QUESTIONS & ANSWERS

DR. TRUESWELL: So just a couple of announcements before we take questions here. One is that if you have a question you should use the mic. There will be people walking around with microphones. I'll stand in front of this microphone here to help you because this is being recorded. We'd like you to use microphones. The other announcement is that we're doing something different this year.

We're having—if you're an undergraduate and would like to talk with Dr. Kuhl later on this afternoon, we're having a discussion session over at the Institute for Research in Cognitive Science. Anybody who is an undergraduate who is interested in this, you're welcome to attend. It's a 3401 Walnut street on the 4th Floor and it's at, I believe, 3:00 P.M. So please come if you can. And then finally there is going to be a short reception outside here as well after the question period. So I'll let you take the questions.

DR. KUHL: Yeah. Lila.

DR. TRUESWELL: Do you have a microphone?

DR. KUHL: It's coming.

LILA: I want to go back first to the first part of your talk.

DR. KUHL: Yeah.

LILA: Where you introduced the problem that the discovery of the phonemic system is made difficult or is made impossible in its own terms by the enormous variability and overlap.

DR. KUHL: Um-hum.

LILA: Okay.

DR. KUHL: Right.

LILA: And yet the mystery is going to be that somehow or other children solve this problem.

DR. KUHL: Right.

LILA: Right. So the problem seems to be unsolvable in its own terms as we know from Bill Gates is a very smart guy—

DR. KUHL: [Interposing] Very smart—
LILA: --and his 500 cohorts working. And people have been working on this for 50 years--

DR. KUHL: [Interposing] Right.

LILA: --as we know. The problem seems unsolvable. And it seems unsolvable for the usual platonic reason, because of the overlap you don't know which items belong in the set.

DR. KUHL: Right.

LILA: Right? Which are the E's [phonetic]?

DR. KUHL: Right.

LILA: So I was interested in your first experiment, very early experiment, in which you gave infants a prototype.

DR. KUHL: Right. Um-hum.

LILA: So here's the central member.

DR. KUHL: Um-hum.

LILA: And given the central member, well then they can do statistical learning.

DR. KUHL: Right.

LILA: But we're faced with the problem here that in natural running speech the children aren't given the central member.

DR. KUHL: Right.

LILA: Right?

DR. KUHL: Right.

LILA: So. Now turning to the second half of your talk, you turn now to the social question.

DR. KUHL: Right.

LILA: Okay. And as I saw in the specially the Spanish guy--

DR. KUHL: [Interposing] Yeah, yeah.

LILA: --with the little kids.

DR. KUHL: Yeah.

LILA: Okay. It seems to me that you see something much more specific than saying well here was the computational and
here's the social.

DR. KUHL: Um-hum. Um-hum.

LILA: But more specifically, social at its margins, yeah, you have to get them interested. But that's just an enabling condition. So I want to try this on you--

DR. KUHL: [Interposing] Yeah.

LILA: --the causal condition is the following: you have to solve this problem top down to solve the Plato problem.

DR. KUHL: Right.

LILA: Right, which are in the system.

DR. KUHL: Um-hum.

LILA: So you have to know which count as the E's.

DR. KUHL: Right.

LILA: Right? Or you can't solve the problem. Just too much variability. Right? But you can solve it if it's peat-a-ble, peat-a-ble [phonetic]; it's always the same object.

DR. KUHL: Right, right, right, right, right.

LILA: So if you're socially interested in the problem--

DR. KUHL: [Interposing] Um-hum.

LILA: --that enables--but that problem's computational too--

DR. KUHL: [Interposing] It is.

LILA: --it means that you have to solve it top down.


LILA: So what do you think of that?

DR. KUHL: I think those are wonderful questions. So let's go to the first one first about the prototype. So where do prototypes come from? I mean okay so are they platonic, well one answer was they're platonic ideals. They're built in. And we say no. You know, they're derived statistically through a process that we don't quite understand but they're derived from statistically distributions. Maybe it's--yeah, that's one possibility. But computers aren't solving it that
way, right.

