that average is shown here as usage over here on the left on the Y-axis. So the more time you spent in a listening environment, the higher up the scale you would go. Like this particular character right here, who was 98% of the time in quiet.

And for those of you who aren't familiar with Log It All, if you ever see one of those, that tells you he left your office, left the hearing aid batteries in, put the hearing aids in the box, set the box on his nightstand, and didn't wear it for the next two weeks despite whatever it was he told you when he came back with the hearing aids at the end of the two weeks. The only way you get 98% absolute quiet listening, you couldn't possibly have worn the hearing aids. So, what we see here is on average, quiet listening is the most common environment, rarely that much. The next one is conversation in a small group. And then the various noise environments are less time. But what's more important is look at the individual scores. And individuals within this group of 50 to 59 year olds actually vary fairly widely. If we look at the same group or the next group, which is the 70 to 79 year olds, what you see is inside the group, they also vary extremely widely. But on average the pattern looks quite similar. You see a little bit more time is spent in quiet listening and a little more time, or about the same amount of time in conversation in small group, a little bit less time conversation in a crowd and a conversation in noise than what you might see over here. But that less time is really due more to the overall range of individuals than it is the average score. If we now look at the averages for people across different ranges and we average together everybody in a group, now what you see is a very familiar pattern. We have each line represents one age group from 40 to 49, up to 90 plus. You notice the 90 plus has this crazy amount of time in conversation in quiet. The 90 plus group unfortunately only consisted of about 12 people, so it was a very small group. The rest of these groups had anywhere from 150 to 250 people in the group. And if you look at the average amount of time they spend in each listening environment, you will see that on average, the pattern stays very similar across all age groups. Although there are slight differences as people get younger they spend a little less time, maybe 8% less time in quiet. And they spend seven or 8% more time in a conversation in noise or conversation in a crowd.

But aside from that, the patterns are quite consistent. So there's a couple of meanings that you could draw from this. What we saw was that it didn't matter what age you were, that pattern was pretty consistent. It didn't matter what country you came from, the pattern was consistent. Didn't matter if you were male or female, the pattern was consistent. It didn't matter if you came from a small town or rural area, a big city, or a suburban environment, the pattern was consistent. Hearing aid experience also had no impact on the pattern. So, the first thing we learned from GLES was the results from the hearing aids were extremely consistent no matter what. Which leads me to the next obvious question. Is that because people's overall experiences on average are roughly the same distribution regardless of all these factors? Or is it that our classifier just sees the same distribution regardless of the person's daily life style? Which gets to the point of this, the next benchmarking study that we did.

And that point is, is our classifier actually classifying it and people really are in very similar environments? Or is there something wrong with our classifier? So to answer that question we needed a controlled environment. What we did was we did a benchmarking study at the University of South Florida. It was a multiyear project. Our two principal investigators were Dr. David Eddins and Dr. Erol Ozmeral. They're both professors at the University of South Florida. We worked closely with them. That picture on the right is for the room with the speaker array where we tested all the hearing aids. That's a speaker array of 24 different speakers in 360 degrees azimuth all the way around. The first thing I'm gonna show will be the classifier performance. We also looked at azimuth detection and we did a take home comparison.

I'm not gonna go into that level of detail initially. So with respect to classifier performance, what you have to do is you have to come up with some sort of a sound parkour or some sort of a simulated listening environment where you have a known quantity of different listening environments that are actually representative of the real world. So let's start by looking at the sound parkour. The sound parkour was a series of individual recordings that were done and we could look at them as what would happen if you walked through an acoustic park where the sound environments changed every couple of minutes? The way we would define this is we had to decide how did we want to set up our sound parkour? And to set up our sound parkour we needed rules for how would you define an individual listening environment?

