globeandmail.com
print edition
Home | National | World | Business & Investing | Sports | Opinions | Arts | Technology | Travel

SCIENCE & IDEAS: LANGUAGE: THE HUMAN SPEECHOME PROJECT

Someone to watch over me

Every move, every burble, every stage of development of their young son is being recorded as two university professors study how humans learn to speak - and whether robots can be taught the same way.
Susan Pinker reports

SUSAN PINKER
December 13, 2008


Globe Focus

IN DARK AND LIGHT

It all started with a dare. The first parry came from Deb Roy, now an associate professor at the Massachusetts Institute of Technology Media Lab but then a PhD student at MIT, who was trying to make a point to a female colleague over dinner. He claimed that, given exactly the same input, a robot could learn the rudiments of language as well as a child.

Rupal Patel, then a doctoral student in speech pathology at the University of Toronto and now an expert in prosody - the melody of human speech, had her doubts. Her response was, "Prove it."

Fast-forward 10 years, and Drs. Roy and Patel - originally from Winnipeg and Calgary respectively - are now university professors in Boston, married to each other and the parents of two preschool children: the first-born, a thoughtful, delicate-looking boy, aged 3, and an outgoing 18-month-old girl.

Using their young family as a model, the couple are currently engaged in the most ambitious account of human-language development ever attempted - with Dr. Patel firmly onside in the testing of Dr. Roy's hypothesis about language.

By wiring up their 1970s suburban bungalow with 11 omni-directional fish-eye cameras and 14 high-performance microphones, Dr. Roy and Dr. Patel are capturing every burble, babble and bye-bye uttered by their son, his parents and caretakers during his first years of life. Eventually, data from these multiple ceiling-mounted cameras and microphones will be fed into a learning machine that Dr. Roy is designing to test his initial proposition: that a robot can "step into my son's shoes" and extrapolate rules from its environment the way a child can.

Day to day, it's like a professional home video that runs nearly all the time, creating not only a dynamic family portrait, but an exceptionally naturalistic one. "These are candid moments that capture the texture of everyday life," Dr. Roy says. And there are a lot of them - about 250,000 hours of moments to date.

The ceiling cameras do not reveal everything: The children's faces are not visible, and the parents turn off the cameras and microphones at their own discretion - during marital discussions, baths or diaper changes, for example. But spontaneous outbursts and sudden shower exits do happen. So, at Dr. Patel's suggestion, the system also includes an "oops!" button that erases recent footage, like the "many times I had forgotten to turn off the system while nursing, changing or dressing," she wrote in an e-mail.

Even so, the information collected so far - about 200 gigabytes, or about two laptops' worth a day - is a fairly complete record of how human interaction begins. The couple have dubbed it the human speechome project, a reference to the broad implications of mapping a single unit. And like the human genome project, it may change our understanding of the ways nature and nurture come together as babies grow.

Scientists - including Charles Darwin and Jean Piaget - have long recorded their observations about their own children. More recently, Ruth Weir and two economists, Sharon Oster and Ray Fair, recorded the sophisticated crib talk of their child, now a young economist named Emily Oster, who was subsequently immortalized by Malcom Gladwell in The Tipping Point, and by Freakonomics author Stephen Dubner, in Slate.

The practice of recording children's language has continued, with a public database of parent-child interactions called the Child Language Data Exchange System (CHILDES) that was created, for the most part, by psychologists taping and transcribing their children's speech. The online project, started in the early eighties by psychologists Brian MacWhinney, at Carnegie Mellon University in Pittsburgh, and Catherine Snow, now at Harvard University, has accumulated parent-child interactions in 20 languages and at several sites, including one at Memorial University of Newfoundland.

Still, Dr. Roy and Dr. Patel's efforts will be the most complete and accurate record of early child development to date. Because the technology is as unobtrusive as pot lighting, there is no "observer effect" - that is, the subtle self-monitoring that often occurs when subjects know they are being watched.

Dr. MacWhinney, who is one of the co-directors of the CHILDES speech transcript archive, estimates that researchers who tape children's speech capture perhaps 1.5 per cent of a child's interactions - usually a total of one hour a week. That's why he calls the speechome project groundbreaking. "It's the technology that drives the science," not the researcher's availability or expectations, he told Wired magazine.

But for those not stirred by high-tech toys or the disk farm in Dr. Patel and Dr. Roy's basement, what's really impressive about the speechome project is its bird's-eye view of the development of their son. Its crisp snapshots of early childhood resonate with any parent. When Dr. Roy played a video that captured his son's tentative first few steps at Idea City in Toronto in June, 500 people watched spellbound as the tiny boy tottered toward his father, paused, then whispered, "Wow!" under his breath.

