WORDS ON TRIAL

Can linguists solve crimes that stump the police?

BY JACK HITT

In the early weeks of 2009, Chris Coleman began telling friends and associates in Columbia, Illinois, that he was worried about the safety of his family. He had been receiving death threats from an online stalker, and the e-mails had begun to mention his wife, Sheri, and his sons, Garret and Gavin, who were nine and eleven. Coleman asked his neighbor across the street, a police officer, to train a security camera on the front of his house.

Coleman understood surveillance better than most. He worked as the security chief for Joyce Meyer, whose cable television program, "Enjoying Everyday Life," is at the center of an evangelical empire estimated to be worth more than a hundred million dollars a year; it includes a radio program, self-help and children's books, CDs, podcasts, overseas missions, and motivational conventions. Initially, the threats focussed on Meyer, warning that if she didn't quit preaching she'd pay the price, but the stalker soon turned his wrath on Coleman and his family. One note to Sheri read "Fuck you! Deny your God publicly or else!" Another read "Time is running out for you and your family."

On May 5th, Coleman left his home early to work out at the gym. Afterward, when he called his wife and got no answer, he asked his neighbor the policeman to check on her. The officer found a horrifying scene. Red graffiti—"Fuck you" and "U have paid!"—was scrawled on the walls and on the sheets of the beds in which Sheri, Gavin, and Garret lay strangled to death. A back window was open, suggesting that someone had entered the house out of view of the camera.

The police quickly came to suspect Coleman himself. A long trail of text messages on his phone made it clear that he was having an affair with a cocktail waitress. And if Coleman had left Sheri he would have risked losing his salary of a hundred thousand dollars a year. Meyer maintained a strict no-divorce policy among her employees, in keeping with her understanding of Scripture. In February, Sheri had told a friend that she was afraid of her husband and "said that if anything happened to her, Chris did it," the friend testified.

Still, Coleman maintained his innocence, and the evidence against him was circumstantial. His DNA was not found anywhere that would connect him directly to the crime. At the trial last spring, nearly forty witnesses and evidentiary specialists took the stand. The modern science of forensics has spawned numerous specialties and subspecialties, including forensic branches of dentistry, anthropology, sculpting, and even entomology (interpreting the presence of certain insects to fix the time of death). At the trial, experts could show that some of the threatening e-mails had been sent from Coleman's computer, but they couldn't prove that it hadn't been hacked. They could tell the jury that Coleman had purchased a can of red spray paint months before, but they couldn't link Coleman to the can used in the crime. Toward the end of the trial, prosecutors asked for the testimony of Robert Leonard, a practitioner of one of the lesser-known subspecialties: forensic linguistics.

These days, the word "forensic" conjures an image of a technician on a "C.S.I." episode who delicately retrieves a single hair or a chip of paint from a crime scene, surmises the unlikeliest facts, and presents them to the authorities as incontrovertible evidence. If "forensic linguist" brings to mind a verbal specialist who plucks slivers of meaning from old letters and segments of audiotape before announcing that the perpetrator is, say, a middle-aged insurance salesman from Philadelphia, that's not far from the truth. In the Coleman case, Leonard, the head of the linguistics program at Hofstra University, on Long Island, was punctilious in his presentation. Relying largely on word choice and spelling, he suggested that the same person had written the threatening e-mails and sprayed the graffiti, and that those specimens bore similarities to Coleman's prose style.

In preparing his testimony, Leonard consulted with his frequent partner, James Fitzgerald, a retired F.B.I. forensic linguist. Fitzgerald brought the field to prominence in 1996 with his work in the case of the Unabomber, who had sent a series of letter bombs to professors over several years. Fitzgerald had successfully urged the F.B.I. to publish the Unabomber's "manifesto"—a rambling thirty-five-thousand-word declaration of the perpetrator's philosophy. Many people called the Bureau to say they recognized the writing style. By analyzing syntax, word choice, and other linguistic patterns, Fitzgerald narrowed down the range of possible authors and finally linked the manifesto to the writings of Ted Kaczynski, a reclusive former mathematician. For instance, the bomber's use of the terms "broad" and "negro," for women and African-Americans, enabled Fitzgerald roughly to calculate his age. Both Kaczynski and the Unabomber also showed a preference for dozens of unusual words and expressions, such as "chimerical," "anomic," and "cool-headed logicians," as well as the less familiar version of the cliché "You can't eat your cake and have it too." A judge ruled that the linguistic evidence

was strong enough to prompt him to issue a search warrant for Kaczynski's cabin in Montana; what was found there put him in prison for life.

