Nicholas Christakis

NETWORKED

Exploring the weblike structures that underlie everything from friendship to cellular behavior

If a campaign volunteer shows up at your door, urging you to vote in an upcoming election, you are 10 percent more likely to go to the polls—and others in your household are 6 percent more likely to vote. When you try to recall an unfamiliar word, the likelihood you'll remember it depends partly on its position in a network of words that sound similar. And when a cell in your body develops a cancerous mutation, its daughter cells will carry that mutation; whether you get cancer depends largely on that cell's position in the network of cellular reproduction.

However unrelated these phenomena may seem, a single scholarly field has helped illuminate all of them. The study of networks can illustrate how viruses, opinions, and news spread from person to person—and can make it possible to track the spread of obesity, suicide, and back pain. Network science points toward tools for predicting stock-price trends, designing transportation systems, and detecting cancer.

It used to be that sociologists studied networks of people, while physicists and computer scientists studied different kinds of networks in their own fields. But as social scientists sought to understand larger, more sophisticated networks, they looked to physics for methods suited to this complexity. And it is a two-way street: network science "is one of the rare areas where you see physicists and molecular biologists respectfully citing the work of social scientists and borrowing their ideas," says Nicholas Christakis, a physician and medical sociologist and coauthor of *Connected: The Surprising Power of Our Social Networks and How They Shape Our Lives* (2009).

The basic elements of a network are simple: it consists of nodes

BY ELIZABETH GUDRAIS

connected by links (also called ties). But as the numbers of nodes and links increase, the number of possible configurations grows exponentially. Likewise, there are innumerable possibilities for what a node and a link can represent: a word, a gene, or a person, in the first case; phonetic similarity, coincident expression, or a conversation, in the second. Structurally simple, yet analytically incredibly complex, networks hold the answers to so many questions that at Harvard alone, the number of researchers studying them may reach three digits. Here is a sampling of the newest work in this dynamic field.

"STUFF SPREADS" IN MYSTERIOUS WAYS

Christakis, professor of medicine and medical sociology at Harvard Medical School (HMS) and professor of sociology in the Faculty of Arts and Sciences, and University of California political scientist James H. Fowler '92, Ph.D. '03, wrote *Connected* after discovering that each was working on a special case of network effects (the effect of a spouse's death on one's own health, for Christakis; the spread of voting behavior, for Fowler) and realizing they shared an interest in what else could be spreading through networks.

The book is an exuberant romp through the field, presenting findings from medicine, epidemiology, evolutionary biology, sociology, anthropology, political science, economics, mathematics, and beyond. The authors discuss the spread of laughter, tastes in music, sexual behavior, and anxiety over nut allergies. They note one study that rigorously compared the structure of networks of myriad phenomena and found a strong similarity between the bill-sponsoring patterns of U.S. senators and social licking in cows. They report that *Physarum polycephalum*—slime mold—is more efficient than Japanese graduate students in finding the shortest route through a maze (the fungus can "collaborate" by fanning out in the form of a network to explore all possible paths); and share Japanese mycologist Toshiyuki Nakagaki's follow-up studies, in which the fungus was as good as or better than humans at devising maps for railway systems in Great Britain and Japan. These studies, they say, demonstrate the problem-solving power inherent in networks.

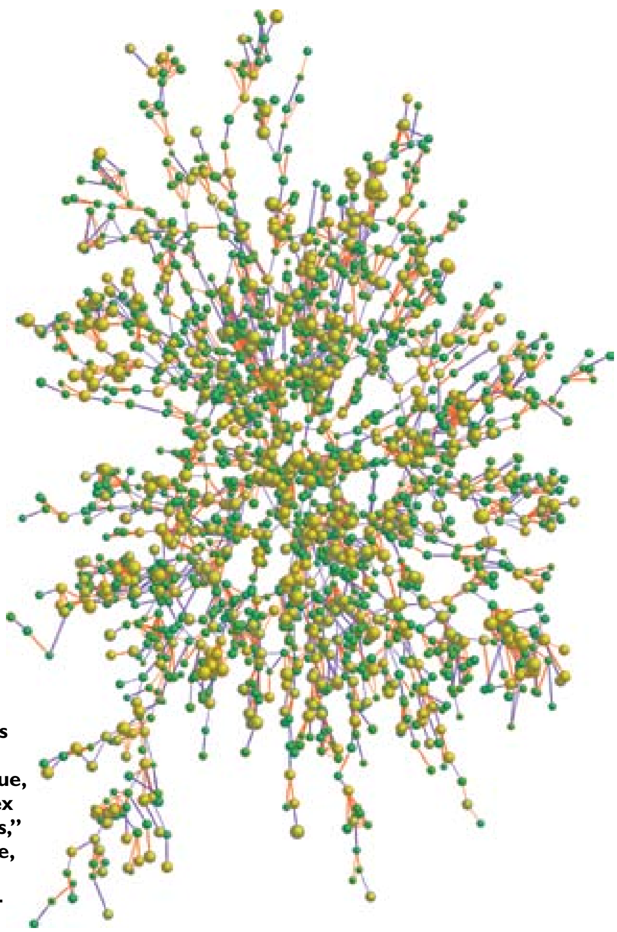
Visit harvardmag.com/extras to see a video of slime mold solving a laboratory maze, and one of mold forming a Japanese railway system map