So we defined it first on these two axes. We went from no speech in quiet, which is the bottom left corner of this axis. Up to no speech, but noise. So from left to right, you go from quiet to noisy. And from top to bottom we start with no speech and at the top we have our speech environments. Conversation in quiet, conversation in small group, conversation in a crowd, and conversation in noise. In other words we go from a listening environment with no speech in it from quiet to noise. And with speech from quiet to noise. Another way to see these axes are to go from no conversation to conversation. And from a relatively easy environment to a more challenging environment. So the way that I like to look at it is from simple to complex. The simplest environments are in the lower left hand corner. As the environment get more complex, they go to the upper right hand corner. What you do not see on this graph is music. So we've left the music environment off. But if you see the rest of these other six environments, these are the environments that are in our hearing aid as well. This is the way our classifier classifies sound with the exception of also music.

So if we look at our sound parkour, the first thing you see is we start at the lower left hand corner, which is a quiet fan, nobody talking. Quiet conversation is our next listening environments with one talker or three talkers. And we start in the quiet listening environment with a fan noise and no speech. Here we have no speech and actually, or sorry, one speaker at zero degrees azimuth, right in front of the hearing aids with no noise. This environment is no noise and three speakers. One at zero, one at 300 degrees, and one at 60 degrees. We did this to represent a real world environment, not just by speech and noise and SNR, but also by where the azimuths would be that people would actually be speaking to you from. So if I'm in a one on one conversation I'm almost certainly looking at the person that I'm talking to. But if I'm sitting around a table and I'm talking to three different people in a quiet kitchen, one of them will be at about 300 degrees, one of them will be at zero right in front of me, and one of them will be at about 60 degrees off to my right. So we wanted the distribution of where the signals came from to also be represented in our sound parkour. The noise distribution in this case for the fan is from four different azimuths. Left, right, front, back. There's no noise in the next two. The quiet conversation with one talker or two talkers. There is no SNR in any of the three because it's either noise or speech. The overall level is 40 dB for the fan, which is extremely quiet.

This is all done again in that sound booth. And for the speech it was at 55 dB, which is rather soft conversational speech. So each one of those rows represents one sound file in our sound parkour and we ended up with about, for this example, we ended up with about 35 sound files. So here's how we have changed some of them. If you look at the next two lines, instead of having noise, we've added music. So here we have one talker, three talkers, same as the two above. But now we've added a music. And the talker distribution is the same. Either directly in front or around a table. In this case the music is coming from the right at 90 degrees at a -3 dB signal to music ratio. But still very soft. In other words, what this is is background music from a speaker or something off to the right. And the music is quieter than the speech. So we're sticking with our going from the lower left hand corner of the original example of our theoretical box, to the upper right hand corner.

Now, we add small group conversation with noise and a small group conversation with noise plus music. This time we have multiple talkers. We have different types of noise. A subway, which comes and goes. Traffic. Car noise. Or a food court, which was a recording that we did at the local mall. Or just traffic with music. And here's all the distributions. I'm not gonna go into everything 'cause it'll take me a half an hour just to get through that. The next we have a large group conversation. At this time we have added speech babble in the background rather than traffic or food court noise. Again, the signal levels now are a little higher. We're up to 65, 70, 75 dB. And we have the signal to noise ratio that is positive. That is zero, or that is at -5. Then we went to a large group conversation like before with babble in the background but not with music. And then the last two sound files in our parkour that you, really they got kinda cut off in this slide, but they're both TV listening.

All right. Now let me give you an example of what happens in a sound parkour. What you see here, there's a few things going on. So I'll just start from the top and work my way down. The quiet fan that I mentioned about at the very upper left hand corner of the table before, this is a quiet fan in a sound booth at 40 dB SPL. This is the wave file that shows you the, or sorry, the wave form in the time domain over the last 30 seconds of the sound file. The sound file actually ran for two full minutes. I'm just showing you the wave form for the last 30 seconds. Then we switch to our next row which is the quiet speech with one talker from directly in front. And now you can see the speech wave form. Now this is only 50 dB or 55 dB SPL. This is 40 so the fan is, you can't even see it. It's so quiet. What we have down here, this is the output of our classifier. So there are seven destinations. Speech in quiet, small group, large group, conversation in noise, noise, music, or quiet. What you see when the fan is playing in the background is it's so quiet that you have a 100% probability for that last 30 seconds that you're listening to a fan in quiet. This is as perfect as it's gonna ever get with a classifier. One sound. Nothing else. Extremely simple.

