

This unedited transcript of a live webinar is provided in order to facilitate communication accessibility for the viewer and may not be a totally verbatim record of the proceedings. This transcript may contain errors. Copying or distributing this transcript without the express written consent of continued is strictly prohibited.

Maturation of Speech Recognition for Noisy Environments, in partnership with American Auditory Society

Presenter: Lori Leibold, PhD

At this time, it's my pleasure to officially kick off today's course. The title is Maturation of Speech Recognition for Noisy Environments. And we are so grateful to the American Auditory Society. We've been partnering them for many, many years to feature a course that's typically based on a presentation from their conference. And that's where today's course came from.

I invite you to check out American Auditory Society if you're not already a member. They have a wonderful conference in Scottsdale that'll be happening next April. Lot, a lot, a lot of great research with practical application that they share there. And again, very thankful for this partnership at this time. Let me introduce our presenter, Dr. Lori Leibold.

She is senior director of the center for Hearing Research and the director of the Human Auditory Development Laboratory at Boys Town National Research Hospital in Omaha, Nebraska. You can read more about Dr. Leibold on our course page and in our handout. Dr. Leibold, I'll turn the platform over to you. Great. Well, welcome everybody.

I'm honored to talk to you all today about my favorite topic, which is how infants and children learn to hear in noisy environments.

Here are my disclosures. Feel free to read my disclosures. And also here are limitations, risks that you should take a look at as well. And we'll get right into the presentation. Several learning outcomes I want you to be able to describe speech recognition in the presence of competing sounds and how it develops.

And also throughout the talk, we'll be discussing how factors like age, language skills, hearing loss, some other factors might influence the ability to hear in background noise. And finally, at the end of the talk, we're going to talk about development of a clinical tool that might be useful for those of you who work in clinical settings.

So before I begin, I want to acknowledge my great collaborators. I don't really do anything in isolation. I get to work with scientists ranging from engineers to psychologists as well as clinicians.

And here are the key collaborators of the work that I'm going to present today. I have more collaborators, but these are the ones that we've highlighted their work today. I also get generous support both from my institution and my former institution, which was the University of North Carolina at Chapel Hill Boys Town National Research Hospital. And then the funding for the experiments that I'm going to describe today was provided by the National Institutes of Health. Okay, so let's get started here.

The general overview of the talk is I want to start by highlighting several kind of big picture questions that really drive a lot of the research that we do. I think that it will help to provide that structure. So as we go through some of the slides, I've tried not to get into too much nitty gritty detail, but I think it's important to think all the time about what is it exactly that we're trying to answer. A primary question of my research really for 25 years has been to characterize the time course of maturation for speech recognition in noisy environments. One of the things that we've learned is that what a baby hears is not what a 4 year old hears, which is not what a teenager hears, and of course not what an adult hears.

But it turns out that the details really matter. Factors related to both the listener and the stimuli that we use to measure outcomes can really influence those results. And that's one of the main things that I'll get into today. A related question is, can we identify then and evaluate factors that either provide resilience and help people in noisy environments, or might limit how well they can understand speech and other important sounds in noisy environments? And we'll get into this in a little more detail.

But a lot of the factors we think about is age, hearing loss, intellectual disability, how hearing loss is treated, language proficiency.

I think you kind of get the idea of some of the things that we look at and whether those factors are malleable meaning can we change them and maybe improve performance, or whether they're relatively fixed.

Then the final question that we think about a lot, and maybe not as much as we should, is can we balance scientific rigor with real world validity? We conduct and we design experiments in the lab. People like me who are scientists, but I'm not in the clinic with patients like many of you are. And we know from countless studies and observations and reports from clinicians that our measures don't always capture the difficulties that people have in real world environments where it's often noisy and you have all these different sources of sound. From a research standpoint, you know, we really have to balance precision and rigor.

It's one thing to try to design an experiment that captures those factors, but if the end result is just so messy that you can't really interpret the findings, that's not going to be helpful to anyone either. It's pertinent to that. So it's pertinent to audiology clinics, it's pertinent to scientists, and it's a real struggle. And I think we've made a lot of advances in the last few years getting us moving us a bit closer to accomplishing this goal. All right, so the General framework of how I'm going to attack this presentation is I want to provide sort of a conceptual framework and review about 50 years of research, but do that all in three slides.

I'm going to try really hard to make it big picture focused, but I want to set the stage for really experiments we've done in my lab over the past, I guess, 25 years or so. Then I'm going to switch gears and talk about what I consider to be the key scientific findings from our work. So I'm highlighting kind of big, broad themes and trying not to get into the nitty gritty of specific experiments. But of course, if people have questions, we'll be able to answer those at the end. Clinical implications of our work and application.

Finally, I'm going to highlight a few unresolved questions. I don't have answers to them, that's why they're unresolved. But things that, that we're thinking of, things that I think are important for both clinicians and researchers to think of, and then just a summary of the findings and some future directions that we're pursuing in our own lab here. Okay, So I think it's important and hard though is to think about, especially sometimes, I think for scientists, we really like puzzles and sometimes I think we lose sight of what exactly it is we're trying to understand. So I make myself do this exercise every year.

