

The relative power of alternative phrasings

“Parallel universes” in movies, Twitter, & ChangeMyView



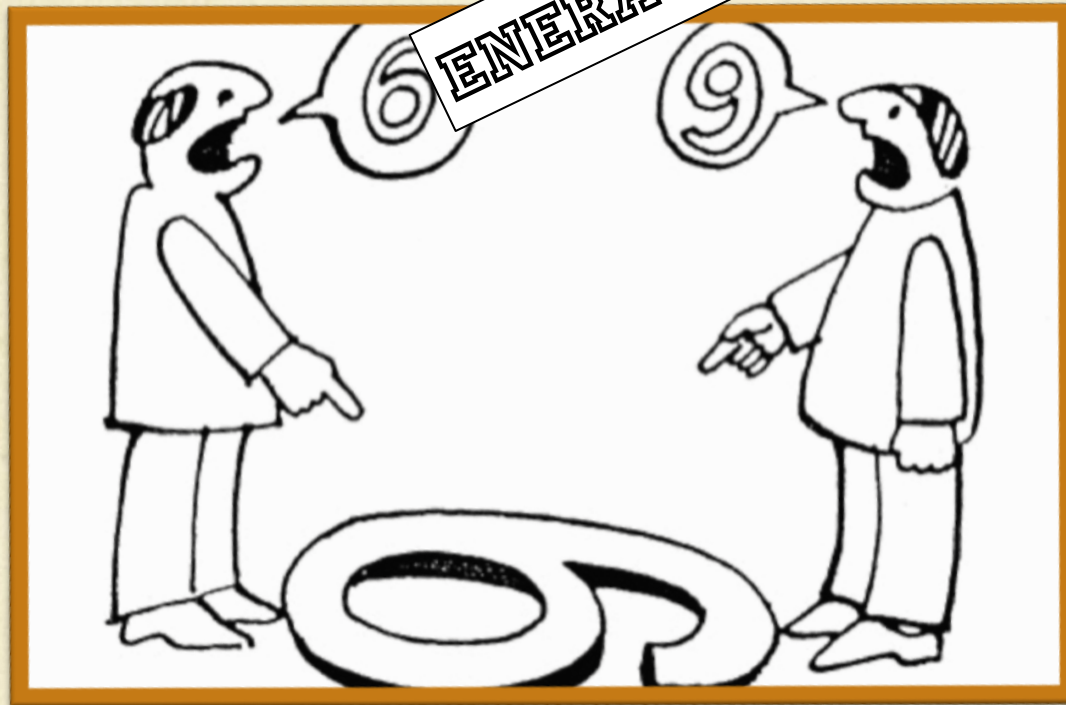
Lillian Lee

Dept. of Computer Science, Cornell University

<http://www.cs.cornell.edu/home/llee>

The relative power of alternative phrasings

“Parallel universes” in movies, Twitter & ChangeMyView



Lillian Lee

Dept. of Computer Science, Cornell University

<http://www.cs.cornell.edu/home/llee>

The relative power of alternative phrasings

“Parallel universes” in movies, Twitter & ChangeMyView



Lillian Lee

Dept. of Computer Science, Cornell University

<http://www.cs.cornell.edu/home/llee>

The relative power of alternative phrasings

“Parallel universes” in movies, Twitter, & ChangeMyView



Lillian Lee

Dept. of Computer Science, Cornell University

<http://www.cs.cornell.edu/home/llee>

To phrasing from framing

The framing of an argument emphasizes certain principles or perspectives.

“One of the most important concepts in the study of public opinion” (Druckman, 2001)

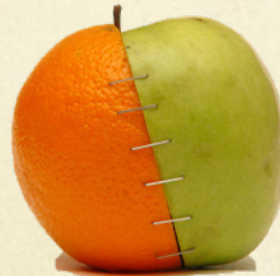
“CL” framing work includes: Choi, Tan, Lee, Danescu-Niculescu-Mizil, Spindel 2012; Nguyen, Boyd-Graber, Resnik 2013; Baumer, Elovic, Qin, Polletta, Gay, 2015; Tsur, Calacci, Lazer 2015; Card, Gross, Boydston, Smith, 2016; Field, Kliger, Wintner, Pan, Jurafsky, Tsvetkov 2018; August, Oliveira, Tan, Smith 2018

To phrasing from framing

The framing of an argument emphasizes certain principles or perspectives.

“One of the most important concepts in the study of public opinion” (Druckman, 2001)

Framing in GMO debates:



“CL” framing work includes: Choi, Tan, Lee, Danescu-Niculescu-Mizil, Spindel 2012; Nguyen, Boyd-Graber, Resnik 2013; Baumer, Elovic, Qin, Polletta, Gay, 2015; Tsur, Calacci, Lazer 2015; Card, Gross, Boydston, Smith, 2016; Field, Klinger, Wintner, Pan, Jurafsky, Tsvetkov 2018; August, Oliveira, Tan, Smith 2018

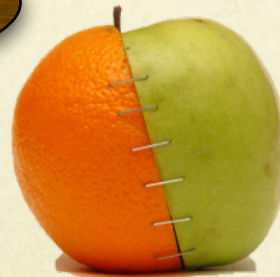
To phrasing from framing

The framing of an argument emphasizes certain principles or perspectives.