Fitzgerald went on to formalize some of the tools used in forensic linguistics, and started the Communicated Threat Assessment Database. The CTAD is the most comprehensive collection of linguistic patterns in written threats, containing more than a million words and some four thousand "criminally oriented communications." At the Coleman trial, Leonard noted that many of the killer's spray-painted sentences began with the word "fuck," as did the e-mails and letters—usage that at first might seem common. But he explained that a CTAD search showed that only 0.8 per cent of threatening notes used "fuck" as the first word in a sentence. In addition, both the graffiti and the threatening notes relied on two obscenities, "fuck" and "bitch," to the exclusion of all others. Leonard went on to compare the graffiti with two hundred and twenty-one e-mails known to have been written by Coleman. He noted that the shorthand "U" for "you" is often found in cell-phone text messages but rarely in e-mails. Both the killer and Coleman used "U" in e-mails. Coleman also consistently put the apostrophe of a contraction in the wrong place—"doesnt'" and "cant'"—as did the killer. Leonard's testimony was disputed in the courtroom, but, in a case with no physical evidence firmly linking Coleman to the crime, Leonard's words—and Coleman's—took on added weight. After a day of deliberations, the jury found Coleman guilty, which made him eligible for the death penalty; the judge sentenced Coleman to three life terms in prison.

Most people assume that meaning is embedded in the words they speak. But, according to forensic linguists, meaning is far more vaporous, teased into existence through vocalized puffs of air, hand gestures, body tilts, dancing eyebrows, and nuanced nostril flares. The transmission of meaning still involves pri-

mate mechanics worked out during the Pliocene era. And context is crucial; when we try to record a conversation, we are capturing only part of the gestalt of that moment. What might appear to be a solid audio recording can easily morph into an acoustic Rorschach test. The plot of "The Conversation," Francis Ford Coppola's

A suspect's conversations and writings can be analyzed for patterns and peculiarities.

classic film, turns on this very point: the protagonist spends the entire movie mishearing a tiny clip of audiotape. Such errors and misunderstandings pervade our lives, in ways that modern language detectives are only beginning to make clear.

The pioneer of forensic linguistics is widely considered to be Roger Shuy, a retired Georgetown University professor and the author of such fundamental textbooks as "Language Crimes: The Use and Abuse of Language Evidence

in the Courtroom." Shuy is now eighty-one years old and lives in Montana. When I asked him to describe the origins of forensic linguistics, he referred me to an Old Testament story. After a confusing battle with the Ephraimites, the Gileadites were able to identify the enemy by asking them each to pronounce the Hebrew word "shibboleth." If they pronounced the first syllable in the Ephraimic dialect, "sib," instead of in the Gilead dialect, "shib," they were killed. According to Judges 12:6, some forty-two thousand Ephraimites failed that first linguistic test.

The field's more recent origins might be traced to an airplane flight in 1979, when Shuy found himself sitting next to a lawyer. By the end of the flight, Shuy had a recommendation as an expert wit-

ness in his first murder case. Since then, he's been involved in numerous cases in which forensic analysis revealed how meaning had been distorted by the process of writing or recording. In a bribery trial in the nineteen-eighties, two Nevada brothel commissioners were caught on tape in a crucial exchange. When they were offered a bribe, one turned to the other and, according to the police transcript, said, "I would take a bribe, wouldn't you?" Shuy analyzed the tape and, on the stand, testified that the defendant had actually said the opposite: "I wouldn't take a bribe, would you?" The tape was scratchy. Moreover, in conversational speech, the "n't" of a contraction is barely vocalized. It was hard to hear—or, rather, easy to hear what the listener was primed to hear. But two facts were indisputable, Shuy noted: both versions of the sentence had exactly eight syllables, and the pause fell just before the last two syllables. Thus, Shuy testified, only one reading of the sentence made sense: "I wouldn't take a bribe, would you?" The trial resulted in a hung jury.

