Rethinking Speech Perception Challenges in Adults Without Hearing Loss, presented in partnership with AVAA

Presenter: Jacie McHaney, PhD

It's my pleasure to welcome to you our guest speaker, Dr. Jacie mcHaney. This course is in partnership with the association of VA Audiologists, and we'd like to thank the current President of AVAA, Dr. Daniel Crawford, for coordinating this wonderful partnership and this webinar with Dr.

McHaney. If you'd like to learn more about Dr. McHaney, you can read about her bio on the course registration page. But at this time, I'll hand the mic over to you. Thanks, Christy.

So, like Christy said, My name is Jacie mcKaney. I'm a research assistant professor in communication sciences and disorders at Northwestern University. This is my first time presenting here in this forum, so I'm really excited to be here. And today I'm going to talk about speech perception challenges in adults who have normal hearing thresholds.

Okay. These are my disclosures. These are my the limitations and risks that are inherent to today's course, and these are our learning outcomes for today. Okay. So, as probably many of you know, as we get older, our ability to understand speech in noisy environments becomes more difficult.

So, you know, imagine you're in this loud outdoor market or practice, a restaurant, a bar, maybe you're walking by a construction site, things like that. It can be really difficult to understand your communication partner. And that becomes increasingly more difficult with age. There are a number of known factors that can contribute to this. One obvious one is peripheral hearing loss, which is something that happens with increasing age.

There's also some changes in cognitive function that can contribute to speech perception challenges. There's a lot of data showing that hearing loss and cognitive function together contribute to speech perception. Then inherently with age, there's changes in auditory processing abilities that can contribute to these problems as well. For those of you who are audiologists, maybe oftentimes you have an adult patient, they come in with a chief complaint of speech perception challenges. And so you

may put them through your clinical battery, and maybe they exhibit prescusis or some other form of hearing loss with elevated thresholds.

And then maybe they also show poor scores on, like, words and noise, speech reception thresholds or quicksin. However, there have now been many, many studies that have found that 1 in 10 adult patients that visit audiology clinics with that chief complaint of speech perception problems, they are end up, you know, given a clean bill of hearing health. You know, they show normal hearing thresholds and they're not showing any issues on our typical test. And so, you know, maybe you can attest to this anecdotally with your own patient population coming in and just saying that they have problems, but then you run all the tests and they look fine.

So, you know, in this case, these 1 in 10 adults are often told that there isn't anything wrong with their hearing. And, you know, maybe they receive some counseling on effective listening techniques, techniques to help improve some of their speech reception challenges. But being told that there's nothing wrong, you know, their hearing thresholds are fine, that has to be frustrating for the patient. Right? They clearly perceive their speech perception difficulties to be so profound to the point that they're actually going through the hassle of scheduling an appointment or professional help.

Just be told that they're fine across the board. Oftentimes, the patients who have these complaints are around middle age. This is about 40, 55, 60 years old. Middle age is a time when speech perception challenges tend to arise even in the absence of hearing loss. These listening difficulties have significant implications on quality of life.

There's a lot of data to indicate that listening challenges will increase rates of social isolation. Depression and listening challenges are even related to an increased risk of cognitive decline. And many of you may have seen this report from Lancet from a few years ago, where they mapped out all the potentially modifiable risk factors for dementia from early life to late life. And they found that in

midlife that hearing loss was the number one modifiable risk factor for dementia. It's critical for us to understand what might be causing speech perception challenges in the absence of elevated hearing thresholds and how this report of problems may be a form of hidden hearing loss.

When we think about hitting hearing loss, oftentimes the main term that might come to mind is cochlear synaptopathy. Cochlear synaptopathy could certainly play a role in these speech perception challenges reported by these adults with normal hearing. However, our understanding is limited by the lack of obJacieive diagnostics and uncertainty regarding the perceptual consequences of cochlear synaptopathy in humans. So, cochlear synaptopathy is the loss of neural synapses between auditory nerve fibers and inner hair cells, primarily beginning at synapses. At the higher frequencies, there's this loss of synapse and this kind of, like, loss of communication from the peripheral auditory system into the central auditory nervous system.

There's plenty of anatomical evidence showing that cochlear synaptopathy occurs with aging in animal models. And there have been some histological studies of postmortem human temporal bones, where they looked at temporal bones in humans who have passed, and they've seen cochlear synaptopathy that has occurred with age. But there's mixed evidence kind of linking cochlear synaptopathy with speech perception in noise challenges in humans. So there's a lot more work that needs to be done to definitively state that cochlear synaptopathy is underlying these speech perception challenges in the absence of measurable hearing loss in adults. I do want to be clear that I do not study cochlear synaptopathy directly, but I do feel it's important to make mention of it.

The core of my work focuses on understanding what is happening at the brain level in middle aged adults with normal hearing and normal cognition for their age. To understand possible neurobiological sources of speech perception problems. My long term goals are geared towards developing diagnostic tests to complement the current audiological battery to better capture these 1 in 10 adult patients with

normal hearing that are reporting speech perception problems. So first I'll discuss some findings using electroencephalography. From here on, I'll have it abbreviated as EEG on some of the other slides.

And so I'll look at how EEG responses can examine how speech sounds are represented in the middle aged adult brain. Next, I'll talk about using pupillometry or the measurement of pupil size. You know, how the pupil is dilating and constricting and how that can be a metric of effortful listening. And finally, I'll talk about some computational methods that use accuracies and response times from behavioral tests to understand cognitive aspects of decision making involved in speech perception. So traditionally with electrical, electrical electrophysiological test in a clinical setting, we'll use isolated stimuli such as tone bursts or click or discrete syllables like a da.