Then we switch from the fan to the one talker. And what you see here at 30 seconds is the probability that this is a fan noise begins to drop off rapidly over the course of the next seven seconds. And the probability that it is a speaker, one person talking, goes up at exactly the same rate because there is only two probabilities at work here. You're either listening to a fan or you're listening to one speaker. And now this is up at 100%. So as one drops down to 0%, the other one goes up to 100%. The reason you don't see any of these other ones showing at all on the graph is because they're all at 0%. This is the simplest case possible. Either a noise or a person talking. This is a transition time frame from when it changes at the end of that 30 seconds to the first 30 seconds of the talker. And although the classifier recognizes it immediately, the output of the hearing aid blends into the next listening environment over the course of about four seconds. At this point we have blended from a soft noise to one speaker one on one going from left to right.

So if you look at the last 30 seconds of the first part, and the first 30 seconds of the second part, that's what it looks like. If we average out the entire two minutes, you will see these two bars. The first two minutes of the fan, the average is ah, where are we here? Noise only. And yet, the color is incorrect. Sorry about that. The color here is incorrect. This is actually a noise only. It shouldn't be red. It should be green. And somewhere along the line this got screwed up when I transferred it into Connect. However the other other is right. This one here shows you a strong probability of speech in quiet with a small probability of what looks like should be noise only, but it actually for some reason, oh no, I'm sorry, it's quiet. It's quiet only because the fan is so quiet that it's registering as quiet only. I was right. I didn't get the color wrong. The reason why this is partially a probability of something other than speech in quiet is this time frame right here. During this time frame you don't have a 100% probability that you're in speech in quiet, you have also a high probability of quiet. So it shows up on a bar graph.

The next few slides that you see will all look like these. So this is a simple case. If we go to a more complex case now, a small group with music, this time what you see is a high probability of a small group for the first, or the last 30 seconds of that wave file. That then drops off to about a 50% probability, meanwhile a large group conversation, the probability goes up. And the reason why this is a small group in a subway is registered partly as a large group is because we named it small group. That doesn't mean that it's actually a small group. We designed it to act like a small group, but there's a lot of noise as the subway car comes and goes. So we call it a small group in our naming of the sound parkour, but the hearing aid is saying this is not entirely a small group, this is some sort of 50% small group, and 50% larger group. And probably the larger group probability comes in when the trains are coming and going during the subway thing. There's also a lot of reverberation and reverberation tends to be characterized as a larger group because it's the same sound coming from a different direction a little bit later in time. So it has more of a characteristic of a large group.

You have to remember the way a human person thinks of a small group is based on the number of people. The way a hearing aid sees a small group or a large group is based on the acoustics of the environment. And since this was down inside of a subway where there's a lot of reverberation and a lot of noise, even though it was only three people talking, it has the characteristics of a larger group as well. So you see that mixed classification. And the hearing aid will react to that as such. The next one is two talkers in traffic versus three talkers in a car. The reason I'm showing you this is now you see there's a lot more variability. Because traffic tends to come and go. Cars come by and then they stop because it was a traffic light and then they speed up again. During this time, over that last 30 seconds, the classifier actually goes up and down a great deal. As it does for the first 30 seconds even inside the car. But you notice the car is a lot more stable because as the car noise rises, it tends to be going in one direction with respect to speech and noise whereas this one, the car noises are coming and going. So you see a lot more variability in the way it's classified over 30 seconds.

The other thing you see when we look at the bars is there's a much wider range of classifications over the full two minutes. Everything from a large group down here to a conversation in noise to a noise only environment when it's quite loud right at this point. The other one here is conversation in a small group. Conversation in a large group. Conversation in noise. And noise only. Although you don't see it here, because it probably happened during the first 30 seconds, and not during the last 30. This one here is what happens in a car. And this kinda shows you where we made a little bit of a mistake in our parkour. We started off with three talkers in a car. And this was average level, which was probably 65 dB. This is three talkers in a car where it's loud. What you would expect would be some sort of combination of the car noise here versus a large group here. But you see mostly car noise and the large group is much less of a factor. Here you see small group conversation with less car noise.

