

## Casting Call: Forensic Linguist for Facebook Lawsuit

By [Alan Perlman](#)

It's not often that forensic linguistics makes the news. It's not nearly as sexy or yucky as the forensics that originates at the murder site or in the pathologist's lab. There's actually a scientific book, called *Men, Murder, and Maggots*, that tells you how to determine when someone was killed, on the basis of the type of parasites that are now feasting on the corpse.

Compared to this, nailing someone by their writing habits is not great entertainment. But it does happen, as reported recently in [a \*New York Times\* article](#) regarding the authorship of e-mails alleged to have been written by Mark Zuckerberg.

The legal battle between Paul Ceglia and Mark Zuckerberg has taken numerous twists and turns, starting in July 2010 with a claim by Ceglia that he is entitled to more than half of Facebook. This claim was quickly followed by claims of a forged contract and, a year after the suit began, Ceglia came forward with emails to further support his claim.

And thus the forensic linguist was introduced to the drama. The consulting linguist, Gerald McMenamin, emeritus professor of linguistics at California State University, Fresno, studied the e-mails, found 11 different style markers, across punctuation spelling and grammar, and opined that indeed Zuckerberg could not have been the author.

However, this case turns on an issue still debated among forensic linguists: the reliability of author identification by qualitative methods. The highly credentialed Ronald Butters questions whether the (non-)capitalization of *Internet* has any value in the identification of an anonymous or suspected author, given that there are only three examples in the Facebook emails. But another linguist, Carole Chaski, notes, correctly, that there are hundreds of possible authorship features in punctuation, grammar, word choice – in every aspect of language.

Those forensic linguists (like myself) who do believe that an individual's "idiolect" can be identified and described, begin with the premise that every author's individual writing practices, all those choices made on the fly, do have a pattern. At the very simplest, a writer chooses to do one thing, but not the other, many times each sentence: to use one word but not another to express a particular meaning, to write dates and numbers in a certain way...and so on. You can see why the number of features can be very large and why even as few as 11 significant differences could point to the answer to questions of authorship.

So the answer to the question if every person has a "linguistic fingerprint" is yes and no (yes, if the quantity and quality of the data are sufficient). People are largely unaware that they have a writing style - and of how many decisions are made on the fly. Indeed, enough of these decisions (even allowing for intra-writer variants), can define an individual's writing style; the profiler just needs enough information.

Sometimes a single feature can carry considerable weight. In one case, the writer repeatedly used *puke* as noun, verb and adjective; it appeared in everything he wrote. Another writer used a grammatical construction which occurs in British and Canadian (but not American) English.

There are broadly defined choices of which the writer is unaware. One indicator that I like to use is the writer's preferred method for linking clauses, whether by starting a new sentence, or by coordination (*and, but, etc.*) subordination (*that/which* or adverb plus clause), or some combination of the three. This decision on how to end a rather large amount of information and begin another is, again, made on the fly, and writers tend to stick to the same pattern.

The *Times* article closes with the hope that technology will save us from the supposed inaccuracies of the experienced analyst, but I contend that for qualitative analysis, there is no technology nearly equal to the sublime powers of the human brain.

**Tell us:** Do you think you have a distinct writing style that would be noticed by a forensic linguist?

*This article is a rewrite of a publication on Alan Perlman's website. It was later posted on [BullsEye](#), a legal blog published by [IMS ExpertServices](#).*