And these stimuli have given us a great idea and continue to give us a great picture of the general health of one's auditory system. But this type of stimuli really lacks a naturalistic context. You know, I'm not talking in clicks and I'm not talking in single syllables that are just completely isolated from other sounds. And so there's been some evidence to suggest that, you know, the research that we have done using these isolated syllables, the, the findings from that don't really apply to more naturalistic contexts that are using more naturalistic stimuli like audiobooks or content from like watching a movie. So over the last decade or so, there's been a paradigmatic shift towards using more naturalistic continuous speech to better understand auditory and speech processing.

In the lab that I work in at Northwestern University, we like to collect EEG signals. While participants are listening to continuous speech from audiobooks, we can then take those EEG signals and break them down to look at a number of different components. Things like how are the brain signals locking onto the speech envelope and how are they tracking the speech envelope when they're listening to someone speak? How is the brain locking onto voice pitch in multi talker listening situations? And we can also look at how different linguistic units are encoded in the brain, such as phonemes, words and

sentences.

Recently, our goal has been to understand whether any of the metrics that I just mentioned may be indicative of, like, a more hidden hearing loss in these patients with normal hearing. So I'd like to talk about some relatively recent findings we have using EEG to naturalistic speed speech. So we had a group of younger adults who were between the ages of 18 to 25 years and another group of middle aged adults who were between the ages of 40 to 55. They came into the lab and we ran a number of different tests on them. To be included in the study, they had to have thresholds less than or equal to 25 dB HL at octave frequencies between 250 to 4000 Hz.

We also collected some more extended high frequencies. And the middle aged adults were showing some elevated thresholds at these higher frequencies, which is kind of to be expected with that age group. But for the most part, for those lower frequencies, from 250 to 4,000 Hz, the age groups were not statistically different from each other. So they were showing similar hearing sensitivities. So that was their audiograms.

We also had them complete the words and noise test. And we found that the middle aged adults had significantly worse words and noise performance than the younger adults. So this tells us that they were showing some speech and noise challenges with the increasing age. So I mentioned that one aspect of continuous speech that we look at are phonemes. So we examine how phonemes, which are the smallest unit of speech, these are just your individual sounds that make up syllables and words, such as the S sound, the sh.

Huh, the things like that. And so we are able to examine how those phonemes, those individual sounds, are represented in the cortex of our brain during listening. So to do this, we had the participants, the younger and middle aged adults, wear an EEG cap on their head. And this EEG cap had anywhere from 32 to 64 electrodes. So, you know, it's covering pretty much the entire scalp.

And while they're wearing this and we're collecting the EEG responses, we're playing continuous Speech from the audiobook of Alice's Adventures in Wonderland. So hopefully you can hear this sound. Alice was beginning to get very tired of sitting by her sister on the bank and of having nothing to do, you know, so they're listening to this for about 15 minutes. And during this entire 15 minutes of them listening to the story, we're continuously collecting EEG signals. So when they're done, we take these EEG responses and we will time lock the EEG to each and every instance of every single single phony that's in the audiobook.

So, you know, I don't know how many people have actually read or listened to the Alice of Adventures in Wonderland book, but it is, there's a lot going on. If you can, you know, think about the Disney movie that they made. But there's a lot going on. So we get lots of rich information with many, many sounds, many, many phonemes. And so it's a really good skinless set for us to get a lot of data from.

So we time lock their EEG responses to every single instance of every phoneme. And in the example I'm showing here, I'm showing the S phoneme. So for every S that was in the story, we, you know, like cut the EEG and then we get a response to that S sound and we call this a phoneme related potential. So for any of you who have done any kind of EEG measure or familiar with it at all, you might know the term event related potential or erp. And this phoneme related potential is kind of the same thing, except in this instance, the event that we're looking at is the phoneme onset.

And so we take, you know, every instance that the S was presented and you know, all the other phonemes that are in the book, and we get an averaged phoneme related potential to that particular phoneme. And then I'll explain a little bit more on the next page. It gets a bit technical and I'm going to gloss over some more of those technical aspects. But what I want you to take away from this is that we're looking at the brain's response to phonemes and naturalistic, continuously running speech. And then, you know, later on in the Q and A portion, I'm happy to talk about this in more detail if anyone has

any questions, but for the sake of time, you know, I am going to clock gloss over a little bit of those more technical details.

So we take those phoneme related potentials and my colleague, Dr. Jianguo, he built a neural network model. And so this is Essentially like a machine learning model that learns what EEG responses look like to all the phonemes. And so we take these phoneme related potentials and we trained the model to classify those phoneme related potentials into the correct phoneme. So continuing on with the example of the S phoneme, every participant had a single averaged response to the S sound.

And we feed in a portion of that data and the model starts to learn, okay, this is what an EEG response to the S sound looks like. And then we feed it the model unseen data, so unseen responses into it. And the model has to guess, was this an S phoneme or was it an A phoneme? You know, it's having to predict what that EEG response was. And what we found was that the prediction accuracy of that model was significantly worse in those middle aged adults there in red.