But what, what do I want to understand? If I could understand this in my career or even make a small contribution to this, I would be happy. And I've tried to illustrate this here with kind of bad AI, a bad AI picture. But in this example, we have a child sitting at the forefront. So this is the child sitting cross legged and he's being asked by his father to set the table.

And that's what we call the target because that's the thing that the child needs to do. Or there's going to be some kind of consequence. And I don't know what that would be in this family, but probably maybe no screen time if he doesn't set the table. The challenge here is that there's all kinds of other sources in the environment. So you can see the sisters playing a game.

The little brother is sitting behind him and he's playing with toys and making noise. The dog's barking. And then you can see that there's all kinds of noises associated with cooking. Right? And so a successful outcome for this child requires that he disentangle the message that his father's saying from all of these other sources of information.

Of sound that are produced in the environment. So this is what we're trying to understand. It's quite a hard problem, as it turns out.

And I've started with sort of a conceptual framework. It's really a simple framework of this problem Just to illustrate some of the stages of processing that have to occur for this to be successful.

What we know in these environments is that the sounds produced by all of these objects in the environment, the dog barking, the father talking, the. The crackling of the stove, these all combine at the child's ear as a combination of acoustic waveform. So if you look at this on something like an oscilloscope or any kind of analysis device, it's really hard to look at it and say, well, what waveforms belong to the dad and what waveforms belong to the dog. It's really difficult. But we call this combination of sounds that reach the ear the auditory scene.

So if you hear me say auditory scene, that's what I'm talking about. And one of the things we know is critically important is that the peripheral encoding of the sounds that come into the ear, both the targets and the masker sounds, right? All the target and the distracting sounds. And by peripheral encoding, I mean the peripheral, the intensity, the timing information, it must be encoded by the peripheral auditory system with enough fidelity that the brain can actually disentangle those different waveforms. And for people with hearing loss, for example, or people particularly something like a cochlear implant, we know that that precision is reduced dramatically, and that affects your ability to hear in these complex auditory scenes.

And the main question for us is, well, how does that develop over time? Does a child have mature ability to represent sounds? Because if they don't, they're going to really struggle in these syndrome environments. So then once information gets sent to the brain, we often talk about object formation and selection. And object formation, sort of in broad terms, involves grouping the sounds into auditory objects.

So ideally, the brain will know which parts of the sounds belong with the dad's speech, with the dogs barking, with the brothers banging blocks together. So that's object formation. Like, how do you know

things go together? And then once they go together and you form these objects, can you attend to the information in the target and disregard the other information? Right.

So you can kind of see those two things are highly related, but both are important. You need good selective attention skills to be able to attend to something and ignore background sounds. And all of that is required to recognize speech, along with language skills and cognitive functions and other kinds of skills that we'll talk about the take home message really here, and maybe for the whole presentation, is that immaturity at any stage of processing can influence how well children perceive speech in their everyday lives. From a clinical standpoint, we think a lot about this peripheral encoding stage, and I would argue that we don't think as much about the other stages, although I think that's changed a lot over the last, you know, 5, 10, 15 years.

Okay, so I'm just going to break this down a little bit more to provide a foundation for when we get into our studies. But this is an example of the auditory scene. So you have all these sources of sound. I've broken it down from the father's voice, which is second to the left. And we call that the.

We call the sounds produced by the target of interest, by the talker of interest here, the father's voice, the target. And really the target is what is the important message. And that can vary. Sometimes the target might be the dog barking. Right.

And sometimes it might be the father. But in this case it's clearly the father telling the son to set the table. But there are other sources of other sources in the environment also create sounds. And we refer to those as maskers, which I think for audiologists and hearing researchers are fairly common terms. But basically the target is the message of interest.

The maskers are everything else. And the combination of these scenes, which you can see at the bottom here, we refer to as an auditory scene. And I think it's important. I know I keep. I've said this

more than once.

But to reiterate what reaches the ear is the totality of all the sounds in their environment at a particular moment in time. So this is what received, which is really a pretty phenomenal problem. And people often make comparisons to like figure ground types of experiments in vision. But one of the things about visual perception is you have this constancy and spatial representation. In many cases, the sounds appear and they're gone by the time we have to figure out what the message was.

Right. There's a temporal property of sound that makes it a very challenging listening situation. Okay. The other thing I want to highlight that I think people are starting to think a lot more about. But what turns out to be critically important is that not all sounds in a scene are the same.

So if you look here on the left, some sound sources produce steady sound. And by steady sound I mean that it's relatively unchanging over time. There are, of course, fluctuations, but things that we think about a lot when we develop classroom standards. For acoustics. And when people measure how well, how noisy classrooms are, they almost always are looking at these steady sources of sound.

Often in an unoccupied classroom, things like H vac systems are notorious. If you know a classroom with a fish tank, tell them to take it out. Very few classrooms have that anymore, but those tend to be pretty noisy. But things like vacuum cleaners, traffic noise outside, you know, you can think of a lot of examples where the it's noisy, it's sh. It's pretty consistent over time.