So the second one which comes from the social piece is to say that it's actually mediated through words. That knowing that something is referred to by different people in the same way allows you to decide, okay, that everything that is referred to in the same way must be part of that statistical distribution. However, you know, it's hard to imagine that the kids at 6 months or the data show that by 6 months the babies in Stockholm and the United States are solving it already. So either they're cracking the word code much earlier than we think they are, possible, possible, or there's something else going on that we don't quite understand.

Oh yes, you may only need a few words. You may only need a few words. And you know, the mother-ese phenomenon does do this in spades by not only naming objects in the presence of kids and providing more prototypical examples, right? Much more prototypical in mother-ese as we saw from the stretched triangles, but also varying--they do two things. They develop a more prototypical average but they also have greater extent of variation. It's almost as though mother-ese tries to be multiple talkers. So it's possible that's a leg up on the solution which would explain why do computers not do it. Why isn't it just statistical? What's the piece of the social? You know, the word, maybe the vehicle in.

The Spanish experiment's interesting because we measured both words and phoneme learning at the same time. Kids are exposed to Spanish, 12 sessions, post session we measured both phonemes and words. We asked the question. At the individual level do all kids show 1 of the 2 first? So some kids may learn 1 and not the other. Show learning of 1 and not the other. And some may show learning of both. The kids who learn both or neither are not as interesting as the kids who show learning of 1 and not the other.

Interestingly, we see it almost--we're not done with this analysis. This is not--so don't hold me to this. Kids seem to do it differently. Some kids are showing phoneme learning, not word learning yet. Other kids are showing word learning, not phoneme learning yet. Suggesting that different kids may go at this differently which would be very interesting. But we'd have to do more analyses; we'd probably have to do more experiments to understand exactly
what they are. But you're absolutely right. They've got this problem.

They've got a big problem, variation all over the place. Social mediation to help steer them in the right direction. Great computational skills but you still have to know over what to compute. How do I compute? So categorical perception will help them somewhat, right, because they are hearing these boundaries. So they've got some gross partitioning. That would say gather the statistics on this acoustic part of the map, put them all together. So it could be something sloppily like that until they hone it to objects. We don't know. Keep us in business for a while, right? Go ahead.

FEMALE VOICE 1: So you said that Autistic kids prefer robotic sounds?

DR. KUHL: Yeah.

FEMALE VOICE 1: Is it because it's just disturbing to them or can robotic language actually help them learn language better?

DR. KUHL: Well it may be both of those two. So on the one hand what we see is behavior that looks like they're afraid of the face and the voice. They will never choose the face when they have a choice. So they cover their eyes if we don't give them a choice. And they also, if you give them a listening choice, they will steer clear, the more emotive the signal sounds the less they like it.

So I think, you know, MEG data, again, you know, what's happening to the amygdala when they hear faces—when they see faces that are very animated as opposed to an object. They always prefer objects. So it's possible that fear centers in the brain will illustrate some oddity of their brain's processing of signals. Somehow what typical babies find pleasurable, speech is the hands down choice of children given a choice, and faces, most interesting visual stimulus. That's different in children with Autism. And I think that could prevent their learning.

The second question is what if you presented speech to children with Autism through the robot with a robotic voice? We haven't tried it yet but I think it's extremely interesting to think about that. It might help this initial mapping process.
FEMALE VOICE 1: Thanks.

DR. KUHL: Yeah. And we'll come back over here.

MALE VOICE 1: Hi. I wanted to stick up for the machines a little bit.

[Laughter]

DR. KUHL: Okay go for it.

MALE VOICE 1: The scatter plot that you showed of English vowels which I think was published in your 2004 Nature Reviews Neuroscience--

DR. KUHL: [Interposing] Yeah.


DR. KUHL: Right.

MALE VOICE 1: And they did--

DR. KUHL: [Interposing] And earlier from, you know--

MALE VOICE 1: --they did--

DR. KUHL: --Peterson and Barney.