This is a good example of where changing the car noise, but keeping the talkers the same has a completely different impact on the overall classification. But the hearing aids do definitely react appropriately. So that was some examples of the moment to moment classification and how the classifier blends different listening environments. And how not everything is nice and clear cut the way you might like to see it. So let's go on to the actual benchmarking. So when we look at the classifier benchmarking, there are two ways you can benchmark stuff. You can use a gold standard or you can compare against similar devices. We did both. Our gold standard was 16 normal hearing listeners. And our similar devices were four of the top competitor's to Unitron. I will tell you now I'm not gonna tell you who the other competitors are. This is not about who's better than what. This is more about showing you the different philosophies from different manufacturers can yield different results. I'm not gonna say one is definitely better than the other or one is superior to the other, or any of that stuff. That's not the point of benchmarking. The point of benchmarking is to see how you do when you compare yourself against other competitors who have really good results. So, the classifier performance. We did our sound parkour first.

Now, the next thing we did was we asked 16 normal hearing listeners to classify each of the wave files so that they could tell us what they thought they heard. Then we classified our hearing aids. We had our hearing aids and our competitor's hearing aids classify exactly the same sound files and we see how they compare. So let's start with the beginning is the human. Let me just get that out of the way. So basic classification. This is the simple soft speech listening environment. On the left, this bar shows us how our normal hearing college students classified the soft speech, one talker listening environment with 55 dB. So this bar represents our young, normal hearing listeners. And what they told us what it's about 95% speech in quiet and about well maybe four, 5% speech in noise.

For the record, I forgot to mention this before, all the different manufacturers use different names and different listening environments. And they blend them differently. But if you look at all hearing aids, there's four major categories that you can pretty much count on everybody classifying into. Regardless of what they call the different listening environments or how they're grouped, there's speech in quiet, there is speech in noise, there is noise only, and there is music. Although they may classify other things, those four will always be in there. So you can always see when manufacturer's have classified them. So we brought everything from our parkour that we could down into those four listening environments so we have a reasonably good comparison across all the different five hearing aids. So when I say our young normal said it was speech in quiet or quiet listening, that's kind of, we have to use those sort of general terms so that we can compare across hearing aids later on. The Unitron hearing aid classified all of that as speech in quiet or quiet listening. It did not recognize that there was speech in noise there. Hearing aid A, B and C and D all were almost exclusively conversation in quiet or speech in quiet or quiet listening. Hearing aid A also said that there was noise alone, which is how hearing aid A probably saw what the young normals were saying was speech in noise. Hearing aid C classified some of that as music. There was no music in the sound file, so we're not sure about that. But that was what it was. So overall, when you look at soft speech, the simplest case environment, virtually all the manufacturers do a pretty good job of getting this one right.

If we then take the next level and we go from soft speech to loud speech, now our young normals called this quiet listening or speech in quiet. It's not the level of the speech they're rating, it's whether there's any noise. And they're all basically saying this is quiet speech or speech in quiet. Unitron said pretty much the same about 92% of the time, speech in noise about 8% of the time. Hearing aid A classified this almost exactly as it did the soft speech. Hearing aid B almost exactly as it did the soft speech. Hearing aid D, once again, almost exactly like the soft speech. Once again, all four Unitron, A, B, and D classified it the same. Hearing aid C went to speech and noise. Now I don't know exactly why that is. My guess would be that the reason why it went to speech in noise is that we went from soft speech at 55 dB to loud speech at 70 dB. It was exactly the same sound file, we just raised the speech by 15 dB. So this particular device, hearing aid C, apparently may be very sensitive to overall speech level, and may go to speech in noise when the speech level is louder regardless of the presence or absence of noise. Once again, not necessarily the wrong thing to do, just a different thing to do. Now if we look at something a little more complex, this time we have the subway. So we have an 80 dB overall signal with a zero dB SNR and three talkers in the subway. Now a zero dB SNR in a subway is not like setting up a HINT test and doing long term average speech factor. The subway cars come and go. When the subway cars exit, the overall level drops a bit, drops quite a bit. When the subway cars are coming into the station, the overall level goes up a lot. So the average level at 80 dB was zero SNR. However, the long term SNR, if you looked at it over the full two minutes went up and down quite a bit. And this is what our young normals were responding to. Overall they said this is a speech in noise environment about 80% of the time.