Important to note that we have these classroom standards for unoccupied classrooms, and they're really important benchmarks because they tell us about what are appropriate levels of noise for optimal speech understanding. One thing to highlight is in the US at least, and I think this is probably true in most countries, most unoccupied classrooms exceed these recommended levels. We already know that classrooms are noisy, and that's where that data comes from. One of the things we think a lot less

about is that there's a lot of other sources of sound in classrooms, and they're dynamic, so they vary in time. They're sometimes unpredictable.

And the biggest contributor to these kinds of sounds are people talking. In this example, we have, say, a classroom where an adult and a child are reading in a corner. But there's another lesson being taught in the other corner. In this example, if you're listening to the teacher, you want to ignore the people reading in the background. That's a masker, but it's a very different type of masker, as it turns out.

And then this example here is a cartoon character on tv. So media and other sources of or TV and other sources of electronic media produce a lot of background speech as well. And far fewer studies in the literature that characterize exposure to competing speech and outcomes. But the data we do have suggests that kids hear speech more than half of their waking hours, and that's at home and away from home. So things like the TV running in the background all the time.

But even for houses, homes where TV is not super prevalent, competing speech is kind of pervasive. So hold that thought as we move forward that it's a really pervasive source of sound that we've understudied. All right, so why is it important to think about these different sources of sound? Well, from a hearing science point of view and a real life point of view, that these different sounds impacts speech recognition, but in different ways, the processes responsible for masking with steady noise and competing speech aren't the same. I've tried to illustrate this, in this example, you have a target talker.

It's the teacher, and she's giving a math instruction. And our examples here are steady sounds. We have the vacuum cleaner outside the hallway. They're vacuuming, I guess. And the H VAC system.

And the way that we are visually representing the interference that these background sounds cause with hearing the spoken message is kind of with these opaque lines. So we often think about steady noise as primarily reducing the fidelity with which the speech cues that the target speaker produces are

represented by the periphery. So adding this kind of noise, it degrades the representation of speech. You don't get all the parts of it. It's almost like having hearing loss.

In fact, there's models of hearing loss where they try to make noise, degrade speech in a similar way. But this is the problem is you're not giving the input into. The system isn't complete, right? So you have missing speech information. And it really depends on how far away the noise is from the target, how loud it is.

These kinds of factors. Now, the other type of speech, same target, same listening situation. But now let's say that the competing sounds are these other people talking. We've got a TV on in a classroom, probably bad form, but it's happening. And then we have the two people talking.

Now, this is a different situation. You still have interference with the target on the periphery. Like having competing speech does interfere and overlap on the basilar membrane. But it does, it tends to do so a lot less than steady noise. But what it does is it makes it really difficult to disentangle at a kind of central auditory level, like what's the target?

What's the masker? Can I pay attention to one and pay attention to the other? And you can see it here. It's like a figure ground problem. You can, if you're really good, you can make out the teacher, you might have to pay a bit more attention.

As it turns out, it can make it really difficult to recognize speech, even though from a peripheral auditory standpoint it should be easier than when noise is present.

Okay, still on conceptual framework. So I keep mentioning that if you don't have a good representation of sound, it's going to impact your ability to hear in noisy environments. And so a question that people spent probably 30 years trying to address is.

When does peripheral resolution mature? Right. If it takes a decade for a child, even with normal hearing, who's otherwise neurotypical, to be able to represent sounds at the periphery in a mature way, then they're not going to hear as well in these kind of environments. And so what I'm referring to here is the spectral intensity and temporal characteristics of the sound mixture must be adequately represented by the peripheral auditory system. And as it turns out, Again, this is 30 years of research in one slide.

We think that this is mature by probably six months of age. There's some indication that there's some fuzzier representation at the level of the brainstem in early, in later infancy. But really by six months of age, we think that the encoding of speech is quite good. And in some cases, you see sharper kind of peripheral resolution in an infant than you might in an adult, which kind of makes sense if you think of it from an aging perspective.

So again, people spent a lot of effort and fortunately for me, I kind of came in right as this work was finishing up. Otherwise it would have been probably impossible for me to answer or to investigate some of the questions that we were interested in. Okay, I'm switching now to key findings. So we've set the stage. I've talked about how we don't think the main issue is peripheral resolution, but how good are children at hearing speech?

So the answer to the first question, in part this was answered well before we started doing work in that we've known for a long time that children are worse at recognizing speech that's masked by competing sounds. And up until 20, the early 2000s, almost all of the time we were thinking of speech that's masked by noise. So in the clinic, it would be, you know, the quick, that's a bad example.

It would be something like speech recognition and noise, where you're playing kind of narrow band noise or wideband noise from the audiometer. But lots of lab studies, and this was a study though,

where we showed that, yes, this is true, but it's really pronounced when the competing sounds are background speech. So in this study, which was done by my colleague Nicole Corbin, we looked at a fairly large sample of children and adults with normal hearing. They ranged in age from 5 to 16. And we also had adults, a smaller number, but we usually include adults, just so we get an estimate of what's mature performance.

They, they did an adaptive speech recognition task, kind of like the words you might hear in word discrimination, but they were single syllable words that were in the lexicon of a young child, but they were open set meaning you heard a word, you had no context and you had to repeat back what you heard. We looked at this in noise and in two streams of competing speech. These were two people reading from a familiar Children's book, different passages. In most of our studies, and I'm going to report here, we use an adaptive procedure. What that means is we basically try to estimate what the signal to noise ratio is to get a certain percent correct.