MALE VOICE 1: --well they replicated Peterson and Barney--

DR. KUHL: [Interposing] Yeah.

MALE VOICE 1: --but they did discriminate analysis on the measurements for those vowels. And they got 97% correct of--

DR. KUHL: [Interposing] That's--

MALE VOICE 1: --all identification--

DR. KUHL: [Interposing] Yeah.

MALE VOICE 1: --92% to 97% depending on exactly how they set it up. So in fact that plot showed F1 and F2, they threw in F0 in duration.

DR. KUHL: Right which was helpful. Yeah.

MALE VOICE 1: which is helpful because--

DR. KUHL: [Interposing] Yeah.

MALE VOICE 1: --humans also get that--
DR. KUHL: [Interposing] Right.

MALE VOICE 1: --so I'm not trying to suggest--

DR. KUHL: [Interposing] Yeah.

MALE VOICE 1: --that the speech recognition problem by machines has been solved but I think you--

DR. KUHL: [Interposing] Yeah.

MALE VOICE 1: --radically exaggerated--

DR. KUHL: [Interposing] Yeah.

MALE VOICE 1: --how far away--

DR. KUHL: [Interposing] Okay.

MALE VOICE 1: --the machines are--

DR. KUHL: [Interposing] No that's fair. Yeah that's fair.

MALE VOICE 1: --so you also said that speech recognition requires words to be separated and that actually hasn't been true for quite a long time. It certainly doesn't get perfect transcription. And there's certainly plenty of research to be done in that area but I think there might actually be an opportunity for greater confluence of research--

DR. KUHL: [Interposing] Yeah.

MALE VOICE 1: --interests and results--

DR. KUHL: [Interposing] Right.

MALE VOICE 1: --for the engineers--results that the engineers--

DR. KUHL: [Interposing] Yeah I'm excited about the--

MALE VOICE 1: --from your research as well.

DR. KUHL: --okay, thanks for making that clarification. I'm excited about the engineering perspective and think that, you know, greater experimental sharing between human learning systems and machine learning systems is really important. So a post-doc of mine, Debo [phonetic] and I actually looked at machine learning for mother-ese and adult directed speech and could see the value of mother-ese signals in that kind of thing. Especially with regard to the robustness after training to new data sets.
So I think there has been a lot of progress. I think the progress may not be happening so much at the phonetic level and I was only trying to illustrate with that diagram from Peterson and Barney, I mean, what my understanding is we have not got the ability to play natural language into a machine learning system and have it derive the categories of the language. I mean that's the ultimate test of whether or not the machine can process natural language in a way that allows it to sort. So when speakers like in the Hillenbrand or Peterson and Barney are producing he'd, hid, had, hood, you know, highly constrained CVC, prototypical examples, that's not really a fair test that compares them to the babies.

So but I do think that the mother-ese productions, not the kind of stylistic stuff we do when we are trying to be made understood by a machine, so sometimes when the machine is learning the algorithms, make a person repeat, they do just the wrong stuff, right? They do something that's not what we do when we pronounce these clear instances of mother-ese. I think there's a lot to be learned by using input into machines that really constrain the syntax, constrain the vocabulary and enhance the acoustics in the way that mother-ese does.

So I think that it will be a very promising line of studies. And if we ever get to the place where robots are, you know, capable of teaching foreign languages to kids I think, again, the engineering psychology links will be stronger still to try to understand what kinds of learnings can we do from machines, what kinds of learning can machines do and humans do that are comparable or not. That remains an interesting question.

I used to tell Bill Gates that his computers need a life and a brain. But I've modulated that a little bit over the years.

FEMALE VOICE 2: I'd just like to comment on your work with the Autism. Conventionally it's difficult to diagnose an Autistic child until they are about 18 months or so--

DR. KUHL: [Interposing] Right.

FEMALE VOICE 2: --and yet there's another test that works completely in agreement with your findings. If you project on a screen the picture of the mother, the Autistic kids will look at the eyes.
DR. KUHL: Um-hum.