At times they said it was quiet listening. At times they said it was just noise. Probably when the cars were coming in and you really couldn't even hear the people talking. And for some reason, they also said there was music. So maybe there was a little music playing in there. I didn't hear it. Unitron reacted to this by saying our hearing aid is probably 85, 86% speech in noise, very similar to the young normal. The rest of the time it was just noise. So we saw two things. Hearing aid A speech in noise about 75% of the time. Speech in quiet maybe 10% in noise only. Hearing aid B and hearing aid D were a little more likely to say more noise and less speech in noise, particularly hearing aid B. Hearing aid C was speech in noise all the way. So when I look at a lot of these hearing aids, they definitely do seem to be more sensitive to signal to noise ratio than they are to overall level, except for hearing aid C. Where you really see that is here. This time we went from 80 dB at a zero SNR to 80 dB at a +10 SNR. At +10, even when the trains are coming in, the speech is very easy to hear if you're a young normal hearing listener. Our young normal hearing listeners still said the same thing. They said it was speech in noise about 70% of the time, speech in quiet about 5% of the time. And noise only about 85% of the time. Or I'm sorry, about 25% of the time. Which is kinda the opposite of what you would expect. But that was how they rated it. We had it down to speech in noise mostly. Hearing aids B and D had this as quiet listening. As hearing aid D did a little bit more than the young normals, but not a lot more. They also had noise and speech in noise. Where B and C fit into this is unlike hearing aid C, which appears to be very sensitive to overall level, what you seem to see here with hearing aids B and D is that they're very sensitive to signal to noise ratio. At a more favorable signal to noise ratio they will switch to processing that is much more consistent with speech only than speech in noise. So comparison to normal listeners and competitors show in a simple acoustic environment all manufacturers do a pretty good job of classification. However you can see the difference in philosophy. Some hearing aids are more sensitive to overall level. Some more sensitive to SNR. And some lock in on speech. The Log It All results are valid as a result of all the sound parkour stuff we did, which is much more than what I showed you.

Our goal in this test was to find out, hey does Log It All actually give you accurate results? And the answer to that question by our reckoning is yes indeed in fact it does. Log It All is very accurate. Most hearing aids are actually pretty good in almost all listening environments. They're sensitive to different things. And they react a bit differently, but they don't do a bad job. Until you get into a real complex environment where things can get difficult for all manufacturers. But the key ingredient for Unitron is we can now rely on Log It All as a very, very reliable tool for acoustic analysis of the environment. So let's look at some real world data. This is a very different study. And this one we did at the University of Western Ontario. This is an ecological momentary assessment. By that I mean, and you're gonna hear a lot more about this out in the world in the next few years, what the ecological momentary assessment is is you wear the hearing aids out in the world, and you actually indicate in the moment how you're doing.

So in the moment I know whether you are doing well or doing not so well. And it's not a question of you having to come back and fill out a sheet in a diary, you're actually putting into an app on your phone hey I'm having trouble. I'm in a conversation in noise, I'm in a restaurant, and now I'm not doing well. And you give me a frowny face or something like that. So that's an ecological momentary assessment. And as hearing aids become more capable with apps, you see a lot more ecological momentary assessments. Unitron, we do it with all hearing aid fittings. Anytime you use patient ratings, you're doing an ecological momentary assessment for those of you who are familiar with Unitron products. So this was a joint effort between UWO under a large grant called the Ontario Research Foundation Grant. It's a five million dollar, or sorry, nine million dollar, five year grant that we participate in with Western. Form the University of Western at the National Center for Audiology we had Dr. Susan Scollie and Dr. Danielle Glista. From the geography department we had Dr. Jason Gilliland and Dr. Tayyab Shah, who has recently obtained his PhD in part due to the work that he did here.