Shuy has become famous in his discipline for some of the field's finest Holmesian aperçus. Early in his career, the police in Illinois approached him regarding a notorious kidnapping case; they had several suspects, and they hoped his reading of the ransom notes might help narrow down the list of suspects. In each note, the kidnapper demanded money in a semiliterate rant: "No kops! Come alone!!," followed by a terse instruction—"Put it in the green trash kan on the devil strip at the corner 18th and Carlson." Shuy studied the letters and then asked, "Is one of your suspects an educated man born in Akron, Ohio?" The cops were stunned. There was one who matched that description perfectly, and when confronted he confessed. As Shuy subsequently explained, "kop" and "kan" most likely were intentional misspellings by someone posing as illiterate. And he knew from his research that the patch of grass between the sidewalk and the street—sometimes

known as the "tree belt," "tree lawn," or "sidewalk buffer"—is called the "devil's strip" only in Akron, Ohio.

In recent years, following Shuy's lead, a growing number of linguists have applied their techniques in criminal cases, such as Chris Coleman's, and even in major commercial lawsuits. An upcoming suit between Apple and Microsoft, slated to go before the Trademark Trial and Appeal Board, features two stars of the field, Rob Leonard and Ronald Butters, a retired Duke University linguist. At issue is: What part of speech is the phrase "app store"? Leonard, siding with Apple, contends that it is a proper noun, which is to say a trademarked expression that should be capitalized. Butters's work upholds Microsoft's view: the term consists of two common nouns and is not proprietary at all.

Butters is a past president of the International Association of Forensic Linguists, which has some two hundred and fifty members. Most of them, he said, are in the United States, England, and Spain, but interest has spread to Australia, Japan, and China. Today, one can study forensic linguistics at several schools, and last year Leonard inaugurated the first graduate program in forensic linguistics, at Hofstra. For those earning a master's degree, the field offers job prospects outside the courtroom. Immigration and Customs Enforcement hires language detectives to assist agents in evaluating asylum seekers. In such cases, forensic linguists interview applicants to verify that their accents and their use of idiom and slang match those of the country they claim to have fled.

Increasingly in the courtroom, however, forensic linguists have been asked to weigh in on matters of "author identification"—not to determine the grammatical significance of certain words but to identify who said or wrote them. This trend has widened an old schism in the field. Given the stakes in, say, the Coleman case—a felony murder potentially involving the death sentence—some linguists hold the view that Leonard is tak-

ing forensic linguists into groundbreaking territory. Others, including Butters, wonder if he isn't leading them over a cliff.

When I visited Leonard one afternoon at Hofstra, he was reviewing a range of cases: another murder involving the killer's letters; a libel suit that turned on a single, ambiguous sound; an attempt to identify a potential assassin of a prominent politician; and a Whirlpool Corporation lawsuit involving the meaning of the word "steam." In a modest office walled with books, I found Leonard working at a laptop. He was noticeably kempt, in pressed slacks and a crisp blue button-down shirt—a Sam Spade of semantics. His hair was surprisingly dark for a man in his sixties; his eyes were playful and his smile fetching, a little bit show biz.

Long before he emerged as one of the foremost language detectives in the country, Leonard had achieved a different kind of celebrity. As an undergraduate at Columbia in the nineteen-sixties, he and his brother George revolutionized the school's a-cappella group by having everyone dress as faux Brooklyn thugs (white T-shirts, greased-back hair) and sing up-tempo arrangements of such nineteen-fifties doowop classics as "Duke of Earl" and "At the Hop." They named the group Sha Na Na and became wildly popular. One of their hits was "Teen Angel," which Leonard sang at Woodstock just before Jimi Hendrix, who had invited Sha Na Na, débuted his version of "The Star-Spangled Banner."