In this case it was 50% correct, like the SRT. And the reason that we like using an adaptive procedure when we cover this wide age range is it's really hard otherwise to use a fixed signal to noise ratio because you usually run into issues where it's really too hard for one age group or too easy for another. And this often gives us a good first approximation. The punchline of this study is that like other people have found before us, our estimates of speech reception thresholds, young children, the five to seven year olds, needed about 4dB more, a more advantageous signal to noise ratio to do, as well as older kids and adults. Again, this is something other people have found, which was really surprising was that this effect was almost 7dB in competing speech, which is a really large difference.

Remember that DB is a logarithmic scale. It's a pretty whopping difference and something that we were a little surprised about, honestly, at the time. So this was a group difference. The next thing we wanted to look at is, okay, so we know that children are at a disadvantage, but when does this ability mature?

And up until Nicole's study, really, we often stop testing at 9 or 10 years of age.

And that's because most previous studies had shown that by nine or ten children looked like adults for speech and noise measures. There's a few exceptions, but in general that's what we found. We didn't see that when we measured speech and speech. So over the years we've slowly increased our age range. Now we basically test up until continuously from five into early adulthood, and we no longer kind of group kids into 5 to 7 year olds, 8 to 10 year olds.

We now look at age continuously and what that allows us to do is estimate when mature performance is observed for different skills. That's been really valuable. This is a study that my colleague Emily Buss did. Again, similar adaptive procedure, but instead of this open set, repeat what you heard. It's an adaptive closed set procedure.

They hear one word, there's four words they can pick from. They don't have to say anything, they just pick the word they think they heard. The words are in the lexicon of young children. And we've had success with this task on children as young as 3 and 4 years of age and on individuals with down syndrome across the lifespan. So we like it.

There's some pluses and minuses of any procedure. But one of the things we like about this procedure is it's something that people can do and we can use the same procedure across a wide age range. This time the masker was just one stream of speech. So this is considerably easier than when it's two streams of speech. And what we found is that over the course of five to when we see mature performance, we see a 16 decibels improvement.

That's really large. But what's. I think the most interesting finding from this study is we didn't observe mature performance until after 13 or so so into adolescence. And since that time, us and others have shown this to be true for most of the tasks we've looked at. Even things like detecting a word in the

presence of one or two talkers, certainly sentence recognition, word recognition.

This seems to be a really stable finding, not just a something we saw with our specific task. This is important. What it means is this is a really late developing auditory skill. I think when a child hears in an environment where people are talking is really different from what we hear as adults, presumably adults with normal hearing. I should say two more key findings.

One is that these challenges that children face when they listen to speech in the presence of speech seem to be driven by central auditory factors, not peripheral development. And those factors, as I mentioned before, could be grouping sounds into objects. And I've illustrated that here. Like you might have sounds from all these four objects in the environment. They combine, but the brain has to then try to reconstruct them into the con, into the objects that they are in the environment.

And most people do pretty good at that. If you ask somebody, tell me what's in the environment, you can tell me, there's some dude talking, there's a dog, I hear blocks. Like you can reconstruct that, which is honestly phenomenal in my opinion. And then they have to selectively attend to one of the objects and ignore the sounds from the other objects. Okay.

And one of the ways we've kind of provide one of the ways we've obtained evidence that some of the issues that children face are central comes from studies like this, where there are acoustic cues that can actually make the task of listening to one talker in the presence of other talkers easier.

A really obvious example is if you do this experiment and you have all female talkers for your stimuli. So you might have a target produced by a woman. And the maskers are speech produced by other women or sometimes even the same woman, which is really hard. But if you switch that and you have a mismatch. So maybe the target speech is produced by a woman, but the background speech is produced by two men.

It's a lot easier. And we looked at that because we wanted to look at when that ability develops. Like, is that something that might be limiting performance? So in this study, we had five to ten year old children. We wanted to look at an age range where we knew children struggled with speech and speech recognition.

We had adults, same procedure I mentioned before, but this time we have words that were either produced by a male or female. And the masker was always male. So we had a matched or mismatch condition, we called it. And the results suggested that children really struggled at this task overall, but they could benefit from this cue, which was really exciting. You know, it's, it's, here's something that might actually make performance better.

We know they're worse overall, they continue to be worse overall. So that's an interesting finding. But this cue was a benefit. But in the same general study, we had another experiment using the same stimuli. This time one di, one word out of the corpus.

And we measured speech detection in a speech masker. And we looked at infants and adults. So really simple tasks. They're not even having to recognize speech, they just have to know it was presented. And what's interesting is infants didn't benefit from this cue, which suggests that the ability to rely on some of these acoustic cues, whether it's to form an object or to attend to an object, seems to take time.

Okay. So the second related issue is that children listen less selectively than adults. And I've showed a really simple example. And I want to highlight that this has been replicated in the speech perception literature as well. But it's easier for me to kind of give you an example using tones and speech.