So the dashed line just below here is chance. And so if the data were below this line, this would tell us that our model was not very good at predicting it at all. But because all the data is above that line tells us that our model was working correctly and well. But the model was much more accurate at predicting the phonemes from the EEG response for younger adults. Another metric that we can get from this model is called entropy.

So entropy is a level of, it's a metric of uncertainty. So was the model like, absolutely, that's an S, or was the model like, I don't know, it kind of looks like an S, but it could also be an S sound. And so that's what entropy is telling us how confident was the model when it was making its prediction? And for this we found that middle aged adults had much higher entropy. And so here higher entropy means more uncertainty.

So whenever it saw a EEG response in the middle aged adults, it was a lot more uncertain. It was kind of having to, like, guess a little bit about what phoneme that was. This tells us that the model was much more uncertain for classifying middle aged adult responses. So taken together with these findings here, they tell us that middle aged adults have less distinct cortical encoding of phonemes, or that these cortical representations of phonemes are fuzzier in middle aged adults, these fuzzier representations might cause more confusion during listening. So imagine, you know, you're listening to me now, and if my individual sounds that make up my SPE are not getting encoded with high fidelity in your brain, you might have more confusion.

You might wonder, you know, did she say, or did she say sh? And so that can, you know, really contribute to some Sources of speech perception difficulties. So we posit based on this, that by middle age, you know, these cortical representations of phonemes are fuzzier and are contributing to some sources of speech perception challenges in the absence of overt hear loss. So I'm going to switch gears a little bit now and talk about voice pitch. So voice pitch, you know, I'm talking in a relatively higher pitch than maybe a male counterpart.

This pitch is primarily driven by the fundamental frequency. That's the acoustic correlate of pitch, I should say. So listeners will use these fundamental frequency cues to help with speech perception in a number of ways. So in multitalker listening scenarios, we can tag onto our target speakers fundamental frequency to help filter out voices that don't match that fundamental frequency. So if I was surrounded by a bunch of male talkers right now, he would probably latch onto my fundamental frequency as being higher in pitch and know, like, okay, I don't need to listen to the ones that are lower in pitch.

And that's one of those kind of like subconscious subtle cues that we use in multi talker listening scenarios.

The fundamental frequency is also used to communicate emotion. So as I'm talking, if I'm talking about something really happy, I'm going to get higher in pitch and, you know, if I'm sad, I might get lower in pitch. And then another thing we use fundamental frequency for is to signify linguistic pitch accents. And I'll explain more on that in a little bit. So in addition to the phoneme classification results from that study that I just showed you, we also took those EEG responses from the Alice's Adventures in Wonderland part.

And we wanted to see how our pitch accents represented in the brain during listening. So you might be asking yourself, you know, well, I don't know what pitch accents are. Well, let me show you. So in English, there are four primary pitch accents that we as speakers use to communicate important information. So I kind of mapped out what these four pitch accents look like.

And in each of these examples here, this blue line is the fundamental frequency contour. So imagine here, in this first one, the fundamental frequency is starting off low, it kind of has this shallow rise, and then it starts to fall again. So in this first example, in example A, you know, someone asks who has the book? And I'm like, malcolm has the book. And I'm emphasizing the name Malcolm with this type of fundamental frequency contour to indicate that that is, who has the book.

And then in example B, here, someone says, malcolm has the book. And then I can use a Different fundamental frequency contour, which is this pitch accent, to say, Malcolm, he doesn't have it. And so by changing my fundamental frequency within that word Malcolm, it helps kind of signify uncertainty in this instance. So there's, you know, two other examples here that I won't delve into really. But essentially, these pitch accent categories carry distinct information linked to the speaker's intention.

So we speakers use this to communicate critical information. And the listeners, you know, over time, as we're growing up, we learn to kind of pay attention to this intonation to help understand speech.

With the Alice in Adventures in Wonderland data, we wanted to see how voice pitch in general is encoded in the brain between these two age groups. There have been other studies, and we replicated this here, showing that the neural coding of the fundamental frequency is reduced in middle age. And that's what we see here as well. But our question for this further part of the study was if the neural coding of the fundamental frequency is reduced in middle age, then does that impact how these pitch accents are represented in the brain? And does that maybe underlie some of their speech perception problems?

And again, these are the same participants who had relatively normal hearing threshold. So as with the phoneme part of the study, we used a very similar approach, but this time, instead of time locking the EEG to the phoneme instances, we had some students manually transcribe the audiobook for pitch accents. So they went through and, you know, here in blue are these pitch accents. Here they went through and indicated, okay, here's this type of pitch accent, here's the onset of this type of pitch accent, and so on and so on and so on. So now we are time locking the EEG responses to each and every instance of pitch accent of those four categories that I showed on the previous slide, and they give we get one single average for each of those four pitch accents for each participant.

And then we trained a similar neural network model, again, this machine learning model that is trained to classify these EEG responses into, okay, it was this H star pitch accent. No, it was this L&H; star pitch accent. So basically, the model is just trying to predict which pitch accent category that EEG response belonged to. For this part, however, we found that the prediction accuracy for the model was similar between age groups. So middle aged adults and younger adults, the model was able to accurately identify these pitch accent categories with similar levels of accuracy, I should say.

But what we did find that is interesting is the middle aged adults had much higher entropy. And as I mentioned earlier, entropy is this metric of uncertainty where higher entropy indicates greater

uncertainty. Even though the model was able to accurately predict middle aged adult responses, the model was much less confident in doing so for middle aged adults. And it seems contradictory at first, but it's just healthy the way this model is trained and built. And I'm happy again to dive into more details in the Q and A if anybody's interested in this nitty gritty stuff.