By 1970, Leonard the heartthrob had to choose between academia and show business. "All of our good friends were dying of drug overdoses," he said. "I just decided to move on." Leonard finished his undergraduate studies at Columbia; William Labov, a prominent linguist who had introduced him to the field, helped him earn a fellowship. Leonard pursued a scholarly career until 2000, when he heard Shuy give a lecture urging linguists to apply their training in the real world—especially in the courtroom, as language detectives. Leonard struck up a professional friendship with Shuy and has been con-

sulting on cases ever since.

As we sat in his office, Leonard described his recent involvement in the tabloid saga of Natalee Holloway. In 2005, after graduating from high school in Alabama, Holloway went with her friends on a chaperoned trip to Aruba and disappeared. The case remains unsolved. The chief suspect is a young Dutchman named Joran van der Sloot, who pleaded guilty in 2012 to charges of murdering a twenty-one-year-old woman in Peru. In Aruba, two young brothers, Deepak and Satish Kalpoe, were initially arrested (they and van der Sloot had partied with Holloway the night before she disappeared), but were released in the first weeks of the investigation. After being the subjects of a television exposé, the brothers are suing Dr. Phil McGraw and CBS for defamation. The Kalpoe legal team has hired Leonard as their expert witness in a lawsuit that could turn on the pronunciation of a single syllable.

The "Dr. Phil" show promoted the exposé by claiming, "You are going to find out what he"—Deepak—"says he did with Natalee the night she disappeared." An announcer adds, "What he said brought Natalee's mother to tears." On the show, viewers listen to the audio of Deepak being secretly videotaped by a private investigator named Jamie Skeeters and making an astonishing confession:

SKEETERS: I'm sure she had sex with all of you.
KALPOE: She did. You'd be surprised how simple it was.

Leonard examined the uncut version of the exchange. In it, Kalpoe denies having sex with Holloway. "Simple" refers to the fact that, from his point of view, the evening was uneventful:

SKEETERS: I'm sure she had sex with all of you, and . . . good . . .
KALPOE: No, she didn't.
SKEETERS: O.K., well, I mean, good. If she did, fine.

KALPOE: You'd be surprised how simple it was that night.

Watching an unedited piece of footage doesn't require a linguistics expert, but Leonard realized that there were other issues at play. During the covert interview, the microphone generated a great deal of confusing ambient sound. Moreover, the hidden camera captured only the top of Kalpoe's head, so his face and lips weren't visible. Amid the muffled noises, and before Kalpoe speaks, there is an odd sound—*sha!*—which Kalpoe appears to make just before "No." When I met with Leonard, he had been concentrating on this sound. An expert hired by the opposing counsel was taking the position that the *sha!* might not be a throat-clearing or some other stray sound, as Leonard contends, but a "voiceless vowel with 'r' coloration."

Leonard explained: "Vowels are the most open of sounds, and when you come off a vowel and cease saying it you switch the vocal apparatus to pronounce the next sound." In some words, like "forth," the "r" gets full phonetic treatment, but in many words, like "bird" and "sure," the "r" isn't fully voiced and instead becomes a shadow of the vowel, just because it's easier to say that way. If the lawyers for "Dr. Phil" can show that the first word Kalpoe spoke in that sentence was "sure," and that there is no audible "n't" at the end of "did," then the transcript of Kalpoe's first utterance changes from "No, she didn't" to its opposite, "Sure, no, she did."

Like all linguists, Leonard starts from the position that meaning is delicately contingent, and that the most common way we compensate for this frailty is "redundancy." We say the same thing more than once, or in more than one way. In his written report to the court on this case, Leonard notes that the original video of the meeting between Deepak Kalpoe and Skeeters shows Deepak "shaking his head 'no' from side

to side," as if to deny the accusations. The program, though, aired only a still photo of Deepak. The case has yet to go to trial, but when it does, Leonard says, he will argue that there is enough redundancy in the semiotic detritus of these sounds to conclude that Kalpoe's meaning is clear: he is stating that he did not have sex with Holloway that night.