But in this study, we looked at young children, so four, and we split them into two groups, four to six year olds, and then we looked at seven to nine year olds. And then we also looked at adults. And the task was really easy. All it was is we presented a pure tone, a thousand hertz pure tone in quiet. And we measured a threshold.

We used this two alternative force choice detection tasks. But you could do this in the clinic using a standard task as well. It's a little trickier to control for a, for a few things that we'd like to control for, but you could do that. And then we did the same thing. Well, actually we randomized the order, but we measured thresholds for the Tone and quiet.

And then we measured it in noise. But it isn't noise like you would have in the audiology clinic that overlaps in frequency, which would overlap in the basilar membrane. In this case, the noise was 4 to 10 kilohertz, so it was two octaves above the signal frequency. And if you do this in adults, they act like it's not there. They just ignore it and it doesn't interfere on the level of the periphery, so the thresholds don't change.

So we basically say there's no masking. They can ignore the noise, basically. And what we found is that four to six year olds, when you put in the noise, it affects their performance. They aren't as good at selectively attending to the tone and ignoring the noise. And at the time, Lynn Werner, when she discovered this phenomenon in infants, called it distraction masking.

I think the idea was that it pulled your attention away. I think later models would suggest that no, kids just are always kind of unselective in how they're listening. So they're more susceptible to noise because they're always kind of a little bit unselective. It takes a while to fine tune your perception. People call this perceptual narrowing in other contexts, but it's a really interesting phenomenon.

And as you get more exposure to sound, you start knowing what's the most important to attend to. But you also get better executive function and skills and those kinds of things. By the time the children were 7 and 9, they didn't show these effects. Again, I mentioned that we have some colleagues at the Ohio State University who did the same general experiment, but with speech, and found the same kind of age pattern. The last key point, immature language skills impact speech recognition.

Speech and speech recognition. So I think the easiest way to do this is just to describe the experiment. Same task. Now we're looking at preschoolers who can do this task, but sometimes they don't know all the words you're presenting or they may not know them really well. So we have this active training process where if they don't, we pick words that are in the lexicon of like a 4 and a 5 year old.

We train children to see how well they know the words. If they don't know them, then we keep training until they do them and they can recognize them as a picture. And most of the kids could do the task. We measure thresholds again. This time it's just speech in the presence of speech.

And really what we found is that children with larger vocabularies so better Receptive language, certainly better.

They know more words. It doesn't reflect all aspects of receptive language, but they showed better performance on the task, and that's after controlling for age. So certainly people get better when they get older. But receptive vocabulary matters. Now, it doesn't necessarily matter for older kids, in part because the words are really acquired early.

But if you test children who are acquiring these words, how well they know them and how well they're represented in their lexicon really matters. And that's a simple example. But other people have shown this kind of general phenomenon. Okay, so I'm going to switch to clinical implications and some application before finishing the talk and answering questions. I guess about 15 years ago or so, we

started to feel like we knew enough about typical development to look at these phenomenon in children with hearing loss.

So what's interesting is we've known for a long time that people with hearing loss, children in particular, really struggle in noisy environments, and more so than their peers who do not have hearing loss. So we know all children struggle, but their children with hearing loss struggle more. And of course, some of that is related to how well sounds are represented by the peripheral auditory system. So some of that's reduced audibility. Right.

But even with hearing aids, there's reduced representation for a lot of children because we get, you know, less precise frequency and timing information due to damage to hair cells. If you have a cochlear implant, you've basically created a system where you have very limited spectral resolution. Right? And so people can perform really well, but they tend to perform pretty poorly when it's noise. Okay, but the other thing we haven't thought nearly as much up until this point was it isn't just peripheral.

We think, or we hypothesized, if we think that you need time to learn the properties of sound that make you a better listener in noise, and you need to learn strategies. You need better executive function and language skills. It's going to depend on how much experience you've had with sound. And people like Beth Walker and Ryan McCreary, my colleague here at Boys Town, have shown that children vary a lot with hearing loss who wear hearing aids and cochlear implants. But they vary a lot in terms of how much they wear their devices, but also how well those devices are fitted.

And, of course, people have different degrees of hearing loss, and those metrics seem to predict outcomes, language outcomes, speech and noise outcomes. So we wonder, do these also predict how well somebody listens to speech in the presence of other people talking?

So we know this is not a new finding. We weren't the first people to find this. But again, we looked at this four alternative force choice task and now we had a fairly large or no, this is a smaller sample of children with moderate to severe hearing loss tested wearing their hearing aids. So that's important. And they were well fitted hearing aids in this case.

And we compared them to peers with normal hearing in this small preliminary study. And what we found that on average children with hearing loss needed 3.5 dB more favorable SNR than their peers with normal hearing to recognize speech in steady noise. What's interesting is that their performance was related to audibility. The kids who had better aided audibility did better on the task. But the outcomes on our task, our speech and noise tasks, weren't related to what parents were reporting their children's listening difficulties were.

Other people have started to think about this a lot in the audiology clinic. A lot of our speech and noise measures aren't really telling us which people are struggling and which people are doing really well in the real world. So this is all in line with that. We also though looked at performance in masker composed of two streams of speech, and we get a slightly different answer. The first one is sort of whopping and really surprising to us at the time is how poor children with hearing loss do in these conditions, even relative to peers with normal hearing who also do really poor compared to adults.