FEMALE VOICE 2: The normal ones will look at the mouth.

DR. KUHL: Yes.

FEMALE VOICE 2: And that seems to mesh very well--

DR. KUHL: [Interposing] Right.

FEMALE VOICE 2: --with your findings.

DR. KUHL: So on McClune's data.

FEMALE VOICE 2: Yeah.

DR. KUHL: Yeah. I think that's really interesting. And, you know, from the standpoint of the clinicians as well as the scientists, the earlier that you could diagnose these children and the earlier you could attempt treatments that are getting at what you think the fundamental cause is, the better we would be both scientifically and from a clinical standpoint. I think it's fascinating. I think these early precursors, the eye gaze, the social work that McClune's doing and this auditory, this sort of early language precursors in the auditory domain have real promise. Again tapping the social and the linguistic preparation for language.

Thank you.

Other comments? Yeah.

MALE VOICE 2: Just a basic question about for children who are exposed to Mandarin Chinese between the 8 to 10 months, so have you studied them after they became older--

DR. KUHL: [Interposing] Yeah.

MALE VOICE 2: --for--do they keep the ability or do they lose it?

DR. KUHL: Well we're really interested in the forgetting function for adults and children, thinking that it's quite different. That children, as we saw in these slides, learned so very well and we think that it has some staying power but not forever. So we brought the kids back in at 14 months and we could see the ERP signatures of that learning are still there. That's, you know, quite a few months after, without any intervening Mandarin experience. But then by 18 months it's gone. So we don't see the neural signature any more.
The hypothesis would be, but we couldn't do that with these kids, if you brought them back at 3 or at 5, that they would look different when exposed again to Mandarin, than the kids would if they had never been exposed, like the control group. We didn't have enough of these children we could bring back in. But we'd love to do experiments where we demonstrate that there is some staying power. This early experience that you have really has a kind of potency with regard to the neural structure and the vestiges of that would remain quite a long time.

Whereas in adults we know that when we're exposed to foreign languages we sometimes learn but it doesn't stay. So there's something different about the short term memory system and long term memory systems for language and speech that's going to link to the critical period that protects us from being written over, right, if you go to Japan for 2 months, it wouldn't be good if your English skills were written over by your newly forming Japanese skills. So there's something about memory systems and the critical period that we'd like to understand. And maybe the tractable experiment are these forgetting functions after interventions to see, you know, what's happening. And of course brain imaging with that would tell us something about what's happening in the auditory areas, in Broca's, and in executive function areas.

MALE VOICE 2: Thank you.

DR. KUHL: Yep.

MALE VOICE 3: Well there's a phenomenon that's been quite important in my own work which is quite puzzling but all linguists are familiar with it. Children, there's massive evidence to indicate that children do not acquire the foreign accent of their parents.

DR. KUHL: Yes.

MALE VOICE 3: And this happens quite lately in a period that seems to be considerably later than the intensive learning program--

DR. KUHL: [Interposing] Than the really early stuff.

MALE VOICE 3: So I was wondering if that later liability can be modified by social factors of a sort that we don't yet understand.
DR. KUHL: I totally agree with that. I think that we may see that the social factors are so potent that when you take—we've been looking at toddlers. Toddlers who come into the country say from Japan or from Taiwan, speaking either Mandarin, their families have spoken Mandarin or Japanese. The parents are not good English speakers and so they are attempting a little bit of English at home but basically the children are exposed to one language at home and a different language in preschool.

And what we see is that as soon as they develop—this is mostly anecdotal, ethnographic kinds of research, that as soon as they develop a social group, a group of pals, a group of friends, the 3-year olds, 4-year olds, that they completely lock onto the language style of their peers. So I interpret that as a social effect. We don't have it made tractable.

I'm more interested in teenagers who are, you know, edging towards the end of the critical period and there's also there a potency with regard to social factors and a group of emotionally close kids and a mimicry, this is in speech production data, a mimicry of the kids in their social group. So I think the social factors are really strong.