It may be that the changes made to the edited interview were deliberately damaging, but forensic linguists offer another possibility: that a subtle presumption of guilt unconsciously overwhelmed the editing process and inverted the meaning of the exchange. Such inversions, linguists say, happen far more often than we might like to believe.

According to Leonard, words serve as catalysts, setting off sparks of potential meaning that the listener organizes into more specific meaning by observing facial expressions, body language, and other redundant cues. We then employ another powerful tool: prior experience and the storehouse of narratives that each of us carries—what linguists call "schema." To every exchange we bring unconscious scripts; as any given sentence unspools, we readjust the schema to make better sense of what we are hearing.

One afternoon at Hofstra, Leonard explained to the twenty students in his introductory course how this works. He wrote a sentence on the board: "John was on his way to school last Friday and was really worried about the math lesson." He quizzed the students on what they might presume about this story. John is a student, one called out; he is either on a bus or walking.

"So we can just close our eyes and imagine John the schoolboy on the bus," Leonard said. "But are we all imagining John with the same height, the same hair color?" Nothing in the sentence signals any of that information, yet each of us supplies our own variant, which awaits further verbal data for confirmation.

Leonard wrote another sentence beneath the first: "Last week, he had been

unable to control the class." Who is John now? "A teacher!" someone shouted. And how is John getting to school? "A car!"

Leonard wrote a third sentence: "It was not fair for the math teacher to leave him in charge." Instantly, the students revelled in John's new identity as a janitor or a substitute teacher.

Meaning, Leonard noted, is constantly bent by expectation, and can be grossly distorted. Indeed, one of Shuy's first studies, of the Abscam trials of the nineteen-eighties, reveals just how easily the meaning of linguistic evidence can be twisted by a background assumption of guilt. Abscam was an F.B.I. sting operation in which nine United States congressmen were lured to meetings with a government agent posing as an Arab oil sheikh with "Abdul Enterprises." The initial meeting was described as a legitimate business deal. At one point, though, the agent playing the sheikh would offer the congressmen an outright bribe. Their conversations were videotaped, and some of the evidence was breathtakingly unambiguous. Representative John Jenrette, of South Carolina, accepted the money cheerfully and chirped on tape, "I've got larceny in my blood!"

The sting resulted in seven indictments. Toward the end came the trial of Senator Harrison (Pete) Williams, of New Jersey. Shuy listened to those tapes and became convinced that the Senator was innocent. Whenever the sheikh raised the issue of bribery or illegality, Williams steered the conversation to legal ground. At one point, the sheikh put the bribe directly to Williams: "I would like to give you . . . some money for, for permanent residence." The first four words of Williams's reply were "No, no, no, no."

A prosecution memo at the time stated that there was no case against Williams, but the judge, who, in his ruling, decried "the cynicism and hypocrisy of corrupt public officials," set it aside; Williams was found guilty and sentenced to three years in prison. Shuy later noted that, with such attitudes prevalent, the schema of the "corrupt congressman" overwhelmed even the plainest facts pointing to Williams's innocence. After the trial, the lead juror confessed that had he known all the facts he would not have found Williams guilty. The Senator was forced to resign his seat, though he declared his innocence at every opportunity. He was the first senator in eighty years to go to prison; President Bill Clinton refused to pardon him.

Shortly after the Unabomber case was cracked, in 1996, forensic linguistics gained another public boost. Donald Foster, an English professor at Vassar, employed the most basic forensic technique—tallying word frequency—to unmask the anonymous author of "Primary Colors," the best-selling novel about Clinton's first Presidential campaign: Joe Klein. Foster analyzed dozens of pages of writing from several suspects, including *Time's* Walter Shapiro and the former Deputy Treasury Secretary Roger Altman. He compiled a concordance that showed how frequently each writer used certain words, compared this information with a database of word frequency in the novel, and was able to identify the author.