But what this is telling us is that, you know, these middle aged adults had reduced neural coding of voice pitch in general. Not just the pitch accents, but voice pitch. And so because of that, it may lead to troubles in coding pitch accents, which are critical to understanding speech, because I said that these pitch accents are communicating important information from the speaker. So these pitch accents are not coded well, then the listener may have trouble identifying the important information, they may have trouble identifying the important information in the sentence or the narrative, leading to some speech perception problems. I'm just going to summarize this EEG portion of my talk.

Middle aged adults showed fuzzier representations of phonemes in their brain, and they also showed reduced neural encoding of linguistic pitch accents compared to younger adults, despite the younger and middle aged adults showing similar hearing sensitivity. And something I didn't mention in the earlier slides was these adult, all of the participants in the study, they went through cognitive screeners. We assessed them for tinnitus. We had, you know, auditory brainstem responses. We basically did an entire full battery to show that these age groups were not different on these metrics.

But we are seeing kind of more at the brain level that middle aged adults are showing some changes relative to younger adults. So why might this be happening in middle age prior to changes in hearing thresholds? One possible mechanism that we're exploring is that it's called this age related neural de differentiation. So age related neural de differentiation is a hypothesis that has been primarily applied only to older adults. And so this hypothesis basically states that as we age, particularly in older age, our brain has reduced specialty in the sensory regions of our brain.

Instead of having high sensitivity or high specialization in these sensory regions, it spreads out that specialization across our brain. This drives some of these age related processing challenges, because rather than, you know, visual or sound processing occurring in a specific region, it's kind of recruiting the entire brain to help with that. And as I said, it's primarily been applied to older adults. But our two studies that I just showed you are suggesting that these this age related neural de differentiation may occur as early as middle age, because we're seeing that the brain is sort of not encoding these things with high fidelity anymore. And we believe that it could be contributing to some speech processing challenges.

And so with further validation with our two measures that I presented here with the phonemes and the pitch encoding, we hope to kind of validate these metrics to serve as an early indicator of this, you know, hidden hearing loss in middle aged adults that are complaining of speech processing difficulty. The other showing normal hearing through a traditional audiological battery. Okay, so taking what we just learned with all the EEG information I just dumped on you, if we think about speech sounds not being fully represented in the brain during listening, how might that affect your listening? One such effect could be increased listening effort. So listening effort is defined as the intentional and the keyword is an intentional, the intentional reallocation of cognitive resources to help with listening.

So let me break this down for you with this diagram here. Imagine in any given listening situation where there is an increased acoustic challenge. This acoustic challenge could be coming from you as the listener. Maybe you have some hearing loss, which is creating some challenge in understanding the speech. Maybe it's coming from the speaker themselves, where maybe they're under articulating their words, or maybe they have an unfamiliar accent that you're not used to, just making it harder for you to understand their speech.

Or maybe this acoustic challenge is coming from the environment through background noise, your typical speech and noise scenario. These acoustic challenges increase the cognitive demand for listening. So there's a, there's an acoustic challenge, and it means it's harder for you to hear. So listening effort is moderated by how motivated are you to overcome that challenge? And so, you know, we have the acoustic challenge increases the cognitive demand, and then you have to be motivated to overcome that cognitive cognitive demand in order for there to be listening effort per se.

And we can measure listening effort through a number of ways, the most basic being behavior responses. So this could be subJacieive reports, you know, them just saying, yeah, that was effortful. We can also look at like response times on tasks to understand how effortful it was. We can look at, or we can use neuroimaging also to look at listening effort. Oftentimes this is used with like functional magnetic resonance imaging studies.

And we can also use physiological measures. So that is EEG and pupil dilation. And I. Sorry, wrong way. I am going to focus the next few slides on pupil dilation and talk about how we can use pupil dilation to index effort, specifically where larger pupil responses are suggesting that it's more effortful.

Okay, so when I say that pupil size is related to Listening effort. It doesn't necessarily mean that it's this linear relationship. So it doesn't mean that because you're putting in more effort that you're going to do better with speech and noise. So pupil size, effort and performance. And in this case, I'm talking about speech performance, you know, your ability to understand speech.

That relationship actually follows this inverted U shaped curve here. So as the acoustic challenge increases and you're having to exert more effort to overcome that challenge, your pupil dilates. So as we move along this X axis here, the pupil is getting larger, right? Well, there becomes a point where you're putting in so much effort because the acoustic challenge is so hard. You know, there's too much noise in the environment, but you're really, really motivated.

So you're putting in a lot of effort and your pupil is going to be really large. But the challenge is just so difficult, you're not going to do well. Speech and noisy performances. You know, I study this now and sometimes I'm in environments where I'm putting in effort and I'm just like, yep, nope, this isn't going to happen. I know my pupil's big.

I am not going to be able to understand what this person's saying to me. So there's actually this sweet spot now, I'm going to refer to this sweet spot on several slides from now there's this sweet spot for listening effort and performance. And it's actually at intermediate levels. Effort. So the acoustic challenge is difficult.