“One of the most important concepts in the study of public opinion” (Druckman, 2001)

Framing in GMO debates:

"green revolution"



“CL” framing work includes: Choi, Tan, Lee, Danescu-Niculescu-Mizil, Spindel 2012; Nguyen, Boyd-Graber, Resnik 2013; Baumer, Elovic, Qin, Polletta, Gay, 2015; Tsur, Calacci, Lazer 2015; Card, Gross, Boydston, Smith, 2016; Field, Klinger, Wintner, Pan, Jurafsky, Tsvetkov 2018; August, Oliveira, Tan, Smith 2018

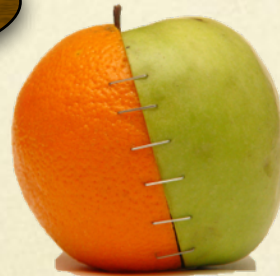
To phrasing from framing

The framing of an argument emphasizes certain principles or perspectives.

“One of the most important concepts in the study of public opinion” (Druckman, 2001)

Framing in GMO debates:

"green revolution"



"Frankenfood"

“CL” framing work includes: Choi, Tan, Lee, Danescu-Niculescu-Mizil, Spindel 2012; Nguyen, Boyd-Graber, Resnik 2013; Baumer, Elovic, Qin, Polletta, Gay, 2015; Tsur, Calacci, Lazer 2015; Card, Gross, Boydston, Smith, 2016; Field, Kliger, Wintner, Pan, Jurafsky, Tsvetkov 2018; August, Oliveira, Tan, Smith 2018

Past research: the limits of phrasing?

Past research: the limits of phrasing?

Political scientists: Justin Grimmer, Solomon Messing, Sean Westwood
book, 2014: message frequency matters more than \$s mentioned in the
messages.

Past research: the limits of phrasing?

Political scientists: Justin Grimmer, Solomon Messing, Sean Westwood book, 2014: message frequency matters more than \$s mentioned in the messages.

Sociologists: Matt Salganik, Peter Dodds, Duncan Watts, “MusicLab“, Science 2006: “when individual decisions are subject to social influence ... there are inherent limits on the predictability of outcomes” based on item quality.

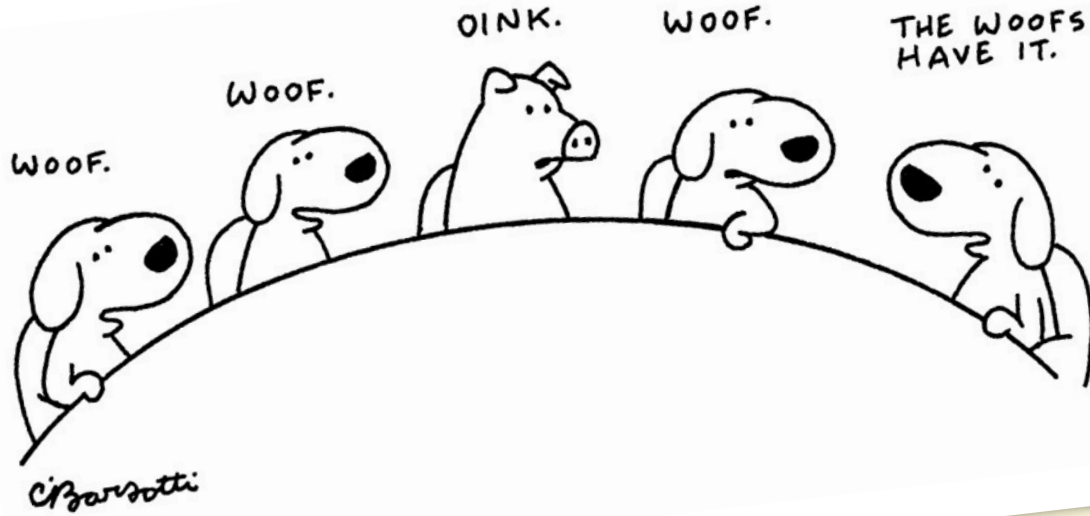
Past research: the limits of phrasing?

Political scientists: Justin Grimmer, Solomon Messing, Sean Westwood book, 2014: message frequency matters more than \$s mentioned in the messages.

Sociologists: Matt Salganik, Peter Dodds, Duncan Watts, “MusicLab“, Science 2006: “when individual decisions are subject to social influence ... there are inherent limits on the predictability of outcomes” based on item quality.

Computer scientists: Miles Osborne, 2014: “...a famous person can write anything and it will be retweeted. An unknown person can write the same tweet and it will be ignored.”