These wide-ranging, sometimes wacky findings reflect the field today. The boundaries between disciplines are becoming all but meaningless in network analysis; Christakis's lab group includes scholars of physics, economics, anthropology, computational biology, sociology, and healthcare policy. "Often new knowledge is produced at the intersection of disciplines," he says, "and in network science this is happening in spades."

But the core of the Christakis-Fowler collaboration is original research on what spreads through human social networks. With data from the Framingham Heart Study, under way since 1948, they mapped more than 50,000 social ties among 5,124 people (who connected outward in turn to a network of more than 12,000 people). Because the study tracked all manner of health markers and asked subjects about an exhaustive list of behaviors—diet and exercise, medications, recreational substance use, emotions—it was a rich lode of data.

The two men started publishing their findings with a splash: a 2007 article in the *New England Journal of Medicine* reporting that *obesity* spreads through social networks, as people are apparently influenced by friends' weight gain to become obese themselves. More perplexing is their finding that obesity spreads through up to three degrees of separation. If a subject named a friend who was also in the study, and that friend's friend became obese, the first subject's chances of becoming obese were roughly 20 percent greater. Across one more degree of influence (husband's friend's friend or friend's sibling's friend—i.e., three degrees away), the risk was 10 percent greater. Weight gain appears to ripple through friend groups via some unseen mechanism such as altered eating or exercise behavior, or adjustment of social norms regarding weight.

The authors found similar patterns for happiness, loneliness, depression, alcohol consumption, the decision to stop smoking, and even divorce. "Our health depends on more than our own biology or even our own choices and actions," they write in *Connected*. "Our health also depends quite literally on the biology, choices, and actions of those around us."