It's not too difficult, and you're able to put just enough effort to reach your peak performance. Our changes in pupil size are considered to be primarily driven by the release of norepinephrine by the locus coeruleus in the brain. There are some other neurotransmitters that can be involved, but norepinephrine is the primary one that drives this dilation in our pupil sizes. And so you might be wondering, well, why does listening effort really matter? Let's think back to the adult patients that I've been talking about throughout who, you know, they have normal hearing, but they're complaining of speech and noise difficult.

We know that their hearing thresholds are technically within normal hearing limits. But say they come into the clinic and you're like, I'm just going to run you on quicksand anyways just to see. And they come out with like a 1dB, X and R lock on quicksand. And that's well within normal limits. And so you're like, you know, I really, I Think you're fine.

But let's also compare, you know, patient B. Patient B doesn't have specific complaints about speech and noise. They also show normal hearing thresholds on an audiogram. And then when we run them on quicksand, they're also showing one db, SNR loss, and quicksand. So what's the difference here?

Listening effort. So in my example here, person A in blue likely had to exert way more listening effort to achieve that 1dB score compared to patient B who didn't have speech and noise complaints. And so this accuracy, you know, this score on a test based on, you know, behavior and subJacieive responses really isn't telling us how much effort they had to put forth to achieve that score. And so I dived into this quite a bit, and I'm going to spend the next few slides showing how increased effort and behavioral performance has a sweet spot. There's some empirical data that I have in adults with normal hearing.

So in another study, I had younger adults come in, they all showed normal hearing thresholds. And we had them complete four quicksand lists. And we calculated the traditional SNR loss score in db. And this is, you know, typically what's in clinic, you run the quicksand, and it outputs their SNR loss score. So as a quick review of how this works, you know, in quicksand, each sentence has a number of keywords that need to be identified.

And so when you run them through a list, there's six sentences. And after that list, it takes the sum of keywords that were correctly identified across that list, then it subtracts it, and it gets this DBS and our loss. And what this SNR loss is telling us is the SNR level at which the person can accurately hear the words 50% of the time. And so, you know, here I've plotted everybody's SNR loss on this quick send test, and they were all within normal limits. Normal is considered less than 3 DB SNR loss.

And so they all looked really great. And so we would take that initially, basically from this behavior, that their speech and noise performance is fine. Right. However, if we look at individual performances at each SNR level in the quicksand list, we see a slightly different story. And this is not typically how we look at it in a clinical setting.

So at these easier SNRs here, I'm not showing 25, and I can explain that later. Basically, another metric we had just told us to take out 25. So we're only looking at these SNR levels here. 20, 15, 10, 5 and 0.

So for these easier SNR levels at 20, 15 and 10, everybody's performing, you know, pretty near ceiling.

They're almost at like 100% accuracy for identifying keywords. But as we get into more harder SNR levels at SNR5, there's like a slight drop in overall performance, but they're still doing really, really well at SNR5. Then when we look at the hardest SNR level of 0, there's a ton of individual variability. And maybe you as a clinician have, you know, anecdotally noticed this as well, that like when they get to that zero SNR sentence, it's either they're like, I didn't hear anything, or maybe they heard like a word or two. And then maybe you have that random patient that hears the sentence perfectly.

So there's a lot of individual variability at that zero SNR level. And so this here isn't necessarily captured by the SNR loss score. What I haven't told you yet is that while they were listening to quick sense sentences, we were recording their pupillary responses. And I've plotted here their pupillary response over time. So here is time in seconds on the x axis with the percent change in pupil size.

This is the, the diameter of the pupil, is it getting larger as we go up or down towards zero would be getting smaller. And so with quicksand, the background noise starts two seconds before the speaker begins. And so we are time locking our responses to when the target speaker began. And so when the background noise begins for each SNR level, which is coded by these different colors, when the background speakers start, you know, there's a gradual increase in the pupil response. And then when the target speaker starts, that's where things start to split.

So for the easier SNR levels at 20, 15 and 10, there's like this little kind of variability here, but for the most part, if you look at the average, it's pretty flat. Like they didn't have much change in their pupil size for those easier SNR levels. Now when we get to 5 and 0, which is orange for 5 and red for 0, orange starts off kind of low and then it gets larger. So there is an increase and then zero, there's a definite clear increase. You know, I don't have to tell you that you can see it visually with your eyes.

So based on these people responses, we get a single metric to tell us, you know, the rate of dilation over time. And so this is called the slope of dilation or slope of pupil change. And so we got a Slope of pupil change for each participant at each SNR level. And we overlaid that slope of pupil change onto their quicksand scores. So let me break down this graph for you in gray.

Up here is their Quick Sense score, which is also on the right hand side. So I showed you this on the previous slide where at these easier SNR levels of 20, 15 and 10, they were performing really, really well. Now here in black is the slope of people change and it's at zero for the most part. There's not much change during listening at these easier SNR levels. And that tells us they really didn't have to put much effort into overcoming that acoustic challenge to achieve their good performance.

When we look at SNR5 again, their quicksand performance dropped slightly. When we look at their pupil size, it actually increased quite significantly. So this is telling us that they had to increase their effort to try to maintain that maximal performance at SNR5. Totally different effect here. For SNR0, performance tanks pretty dramatically.

And then there's an even further increase in effort for zero. This is telling us that by SNR0, younger adults, specifically for younger adults. Keeping that in mind, the acoustic challenge at a zero db SNR is too difficult to maintain performance regardless of how much effort they're putting forth. So as part of this study, we also collected an envelope following response measure. This is like an EEG metric and it's referred to as the EFR.