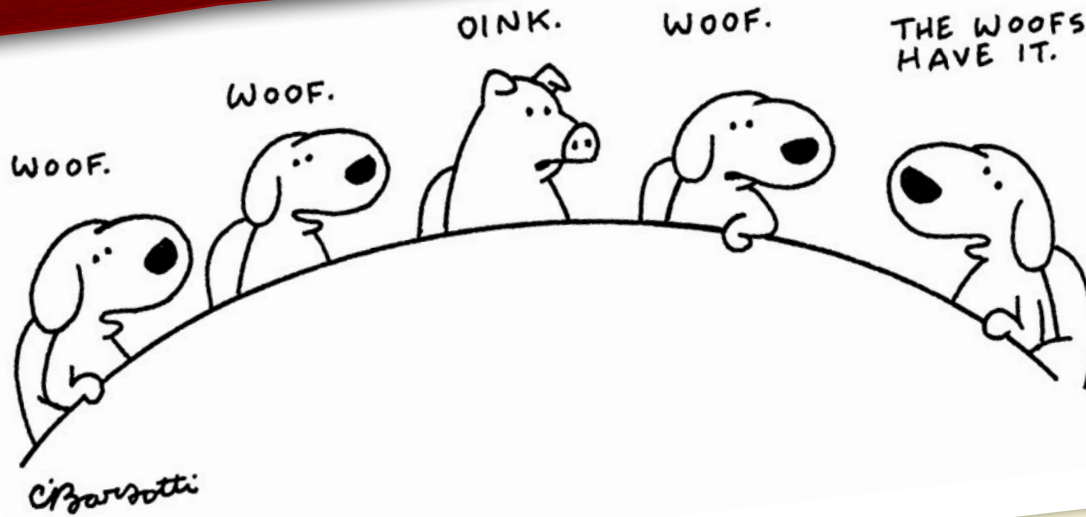
Still, can wording have any influence?



Still, can wording have any influence?

Purely hypothetical faculty meeting

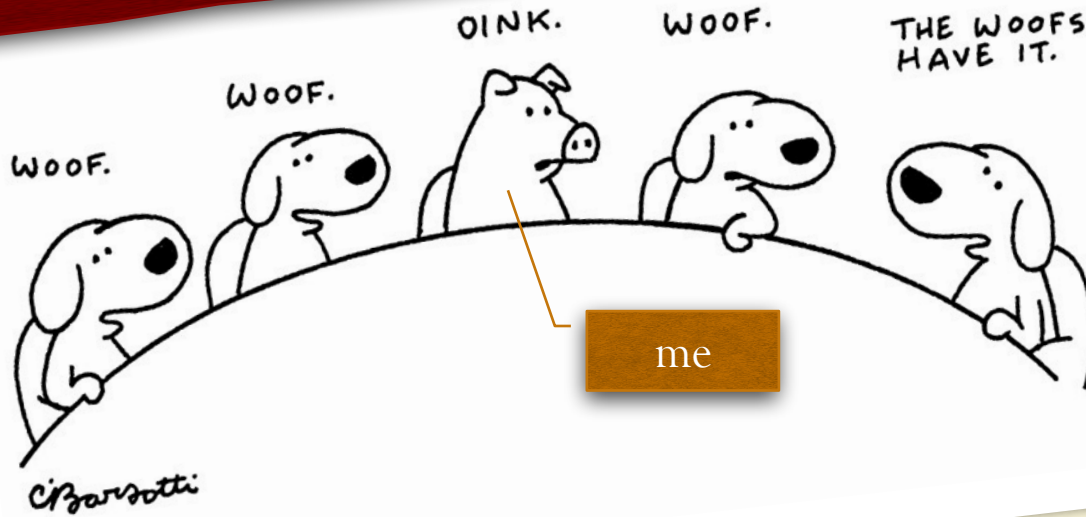
inspired from the Cartoon Bank



Still, can wording have any influence?

Purely hypothetical faculty meeting

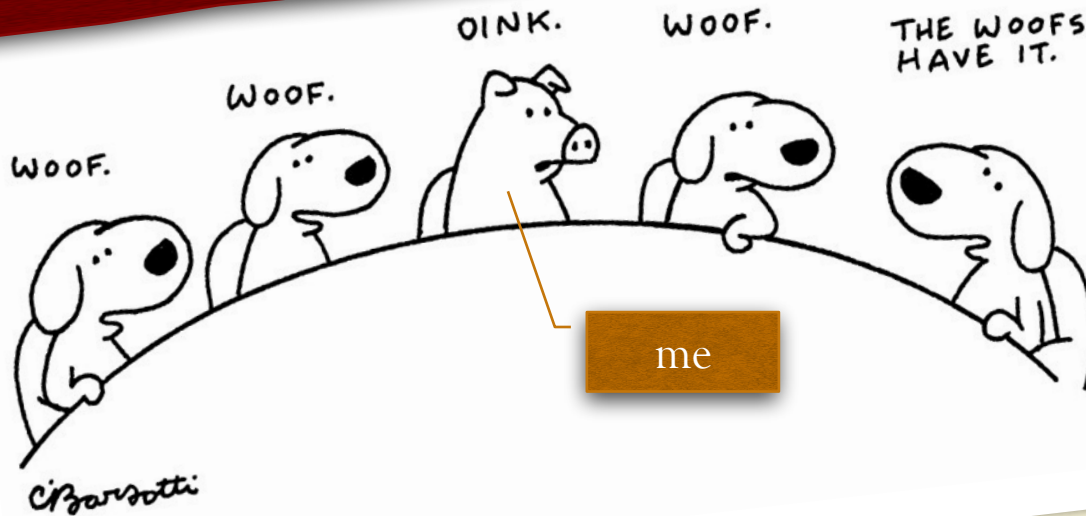
sourced from the Cartoon Bank



Still, can wording have any influence?