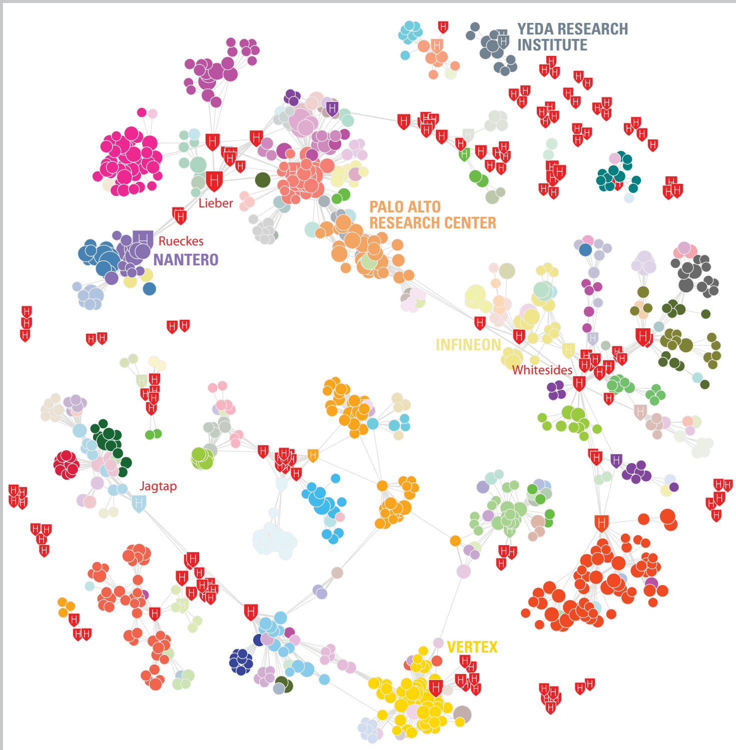


COURTESY OF NICHOLAS CHRISTAKIS AND JAMES FOWLER

THIS MAP is a typical example of the computer-generated images used to help understand networks and network effects. A subset of the "obesity network" mapped by Nicholas Christakis and James Fowler, the image shows 2,200 subjects from the Framingham Heart Study (from a total of 12,067). Each dot, or "node," represents one person (red borders indicate women; blue, men). The yellow dots represent obese people—those with a body mass index (BMI) of 30 or more—and node sizes are proportional to BMI. Colors of "ties," or links between nodes, indicate relationship type: purple for friend or spouse, orange for family. Note the visible clusters of obese people; Christakis and Fowler report that network effects—in which one subject's weight gain influences the BMI of those around him—help to explain these obesity clusters.

Many people want to achieve success as inventors; few actually do. Lee Fleming, Weatherhead professor of business administration at Harvard Business School, is studying whether someone's position in a network seems to matter when it comes to determining who succeeds.

Using data on all U.S. patents since 1975, his team mapped innovation networks for inventors across the country. The network of Harvard inventors who applied for patents between 2003 and 2008 is shown here, with larger nodes indicating that an inventor's patents are frequently cited, i.e., influential; a connection between nodes indicates collaboration on an invention, measured when the inventors list themselves jointly on a patent application.



RONALD LAI AND ALEX D'AMOUR, INSTITUTE FOR QUANTITATIVE SOCIAL SCIENCE

At a basic level, this map illustrates the global connections of Harvard inventors. They collaborate not only with each other, but with the distinguished Palo Alto Research Center in California; the multinational General Electric; and the Yeda Research Institute in Israel. They collaborate in a variety of domains: nanotechnology (Nantero), electronics (Infineon Technologies), and pharmaceuticals (Vertex).

This method also allowed Fleming to examine the question of what network structures seem to bolster creativity. Taking a cue from the definition of creativity as the combining of familiar ideas in unexpected ways, Fleming parsed "novel combinations"—the first time a single patent combines two subclasses of technology—from the database, which goes all the way back to 1790. He then asked what network structure surrounds people who filed patents with these novel combinations, as well as those whose inventions prompted the U.S. Patent and Trademark Office to create a new subclass. Comparing the models of the "broker"—an influential person connected to many others who don't know each other—and the "connector"—an influential individual with a habit of introducing his collaborators to each other—he found that brokers are more likely to come up with new ideas, because they are situated at the center of a group and communication goes through them. (Flowers University Professor George Whitesides, a chemist whose work has spawned more than a dozen startups, and Hyman professor of chemistry Charles Lieber, whose projects include small-scale devices for communicating with neurons in novel ways, are clear examples of brokers in the diagram.) But brokers have a harder time getting their ideas publicized, relative to connectors. Fleming found that brokers whose ideas became influential most often were connected to a "gatekeeper" who was part of a more highly integrated network and could disseminate the idea there. (Gatekeepers in the diagram include Prakash Jagtap, who worked as a scientist at Harvard Medical School in 2001 and 2002, and now directs drug discovery at the Lexington, Massachusetts, firm Inotek; and Thomas Rueckes, Ph.D. '01, cofounder and chief technology officer of Nantero in Woburn, Massachusetts.)

The database also enabled tracing inventor mobility, from firm to firm or university to university, across the last 35 years. Fleming and Matthew Marx, M.B.A. '05, D.B.A. '09 (now an assistant professor at MIT's Sloan School of Management) determined that statewide enforcement of noncompete clauses—where companies bar employees from working for a competitor for a set period of time after leaving the employer in question—induce brain drain. States that enforce such clauses are particularly likely to lose their most productive and well-connected inventors, for whom opportunities in other states are easy to come by. An effort is under way to change Massachusetts law to prohibit the enforcement of these clauses—a change Fleming and Marx support.

In a new project, with Vetle Torvik of the University of Illinois, Fleming will integrate patent information with information on publication and collaboration from the PubMed database (which contains more than 15 million scientific journal articles), with information on government grants (for instance, from the National Institutes of Health), and with information about commercial outcomes, in a database that will be publicly available online. "We're going to be able to trace the process of knowledge generation," he says, "all the way from government funding, through scientific publishing and patenting, to what firms were founded and how successful they were."