The questions are why aren't they strong for adults. Why don't social factors—or do we avoid social factors? As adults we protect ourselves socially by kind of—we don't move to foreign countries and develop close friends to the same degree as kids would do when they're kind of desperate for that contact. What is it about the social that might change over time if this is a real fundamental explanation to what's going on with learning mechanisms? So I think that's a puzzle. But it's absolutely true for the kids that this peer group will, you know, win out over the parents hands down, very potent. John.

DR. TRUESWELL: So this is somewhat related to that. I think really, I don't know the developmental literature well enough to say this for sure but it seems like an understudied topic in this area is the adaptation to particular speakers.

DR. KUHL: [Interposing] Um-hum. Um-hum.

DR. TRUESWELL: --speakers, adapting to speaker variability. It would seem especially for these computational problems that,
you know, one of the things that the child has to learn is, and is a model of an individual speaker, how quickly can you develop--

DR. KUHL: [Interposing] Um-hum.

DR. TRUESWELL: --a model of the priorities of this individual speaker--

DR. KUHL: [Interposing] Right.

DR. TRUESWELL: And how do these map onto the so-called invariance, or maybe that's what the invariants are, developing a model for each--

DR. KUHL: [Interposing] Um-hum. Right.

DR. TRUESWELL: --speaker quickly.

DR. KUHL: Um-hum.

DR. TRUESWELL: And so I guess part of the question is are there studies of that sort going on? You know, how early can infants adapt to different speakers?

DR. KUHL: Yeah.

DR. TRUESWELL: And also, you know, would this be part of the puzzle that, you know, people--

DR. KUHL: [Interposing] Um-hum.

DR. TRUESWELL: --a crucial piece of the puzzle that's being overlooked right now.

DR. KUHL: Yeah. I don't think the right kinds of experiments are going on and these are the ones where the engineering approach would be very interesting to compare to infants.

So what you want to do is model learning with machines and with infants if you could control the kinds of speaker variability that they're exposed to. So on a completely statistical explanation; you'd imagine the kids are attempting to develop the distribution of allowable instances of a particular type. Now you have to decide what you're averaging over, but let's say that you can limit the extent to which there are speaker variances in the input of a child. Maybe a child raised only by the mother for the first 6 months, just to do a thought experiment, right?
So you'd be over-representing in your statistical distributions. Her utterances, they would simply look like her. And so when other females, males and children come into the picture that statistical distribution has to be modified until the very next instance you hear doesn't change the distribution any more. So once you've heard what, a million instances of a vowel and you've heard many different talkers of different ages, is that the point at which you stabilize your distribution doesn't change any more? Because the next instance is not new. It's subsumed by the distribution.

So how that happens, the stability and how many instances does it take, you know, we have absolutely no idea how much it takes. And then we also know that we drift, both in speech production and in speech perception. If you go to Britain you will drift your categories. You'll sound more British with regard to your productions. You'll also perceive things differently. So why is that you don't overthrow your distributions when you're trying to learn a foreign language and you're not completely, you know, mapping a whole new structure when you go to Japan for 2 months?

Yet you can drift when you go to another country. You're beginning to sound different and your perception is shifting if you do micro perceptual tests. So how does the neurobiology protect us from the overthrow problem while still allowing us to adjust, again, socially why would we do it? Well because socially it's sort of interesting to adjust the behavior patterns, the speech patterns, the listening patterns of the social group you're in.

Comparing human learning experiments with machine modeling experiments that look to see what does it take to stabilize a distribution, think of it just statistically, that's an interesting problem. And I don't--I've been thinking about it and I write about it but only as a, you know, codunkin (phonetic), it's not something that I'm solving in experiments yet. It would be nice to design those and conduct those. I think that's a really interesting question. But I don't know how that works. We have a degree of openness but we're not completely open. How is that?

DR. TRUESWELL: Well thank you. That's all the questions. Thanks--

DR. KUHL: [Interposing] Yeah thanks, that's great.
[Applause]

[END 164074.MP3]