Purely hypothetical faculty meeting

sourced from the Cartoon Bank

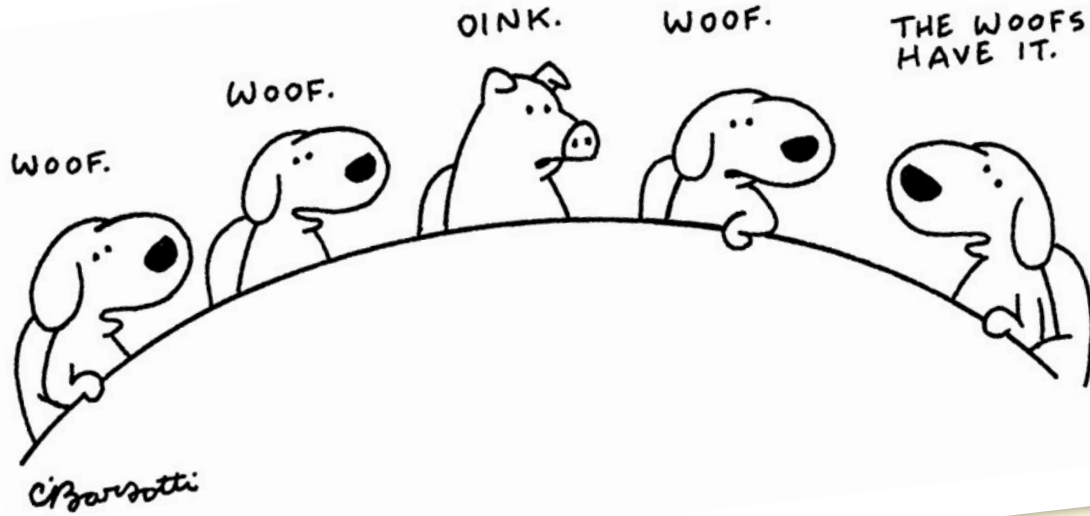


Still, can wording have any influence?

Non-options:

Become alpha dog.

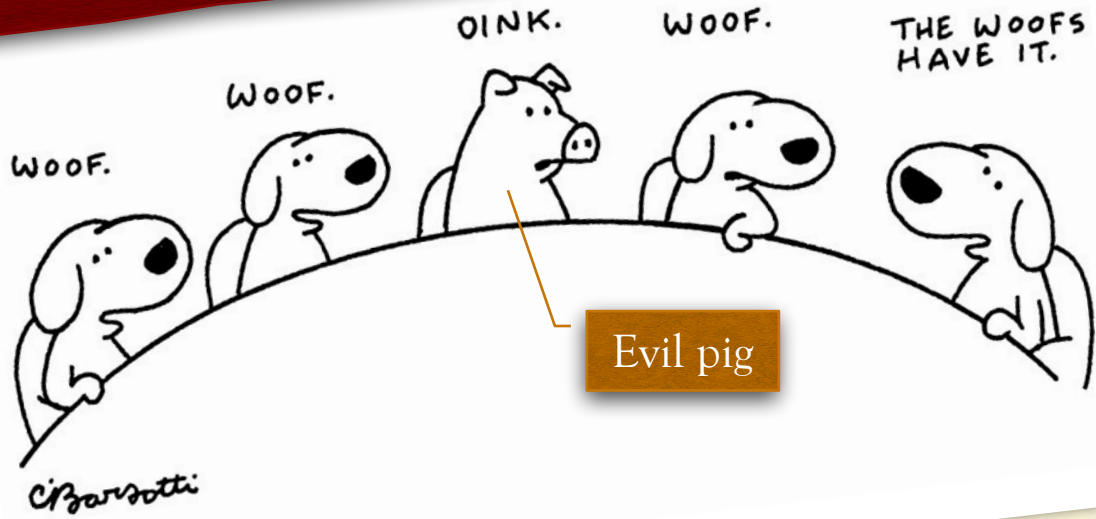
Be a dog at all.



Still, can wording have any influence?