READ MORE ONLINE. Visit harvardmag.com/extras to read three online-only sidebars to this article:
COSTS AND BENEFITS OF CONNECTION. Read more about Nicholas Christakis and James Fowler's research, including their responses to critiques of their work.
NETWORKS, NEOLITHIC TO NOW. The characteristics of human social networks seem to have persisted through time—and to have a genetic basis.
VIRTUAL FRIENDSHIP, FOR REAL. The digital revolution has created new modes of social interaction. How do online and offline friendship differ?

Laura Bogart

For each trait that spreads through networks, Christakis and Fowler (and others working in the field today) meticulously chart how, and between whom, transmission occurs. Does geographic proximity matter? Are family relationships more influential than social relationships? What about people who work together? The answers vary depending on what is being transmitted.

Precise knowledge is needed for the type of network-based public-health interventions they envision. In addition to knowing *what* works—in the case of obesity, perhaps distributing healthy recipes, or posting on Facebook or Twitter that you "feel so great after going for a run" to encourage friends to exercise—such interventions require knowing *who* is most influential, and this may vary from purpose to purpose. Christakis and Fowler write that a network-based vaccination campaign, targeting people with the most social contacts, could be three times more cost-effective than a campaign that aims for universal vaccination. Campaigns of the latter type over-vaccinate; immunizing only people who are hubs in social networks would enable administering a minimum of doses for maximum effect. (Recommendations that healthcare workers receive more vaccinations than average citizens follow a similar model, assuming that such workers will have more contact with sick people and thus are more likely to spread infections.) A network-based *surveillance* campaign, prioritizing well-connected people when monitoring infection's spread, could be *700 times* more efficient than random monitoring.

But when it comes to diet and exercise, is it better to have people with *more* connections float the healthy recipes and exercise messages, or to have the positive signals come from *close* friends of the target? That puzzle has not been solved, even though efficient public-health spending depends on the answers to such questions.

A CONTAGIOUS CONSPIRACY

While fowler and christakis are concerned mainly with tracking the spread of behaviors and mental states through networks, Laura Bogart is interested in how *information* spreads through networks—and then *influences* health behavior. The associate professor of pediatrics at HMS and researcher at Children's Hospital Boston has used a social-network perspective to study the spread of HIV conspiracy beliefs. She led a national telephone survey of African Americans in which more than half the respondents agreed with the statements "There is a cure for AIDS, but it is being withheld from the poor" and "A lot of information about AIDS is being held back from the public." Such beliefs "are a response," she says, "to years of discrimination, and years of reasons to be suspicious about medical treatments, including unethical practices in the medical system and knowledge about wide disparities" in care.

These beliefs may indeed be a natural response under the circumstances, but they are dangerous: male respondents who agreed with the conspiracy statements were significantly less likely to use condoms consistently. In a separate study of HIV-positive patients, Bogart found that giving credence to such beliefs was negatively correlated with adherence to treatment: 25 percent of patients who did not follow treatment recommendations also agreed with the statement, "People who take the new medicines for HIV are guinea pigs for the government," versus just 8 percent of patients who *did* take their medication as directed. And in a third study, Bogart used social-network analysis to determine that the greater the number of friends who had mentioned a conspiracy belief to a study participant, the more strongly the participant believed the statement.

She has now begun a larger-scale study of how HIV information—conspiracy theories and otherwise—flows through social networks. Her team will track 240 HIV-positive African Americans in Los Angeles for one year, as they go in and out of treatment, on and off medications, and friendships form and erode. Later she hopes to test interventions, identifying "opinion leaders" in the community and offering them HIV education in the hope that they will spread scientifically accurate information that will supplant the myths.

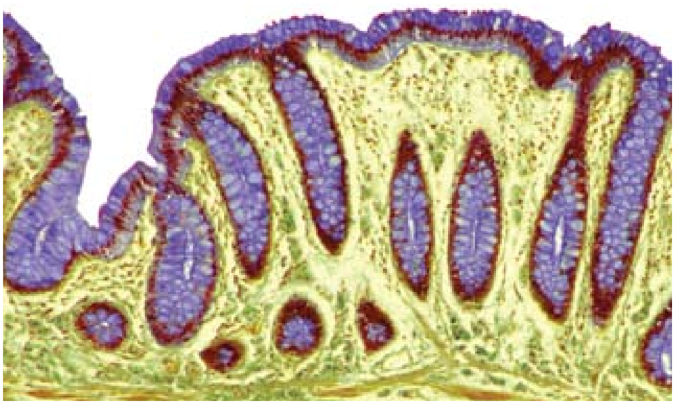
It may come as a surprise that friends' statements of opinion can have such tangible impact, says Bogart, who holds a doctorate in social psychology. "People like to think of themselves as independent thinkers," she explains. "They don't think about social networks as having such an influence on them."

