

Let's TALK

*Listening in on conversations around
the world reveals that human dialogue
follows some universal rules that enable
and enrich our social interactions*

By Mark Dingemans and N. J. Enfield

ILLUSTRATIONS BY JAMES YANG

The Study of animal communication has a long and colorful history. In the 1950s Dutch biologist Niko Tinbergen collected stickleback fish and carefully observed how they interacted. He noticed that the abdomen of male fish would flush bright red during breeding season, as the fish built nests and established their territories. This color served as a warning signal to rivals—so much so that Tinbergen found that territorial males would lunge at any object with a similar hue, including wood blocks he held outside their tank and even a mail van passing by the laboratory window.

Tinbergen's work—which combined the observation of natural behaviors with systematic experimentation—not only earned him a Nobel Prize, but it became a model for the study of animal communication. This classic approach has proved

so successful in understanding how animals interact, it seemed only natural to use it for investigating human discourse. Our goal was to discover what people—from a variety of cultures, in the act of everyday conversation—could tell us

about the structure of human language.

Quite a lot, it turns out. During the past 10 years our team and others have traveled around the world, learning different languages and listening in on conversations. By analyzing our collective data and returning to the field for further exploration, we have learned that human language has a structure that transcends grammar and goes beyond the words we use and the order in which we deploy our nouns and verbs. This conversational “infrastructure” is the same in all cultures, from the rice fields of Laos to the fjords of Iceland. By teasing apart these commonalities in communication, we are coming closer to understanding the





universal principles that form the foundation for language and, ultimately, the fabric of human societies.

Your Turn

Humans spend more than half their waking hours—and a great deal of their mental resources—interacting with one another. And a good portion of those social encounters involves speech. We use words to cement our relationships, exchange information and build social networks. So to better understand the behavior of our species, it seems, we need to study how people use language.

Language comes into play in all our dealings, but perhaps the most fundamental use of language is conversation. Engaging in verbal back and forth is how we first learn to speak and how we car-

ry out the business of social life in our families and our communities. For these reasons, we focused our efforts on the kind of chitchat that makes up commonplace exchanges.

The study of conversation is, in itself, not new. In the 1970s American sociologist Harvey Sacks of the University of California, Irvine, co-established the discipline of “conversation analysis”—the detailed study of how people use language in day-to-day life. Sacks had been involved with a suicide prevention center in Los Angeles, and it was while working with recorded telephone calls to the center that he became intrigued by the orderly structure of conversation. One thing he noted was that the transitions between one speaker and the other were fairly smooth and well coordinated, so

that—for the most part—only one person talked at a time.

How do we manage such fluid give-and-take? Sacks and his colleagues Emanuel Schegloff of the University of California, Los Angeles, and Gail Jefferson, then at U.C. Irvine, pointed out that our understanding of the rules of grammar should allow us to determine when any utterance is finished. For example, “I know the owner” is a complete statement, whereas “I know the” is missing something. Thus, using grammar as a guide, we can predict when our conversation partner’s “turn” will come to an end.

In 2006 one of us (Enfield) joined forces with psychologists Holger Mitterer and J. P. de Ruiter, both then at the Max Planck Institute for Psycholinguistics in Nijmegen, the Netherlands, to explore this model further. We recorded spontaneous conversations between friends on a telephone we had set up. We could then manipulate those recordings to determine what cues people use to anticipate when it is their turn to speak. Some of our subjects listened to the original recording. Others heard a robotic-sounding version in which the pitch of each speaker’s voice was com-

FAST FACTS

A GLOBAL GABFEST

- 1 The study of mundane, everyday conversation is providing insights into the higher structure of human communication—a system that exists across languages and cultures.
- 2 The timing of responses during conversation tends to minimize both silence and overlap. So accustomed are we to anticipating this pattern that a hesitation of even half a second is interpreted as dissent.
- 3 Humans display a unique ability to interrupt conversation to request clarification. In dozens of languages, this basic act is triggered by a word that sounds like “Huh?”

pletely flattened. Yet others heard a version in which we allowed the voices to rise and fall naturally but used a filter to mask what they were actually saying. What we found was that listeners had no trouble predicting when each “robot” was done talking—yet they performed terribly when they heard the conversational lilt but not the words being spoken. The results indicate that grammar is indispensable for conversational navigation.

Everybody Talks

People are not only good at taking turns while speaking, they are also remarkably quick to jump in once they determine it is their time to speak. In the 2006 study with de Ruiter, we made more than 1,500 measurements of the time it took for one person to begin speaking once the other had finished. We found that most of the transitions occur very close to the point at which there is no silence and no overlap: the average lull in the conversation was around 200 milliseconds—less time than it takes to blink an eye. This turnaround time is so rapid that it suggests people must gear up to speak—mentally planning what they will say next—while their partner is still talking. That way we can initiate our next contribution as soon as our partner yields the floor.