Hypothetical disruption attempt

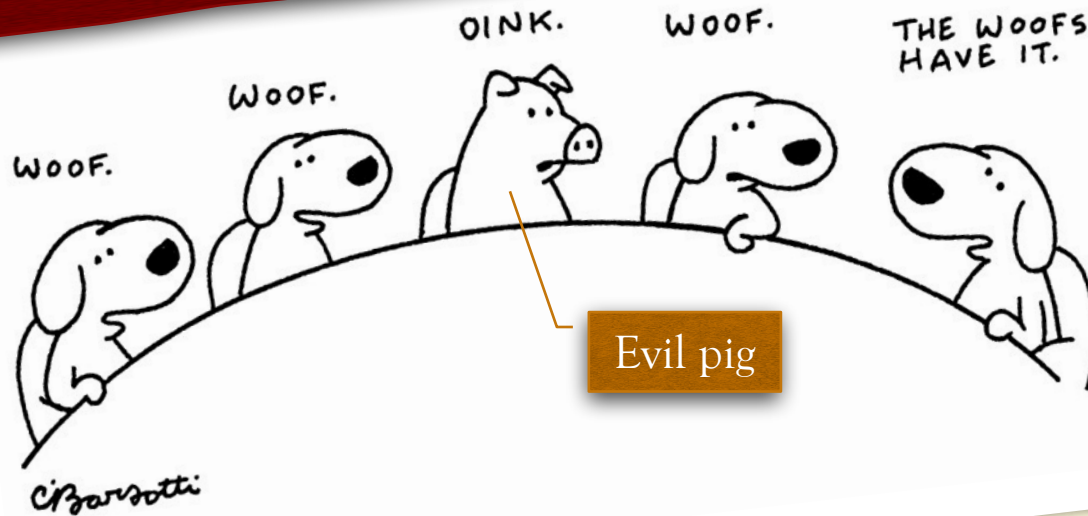
sourced from the Cartoon Bank



Still, can wording have any influence?

Hypothetical disruption attempt

inspired from the Cartoon Bank



Still, can wording have any influence?

(Non)influence can be “good” or “bad”.

Research findings may apply to either case.

“Parallel universe” experimental paradigm

Exploit situations with many instances of:

...the same speaker

...in the same situation, or

conveying the same info...

...**varying their wording**

(beyond a fixed set of lexical choices)

and see the effects.



Outline

Memorability & popular uptake: movie quotes

Cristian Danescu-Niculescu-Mizil, Justin Cheng, Jon Kleinberg, Lillian Lee, ACL 2012



Information sharing in Twitter

Chenhao Tan, Lillian Lee, Bo Pang, ACL 2014

Persuasion in ChangeMyView

Chenhao Tan, Vlad Niculae, Cristian Danescu-Niculescu-Mizil, Lillian Lee, WWW 2016



Aside on presentation style

Paraphrasing Stuart Shieber:

Aside on presentation style

Paraphrasing Stuart Shieber:

Your goal is not to show that you are brilliant, but that your solution is trivial.

Aside on presentation style

Paraphrasing Stuart Shieber:

Your goal is not to show that you are brilliant, but that your solution is trivial.

It takes a certain strength of character to take that as one's goal.

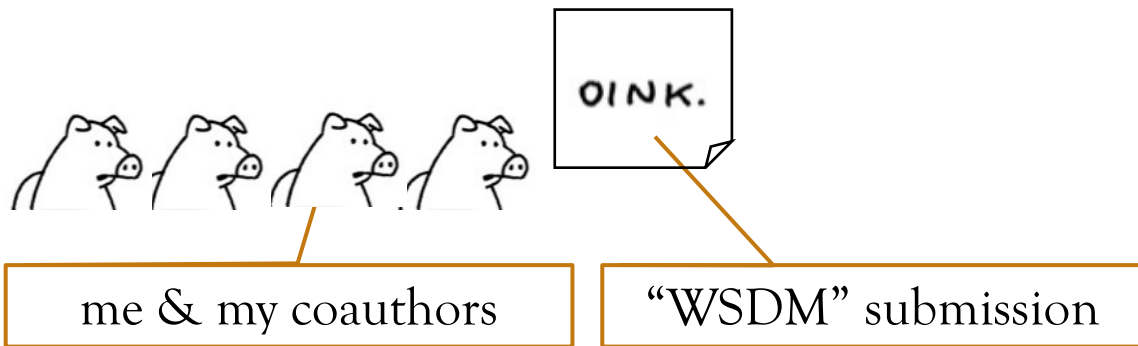
Aside on presentation style

Paraphrasing Stuart Shieber:

Your goal is not to show that you are brilliant, but that your solution is trivial.

It takes a certain strength of character to take that as one's goal.