For a short time, the potential of forensic linguistics seemed limitless. With enough raw data and computing power, a trail of words might betray its author as reliably as a set of fingerprints identifies an individual. Foster, though, was a professor of literature, not a linguist; he was not trained to use the forensic methods that Shuy had mastered, such as listening for unconscious semantic patterns and looking for distinctive phrases or unusual colloquialisms. Overconfident, Foster went on to identify a suspect in the JonBenét Ramsey murder case, only to learn that he had already been cleared by the police. In the days after September 11, 2001, Foster falsely implicated the bioweapons expert Steven Hatfill as the person who had sent several anthrax-laden letters around the country; the accusation wrecked Hatfill's career and resulted in a settled lawsuit. Foster then recanted a previous claim linking a 1612 poem of dubious provenance to Shakespeare; another academic had shown that the analysis was fatally flawed. Foster has since retreated to his campus in Poughkeepsie.

Foster's disgrace left most forensic linguists feeling cautious. Now that Leonard's work is bringing the field back into the realm of author identification, some are worried. Ronald Butters, the Duke linguist, provided expert testimony for the defense at the Coleman trial and challenged every aspect of Leonard's testimony as "linguistically meaningless." Butters argued that even though certain linguistic oddities, such as using "U" for "you" or consistently misplacing the apostrophe in contractions, seemed distinctive, there weren't enough examples to be statistically significant. Moreover, Butters told me, it can be tricky to compare different genres of even a single person's writing. Reading, say, a routine office e-mail alongside rants spray-painted on a wall makes about as much sense as comparing the prose in one of Wallace Stevens's insurance riders with the cadences of his poem "Sunday Morning."

"Really bad linguistic testimony is when you go to court and say you're pretty sure that this person wrote that, and yet you're comparing apples and oranges," Butters said. Leonard argues that he never claims to name a specific author but simply presents comparative evidence for the jury. Butters, Leonard said, "is a specialist in trademark cases, so I'm not sure what his experience was in authorship cases, and they are two quite different applications of linguistics." On the stand, Butters admitted that he hadn't read all Leonard's research on the evidence; his challenge was focussed on Leonard's methodology and its purported usefulness in the identification of individual authors. Butters said, "Forensic linguistics has not come to a place where we are mature enough to answer a lot of these questions."
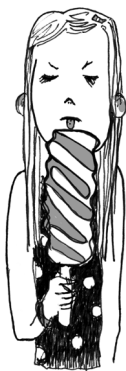
Carole Chaski, the executive director of the Institute for Linguistic Evidence and the president of Alias Technology, in Georgetown, Delaware, which markets linguistic software, agrees. Chaski has been working to perfect a computer algorithm that identifies patterns hidden in syntax. With enough linguistic material to work with, she says, she can run the program and draw accurate linguistic conclusions. Her goal is to develop a standard "validated tool" that police, civil investigators, and linguists can turn to when testifying in crucial cases, such as a capital murder trial. "If this is real, these tools should be so reliable that I can automate them and somebody can use them," she says. Chaski foresees a time when forensic-linguistic "technicians" will do what DNA technicians in crime labs do: "They learn how to run a piece of software or run a Southern blot"—a standard DNA test—"through electrophoresis and then go, 'Here are my results.'"

In Chaski's view, a trail of words can be parsed to reveal its author, but that work is best done quantitatively, through brute computational force, not qualitatively, by subjective scholars. Forensic linguistics, she believes, should not be limited to a few highly credentialled experts who have been approved by the courts to testify. She warned me of the recklessness of an "academic" and an "ex-cop" hanging out a shingle, and said their methodology was "fraught with error." In the small world of forensic linguistics, it was obvious that she meant Leonard and Fitzgerald.

Leonard said that Chaski's computerized approach made him "want to take a nap." His methods and findings are all transparent, he noted, whereas her algorithm is a proprietary "black box." He does not believe that computer software can eliminate the need for human interpretation. "Even those algorithms have to be coded by humans," he said; any good linguist will depend on both quantitative and qualitative analysis. "One thing we have learned about language is that it is a very human form of communication. You have

to have human intelligence, human powers of inference, and human encyclopedic knowledge of the world" to make sense of it. At the end of the day, the scientific findings depend on human interpretation, Leonard said. Computers can crunch reams of words, but only people can decide what the words mean.