We partnered with geography because geography is really good at big data analysis. And these guys generated an enormous amount of data. And I can't cover it all. I'll just give you the highlights. So what we did was we took six grad students and we had them wear the hearing aids for six hours each over the course of one day. Which in total, when we were done it was actually more than six hours for some. And it came out to a total of 54 hours. We sampled what the hearing aids were doing every second for the entire time that they wore them. They were carrying a laptop, wearing hearing aids, using a NOAHLink, they had a GPS system, a LENA system, and cell phones with ratings. So basically we had the GPS, we had the LENA system, we had the hearing aids, and we monitored, every second we monitored what was going on. We looked at a number of different basic things. So if you're ever curious about the listening environments that people are in, we told our subjects to seek out more noisy environments.

So there's a little bit of a, more of a preponderance towards speech and noise and noise only environments. But I just want to show you some of the results that we got. Just basically here's what the environments were because I think this is also interesting because you never really know what the overall level is that people are normally in or what the signal to noise ratio is. So here's the results from that. We'll start with over the 54 hours we look at the signal level by range in dB. So this is a distribution of overall levels. In other words how often or what percentage of time did they spend at less than 20 dB SPL? Or from 20 to 30 dB SPL here. Or from 30 to 40 dB SPL. Or 40 to 50 or 50 to 60 and so on and so on up to 90 to 100 dB SPL. Fortunately this number is zero. Meaning they were virtually never in it. And not surprisingly this number is zero because they were never in that either. Where most of these people spent most of their time, 80% of their time was between 40 dB SPL and 70 dB SPL on average. And they probably spent from the looks of it, over half of their time between 50 and 70 dB SPL. Which is pretty much what you would think an average human being would do. Now this covers everything from brushing your teeth in the morning to riding the subway, to sitting at their desk and working on their computer. The average person, for six people, this is how they spent their day.

Now, if we look at the average level between the ears, because we have the results from the left hearing aid and the right hearing aid, if the average difference between ears is between zero and three dB, that is shown on the backside of this graph. If the average difference between ears was between three and six dB, it's on the front set of bars. Also we show from, again, very soft 10 to 20 dB, up to very loud, 90 to 100 dB. So this is the same range of listening levels. But now it's broken down by the difference between the right and left ear. And what you see is when we get between 30 and about 80 dB, there is on average, somewhere between a zero and three dB difference between the ears probably 75% of the time, almost 80% of the time. So most of the time. These averages are a little weird because they're averaged per listening level. Or overall, but not per listening level. So the difference between the ears being 21% is not what you would necessarily think. But 94% of the time, people are less than three dB difference between the ears. So we have an average between the ears is very, very similar for all listening levels.

And the last one here is the distribution of signal to noise ratios. This is probably the most interesting slide of this bunch. Now these bars are higher at 60 to 70 only because people spent more time in that range. This is the average of the overall amount of time they spent in the range. So this bar is higher. But also what you see is over here is the range of signal to noise ratios from zero to five, which would be extremely difficult to 25 to 30, which is essentially quiet listening. And what you see is very few people spend much time at 25 to 30, but in between 20 and about 50 dB, people actually spend most of their time.

So what this is saying to me is even though we told people go and seek out noisier environments and more complex and difficult environments, most of the time wherever they were they were in a signal to noise ratio that was better than plus five dB. And that's at all input levels. It doesn't matter which input level. Plus five dB is this light blue and that's here on all of these bars. So what you see is people were almost never at a very unfavorable signal to noise ratio. This is something you might want to think about because this is day to day life. Now if you think back to the Log it All results, people only spent about six or 7% of their time in each of the speech in noise environments. So it's not a surprise that they're not in a real difficult situation. But most of the time, people are in fairly favorable SNRs. Which I thought was interesting to get from this data. Now if you look at the SNR ranges by percentage of time, this time now I averaged 'em all together. So what you see is the largest percentage of time, 37.4% of the time, people are between a 10 and 15 SNR. Next to that is either a five to 10, or 15 to 20. So in other words these guys were in relatively favorable SNRs virtually the whole time. Okay.