I remember when we first started doing this study, I can't remember if it was me or Nicole Corbin, who's my graduate student at the time, but like we teared up because it's, you're listening to the speech and other presence of other speech and it's so easy for somebody with normal hearing and who's an adult to hear. And these children really struggle, even children who are doing pretty well. So this is a really, really difficult problem for people with hearing loss. It's even more difficult if you have cochlear implants in general. So very large effect.

The other thing that's interesting is performance didn't relate to audibility. Now again, it's a little bit unfair because all of our kids had good audibility in this study. But unlike performance and noise, the driving factor wasn't audibility. But and I think this is important, parents reports of their children's listening difficulties were associated with performance. So those children who had better SRTs, their parents also reported they were doing better in the real world.

And so we think that they captured different parts of listening that might be really valuable okay. And this is more or less the last topic. I have a few final slides, but we have, we basically were really interested in those findings and with a really large team, developed the Children's English and Spanish Speech Recognition Task. I put the names of these prominent scientists on the task because the name means the Children's English and Spanish Speech Recognition Task. But they also are meant to reflect these giants and pioneers in our field who really laid the foundation for everything we do.

And that's Raymond Carhartt, Joe Hall, Lois Elliot, Judy Gravel, Pat Stomachowicz and Li Feng Shi. And so really I just wanted to make sure that they're acknowledged. But what's the Cheggs test? It's a computer based tool for assessing closed set word recognition in English and Spanish. And we can measure either in noise or competing speech.

So up till now I haven't talked at all about bilinguals or speech perception. What I want to highlight here is the fact that we can measure in noise or competing speech because we think that might give us different information. We also though wanted to recognize that about 15% of our caseloads now are people who are bilingual. Spanish English. We know there's lots of other bilingual communities, but Spanish English is the most prominent.

We also know that very few of our audiologists here in the US actually are proficient enough to administer a Spanish. Myself included a test in Spanish. So we have recorded materials, we have a forced choice response. So we don't need the tester to be fluent in the target language. And we can

measure performance in English and Spanish or in both.

And that might help us as children change their proficiency or maybe to capture how they're doing in different listening environments.

Yeah, I think we have that there. A couple more examples about Cheggs. The Target stimuli are 30 words and they're disyllabic in English and Spanish. And there's lots of details in the literature you can find, but really my colleague Lauren Calandrusio spent a lot of time making sure these were balanced with respect to acoustics, the phonetic characteristics, timing. So the target words were spoken by the same bilingual talker.

And so they're really equivalent acoustically, which turns out to be really important. And one reason why some other tasks that are available in English and Spanish might actually be easier in one language than another, which makes it really hard, right, to use that for criteria for things like deciding whether or not somebody needs a cochlear implant. We knew it was important, especially with these speech in speech conditions, to have really good control acoustically. It's a Picture pointing response. There's no verbal response required, so it makes it appropriate for testing a wide range of ages, abilities, and speech production skills.

We've used it, as I mentioned, I'll show a slide just at the end on a very large sample of people with down syndrome and found it, you know, very feasible. Most people could perform the task. There were a few that could not. But I think surprisingly, very high yield rates for the task. Okay, this is just a highlight, just so you can kind of see where we're at.

But we have extensive normative data, and that's been published under headphones in a study that I did here in 2024. Recently, like, literally days before my talk, my colleague Emily Buss led a paper that has normative data on listening under headphones from these same children. And what we find, which I

think is really reassuring, is a proficiency is good in both languages. You get the same answer, and I think that's what you want. Now, you may have a child who gets a different answer in Spanish and English, but if you look closely, that's usually because they were more proficient in one language than another.

In our sample, people tend to be more proficient in English, even if their parents speak Spanish at home. And I think that reflects living in Omaha. And so we're really curious what results are going to look like in different parts of the country. But we believe this will give us an accurate estimate and that maybe over time you might start off testing in Spanish and switch to testing in English. We are fairly close by.

Fairly close. I mean, within the next month, I think of submitting a paper that describes clinical outcomes from seven audiology clinics across the US on clinical patients. So these would be patients coming in for audiological evaluation, but also patients with hearing aids, cochlear implants. We have a handful of, like, bone conduction and other devices, and we are going to compare those results to the normative data as well. And we are working on implementing the tool as an iPad application that'll be available for free as a download.

It is designed to work in a clinic now, but there are issues with just technology that make it a little hard to share. So we're going to have an iPad implementation that can be routed through any standard audiometer and administered and made available for free. So hopefully by the end of, it still needs to be reviewed, but hopefully that will be coming out soon.

And I mentioned that. Okay, back to unresolved questions. You know, are we capturing the challenges children face in their everyday lives? And I devoted a whole slide to that in the beginning. And the real answer is, I think we're moving closer and incorporating measures like how well children hear when other people are talking is really important.

But I would say we have a long, long way to go. Everyday environments are really complex. And we know that people come into the audiology clinic and report concerns that don't always match their audiogram or that two people can have the same audiometric findings. And we don't see that reflected in the range of concerns they have. So I think that we have a long, long way to go here to reach that goal.