In fact, beliefs and medical mistrust can "have a huge influence on health behaviors," she says. But research and policy discussions tend to focus so much on structural impediments to care—not having transportation to the clinic, lacking insurance—that "we forget that people, in the best of circumstances, sometimes *still* will not take their medication." Her work highlights the importance of social networks in explaining why.

THE EVOLUTIONARY CASCADE

Martin Nowak became interested in networks while trying to develop equations to explain how cancer emerges in the body. "I realized that in order to write down these equations, I had to understand the network of cells in the body," says Nowak, professor of mathematics and biology and director of Harvard's Program for Evolutionary Dynamics.

Nowak began this mathematical quest by studying colon cancer, which arises in *crypts*—tube-shaped glands (below) that produce the cells to renew the colon's lining as old cells slough away. Each crypt comprises 10,000 cells, with one stem cell, or a few at most, at the bottom. Nowak characterized each crypt as a network, with the stem cell(s) at the center and the genetic material passed along through each cell division, from precursor cells to terminally differentiated ones. Each crypt is a dynamic network, changing with time: as the cells at the end of the line—the inner surface of the colon—die away, new cells replace them, although the lineage still begins with the same stem cell.



STEVE GSCHMEISSNER / PHOTO RESEARCHERS, INC

COLON CRYPTS (blue) are visible in this colored light micrograph of a human colon section. Through multiple divisions, the stem cells at the bottom of each crypt give rise to the terminally differentiated cells that form the colon lining (top). The crypts helped Martin Nowak craft a mathematical description of networks that suppress selection: they are organized in a way that discourages the development of cancer. Cells near the surface, which divide often—making cancerous mutations relatively likely—are soon sloughed away, but in stem cells, which pass on mutations to all their daughter cells, division is much less frequent.

As Nowak studied this system, he says, "I realized that our tissues are actually organized in such a way as to protect us from cancer." If a mutation occurs near the surface, "then most likely it's washed out and nothing happens." If a mutation occurs at the stem-cell level, on the other hand, "it will change that crypt into a lesion, a likely site for later development of cancer."

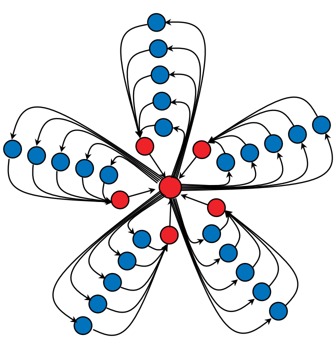
Numerous systems in the body—the hematopoietic system, the skin, and the epithelial layers of the lungs and the breast ducts—behave in the same way. For example, within the entire hematopoietic system—whose stem cells become, through multiple rounds of differentiation, various types of blood and lymph cells—10^{12} cell divisions occur *daily*. "Every one of those cell divisions is a risk for cancer, but most of the divisions happen in cells that don't live long enough to cause cancer," says Nowak. "A stem cell in that system—where it really *would* be dangerous to get a cancer-causing mutation—divides only a few times *per month*."

This realization built on a 1992 Nowak finding. Since Darwin's day, he explains, mathematical representations of evolution had "assumed that populations are well mixed, that everybody is equally likely to bump into everybody else. In reality, populations are not well mixed." Nowak developed a mathematical framework that incorporated the varying likelihood that one organism will meet another (based on living in the same city, sharing a workplace, etc.) and his new research has extended the theory. Just as each human being doesn't have an equal chance of procreating with every other human being, all cells don't have equal likelihood of mutating in a way that promotes cancer, and then dividing and propagating that mutation. Because cancer cells *do* have a fitness advantage relative to normal cells—they live longer and divide more frequently—cancerous mutations would spread quickly in a well-mixed population. But because of the way the cellular network is structured, Nowak says, malignant mutations "have a very small chance of taking over."

