

Women are beautiful, men rational

August 27 2019

Female		Male	
Positive	Negative	Positive	Negative
beautiful	battered	just	unsuitable
lovely	untreated	sound	unreliable
chaste	barren	righteous	lawless
gorgeous	shrewish	rational	inseparable
fertile	sheltered	peaceable	brutish
beauteous	heartbroken	prodigious	idle
sexy	unmarried	brave	unarmed
classy	undernourished	paramount	wounded
exquisite	underweight	reliable	bigoted
vivacious	uncomplaining	sinless	unjust
vibrant	nagging	honorable	brutal

Credit: University of Copenhagen

Men are typically described by words that refer to behavior, while adjectives ascribed to women tend to be associated with physical appearance. This, according to a group of computer scientists from the University of Copenhagen and other universities that deployed machine learning to analyze 3.5 million books.

"Beautiful' and 'sexy' are two of the adjectives most frequently used to describe women. Commonly used descriptors for men include 'righteous,' 'rational' and 'brave.'

A computer scientist from the University of Copenhagen, along with fellow researchers from the United States, trawled through an enormous quantity of [books](#) in an effort to find out whether there is a difference between the types of words used to describe men and women in literature. Using a new computer model, the researchers analyzed a dataset of 3.5 million books, all published in English between 1900 to 2008. The books include a mix of fiction and non-fiction literature.

"We are clearly able to see that the words used for women refer much more to their appearances than the words used to describe men. Thus, we have been able to confirm a widespread perception, only now at a statistical level," says computer scientist and Assistant Professor Isabelle Augenstein of the University of Copenhagen's Department of Computer Science.

The researchers extracted adjectives and verbs associated with gender-specific nouns (e.g. 'daughter' and 'stewardess'). For example, in combinations such as 'sexy stewardess' or 'girls gossiping.' They then analysed whether the words had a positive, negative or neutral sentiment, and subsequently which categories the words could be divided into.

Their analyses demonstrate that negative verbs associated with body and appearance are used with five times the frequency for females than males. The analyses also demonstrate that positive and neutral adjectives relating to the body and appearance occur approximately twice as often in descriptions of females, while males are most frequently described using adjectives that refer to their behaviour and personal qualities.

In the past, linguists typically looked at the prevalence of gendered

language and bias, but using smaller data sets. Now, [computer scientists](#) are able to deploy machine learning algorithms to analyze vast troves of data—in this case, 11 billion words.

New life for old gender stereotypes

Although many of the books were published several decades ago, they still play an active role, points out Isabelle Augenstein. The algorithms used to create machines and applications that can understand [human language](#) are fed with data in the form of text material that is available online. This is the technology that allows smartphones to recognize our voices and enables Google to provide keyword suggestions.

"The algorithms work to identify patterns, and whenever one is observed, it is perceived that something is 'true.' If any of these patterns refer to biased language, the result will also be biased. The systems adopt, so to speak, the language that we people use, and thus, our gender stereotypes and prejudices," says Isabelle Augenstein, and gives an example of where it may be important:

"If the language we use to describe men and women differs, in employee recommendations for example, it will influence who is offered a job when companies use IT systems to sort through job applications."

As [artificial intelligence](#) and language technology become more prominent across society, it is important to be aware of gendered [language](#).

Augenstein continues: "We can try to take this into account when developing machine-learning models by either using less biased text or by forcing models to ignore or counteract bias. All three things are possible."

The researchers point out that the analysis has its limitations, in that it does not take into account who wrote the individual passages and the differences in the degrees of bias depending on whether the books were published during an earlier or later period within the data set timeline. Furthermore, it does not distinguish between genres—e.g. between romance novels and non-fiction. The researchers are currently following up on several of these items.

Provided by University of Copenhagen

Citation: Women are beautiful, men rational (2019, August 27) retrieved 4 April 2024 from <https://techxplore.com/news/2019-08-women-beautiful-men-rational.html>

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.