So related, I think some of our tasks are likely too difficult now. Some of them are too easy, too, probably. But I wanted to provide an example of a task that we've done just to kind of caution that, like, sometimes we design these really interesting experiments in the lab, but they're really hard. And so this was an example my colleague Lauren Collinger, CEO, did. Instead of the four alternative tasks, it was sentence recognition.

So BKB sentences, and they were produced in two streams of speech. So no noisemaster here. And we looked at young kids, older kids, and adults, and what we wanted to look at was whether they perform better when the talker used a clear speech style, which, honestly is what most of our materials in the audiology clinic are at least close to. So clear speech would be. It's a type of speaking style people use when you talk to people with hearing loss or people having trouble understanding understanding you.

But it tends to be a little slower. It exaggerates some pitch contours and generally makes it easier or more intelligibility for adults. And we wanted to know, is that true for children? And what we found is it was true for older children, but young children didn't take advantage of those cues. They performed poor in both conditions, which I think is really interesting.

Right. If it's saying that right. That gives you some insight into the real world. But here's the part about our study being really hard. We had a talker who was a female, and we had background talkers who

were also female, and they were kind of matched to sound similar.

We had this task where everything was presented to both ears of participants, which, again, is not what we listen to in the real world, where sounds might come from different spots in space. So we decided, let's do the experiment again, but now let's make it a bit easier. And so what we. We have this manipulation where you can just under headphones, you can flip the phase of either the target or the masker sound, and it gives you a spatial cue. And it does make the task quite a bit easier.

You know, depending on the study, 4-7-D be easier. It just kind of pops out. And so we did that and we then redid the clear speech experiment. This time we only included the younger kids. And really what we found here is if you give them this extra information now, they can access it.

So there's something about making the task really hard that has cascading effects. And I think a big question is, well, how hard is the real world? Right. In the case of this study, I would argue that our task was probably harder than what kids are experiencing in the real world. In other cases, I think it's probably easier.

And so finding that sweet spot, I think is going to be important for predicting how well children are going to do important. I think we all know this. I think one of the issues a lot of our studies have is that when people come to infra lab studies, we tend to have an unrepresentative sample. They tend to be people who can travel during non work hours to the hospital. We don't tend to represent children who live in poverty.

Rural communities speak a language other than English.

I don't know what the statistics are now, but I know when I was school something like 50% of all children with hearing loss have comorbid developmental disabilities. They are often excluded from

studies. And so I think, consequently, we don't know that. Our studies generalize to everybody. And this is just an example.

I'm not going to go into detail here, but we have used the Cheggs task on people with down syndrome across the lifespan. They can do the task and we get a very different pattern of results. You know, they seem to really struggle with listening to speech throughout their lifespan. And that's probably due to a combination of hearing loss, which is very prevalent in the population, language skills, which tend to be an area of weakness, cognitive abilities. We also see a lot of variability though, like some people can perform as well as their peers who do not have down syndrome.

And I think that we need to do more work in this realm and think about the challenges that all our patients face.

Okay, summary. If you take home anything from this slide, I would say that what a child hears in a noisy environment, it's just not what we hear. One of my colleagues used to say that we have these guidelines for playground equipment at different ages, but we don't for listening and noise. But we know that the ability to recognize speech, especially when multiple people are talking at the same time, requires years and years More than a decade of experience and maturation of central auditory processing skills. Also, that consistent access to high quality input gives a foundation for language development, social functioning, and school success.

But you see here that not everyone gets this access. Here's an example of a classroom where surprisingly few children, like most classrooms, have a lot more. You can see there's even an airplane flying by, which is not uncommon for some of our children, particularly in lower SES backgrounds. Some of these children might need hearing aids and might not be wearing them or might not have appropriately fitted hearing aids. But these are difficult listening situations which require sort of the best possible audibility and support.

And yet we know we're not meeting the mark there. Last slide. We're currently really interested in speech produced by a variety of different talkers and styles. So not everybody listens to really clear speech produced by somebody in Omaha, I guess, would be my example. And one of the ways we're doing that is we have a new mobile vehicle with the booth in it, and we're trying to get out into rural communities and really all over and get a community based sample to do that with.

And I think I went a little, a few minutes longer, but hopefully there's still some time for questions. And thank you very much. I really appreciate your time today.

This is fascinating. Thank you so much. Dr. Leibold, we have one question here from Ann. Let me read it. Do we know if speech recognition and noise differs in Spanish versus English when all other factors are controlled for?

And I'm going to expand on that and just say other languages as well. So I think, yeah, that's a great question. And I think it depends on the age of the listener as well. So there, there are a number of studies that have looked at second language learners of English, and it's sort of, it's. I wouldn't say it's arbitrary, but often the number that's used is they Learned English after 7.

Certainly if they learned English as adults. And in those cases, there really does seem to be better speech recognition in their native language, for sure. And I think those data are really compelling. What's interesting is if you look at what, what people call balanced, simultaneous, balanced bilinguals, those effects go away. Now, they don't go away in all studies, but if you use materials that are balanced both acoustically and linguistically, you tend to see pretty similar performance.