Nowak and colleagues have seized on this attribute as the defining feature of networks that suppress selection: even strongly advantageous mutations do not spread through the population. Their theory, published in *Nature* in 2005 and 2006, offers a mathematical description of some networks that make sense intuitively: for example, a small lake feeding a stream that flows into a larger lake. Within each lake, natural selection will operate freely in the fish population. Mutations that occur in the smaller lake will affect fish in the larger lake—but mutations that occur in the larger lake, no matter how advantageous, will never reach the fish in the smaller lake because the stream flows only one way.

The 2005 *Nature* paper also set forth structures that *amplify* selection, including a star-shaped network with one central individual connected to all other individuals, but without connections among the outlying nodes. These structures can be used to understand not only genetic evolution, but also cultural evolution, says Nowak: "If one person has an idea, which network is best for the spread of this idea?" The star structure works well if all the connections are bi-

Martin Nowak



directional (see diagram at left): a good idea that originates in the center is immediately transmitted to the outlying nodes. A good idea that originates at one of the outlying nodes quickly reaches the center and is disseminated from there. Thus, Nowak's research, which began with cancer-cell biology, has flowered into a broadly applicable theory of network structures that inhibit or promote natural selection.

At harvard and elsewhere, network analysis is evolving at an explosive pace. Scholars are using methods they could never have imagined, with collaborators they might never have envisioned, to analyze, in some cases, entirely new types of data.

Another Christakis collaborator, assistant professor of statistics Edoardo Airoldi, is tracking the spread of news on the Internet; the study's purpose is primarily theoretical (it aims to better understand the concept of diffusion in networks), but it will require inventing new tools that could have practical applications for journalism and public relations. Jukka-Pekka Onnela, a physicist who is a postdoctoral fellow in the Christakis lab, has used network analysis to help understand patterns of movement in stock markets—a method that has been picked up by at least one national central bank.

Separately, Onnela has analyzed the call records of 7.2 million mobile-phone users in an undisclosed European country. The researchers got access to 18 weeks' worth of records, for a total of 22.6 million links between callers (each link representing at least one call placed). They were charmed to find that their results displayed a fundamental property of human social relations, and a tenet of network analysis, formally elaborated in 1973 by Harvard sociologist Mark Granovetter, a pivotal figure in the field. In his theory of "the strength of weak ties," Granovetter found that as tie strength—indicating closeness of friendship—increased, the number of common friends also increased. In the cell-phone study, the more time two people spent on the phone with each other, the more likely they were to have commonalities in the list of other people they called. This suggests that tie strength in the cell-phone study probably does correspond to strength of friendship as Granovetter defined it. "We have lots of theories from philosophy and social science about how society is organized," says Christakis. "Now we have data and methods to test a lot of them. It's phenomenally exciting."

A more lighthearted study led by a member of Christakis's lab group searched for meaning behind users' decisions to make their Facebook profiles public or restrict who can view them. It found that users with public profiles had a higher-than-average chance of listing the Beatles, Pink Floyd, and Led Zeppelin among their favorite musical artists, whereas people who restricted access to their profiles were more likely to list Coldplay, Rage Against the Machine, and Ray Charles. The taste for privacy (or public exposure) may be correlated with a personality type that also prefers certain musicians; the researchers aren't quite sure, yet, of

ON THE TIP OF YOUR TONGUE

In the diagram shown here, connections are drawn between words differentiated by a single sound—a framework that has helped researchers understand how humans process language. In a test designed to take users to the tip-of-the-tongue state (where they are at the edge of being able to recall a word), people had an easier time retrieving words that were "well-connected"—those with a high number of words that sound similar. For example, subjects had a harder time generating the word *pub* when asked, "What is the British word for a drinking establishment?" than they did generating *bail* when asked for the verb that means *remove water from a boat*. (Although some words—*stub, hub, grub*—sound similar to *pub,* many more rhyme with *bail*: *sail, jail, mail, tail, hail, fail, nail, pail*, as well as the numerous like-sounding words ending in *-ale.*)

In subsequent research, Samuel Arbesman, a postdoctoral fellow with a background in computational biology working in Nicholas Christakis's lab, joined psychologist Michael Vitevitch of the University of Kansas and mathematician Steven Strogatz of Cornell in comparing languages and found that this basic principle—words that are well-connected are easier to generate—was universally true, but truer of some languages than others. For example, the pattern was less pronounced in Spanish than in English: having many similar-sounding words *helped less* in Spanish than it did in English. The researchers believe this is due to the intrinsic properties of the languages themselves.

