

The Incredible Thing We Do During Conversations

When we take turns speaking, we chime in after a culturally universal short gap.



Khaled Abdullah Ali Al Mahdi / Reuters

ED YONG | JAN 4, 2016 | SCIENCE

TEXT SIZE
- +

One of the greatest human skills becomes evident during conversations. It's there, not in what we say but in what we don't. It's there in the pauses, the silences, the gaps between the end of my words and the start of yours.

[When we talk we take turns](#), where the “right” to speak flips back and forth between partners. This conversational pitter-patter is so familiar and seemingly unremarkable that we rarely remark on it. But consider the timing: On average, each turn lasts for around 2 seconds, and the typical gap between them is just 200 milliseconds—barely enough time to utter a syllable. That figure is nigh-universal. It exists across cultures, with only slight variations. It's even there in sign-language conversations.

“It's the minimum human response time to anything,” says [Stephen Levinson](#) from the Max Planck Institute for Psycholinguistics. It's the time that runners take to respond to a starting pistol—and that's just a simple signal. If you gave them a two-way choice—say, run on green but stay on red—they'd take longer to pick the right response. Conversations have a far greater number of possible responses, which ought to saddle us with lengthy gaps between turns. Those don't exist because we build our responses *during* our partner's turn. We listen to their words while simultaneously crafting our own, so that when our opportunity comes, we seize it as quickly as it's physically possible to.

“When you take into account the complexity of what's going into these short turns, you start to realize that this is an

elite behavior,” says Levinson. “Dolphins can swim amazingly fast, and eagles can fly as high as a jet, but this is our trick.”

Conversation analysts first started noticing the rapid-fire nature of spoken turns in the 1970s, but had neither interest in quantifying those gaps nor the tools to do so. Levinson had both. A few years ago, his team began recording videos of people casually talking in informal settings. “I went to people who were sitting outside on the patio and asked if it was okay to set up a video camera for a study,” says Tanya Stivers.

While she recorded Americans, her colleagues did the same around the world, for speakers of Italian, Dutch, Danish, Japanese, Korean, Lao, ~~≠~~Akhoe Hai//om (from Namibia), Yélí-Dnye (from Papua New Guinea), and Tzeltal (a Mayan language from Mexico). Despite the vastly different grammars of these ten tongues, and the equally vast cultural variations between their speakers, [the researchers found more similarities than differences](#).

The typical gap was 200 milliseconds long, rising to 470 for the Danish speakers and falling to just 7 for the Japanese. So, yes, there’s some variation, but it’s pretty minuscule, especially when compared to cultural stereotypes. There are plenty of anecdotal reports of minute-long pauses in Scandinavian chat, and virtually simultaneous speech among New York Jews and Antiguan villagers. But Stivers and her colleagues saw none of that.

“Dolphins can swim amazingly fast, and eagles can fly as high as a jet, but this is our trick.”

Instead, they uncovered what Levinson describes as a “basic metabolism of human social life”—a universal tendency to minimize the silence between turns, without overlaps. (Overlaps only happened in 17 percent of turns, typically lasted for just 100 milliseconds, and were mostly slight misfires where one speaker unexpectedly drew out their last syllable.)

The brevity of these silences is doubly astonishing when you consider that it takes at least 600 milliseconds for us to retrieve a single word from memory and get ready to actually say it. For a short clause, that processing time rises to 1500 milliseconds. This means that we *have* to start planning our responses in the middle of a partner’s turn, using everything from grammatical cues to changes in pitch. We continuously predict what the rest of a sentence will contain, while similarly building our hypothetical rejoinder, all using largely overlapping neural circuits.

“It’s amazing, like juggling with one hand,” says Levinson. “It’s been completely ignored by the cognitive sciences because traditionally, people who studied language comprehension were different to the ones who studied language production. They never stopped to think that, in conversations, these things are happening at the same time.”

Pessimists among us might view this as the ultimate indictment of conversation, a sign that we’re spending most of our “listening” time actually prepping what *we* are going to say. (As Chuck Pahlaniuk once wrote, “The only reason why we ask other people how their weekend was is so we can tell them about our own weekend.”) But really, this work shows that even the most chronic interruptor is *really* listening. “Everything points to what astute observers we are of every word choice, every phonetic change,” says Stivers.

And of course, we can change the length of the gaps when we need to. “You don’t want to respond as fast as possible to everything,” says [Stivers](#), now at the University of California, Los Angeles. “If I ask someone to go to a movie with me and they rapidly say no, that doesn’t feel nice. It’s better to have a gap before you turn someone down for something. And if you hesitate, I can say, ‘...or not tonight?’ We’re pretty good at adjusting.”

Levinson now wants to understand how our turn-taking system evolved. It certainly seems to predate language. Great apes like chimps take turns when gesturing to each other and other primates, including several monkeys and

one species of lemur, take turns when calling. One team of researchers recently showed that [pairs of common marmosets](#) leave predictable gaps of 5 to 6 seconds between turns, and will match a partner’s rhythm if it speeds up or slows down. These simian see-saws could be independent innovations, or they could reflect an ancient framework that we humans built upon when we evolved the capacity for speech.

The researchers also want to understand how turn-taking develops throughout our lives. So far, studies have shown that even six-month-old infants respond to their parents very quickly, albeit with more overlaps. At nine months, when they start to grasp that they’re actually communicating with another mind, they slow down. After that, it takes a surprisingly long time to get back to adult speeds. Stivers has found that even 8-year-olds, who have been speaking for many years, are still a few hundred milliseconds slower than adults. “That’s a puzzle that I don’t have answer to,” she says.

Related Video



The Geography of Small Talk

▶ 0:00 / 4:00 🔗 🔊 ↗

Surprising alternatives to “so what do you do?”—from New Orleans to New York

ABOUT THE AUTHOR



ED YONG is a staff writer at *The Atlantic*, where he covers science.