So far the work we have discussed focused on English and Dutch, languages that are fairly closely related. Yet centuries of linguistic research have shown that the languages of the world can vary radically at just about any level, from the sounds they use to the words they have and the manner and order in which words are combined into sentences. Does impeccably timed turn taking occur in all languages? Or are some cultures less hurried in their speech, whereas in others, people trip over one another to make their thoughts known?

The first study to address this question in a systematic way was published in 2009 by Tanya Stivers, then at Max Planck, Enfield and their colleagues. The 10-member team spent years in sites on five different continents, learning the

languages, getting to know the local people and their customs, and videotaping daily activities—including the most mundane conversations. Every team member reviewed their recordings and extracted a set of 350 sequences involving a question and response. When the transitions were measured from one speaker to the next, the findings were strikingly similar to the studies of Dutch and English: people, it seems, try to avoid talking over one another or letting

Making conversation involves more than anticipating when to begin speaking. It is a cooperative venture that requires adherence to the rules of social engagement.

too much time lapse between one utterance and the next. Again, the average gap falls around 200 milliseconds.

Answer Me

The other thing that is striking about human speech is that people expect answers. Making conversation involves more than just anticipating when to begin speaking. It is a cooperative venture that requires adherence to the rules of social engagement. This kind of verbal accountability does not occur in animal communication. Although creatures sometimes engage in a form of call and response, their vocalizations are not as precisely timed or intimately linked as human dialogue. Many animal calls are purely informational—“I am here” or

“Look out: snake!”—and they do not warrant or require a vocal response.

So deeply ingrained is our expectation of a rapid reply that any hitch in the flow of conversation is subject to interpretation. Think of a politician hesitating before replying to a question about the use of illicit drugs. Or how you feel when you ask someone on a date and are met with a silence that feels like it stretches on forever before the person either accepts or declines. In these exchanges, even the slightest pause can feel evasive or seem like a sign of difficulty or doubt.

American linguists Felicia Roberts and Alexander Francis, both at Purdue University, have been examining this phenomenon more closely. In one study, the investigators produced recorded conversations in which one speaker made a request (for example, asking for a ride) and the other answered “sure.” They then experimentally manipulated the length of time that passed between the request and the reply and played these recordings for a group of undergraduates. The students were asked to rate how willingly the respondent seemed to agree to the favor. The results were clear: once the lag in response stretched to about 500 milliseconds—just half a second—listeners began to interpret the delay as a reduced willingness to cooperate, even though the speaker’s answer was “sure.”

That study, published in 2006, was conducted in English. But do people from different cultures across the globe make the same assumptions about the social implications of silence? Roberts

THE AUTHORS

MARK DINGEMANSE and **N. J. ENFIELD** are in the language and cognition department, directed by Stephen C. Levinson, at the Max Planck Institute for Psycholinguistics in Nijmegen, the Netherlands. The authors have worked together since 2010, when Dingemans joined Enfield’s European Research Council project on Human Sociality and Systems of Language Use.

Human conversation has a rhythm, and deviations from that precisely timed patter are rife with social meaning.

and her colleagues subsequently extended their study to include Italian and Japanese. Across all three languages, they found that the longer the pause in conversation, the lower the perceived willingness to comply or agree.

And that perception may indeed be

warranted. In all the languages that Strivers and her colleagues studied, they found that positive responses always came faster than negative ones. Thus, deviations from the average length of time it takes for one person to reply to another may be legitimate indicators of the relative enthusiasm of the speaker. Moreover, the tendency for people to engage in this type of rapid social appraisal is shared across cultures. Thus, it appears that the timing of comments within conversations is part of the structure of our language—not just of one language but of all human communication.

“Say What?”

So human conversation has a rhythm, and deviations from that precisely timed patter are rife with social meaning. Yet not all conversations proceed smoothly, without a glitch. What happens when one

speaker does not catch what the other one said? Such failures in communication could lead to serious misunderstandings if it were not for another important feature of human conversation—our natural tendency to ask for clarification. And the simplest tool we have for that purpose is the humble word “Huh?”

You have no doubt uttered this query countless times—and heard it even more. It turns out, the term may well be universal. In a large-scale study of 200 conversations recorded in a dozen different countries, from Ghana and Laos to Italy, Iceland, Russia and Japan, we found that a word that sounds like “Huh?” occurs in every language we examined. And it always serves the same purpose: it temporarily halts the conversation and prompts the speaker to repeat or rephrase what was just said.

“Huh?” may sound like a random grunt, but our study indicates that it qualifies as a word. Children are not born knowing how to say it—they have to learn it as they learn any other word. Also, it is not a simple reflex. Our closest evolutionary relatives—chimps and other apes—do not appear to grunt “Huh?” although they do sneeze and yelp like we do.

