Language analysis predicts a coming betrayal

August 20 2015, by Bill Steele

Credit: George Hodan/public domain

Being betrayed is worse than just being attacked. Someone you trusted as a friend and ally suddenly stabs you in the back.

According to a team of researchers at Cornell, the University of Maryland and the University of Colorado, there are subtle linguistic clues that predict when a betrayal is coming. Humans are poor at noticing them, but computer analysis can detect them.

"Language does tell us a lot about human relations," said Cristian Danescu-Niculescu-Mizil, Cornell assistant professor of information science. "Not with certainty, but signaling intentions. Language is a social signal we are trying to decode."

Analyzing messages in an online game has identified signals of coming betrayal, Cornell graduate student Vlad Niculae reported at the 53rd Annual Meeting of the Association for Computational Linguistics, held in conjunction with the seventh International Joint Conference on Natural Language Processing, July 26-31 in Beijing. Collaborators on the paper along with Danescu-Niculescu-Mizil are Jordan Boyd-Graber, assistant professor of computer science at the University of Colorado, and Srijan Kumar, a graduate student in computer science at the University of Maryland.

The result is not a recipe for avoiding betrayal in romance and friendship, the researchers said. If it worked, they pointed out, betrayals could not happen. But it might offer ways to help groups work more effectively together, they said, for example in software development teams or online crowdsourcing projects like Wikipedia.

The game they studied was the online version of Diplomacy (also played as a board game), in which players assume the roles of European nations before World War I and try to conquer the entire map. In wars launched in the game, two or more countries will defeat a lone opponent, so players must form alliances to advance. Each round of play begins with a "diplomacy" discussion period for such negotiations. But since the ultimate goal is to rule alone, players inevitably will turn against their allies, an act gamers call "backstabbing" or just "stabbing." Parallels with the TV show "Survivor" come to mind.

The researchers analyzed a dataset of 145,000 messages between pairs of players in 249 online Diplomacy games. They matched pairs of players

whose relationship ended in betrayal with similar pairs where it did not, and compared the language of the messages exchanged. Drawing sometimes on algorithms developed by previous researchers they looked for positive or negative sentiment, politeness and for keywords suggesting such concepts as planning ("next," "then"), claims ("I believe," "I think"), time ("still," "while") or comparison ("as much as," "after").

Among the most obvious findings were that betrayers become more positive and less polite just before a betrayal, and that the language patterns of betrayer and victim become less balanced at that time. Among others, victims used more planning words than their betrayers, further proof that victims were unaware of the coming treachery.

To carry their work forward, the researchers have created a new online game called StreetCrowd, where teams must work together to solve a problem. They hope to study the play to learn how to make group efforts more successful.

Provided by Cornell University