And basically it's a 3000 Hz tone we played in their ears. And this tone is amplitude modulated at different rates so that we can get an idea of how, you know, what's the auditory processing like all the way from the periphery to the cortex. And that's kind of the extent that I'm going to go into here for the sake of time again, but I'm happy to talk about it later. So we had the pupillary responses from quicksand and we also had this EFR metric. And so we wanted to see how do these two metrics kind of

contribute to their overall speech and noise performance at that hardest SNR of zero.

So we built a model to kind of understand how much each of these is contributing to performance. The EFR metric on itself is not contributing much. This Y axis here is called the adjusted R square, which is just kind of basically what percent is it explaining in the SNR? Zero. So it's near zero.

EFR doesn't really contribute that much by itself. Then we added the slope of pupillary change at SNR0 to say, how does that tell us about their performance at zero. And it increased quite a bit to about like 15ish percent. And we, when we combine both the EFR and the pupillary metric at SNR0, we don't really see much change. So this by itself is just kind of telling us, you know, again, like I said on the previous slide, that in younger adults it doesn't matter how much effort they're putting forth at SNR 0, you know, it's just not enough to overcome the challenge.

But then we got thinking about that sweet spot of listening effort that I was telling you about earlier. And so this time I built another model and I said, I still want to know how these metrics are contributing to their speech intelligibility score at the hardest SNR of zero. But this time I want to know how does the pupillary slope of SNR5, which is where they were putting forth a lot more effort, but they were maintaining near maximal performance, is that their sweet spot? And can that tell us about how they're performing at SNR0? And what we found was that that pupillary slope on its own of SNR5 was actually very, you know, it told us a large percentage of what was going on at SNR0.

And when we combined the EFR metric with the pupillary slope at SNR5, it was explaining over 30% of the variance in their quicksand performance at that hardest SNR level. This is telling us that one SNR5 better explains performance at SNR0 than the pupillary effort at SNR0. And so we are taking this to mean that in younger adults, you know, this sweet spot occurred at SNR5. And so basically, if an individual has already reached their sweet spot at SNR5, they're going to do pretty bad at SNR0 performance wise. But if they're kind of still approaching their sweet spot, like they haven't yet reached

their peak at SNR5, then they're going to do pretty okay at SNR0.

And so this is telling us a lot about listening effort in younger adults. We then took this study and extended it to middle aged adults, ran pretty much all of the same measures. So I won't have to spend as much time here. And so, you know, we ran this follow up study and we had the middle aged adults who were 50 or, sorry, 40 to 52 years old. And for both age groups, this is their quicksand SNR loss.

You know, the traditional score, they were all less than 3 decibels, all within normal limits. And then again when we looked at, you know, individual quicksand performance, so here it's in percent. And for each SNR level, we said they do pretty well through from 25 to 10dB SNR, there's a drop in performance at SNR5 and then a further drop at SNR0. And at SNR0 is where we started to see age group differences. So the middle aged adults were doing much worse at SNR0 than the younger adults.

Then when we examine their pupillary responses, this is kind of where we also start to see some differences in age group. So these are the younger adults here. Easier SNRs are down here. There's not much change over time during listening. And then for SNR5 and SNR0, we saw this huge increase in effort.

But for middle aged adults, as the SNR difficulty increased, sort of the listening effort, it's this like nice gradual increase in effort over time with the increasing SNR level. And so when I plot that slope of pupillary change across these SNR levels, we can really see these differences between age groups. Now remember here with their behavior, sorry, it's slightly cut off here. But with their behavior, there were no differences in age groups at these easier SNR levels. And an SNR5 SNR zeros where we saw the age group differences, but with their pupil, we start to see significant differences much earlier than that.

And when we built a similar model to try to understand performance at SNR0, we found that the sweet spot in middle aged adults was at SNR10. So in younger adults it was still the SNR5 as it was with the previous study. But middle aged adults, it occurred much earlier. So with much easier SNR is, they were already hitting their sweet spot. And so anything after that was just becoming way too difficult.

Okay, so to summarize these findings with pupillary measures as a metric of listening effort, effort follows an inverted U shaped curve with maximal performance in speech and noise occurring at intermediate effort levels. We, with increasing age, we see a shift in this sweet spot of effort and performance even without a change in hearing sensitivity. So the middle aged adults and the younger adults had similar hearing sensitivities. They were all within normal limits. But middle aged adults were having to put in a lot more effort much sooner than the younger adults were.

And so we would like to take this to say that listening efforts should really be considered when working with patients. We have normal hearing, but complain of speech and noise difficulties. And I'll kind of expand on that further later on in the talk. Okay, so I've talked about EEG metrics when looking at, you know, speech reception challenges. I've also talked about pupil measures.

These are both really great techniques, but they can be a little challenging to deploy in clinical settings because they kind of require specialized equipment. Maybe not everybody has access to it. So for this last portion, I'm going to shift gears to talk about behavioral responses on speech and noise tests and what we can do with them. So for a lot of the audiometric battery, we're relying on patient responses, more specifically, accuracies. So during an audiogram, did they hear the tone, yes or no?

No. Did they raise their hand? Did they press the button for words and noise? Did they say the correct word, yes or no? You know, same thing for quicksand.

