

# Quantifying Jargon

Shannon D. Willoughby

## INTRO:

We really need to effectively communicate scientific findings to the general public. If they don't understand what we are saying, why should they believe us?

## METHODS:

1. Choose a set of texts to create a scientific corpus, or collection of words.
2. Choose a set of texts to create a contemporary American English corpus.
3. Create both corpora, with stop words, punctuation, numbers, and hyphens removed.
4. Prepare the text for which the jargon is to be calculated, remove punctuation, numbers, and hyphens.
5. Determine how many times each word in the cleaned text occurs in each corpus.
6. Calculate the jargonness value for each word.

$$J = \begin{cases} \log\left(\frac{f_{sci}}{f_{eng}}\right) & 0 < f_{sci} < f_{eng} \\ \frac{f_{eng}}{3} & f_{eng} = 0 \text{ and } f_{sci} > 0 \end{cases}$$

7. Calculate the average jargonness per word for the text as a whole.



This work is supported by NSF grant 1735124.

# Too much jargon makes your communication less effective .

But how much is too much?

This R script will calculate how much jargon you are using so you can change words that are hard to understand.



## RESULTS CALCULATED WITH R SCRIPT

### JARGON (j = 3)

microdroplet	noninvasive
microscale	biofilm
cretaceous	resiliancy
microscale	lipid

### MAYBE JARGON (j > 1)

density	nuclear
oxide	nutrients
boundary	flow

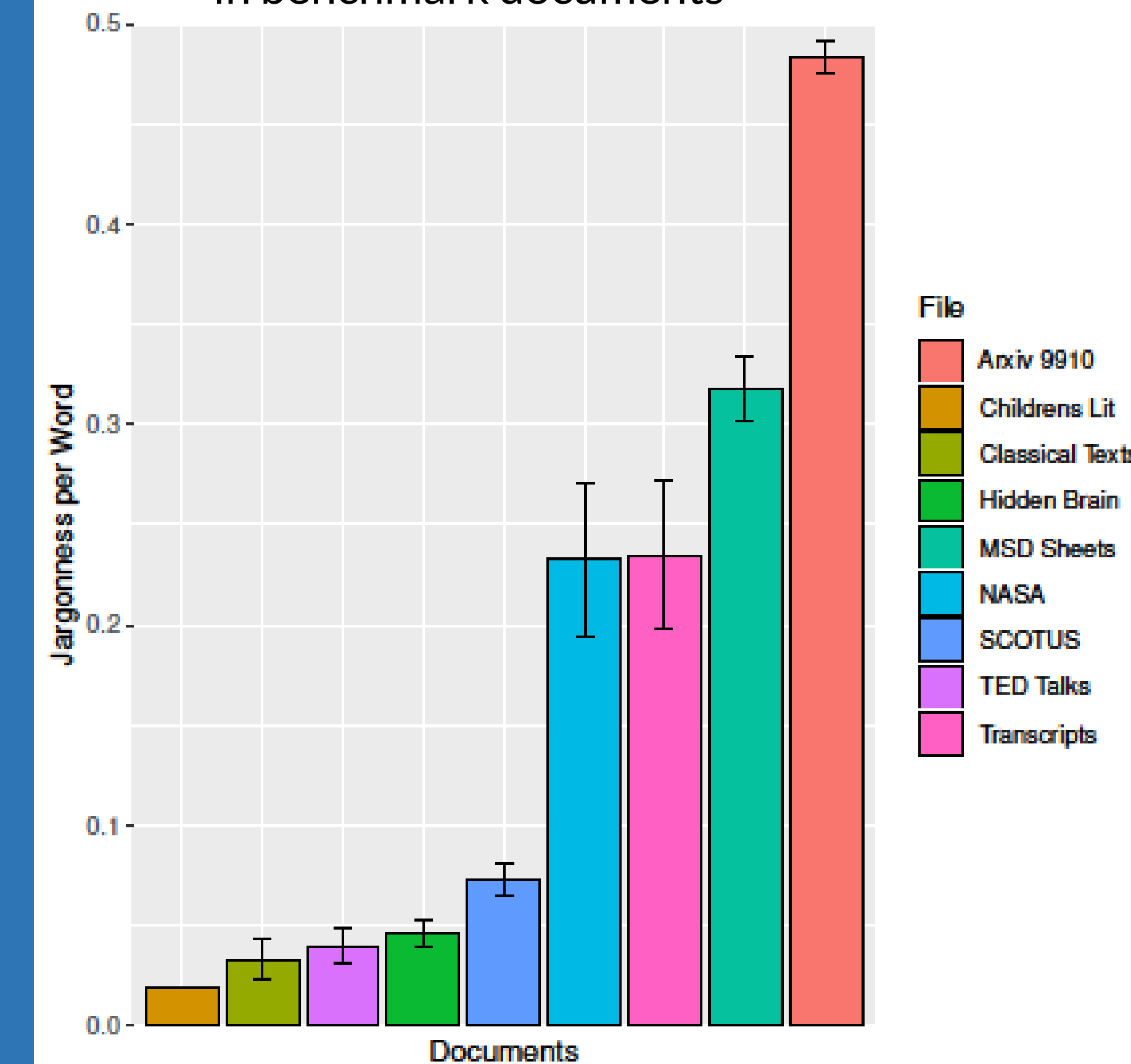
### NOT JARGON (j = 0)

interest	traveler
tiny	friend
time	stomach
worldwide	food
combat	question

### STOP WORDS

these	despite
its	exactly
she	think
five	they'll
going	relatively

Average jargonness per word in benchmark documents



Jenny Green, Bryce Hughes, Chris Organ, Brock LaMeres, Leila Sterman