Purely hypothetical reviewing situation



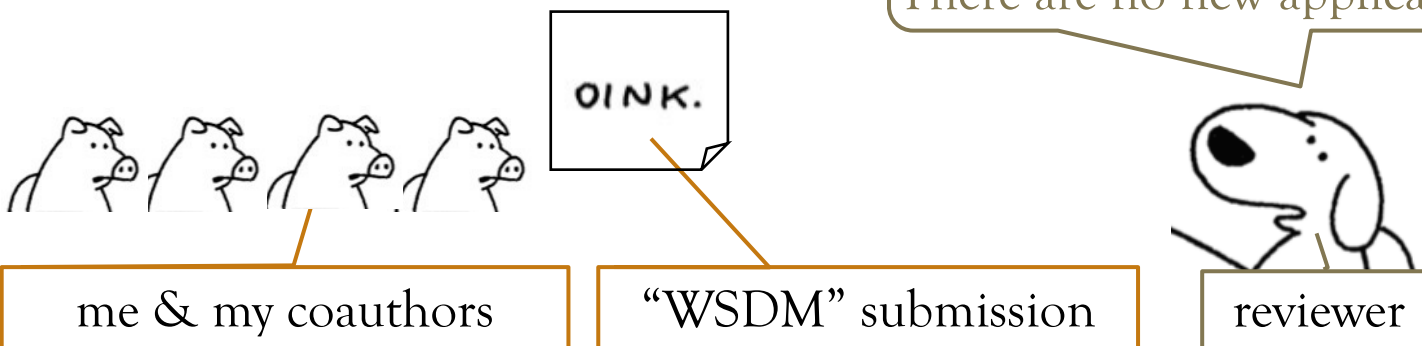
Aside on presentation style

Paraphrasing Stuart Shieber:

Your goal is not to show that you are brilliant, but that your solution is trivial.

It takes a certain strength of character to take that as one's goal.

Purely hypothetical reviewing situation



Aside on presentation style

Paraphrasing Stuart Shieber:

Your goal is not to show that you are brilliant, but that your solution is trivial.

It takes a certain strength of character to take that as one's goal.

But if people think your findings are obvious, they must also think that you are correct.



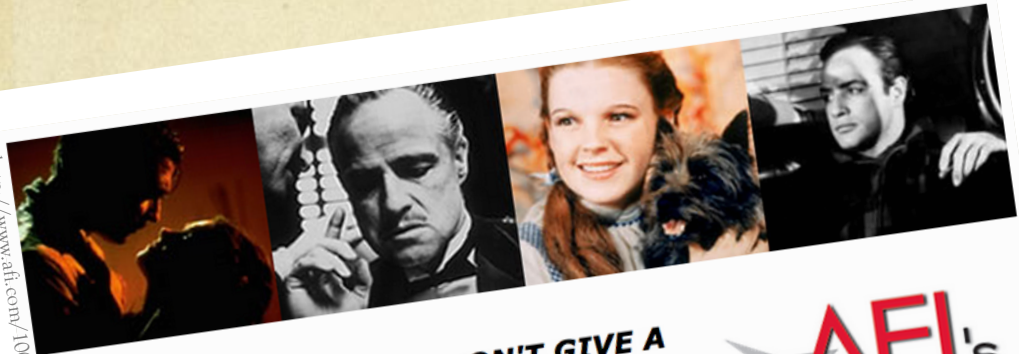
Memorability & popular uptake: Movie quotes

Much related work in many fields; see paper for refs.

Our direct inspiration: Leskovec, Backstrom, Kleinberg 2009

Meme modification: Simmons, Adamic, Adar 2011

http://www.afi.com/100years/quotes.aspx



"FRANKLY, MY DEAR, I DON'T GIVE A DAMN" TOPS AFI'S LIST OF 100 GREATEST MOVIE QUOTES OF ALL TIME

OTHER WINNERS INCLUDE:

THE GODFATHER, **"I'M GOING TO MAKE HIM AN OFFER HE CAN'T REFUSE"**

THE WIZARD OF OZ, **"TOTO, I'VE GOT A FEELING WE'RE NOT IN KANSAS ANYMORE"**

AND CASABLANCA, **"HERE'S LOOKING AT YOU, KID"**

AFI's
100
YEARS
100
"Movie
Quotes"

AFI'S 100 Years...100 Movie Quotes: America's Greatest Quips, Comebacks and Catchphrases

LOS ANGELES, June 22, 2005 — The American Film Institute revealed the top movie quotes of all time in **AFI's 100 Years...100 Movie Quotes**, a three-hour special television event on CBS hosted by actor and action star Pierce Brosnan with commentary from many of Hollywood's most celebrated actors and filmmakers. A jury of 1,500 film artists, critics and historians selected "Frankly, my dear, I don't give a damn," spoken by Clark Gable in the celebrated Civil War epic, GONE WITH THE WIND as the most memorable movie quote of all time.

Movie quotes:
massively,
permanently viral.

Other applications:
advertising,
campaigning, etc.

A parallel universe (far, far away)



Obi-Wan: You don't need to see his identification.