So now let's look at some other stuff from this study. This is the same table I showed you before. It goes from very easy, very quiet up to very difficult, more difficult and noisy, but no speech. Here is conversation with a lot of noise. Here's conversation in quiet. This is exactly the same set of listening environments, just shown a little bit differently. What we did here was we asked the question, we have a classifier with seven listening environments and if we define a primary class as the probability that you're in that listening environment greater than 75% probability, which means that you're in if it's a 75, 80, 85% probability you're in speech in noise, you're in a speech in noise environment based on everything we know about how the classifier works. But as I've showed you before, there are times when you could be in a mixed environment. Conversation in a large group, a conversation in a small group could be 50% each. So we would define that as a blended environment. We would define a primary class as when you are really in one of those seven listening environments. And the probability that you're in that environment is 75% or greater. That occurred 52% of the time. So about half of the time, these guys were in a primary class. If we look at the Log It All results, you can see they follow that same familiar pattern. Quiet listening about just under 20%. Conversation in a small group, about 12%. And then everything else a bit less. The only difference is those percentages are smaller because we're only looking at the Log It All percentages when you're actually in a primary class, not all the time. If we use this rain drop chart to show the percentage of time in any listening environment laid out the same way that it's laid out up here on this table, then you can see these are the two largest listening environments that are primary classes.

Now what happens if we look at the blended classes? Where do they fit in? 48% of the time we were in mixed classes or blended classes. Now when you look at this, we define our mixed classes as background music as opposed to just music. Noise in conversation in a relatively easy situation. Conversation in an easy situation. No conversation in a complex situation. And conversation in a complex situation. Which is consistent with our table from before. Now on the rain drop chart you can see the largest amount of time was actually spent in conversation in a complex listening environment. And the primary classes actually occurred a bit less than that one. We also have an equal split in music between music only and music and noise. Now we use this information in benchmarking to say how can we update our classifier to make it better? So we will be looking to include this kind of stuff in our classifier going forward into the future. All right, so that's the end of classification. We were in clearly defined primary classes about half the time and the rest of time the acoustic environments were complex enough to be mixed or blended. And those portions are going to vary, not just by the listener, but also by the company's philosophy or the priorities that they set for determining how to define a listening environment.

All right, so I got, I'm down to about five, 10 minutes here. I'm just gonna cover spatial detection fairly quickly and go from there. So once again we went back to the University of South Florida, same room to do benchmarking studies of spatial listening. This was again a multi year project. Same professors. What we did was we presented small snips of speech and we asked the listener to indicate what direction it came from. We had 19 listeners once again. These guys all had hearing impairment. We had speech coming from one of eight different directions all the way around them in 360 degrees. They had to respond to one snip of speech every four seconds for 16 seconds so we could see how they did in terms of recognizing the location of speech from different directions over time. So unaided, this is how they did. The better they did, or sorry, the worse they did, the closer to zero. That would be 0% of correct localization. This is an X here up at 45 degrees that says they were almost perfect. They were almost 100% correct at a zero dB signal to noise ratio. From front they were 60% correct. From the back they were 55. Front to back confusions are the most common ones. Left and right confusions are less common. If we go to more difficult minus six SNR where the speech was quieter relative to the noise, now you see it goes from a 60% correct down to less than 40% correct from the front, and down to just barely 40% from behind. If we look at their overall scores aided, we had two scores.

We used processing where we adjusted the microphones to detect speech from different directions using SpeechPro. What SpeechPro does is it locates speech from the front, the right, the left, or the back and it shifts the target area of the directional microphones to match the direction of speech. With these people wearing these hearing aids, what you see is with SpeechPro turned on, they did quite well. Even from the sides with SpeechPro turned off, they did about the same as they did unaided at zero dB SNR. If we go to +16 dB, or -6 dB SNR, once again they did better with SpeechPro turned on, they still did better with the hearing aids with SpeechPro turned off than they did unaided, which is the green line. This goes from the very front through the back and all the way around to the left hand side. So what you see is from the sides there was an improvement with SpeechPro by changing the directional microphones to the side. From the front, there was not much of an improvement or from the back, there was a slight improvement with SpeechPro on. This is exactly the same thing. We've just taken out the unaided results.