But I've already mentioned how purely getting a certain score on a test is kind of hiding some of those underlying issues. What I mean is that accuracies alone do not provide the full picture of one's listening capabilities. So the third arm of my research over the last few years has focused on incorporating patient response times to understand the cognitive aspects of listening. So the easiest example that I can give you is during an audiogram, you probably notice that as you would approach a patient's threshold, their response times get longer. So, you know, when you're starting off at like, 40 decibels, they're hitting the button right away, and then you start to get to their threshold, and maybe they don't press it, or maybe you can physically see them, like, kind of raising their hand but putting it back down.

So I tried to, like, capture how they're responding to kind of understand what's going on at a deeper level of decision making, like those mechanisms for speech understanding. So, sorry, the last thing I want to touch on here is that I've taken a cognitive computational approach using what is called a drift diffusion model, or a ddm. You might see it as DDM on the next slide to incorporate both their accuracy. So did they get it right or not? And their response times, how long did they take to answer to understand kind of those decisional aspects that are underlying speech processing?

So what do I mean by these cognitive decisional aspects? So during listening, especially in noisy environments, or maybe in middle age, when these phonemes are not getting uncoded as well, in the cortex, it's common to miss phonemes in words, and then you take a second to ask yourself, did they say bad or did they say dad? So then you start to use other contextual information to fill in that missing phoneme of the b. Or the D to create the word that you think best fit that sentence. And so you're having to use working memory to like work that out in your mind while you continuously process what the person's saying, because maybe you don't want to stop them to ask them.

You're trying to figure it out yourself. So I'm trying to capture that part. How is that happening in the brain? So the data I'm going to show you here is still incoming. So it's a little bit preliminary, but again

we had younger and older adults come in.

They did a simple computer task that takes like 10 minutes to complete. They heard a BA, a DA and a GA. Sometimes these phonemes, badaga, were in quiet, so just by itself, or sometimes it was masked in speech shaped noise at different signal to noise ratio levels. And basically they hear the sound and they just have to indicate on the keyboard which one it was. So we specifically chose these three syllables because they differ in the place of articulation.

And they were synthetically generated to have the same fundamental frequencies. So the listeners couldn't use any variations in pitch to try to decide between the three. So from this task we take the accuracy. So did they get it right, did they get it wrong? And we take the response times and we implement that into a drift diffusion model.

This drift diffusion model, the chart might look complicated, but I will break it down for you. It basically tells us how efficiently can you extract information from the stimulus and how much information do you need to accumulate? So basically think of an in, you know, in this task they hear the ba at time point zero. Immediately our brains start to accumulate information to say, okay, was that a ba, was that a da or was that a ga? And so we can get a metric from this model to tell us how efficiently they can take that information.

And then as we're accumulating this evidence, so here in red, in the squiggly line, I'm just going to say this is how we're accumulating evidence, you know, very quickly for that ba sound. And once this evidence accumulation reaches its corresponding threshold in red, that's the decision threshold. Oops, wrong way. And that's saying like, how much information do you need before you can be like, yep, that's a bar, like that's how much information I collected. And I know that's ba.

But we're also accumulating this evidence for the DA and the daw. And it's kind of a race, you know, which one can accumulate evidence the fastest and hit its threshold. And that ends up being the decision that we make. So took the accuracies and response times from this task, put it into this model and we found that middle aged adults were much slower to accumulate evidence to make their decision than younger adults. So the Y axis here is basically lower down, is slower evidence accumulation and higher, it's faster.

And for basically all of the signal to noise ratio levels and quiet, they were doing worse. SNR negative 9 is where they were kind of doing the same as younger adults. But this is also just a really difficult SNR level and I think this was just too hard. But for these other ones, you know, the middle aged adults are just slower to accumulate evidence even in quiet, which is interesting because they have normal hearing. So why is it taking them longer to accumulate evidence?

And we also saw that the, the two age groups didn't differ in the amount of information they needed to accumulate. So basically if we think about this bar, the younger adults and middle aged adults both require the same amount of information to make a decision. But the younger adults were much more efficient at reaching that bar than middle aged adults. So to summarize here with this, again this is a work in progress building a computational model to understand the decisional aspects of speech understanding. But middle aged adults with normal hearing are slower to accumulate critical evidence from a stimulus during speech processing than younger adults.

But they don't show any differences in decision thresholds. So for this last couple slides I'll be fairly quick, but I've outlined three sources of so called kind of in hearing loss. You know these, these three different metrics could be kind of capturing this form of hitting hearing loss in adults with normal hearing. I would just like to briefly outline how we're currently working to make these measures more clinically viable. Like how can we scale them to the clinic?

So currently my main priority has been to develop a cortical representation of phonemes. So that was the very first study I showed you where middle aged adults were not encoding phonemes as well as younger adults. We're trying to develop this into a more diagnostic test. So in the lab we use 32 or 64 channel EEG and 15 minutes of speech listening. That is just not feasible for a clinic.

You guys are strapped for time and purely just using like 32 electrodes, just way too many and 15 minutes of listening is just too long. So we are working with a commercial partner to trim this test down to like one or two electrodes. And we're using AI to generate an audiobook that's less than five minutes and that is, you know, the, the audiobook itself is like rich in context with the phonemes we found, like very much segregated the middle age and younger adults. So which phonemes, you know, really separated the two groups? And can we make a five minute audiobook that's heavy on those phonemes so that we can quickly capture individuals who have these fuzzy representations of phonemes using just a couple of electrodes?

We're hoping to, you know, long term deploy this on a device that's already like readily available in most clinics and you know, could hopefully one day be as like universally used as like the ABR.