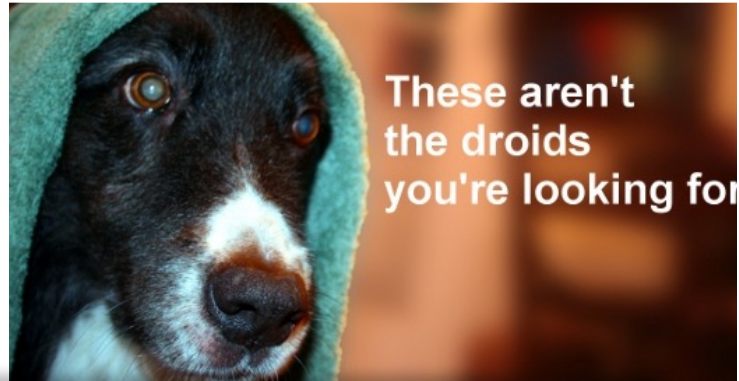
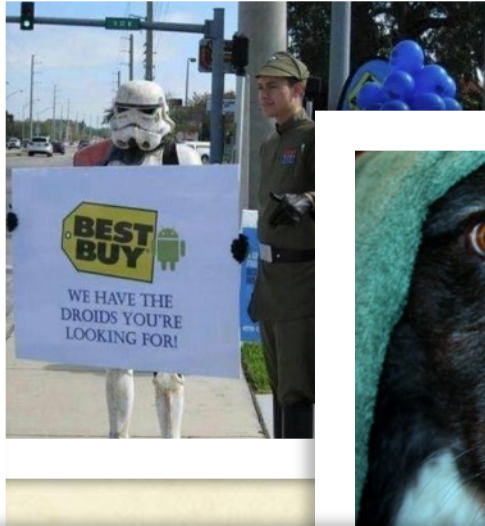
Stormtrooper: [ditto]

Obi-Wan: These aren't the droids you're looking for.

Stormtrooper: [ditto]

Obi-Wan: He can go about his business.

Stormtrooper: [ditto]

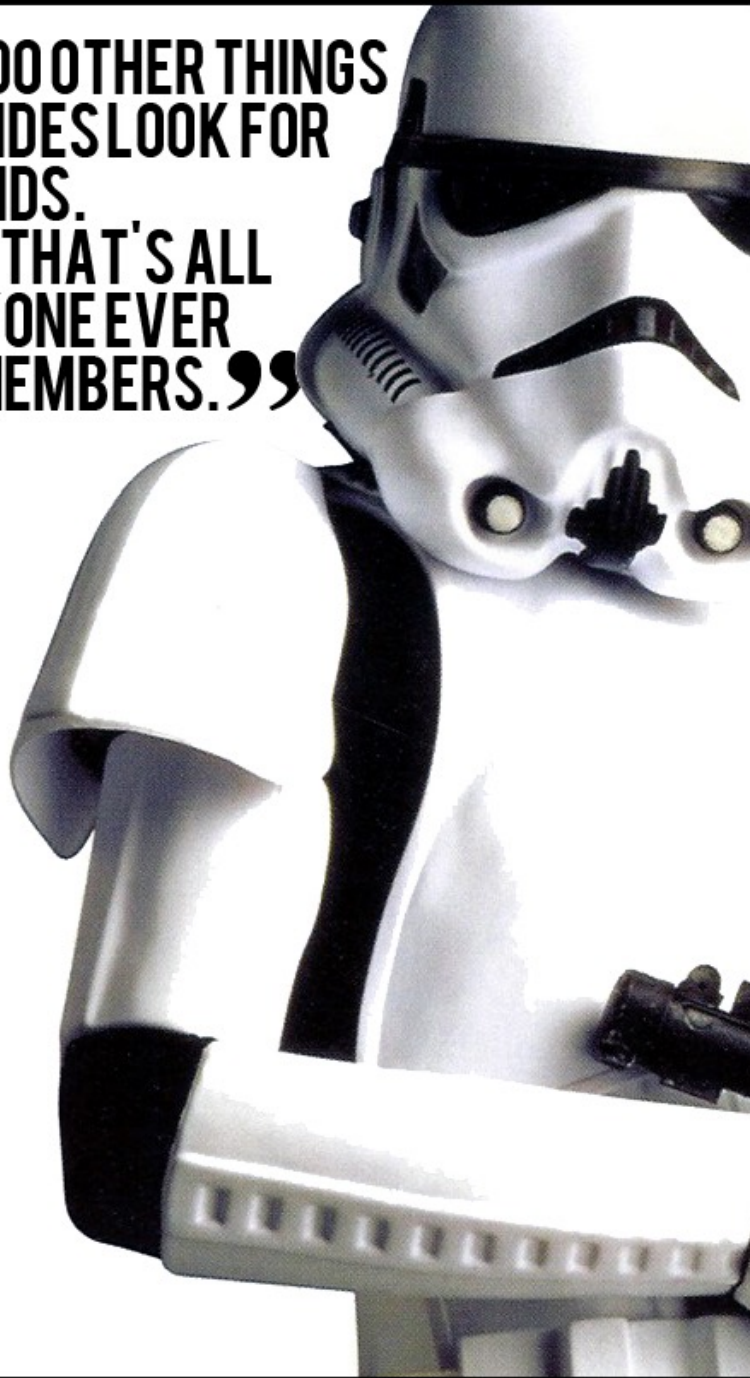


<http://mrkredband.files.wordpress.com/2011/09/froopse.jpg>
http://ic-pics.livejournal.com/1601_speak_471/11608842735440/245440_original.jpg
<http://blackskolor.com/wp-content/uploads/2009/11/final-saw-saw-crash-lookin-for-1.jpg>
<http://www.funnymeme.com/2014/08/10/funny-meme-we-have-the-droids-youre-looking-for/>

Bye, daddy. I hope you find the droids you're looking for.



“WE DO OTHER THINGS BESIDES LOOK FOR DROIDS. BUT THAT'S ALL ANYONE EVER REMEMBERS.”