Shuy told me that he, too, initially had doubts about author identification. "That is how I felt until Rob Leonard started working," he said. "Rob has come up with this competing-hypothesis approach." In the same way that DNA technicians will report only the statistical likelihood that the killer's DNA and the DNA found on the murder weapon are the same, Leonard creates a number of opposing hypotheses and presents the evidence in light of them. In the Coleman trial, Leonard did not declare that Coleman was the author of the red graffiti and the threatening e-mails; rather, he testified that the language in them "is consistent with" the language in Coleman's writings.

"I don't know any forensic linguists who will claim that they can find the answer for you," Shuy said. "Our role is to analyze the data and give it to the triers of the facts, who have to evaluate it or issue the ultimate decision of innocence or guilt. We don't go that far, and shouldn't."

Shuy also noted that it was Leonard who popularized a safeguard against comparing unrelated documents, called a Community of Practice filter. For instance, Coleman's use of "U" for "you" would be of no use in a pool of text messages, but as an unusual abbreviation in an e-mail it becomes another point of data.

Recently, Leonard used this technique to question a charge, levelled at a jailed gang member, of murdering a prison guard. Prosecutors had linked the prisoner, Jarvis Masters, to a note that ultimately led to the guard's murder, based on misspellings such as "has'nt" and "is'nt" and the use of "no" for "know." But in his research Leonard learned that the way Masters's gang, the Black Guerrilla

Family, disciplined its members was to make them copy propaganda by hand. All the gang members had picked up the oddities pinned on Masters, Leonard determined. "Thus, when we examine the corpus of non-murder documents written by other B.G.F. members," Leonard said, "we discover the features that may at first seem to educated writers like the prosecution to be randomly incorrect, highly idiosyncratic features were not random at all but systemic features of the B.G.F. community."

On some level, extracting meaning from linguistic evidence is what we all do intuitively every day. Forensic professionals go about the same work, with better tools and a heightened sense of how easily meaning can be misconstrued. As one forensic-linguistics firm, Testipro, puts it in its online promotional pitch, the field is "the basis of the entire legal system. Both Judges and Juries are using informal or unconscious FL"—forensic linguistics—"every time they weigh a witness statement or testimony document." The field is bound to thrive on the ever-growing piles of what Shuy calls "data." Our embrace of personal media—e-mails, text messages, voice mail, tweets—has created an avalanche of tossed-off language, an evidentiary trail that linguists are getting better and better at following.

Shuy believes that forensic linguistics can do for language crimes, such as bribery, blackmail, and extortion, what DNA has done for violent crimes. It could offer a counterweight to the many old-school methods, like lineups and unrecorded police interrogations, that are heavily relied upon despite serious flaws. "I won't claim that we have anything remotely like DNA in this work," Shuy said, "but we are a whole lot better than a lot of the crazy schemes that cops are being taught."

Leonard offered a sobering statistic: eighty per cent of people who were later exonerated by DNA evidence had falsely confessed to their alleged crimes. "When I got into this business, I figured if there

was an eyewitness or a confession, then case closed, the guy absolutely, one hundred per cent did it. But those are the two shakiest types of evidence, really." He recalled many cases where a confession on paper turned out to be no confession at all. "The way humans perceive language is according to schemas, which lead to misperceptions as much as perceptions." In a sense, investigators who try to extract evidence from confessions are acting as linguists, too, albeit poorly trained ones.

A few weeks ago, Leonard finished testifying in the retrial of Brian Hummert, a Pennsylvania man charged with strangling his wife. After initial suspicions pointed to Hummert, the police received handwritten letters claiming that a serial killer, not Hummert, had committed the murder. Once again, the linguistic evidence was important to the case. The notes bore a resemblance to a series of stalker letters that preceded the killing and to the defendant's writing. As an expert witness, Leonard testified about Hummert's prose style, noting the rare use of what he calls "ironic repetition" in constructions such as "She tried to break it off, so I broke her neck." And all the letters contained a linguistic habit that, Leonard testified, he had found nowhere else: a tendency to use contractions in negative statements ("I can't") but not in positive ones ("I am"). The jury was out for forty-five minutes and returned a verdict of guilty. ♦

THE NEW YORKER SPECIAL REPRINT