

# Science makes an open book of English evolution

July 25 2012

---

"The United States of America" has become entrenched as one of the most frequently printed phrases in the modern era of written English, a study of 500 years of language evolution has shown.

Among the top dozen phrases most-printed in books every year, this one stands out from the other most popular five-word sentence components like "at the end of the", "as a result of the" or "on the part of the".

"'The United States of America' tops the charts quite remarkably if one ignores the more common and by themselves inherently meaningless phrases," Slovenian physicist Matjaz Perc who conducted the research, told AFP.

Perc had made a digital analysis of some 5.2 million books dating from 1520 to 2008, and showed the language going through an erratic period heavily influenced by religion in the 16th and 17th centuries -- a time when [William Shakespeare](#) is also claimed to have coined many [words](#) and phrases.

The printing press was invented in about 1440, spreading rapidly throughout Europe and then beyond.

"During the 16th and 17th centuries, the [popularity](#) (of words) was very fleeting," Perc found. His study was published Wednesday in the Journal of the Royal Society Interface.

"Top words in the year 1600, for example, are no longer top words in the year 1610."

From the 18th and 19th centuries, word rankings became increasingly stable.

"The words that are most common during the year 1950, for example, are also the most common even today," said Perc, who claimed his [computer analysis](#) covered about four percent of all books published up to 2008.

Driven by [curiosity](#) about a [phenomenon](#) known as "preferential attachment", the physicist's analysis led him to compile extensive tables of English words and phrases.

They show the word "the" right at the top of the list throughout the centuries, followed by others like "and", "of", "to", "in" and "a" at varying rankings near the top.

In the 1500s and 1600s, however, "baptized", "hymns", "God", "Christ" and "[pope](#)" also featured prominently among the most-used words, as well as phrases like "baptized in the name of", or "God forbid it should be".

Oddly out of place in the number one spot in 1586: the phrase "A fine old English gentleman".

Phrases like "House of Commons" and words like "Queen" and "Duke" started climbing the list by the mid- to late 17th century.

Top phrases in the 1700s included "the Church of England", "the Law of Nature" and "the Orb of the Sun".

By the 1800s the pattern started looking more as it does today, with formulaic phrases like "at the same time" or "in the midst of" featuring most prominently.

In 1919, the year after World War I ended, the ninth most-published five-word phrase was "for extraordinary heroism in action".

Since 1968, Perc's tables show "the United States of America" consistently among the top 15 most-published phrases, up from the top 20-odd in the 1940s and 50s.

"It seems that the words and phrases we use for writing books have matured, which in turn invites the conclusion that the English language itself is matured over the years," said Perc.

"Today we know what to expect when opening up a book, much more so than we would have if opening a book in the 16th century.

"This of course does not pertain to the content of each book, which will hopefully always surprise us, but rather to the grammatical constructions and certain [phrases](#) that have endured the test of time and are today commonplace."

(c) 2012 AFP

Citation: Science makes an open book of English evolution (2012, July 25) retrieved 3 May 2024 from <https://phys.org/news/2012-07-science-english-evolution.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--