The word is subtly different in each language—depending on the local intonations of that tongue. But it is always a single syllable and sounds like a question. And its short vowel sound—“uh” or “eh”—is extremely easy to pronounce: open your mouth and put a question mark at the end of the simplest sound you can squeeze out, and you come up with “Huh?” These qualities serve its function well: the brevity of the word quickly notifies the speaker that there is a problem, and its questioning quality encourages an equally rapid response.

“Huh?” is not the only word we use to “repair” broken links in a conversational thread. Different cultures also have different phrases to call for clarification, and even in English we often ask, “What?”, “Sorry?”, “Pardon?” or “You mean...?” From the conversations we recorded, speakers called for clarification or explanation an average of once every minute. This frequency—and universality—



Thinking about Talking

Ever been told to “think before you speak”? It is a gentle reminder to say what you mean—and mean what you say. But it also highlights a fundamental property of human communication: language involves both the mental assembly of words and sentences and the sharing of those assemblages with another individual.

Linguists, too, can come at language from either direction. Noam Chomsky of the Massachusetts Institute of Technology, for example, tends to take a thought-based approach. He and his followers are interested in studying

our capacity for language generation—how we build words out of sounds and sentences from words.

To the linguists who adopted Chomsky’s approach, which dates back to the 1950s, how these grammatical compilations form the basis of conversation seemed beside the point. After all, speech can be sloppy. The well-constructed, pristine thoughts we assemble in our brain run the risk of getting garbled as they make their way through our imperfect vocal systems and then get interpreted by a listener who may or may not have heard exactly what we said.

We elected to take on language from the speech side of the equation: analyzing how words and sentences are used to communicate. Our studies are uncovering the social roots of language and showing how the structure of conversation enables us to share a piece of our mind. Together with the studies that explore how we put together words and sentences, this approach is giving us a better understanding not only of what we say but how and why we say it.

—M.D. and N.J.E.



ty—indicates that in some ways human social interaction hinges on the verbal devices we use to make sure we understand what is being said. In a sense, then, it is in what we do when things go wrong in conversation that the uniquely social nature of human language becomes clear.

Why We Speak

What do these findings tell us about the function of human language? First, it is clear that conversations have a certain structure. Participants take turns speaking, prepare their thoughts, anticipate when their input is expected, and call for clarification and correction as needed. This intensely cooperative form of interaction is adhered to across a variety of cultures and is unmatched anywhere in the animal kingdom. These mechanisms—including turn taking, timing and repair—form the foundations of our linguistic abilities. They are like the “fundamental forces” that hold together the words and sentences of conversation, providing them with a certain social weight and flavor. And like physicists studying the composition of matter, we look forward to continuing our search for these fundamental particles and

interactions that inform human speech.

That people can make use of these common structural elements to construct meaningful conversations reflects something psychologists call our social intelligence: a way of thinking in which we intuit one another’s communicative intentions, holding one another accountable for what we say and when and how we say it. This propensity for reading into the minds of fellow individuals reflects the unique sociality of the human spe-

cies. We use language to build our relationships and to work together—in small groups, in larger institutions and at the level of societies. Without the social glue of conversation, these linkages would not exist, and societies might crumble. Learning more about the natural language of our species—by studying how people from a variety of cultures communicate every day—will continue to reveal fundamental insights into the very essence of what it means to be human. **M**

FURTHER READING

- **Projecting the End of a Speaker’s Turn: A Cognitive Cornerstone of Conversation.** J. P. de Ruiter, Holger Mitterer and N. J. Enfield in *Language*, Vol. 82, No. 3, pages 515–535; September 2006.
 - **Universals and Cultural Variation in Turn-Taking in Conversation.** Tanya Stivers et al. in *Proceedings of the National Academy of Sciences USA*, Vol. 106, No. 26, pages 10,587–10,592; June 30, 2009.
 - **Judgments Concerning the Valence of Inter-Turn Silence across Speakers of American English, Italian, and Japanese.** Felicia Roberts, Piera Margutti and Shoji Takano in *Discourse Processes*, Vol. 48, No. 5, pages 331–354; 2011.
 - **Is “Huh?” a Universal Word? Conversational Infrastructure and the Convergent Evolution of Linguistic Items.** Mark Dingemanse, Francisco Torreira and N. J. Enfield in *PLOS ONE*, Vol. 9, No. 4, Article No. e94620; November 8, 2013. Corrected version published April 2, 2014.
- From Our Archives*
- **The Curious Behavior of the Stickleback.** N. Tinbergen; *Scientific American*, December 1952.
 - **Programmed for Speech.** Sebastian Haesler; June/July 2007.
 - **Speaking in Tones.** Diana Deutsch; July/August 2010.