

CONTACT: Steve Bradt at 617-496-8070

EMBARGOED FOR RELEASE: Wednesday, Oct. 10, 2007, at 1 p.m. ET

HARVARD SCIENTISTS PREDICT THE FUTURE OF THE PAST TENSE Mathematicians apply evolutionary models to linguistic standardization

CAMBRIDGE, Mass. -- Verbs evolve and homogenize at a rate inversely proportional to their prevalence in the English language, according to a formula developed by Harvard University mathematicians who've invoked evolutionary principles to study our language over the past 1,200 years, from "Beowulf" to "Canterbury Tales" to "Harry Potter."

Writing this week in the journal *Nature*, Erez Lieberman, Jean-Baptiste Michel, and colleagues in Harvard's Program for Evolutionary Dynamics, led by Martin A. Nowak, conceive of linguistic development as an essentially evolutionary scheme: Just as genes and organisms undergo natural selection, words -- specifically, irregular verbs that do not take an "-ed" ending in the past tense -- are subject to powerful pressure to "regularize" as the language develops.

"Mathematical analysis of this linguistic evolution reveals that irregular verb conjugations behave in an extremely regular way -- one that can yield predictions and insights into the future stages of a verb's evolutionary trajectory," says Lieberman, a graduate student in applied mathematics in Harvard's School of Engineering and Applied Sciences and in the Harvard-MIT Division of Health Sciences and Technology, and an affiliate of Harvard's Program for Evolutionary Dynamics. "We measured something no one really thought could be measured, and got a striking and beautiful result."

"We're really on the front lines of developing the mathematical tools to study evolutionary dynamics," says Michel, a graduate student in systems biology at Harvard Medical School and an affiliate of the Program for Evolutionary Dynamics. "Before, language was considered too messy and difficult a system for mathematical study, but now we're able to successfully quantify an aspect of how language changes and develops."

Lieberman, Michel, and colleagues built upon previous study of seven competing rules for verb conjugation in Old English, six of which have gradually faded from use over time. They found that the one surviving rule, which adds an "-ed" suffix to simple past and past participle forms, contributes to the evolutionary decay of irregular English verbs according to a specific mathematical function: It regularizes them at a rate that is inversely proportional to the square root of their usage frequency.

In other words, a verb used 100 times less frequently will evolve 10 times as fast.

To develop this formula, the researchers tracked the status of 177 irregular verbs in Old English through linguistic changes in Middle English and then modern English. Of these 177 verbs that were irregular 1,200 years ago, 145 stayed irregular in Middle English and just 98 remain irregular today, following the regularization over the centuries of such verbs as help, laugh, reach, walk, and work.

Lieberman and Michel's group computed the "half-lives" of the surviving irregular verbs to predict how long they will take to regularize. The most common ones, such as "be" and "think," have such long half-lives (38,800 years and 14,400 years, respectively) that they will effectively never become regular. Irregular verbs with lower frequencies of use -- such as "shrive" and "smite," with half-lives of 300 and 700 years, respectively -- are much more likely to succumb to regularization.

Lieberman, Michel, and their co-authors project that the next word to regularize will likely be "wed."

"Now may be your last chance to be a 'newly wed'," they quip in the Nature paper. "The married couples of the future can only hope for 'wedded' bliss."

Extant irregular verbs represent the vestiges of long-abandoned rules of conjugation; new verbs entering English, such as "google," are universally regular. Although fewer than 3 percent of modern English verbs are irregular, this number includes the 10 most common verbs: be, have, do, go, say, can, will, see, take, and get. Lieberman, Michel, and colleagues expect that some 15 of the 98 modern irregular verbs they studied -- although likely none of these top 10 -- will regularize in the next 500 years.

The group's Nature paper makes a quantitative, astonishingly precise description of something linguists have suspected for a long time: The most frequently used irregular verbs are repeated so often that they are unlikely to ever go extinct.

"Irregular verbs are fossils that reveal how linguistic rules, and perhaps social rules, are born and die," Michel says.

"If you apply the right mathematical structure to your data, you find that the math also organizes your thinking about the entire process," says Lieberman, whose unorthodox projects as a graduate student have ranged from genomics to bioastronautics. "The data hasn't changed, but suddenly you're able to make powerful predictions about the future."

Lieberman and Michel's co-authors on the Nature paper are Nowak, professor of mathematics and of biology at Harvard and director of the Program for Evolutionary Dynamics, and Harvard undergraduates Joe Jackson and Tina Tang. Their work was sponsored by the John Templeton Foundation, the National Science Foundation, and the National Institutes of Health.

###