So we're working on that this entire 2025. So hopefully maybe next year I have some more exciting results on that and there'll be more to come. In a second approach, we are working to refine this behavioral test that I was talking about with the drift fusion models. Like right now it's like 10 minutes, like I said, still in its early stages, 10, 15 minutes. We've tested it in person, we've tested it in an online version.

We get the exact same findings. And so we're hoping to again trim that down even further to just like five minutes. Can we give this to somebody? And it tells us what's happening at the cognitive level with decision making during speech processing because it doesn't require any specialized equipment. You can just use a computer or a tablet.

And I would also love to see pupillometry integrated into clinics. I'm not working on this directly, but I'm always happy to consult should anyone want to go in this direction. I can just imagine leveraging people as part of a battery when assessing like hearing aid fitting or post the eye implantation. Just to see like, not only are you hearing things better, but is it reducing your effort? Because that can just be exhausting if you're having to insert a lot of effort all day long.

I do know some hearing aid companies are deploying peopleometry in their research and development departments. I don't personally know of anyone clinicians using it currently. Not to say that there aren't any. I just don't know anyone firsthand. But I would love to see it used more in clinic.

Lastly, I'd like to leave you with some key takeaways. Our research shows that cognitive changes during speech perception occur as early as middle age. They have poor cortical representations of phonemes during listening, they have reduced neural coding of pitch and abstract pitch representations. They show greater listening effort, and they have slower evidence accumulation during listening. These changes may contribute to listening challenges reported by 10% of adult patients with normal hearing thresholds.

And it's critical that we continue to investigate sources of hidden hearing losses so we can better treat and advise patients that complain of these problems but aren't showing any changes in auditory thresholds. With that, I'd just like to thank my colleagues who directly contributed to some of the work that I showed, Specifically Bharat Chandrasekaran and Jay Ching Guo and the Sound Brain Lab at Northwestern University. Also Aravindi, who runs the Translational Auditory Neuroscience Lab at the University of Pittsburgh. He was previously at Mass Eye and Ear, where we got some of the data from. And then also my colleague Nike Gunana Teja, who's currently at the University of Wisconsin.

And with that, I'm happy to take questions. Okay, so Bella asked, was APD exclusion criteria. Was APD an exclusion criteria or was it ruled out for any of our study participants? So we did not specifically ask for apd, but we did ask them if they had been diagnosed with any kind of hearing disability, hearing disorders. We made sure that no one was a hearing aid user.

You know, nobody had any cis, and I think I mentioned that, you know, we screened them for tinnitus and we did like, noise exposure questionnaires, things like that. And our participants were balanced across age groups. You know, there was no severe tinnitus and no severe lifetime noise exposure. Okay, so someone asked, which of the findings did I share today that I find most surprising and why? I think there's two.

So one finding was with that drift diffusion model where they're having to categorize badaga, how middle aged adults were much less efficient at accumulating information for their decision, even in quiet, because you would think that there's really no acoustic challenge in quiet. So why would they not be able to efficiently say, oh, that's a ba or that's a DA or a ga compared to younger adults? Now, when I first saw that data, I was like, I don't understand why this is happening. And then when we take it into context with everything else, I've shown that middle aged adults have fuzzier representations of phonemes, even in quiet. And, you know, just everything else, greater listening effort in general, it just, it kind of starts to click into place that there's just something going on with increasing age by middle age, which is only, you know, the 40 to 50 range for our studies, that even in quiet they're having problems.

It seems like it's not just a speech and noise issue anymore. I think the other finding that was interesting to me. That was surprising was when we were classifying how pitch accents were encoded in the brain, that our model was like, young and middle aged adults were able to accurately classify the same, but the middle aged adults had higher uncertainty. And at first we couldn't figure out why the model was

able to do it just as accurately for middle aged olds and younger adults, but was having a harder time with middle age. And, you know, as we dove into kind of more of like, what is this model actually telling us?

You know, I did gloss over it a bit, but that higher uncertainty or the higher entropy metric for middle age, for that is telling us that each of those four pitch accent categories has similar looking EEG responses within middle age. And so it was really kind of having to, like, guess what it was, even though it ended up guessing correctly. Okay, so Erin asks, what is your opinion on hazardous noise exposure as a cause of hidden hearing loss? I certainly think noise exposure is absolutely going to contribute to hidden hearing loss. When I was talking about cochlear synaptopathy, I kind of focused it more from the aging perspective.

But when we think about aging, there's a lifetime of noise exposure there. But these studies in animal models looking at cochlear synaptopathy have shown that even just temporary threshold shifts. We go to a concert and, you know, don't wear hearing protection, and then we leave and everything's muffled. And then maybe by the next day, our hearing is back to normal. But there's evidence from animals showing that that temporary threshold shift actually has lasting damage and is kind of underlying some of this cochlear synaptopathy.

So I think as more obJacieive tests become available to test for cochlear synaptopathy in humans while we're still alive, I think that we'll see that, like, hazardous noise exposure is absolutely causing and hearing loss.

Okay, I just wanted to thank everybody for this, and then I'm going to pass the floor back to Christy. Thank you so much. Dr. McHaney. Thank you for your time and your expertise.

This is just a reminder for anyone who'd like to earn CE credit for today's presentation, please complete the multiple choice exam. Again, thank you so much for the partnership with Ava for producing this presentation. Have a great day, everyone.