So you can see the improvement relative to zero. The other thing we did was a CNIT test. Which is a little bit more interesting because this is a question of how well can you understand speech from different directions with the hearing aid on and off? So to do the CNIT test, there's a paper coming out very soon for the University of South Florida that describes the test. It's a new speech and noise test designed for hearing aid applications. The noise is presented from all four speakers, front, right, left and back. There is running speech presented from one azimuth. In our case we did it either zero, 90, or 180 degrees. Either the front, the right, or the back. And the purpose of that speech is to bias the hearing aid to recognize the direction of speech as being one of those azimuths. It focuses the hearing aid in the desired direction before the test begins. The speech and noise levels are fixed at plus six dB SNR. And then the listener hears from the same direction as the speech they hear a series of integers. Three integers. And the integers are presented adaptively. If they get the three correct, the next presentation drops a few dB. And they get tested with a different set of integers. If they get them wrong, it goes up three dB. So an integer component of the test is an adaptive SRT test much like the HINT. The speech stays fixed at a constant +6 dB SNR. And the purpose of the speech is to keep the hearing aid focused in the correct direction. The purpose of the numbers is to give the person something to listen to that is not speech, that is not text. So here's how our listeners did when we had 19 listeners. Our first 19. A lower score is a better score. And what you see is the zero line is our unaided score for the test from the front, the side, or the back. If they did better with SpeechPro turned on, then it's a lower score. So they had about a 1 1/2 dB on average improvement. They had almost exactly the same score with SpeechPro turned off as they had unaided from the front. From the side they did about 1/2 a dB better. Once again about the same score. From the back they did slightly worse, or they did about the same with SpeechPro turned on as they did unaided. They did slightly worse with SpeechPro turned off. Which makes sense because from the back with SpeechPro turned off we had a front facing directional microphone.

So speech from the back would not be as good with a front facing mic as it would be unaided. These results were mostly significant. But they were only the first 19. We had trouble getting the next group of subjects, another 11 subjects, so we loosened our criteria and we changed the criteria for acceptance by making it more liberal in terms of hearing loss. This graph shows you the results from 15 subjects who had more hearing loss and 15 subjects who had less hearing loss. It's roughly the same kind of data. So we split people by if they had a three frequency pure tone average of higher than 35 or equal to 35 dB HL there on the left graph. The results are shown for the other half of the group which had a better than 35 dB hearing loss pure tone average on the right graph. Once again, lower is better. So what you see is from the front and the side, people with better hearing actually did better aided than people who had worse hearing. Not a big surprise. From the back though they did about the same. From behind, the better hearing group did roughly the same as the worse hearing group. The blue line shows you how they did with SpeechPro turned on and SpeechPro turned off. Sorry the blue line is SpeechPro turned off, and the orange line is SpeechPro turned on. I had that backwards. So what you see is the scores are almost exactly the same aided whether SpeechPro is on or off for people with really good hearing. But for people with more hearing loss, SpeechPro turned on was significantly better than with SpeechPro turned off for those 15 people. The other thing that happened was the zero line is unaided.

So from the front we see that SpeechPro turned on was almost one dB better on average for the worse hearing people than it was with SpeechPro turned off or with no hearing aid or SpeechPro turned off but for the front facing and side facing people who had better hearing, they did better aided than they did unaided by quite a bit. So they were three dB better aided than unaided 'cause unaided is the zero line. And they were two dB better from the side aided than unaided. So wearing the hearing aid actually improved their speech in noise performance, which shouldn't be a big surprise, even though it's a super threshold test. So these are our take home points about azimuth detection. So if we have any questions, I'm more than happy to answer them. I'll read them as they're typed in and then I'll answer them. Well, I either put everybody to sleep or that was so crystal clear that there are absolutely no questions. I'm not ruling out the first one. I certainly hope that you guys got something out of this. The idea of this was to provide as much background on classification as possible. I'm sorry I had to go a little fast on the azimuth detection 'cause I got off to a slow start. I apologize for that. But I hope that this was beneficial to you and everybody does real well on the test. Thanks, bye bye.