In a 20-second audio clip punctuated by the occasional sound of water from the tap, the audience also heard the toddler's exclamations morph from ga-ga to goh-gah to wa-wa to wa-tah - a pattern in the mountains of data that Dr. Roy calls "space-time worms," created by editing together months of his son's attempts to say the word "water."

But it's not just isolated recordings of milestones - almost any parent can do that. The continuous hubbub of family life is picked up on video too - the conversation and clatter of dishes at dinner, the quiet cuddles, and games of peekaboo that characterize life with a baby.

This rich social tableau forms the backdrop for any child's first words, but until now, it has been missing from what has been known about how children develop.

To fund the project and underwrite the vast array of the technology required - hand-ground video lenses, 1,000 metres of cable, the terabyte disk array in the basement and a $2.5-million upgrade of MIT's Media Lab retrofitted to accommodate the monthly rolling suitcases of data Dr. Roy transports from home to work every month - he obtained seed money from the National Science

Foundation and backing from more than 70 corporate sponsors, including Toyota, Google, the Lego Group, PepsiCo and Hallmark.

Dr. Roy, trim, compact and often dressed in black, has a flair for plain speech. "I would often shoot my mouth off about what my technology could do," the 39-year-old computer scientist says, trying to explain how

he got from designing "chatboxes" - talking dictionaries - to a marital wager that includes an omniscient home data recording and retrieval system that he has dubbed Total Recall.

It all started with tinkering, he says. "I've been building things - machines and robots - since I was a kid, and playing with computers as soon as I could get my hands on them. I've long been fascinated by questions about how we think and how we learn, but my natural inclination was not the scientific method of asking questions. What I'm really good at is building things. And in this case I'm building technology that might tell us how we learn in the first place."

In contrast to Dr. Roy's technical focus, Dr. Patel is more interested in how parents and children synchronize their speech, expressing not only meaning but endowing their interactions with emotion and purpose. And it's one reason why she was skeptical when Dr. Roy asserted that "this robot is learning."

Now that their joint project has progressed, Dr. Patel hopes that the data will lead to a deeper understanding of how children use the signals buried in the melody of their parents' speech to understand their intentions. She mentioned just one application - the diagnosis of postpartum depression in new mothers, whose speech rhythms and expression are altered by their damped-down feelings. "That impacts the baby's language acquisition," she says, excited about the possibilities of using this technology to identify an emotional state that is painful and sometimes dangerous to both mother and baby, yet usually invisible.

Much less ambitious in scope but designed to gather similar data as the human speechome project is a portable recording device that has recently been developed by Dr. Roy and his team at MIT. As it resembles a floor lamp, they have called it the speechome lamp, and it will be tried out in two autism clinics early next year, recording family interactions and behaviours that may help experts diagnose autism in infancy, before speech normally appears.

Skepticism accompanies most novel ideas, of course. Janice Goldfarb, a family physician and the parent of a child on the autistic spectrum, wondered about the cost of such a device, who would pay for it and how severe the child's symptoms would have to be for the lamp to detect them. "Right now, in order to get treatment, you'd need a diagnosis by a specialist anyway, such as a pediatric psychiatrist or neurologist," she says. But she adds that it could be useful to reassure parents, many of whom, in her experience, are concerned that their late-talking children might be on the autistic spectrum - and would be willing to pay for a tool that would rule that out.

Discounting the cost, it will take some time to address the perennial questions that dog any screening tool, from mammograms to tests for giftedness or prostate cancer. Would it be sensitive enough, for example, to identify a baby who does not make enough eye contact with his or her parents during the first months of life? Or would it make too wide a sweep, ringing alarm bells when a baby is a bit more reserved than the average child?

Also skeptical about its immediate applications is Simon Baron-Cohen, a professor of developmental psychopathology at Cambridge University and one of the world's authorities on autism.

"Whilst this approach could be fruitful, it could also result in too much data, which may not fit the requirements of a quick, precise, highly sensitive and highly specific early-detection method for autism," he said in an e-mail. Still, he added, projects like the human speechome should be encouraged to see what they produce.

As with the human genome project, one cannot anticipate how a rich database like this one will be used. Critics of that endeavour who initially asked, "Why bother sequencing this junk," never anticipated that genome research would illuminate the ancestry of Neanderthals, as well as the genes for height or language disorders.

The mountains of data emerging from the human speechome are certain to help us understand the complex interplay between nature and nurture - though exactly how is anyone's guess.

In the meantime, Dr. Patel is suggesting that they keep the cameras rolling as their younger child starts to talk.

"I've been the one advocating to keep it on for longer because I see such differences in the two children," she says. "I know it's only a sample of two, but the differences in gender and personality are so striking. It's not at all what I expected, so I'm looking forward to what we'll learn next."

Susan Pinker is a developmental psychologist and the author of The Sexual Paradox. Her next book is about the science of human relationships.