I would say, with the kids, that's what we see. We see differences, but once you account for language ability, it goes away. And in our sample, honestly, we tend to see better performance. In bilingual kids in

English than in Spanish. And I think that's because they're probably a little more proficient in English.

But it's a great question.

I think that it depends, though, on the level of. It depends a lot on how you characterize proficiency, but also the materials you use and how difficult they are. So in the Cheggs task, what I will say is it's both a strength and a limitation of the Cheggs task. But once you have adequate proficiency, we don't see big differences. It could be the case that once you've mastered kind of basic English language proficiency, you can be tested in English, and that's appropriate.

Right. Where we tend to see the biggest differences are younger kids or people who may have immigrated to the United States pretty recently. But for our kids here in Omaha that go to English speaking schools and they're surrounded by English media by the time they're older, we tend to see better, if anything, better performance in English and Spanish. But I think if you. Yeah, so one of the things that's really nice about Cheggs and, you know, there's a lot of details that go in there, but same talkers.

And then Lauren even applied some acoustic manipulation, so they're equated almost exactly for the long term, average speech spectrum, speaking rate, really, really like, really, really rigorous design that the words. So in some of the. I think the hint is a good example where the sentences themselves are not necessarily equivalent linguistically because it's really hard to do that on a sentence level based on differences in the language. But they're also spoken by different talkers. And my understanding is that if you had the same proficiency, you could do a bit better on the hint in Spanish than in English.

Not exactly my area of strength, but that's been something that people have, have noticed. And so. But the benefit of tasks like that is that you have more complex sentence structure and things like that.

Great. Thank you. Thanks, Ann, for that great question. We'll just keep the Q and A open for another minute to see if anybody else has any comments or questions before we wrap up. And as you can see from the slide here, please make sure if you're earning CEUs, you remember to take the course exam.

For those of us today in the live webinar, it will expire after seven days. And for the recorded webinar, it will expire that exam after 30 days. So I don't see anything else coming in through the Q and A. So that'll be a wrap for today's class. Dr. Leibold, it would be great to have you back in the future to see how this research evolves.

And oop, one more did come in. Christina asked, has any work been done on this topic for children with neurodevelopmental disabilities? Another great question. Yes. Okay.

So certainly you mean in addition to hearing loss. I would say a lot of work on down syndrome. We have caught. Let me think about this. I believe we have caught.

Yes, there's been a little bit of work on children who have autism. My colleague Heather Porter does a lot of work on how to test people primarily with intellectual disability, but also with significant motor impairment and has a lot of really clever use of things like social schedules and visual support. The other person is Angela Bonina at Vanderbilt has done a lot of this work. I think it really depends, you know, on one hand, I know people have wondered about people with autism and whether they perceive speech and speech and speech and noise differently, and the results aren't conclusive that they do. But I think some of the issues are providing tasks that are appropriate and maybe targeting feedback and reinforcement to something that on an individual level, we have some tasks we do a lot of work on studying infants that we've applied to people who may not make conventional motor responses.

I didn't talk about that here today, but we have some techniques that don't require like crisp head turns like you might need for vbra. And that, that has really allowed us to test. You know, you might have a

participant who can't raise their hand, but we can still collect data on them. So I would say yes, I think a lot more data need to be collected and we need to be. I was going to say we need to be creative with methods, but I think most really good pediatric audiologists, I know I could learn a lot about testing from them.

Like, I think that getting some of those tools into their hands would be really wonderful because if they can get data like nobody else. Right. I mean, I think that requires some really specific training and experience. But pediatric audiologists are incredible. Yes, for sure.

One more follow up to that. The person was thinking about auditory processing for children with ADHD in the classroom. Yeah. So, you know, it's funny, I always get asked these questions about auditory processing disorder and I guess my comment is like the way it's defined. Every five year old has an auditory processing disorder.

Right. Like it's. I do think there's.

I do think so. ADHD is interesting. And in general, at least if we test kids on their medication, they look no different than their peers.

I do have colleagues who study. They study more like, they use more like, imaging techniques and think about brain responses. I. I don't think that there's any reason to think that issues with, like, object formation and maybe even selective attention are as big of a problem as things like impulsivity. And what would be the. There's another word that I'm thinking of, but it's a really interesting question.

Like, I think that the issue is more attention to the task in general than performing well in the task. I mean, I think you can have children with pretty significant adhd, but if they're focused on the task, can. Can do well. And. And that kind of comes from personal experience with my.

My own children. But people have looked a little bit at that issue, and I think parsing apart what features, what specific executive functions, for example, are responsible, is probably the way to go. I don't think I answer your question very well. So maybe to answer your question, we need more work. But it's very interesting because you have a really strong age effect and a lot of differences across people in how these kind of what I would call auditory processing skills develop.

I like that. So it may not be the answer to the question, but I think you brought up a lot of good discussion points and considerations.

Fantastic. So thank you for those additional questions. That was really interesting. And again, Dr. Leibold, it'd be great to have you back as the research evolves, you and your colleagues. Thank you so much for this presentation.

Again, a big thanks to American Auditory Society for partnering on today's webinar. And we'll close out the webinar today. Thank you. Great. Thank you, everybody.