In Spanish, words that sound similar tend to mean similar things. To take a simple example, *acendrado* and *acendrada* are masculine and feminine versions of the same word, *pure*. In English, on the other hand, words that sound similar often have vastly different meanings (think of *necklace* and *reckless*). This is partly due to the diverse linguistic influences on English (from Romance as well as from related Germanic languages, for example). In addition, English uses inflection less than Spanish—so where English would add an extra word, Spanish often adds just

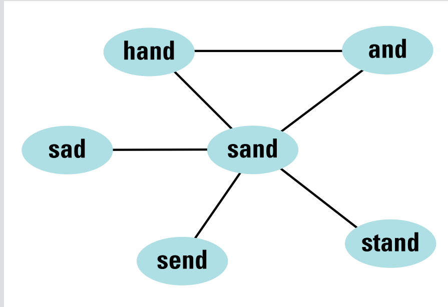


FIGURE COURTESY OF SAMUEL ARBESMAN, MICHAEL VITEVITCH, AND STEVEN STROGATZ

a suffix: for instance, *Ella volverá* for *She will come back* (based on the infinitive *volver,* "to come back").

In English, as one might predict, people had an easier time generating words that lay on "well-traveled" phonetic paths in the brain. But in Spanish, the researchers theorize, there seem to be so many words that are not only similar phonetically but also have similar meanings that the brain simply gets mixed up.

When the researchers moved from speech *production* to speech *comprehension*, they saw the trends reverse. They tested speech comprehension by measuring people's language-processing *speed* (for instance, how quickly they decided whether a snippet of speech was a real word or nonsense) and processing *accuracy* (for instance, people's facility in accurately hearing words played against background static). In these speech-comprehension tests, having lots of similar-sounding words seemed to get in the way of remembering the right word for English speakers—perhaps precisely *because* the meanings were so often different even when the sounds were the same. For Spanish speakers, having lots of similar-sounding words was *helpful* for speech comprehension. The investigators therefore concluded that having lots of similar-sounding words makes both speech production *and* speech comprehension easier in general, but the degree of benefit varies depending on specific attributes of individual languages.

The three researchers also discovered that language networks have another interesting quality. The networks for all the languages they studied (Basque, Mandarin, and Hawaiian, as well as English and Spanish) are *assortatively mixed:* high-degree nodes tend to be connected to other high-degree nodes. In these languages, a given word tends to be linked to other words that are themselves linked to an above-average number of words.

Most of the types of networks that scientists study—the Internet, transportation networks, networks of neurons—are *disassortatively mixed*: nodes that are *dissimilar* in terms of degree are connected to each other. For example, a wireless Internet router may have several computers connected to it, but it's less common for routers themselves to be connected to each other. An air traveler starting out at a small airport can't fly to another small airport without connecting at an airline hub. If your leg itches, you can't scratch it unless the stimulus travels first to your central nervous system and then back out to the periphery, lightning-fast though that process may be—the peripheral neurons are not directly connected to each other.

Human social networks, on the other hand, are one of the few types of well-known networks that *are* assortatively mixed. In other words, people who know a lot of people are also likely to know each other. (With disassortative mixing, a person with many friends would mostly be friendly with hermits.)

Linguistic networks' structure, then, coincides neatly with language's status bridging the biological and the social. And in addition to helping scientists understand language production and comprehension, the network approach is helping them understand the effects of events that damage the brain, such as strokes. Using a method common in network analysis—removing nodes and testing how "robust" the rest of the network is, i.e., what proportion of its nodes are still linked to each other—points toward approaches to rehabilitation. At heart, says Arbesman, network methods are helping scientists understand the evolutionary pressures that shaped the human mind.

the deeper meaning (if any) behind the predilections and associations they observed. But in the online world and the other frontiers that network scientists are exploring, making sense of new data is a crucial task; the Christakis team is diving into genetic, as well as digital, data in search of new insights about human social tendencies (see online sidebar, described on page 46).

"In some ways, the availability of these new kinds of data is like what the microscope was to Van Leeuwenhoek or the telescope to Galileo," says Christakis. "When the telescope was invented, Galileo just started looking at stuff. He looked at the moon and he saw mountains. He looked at Jupiter and found moons encircling it. He looked at the sun and found sun spots. There's this huge part of science which is just about careful observation and curiosity about the world."

Elizabeth Gudrais '01 is associate editor of this magazine.