These aren't the droids you're looking for



A parallel universe (far, far away)



Obi-Wan: You don't need to see his identification.

Stormtrooper: [ditto]

Obi-Wan: These aren't the droids you're looking for.

Stormtrooper: [ditto]

Obi-Wan: He can go about his business.

Stormtrooper: [ditto]

Data

From ~ 1000 movie scripts (many lines long),
pair IMDB “memorable quotes” with
 \sim adjacent,
same-length,
same-speaker
“non-memorable quotes”.
Filter with Google/Bing counts: 2200 pairs.

Pilot study

Subjects were shown 12 pairs from movies they hadn't seen.

<http://www.cs.cornell.edu/~cristian/memorability.html>

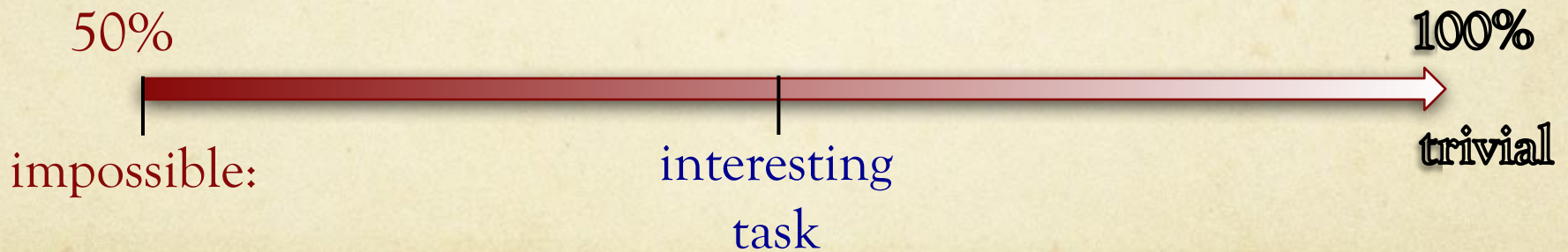
First quote	Second quote
Half a million dollars will always be missed	I know the type, trust me on this.
I think it's time to try some unsafe velocities.	No cold feet, or any other parts of our anatomy.
A little advice about feelings kiddo; don't expect it always to tickle.	I mean there's someone besides your mother you've got to forgive.

Pilot study

Subjects were shown 12 pairs from movies they hadn't seen.

<http://www.cs.cornell.edu/~cristian/memorability.html>

First quote	Second quote
Half a million dollars will always be missed	I know the type, trust me on this.
I think it's time to try some unsafe velocities.	No cold feet, or any other parts of our anatomy.
A little advice about feelings kiddo; don't expect it always to tickle.	I mean there's someone besides your mother you've got to forgive.



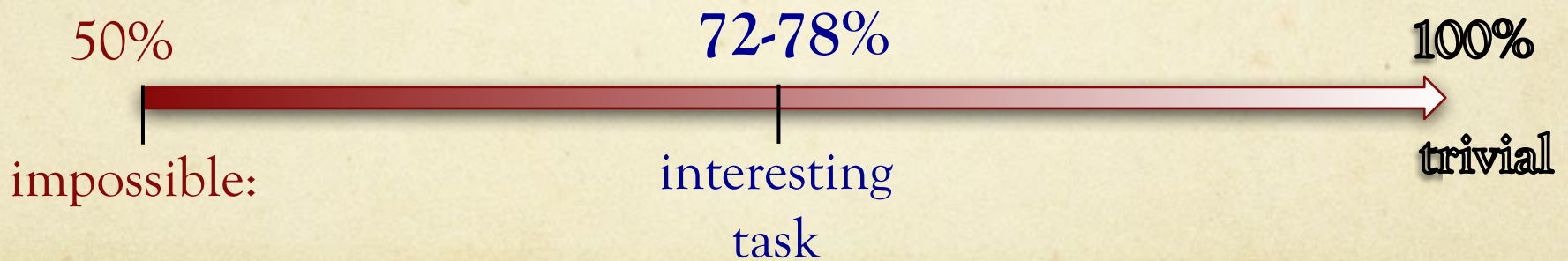
(context/actor effects
explain all, bad labels, etc.)

Pilot study

Subjects were shown 12 pairs from movies they hadn't seen.

<http://www.cs.cornell.edu/~cristian/memorability.html>

First quote	Second quote
Half a million dollars will always be missed	I know the type, trust me on this.
I think it's time to try some unsafe velocities.	No cold feet, or any other parts of our anatomy.
A little advice about feelings kiddo; don't expect it always to tickle.	I mean there's someone besides your mother you've got to forgive.



(context/actor effects explain all, bad labels, etc.)

On average, memorable quotes (significantly)...

On average, memorable quotes (significantly)...

... contain more **surprising** combinations of words

according to 1-,2-,3-gram lexical language models trained on the Brown corpus

“...aren't the droids...”

On average, memorable quotes (significantly)...

... contain more **surprising** combinations of words
according to 1-,2-,3-gram lexical language models trained on the Brown corpus

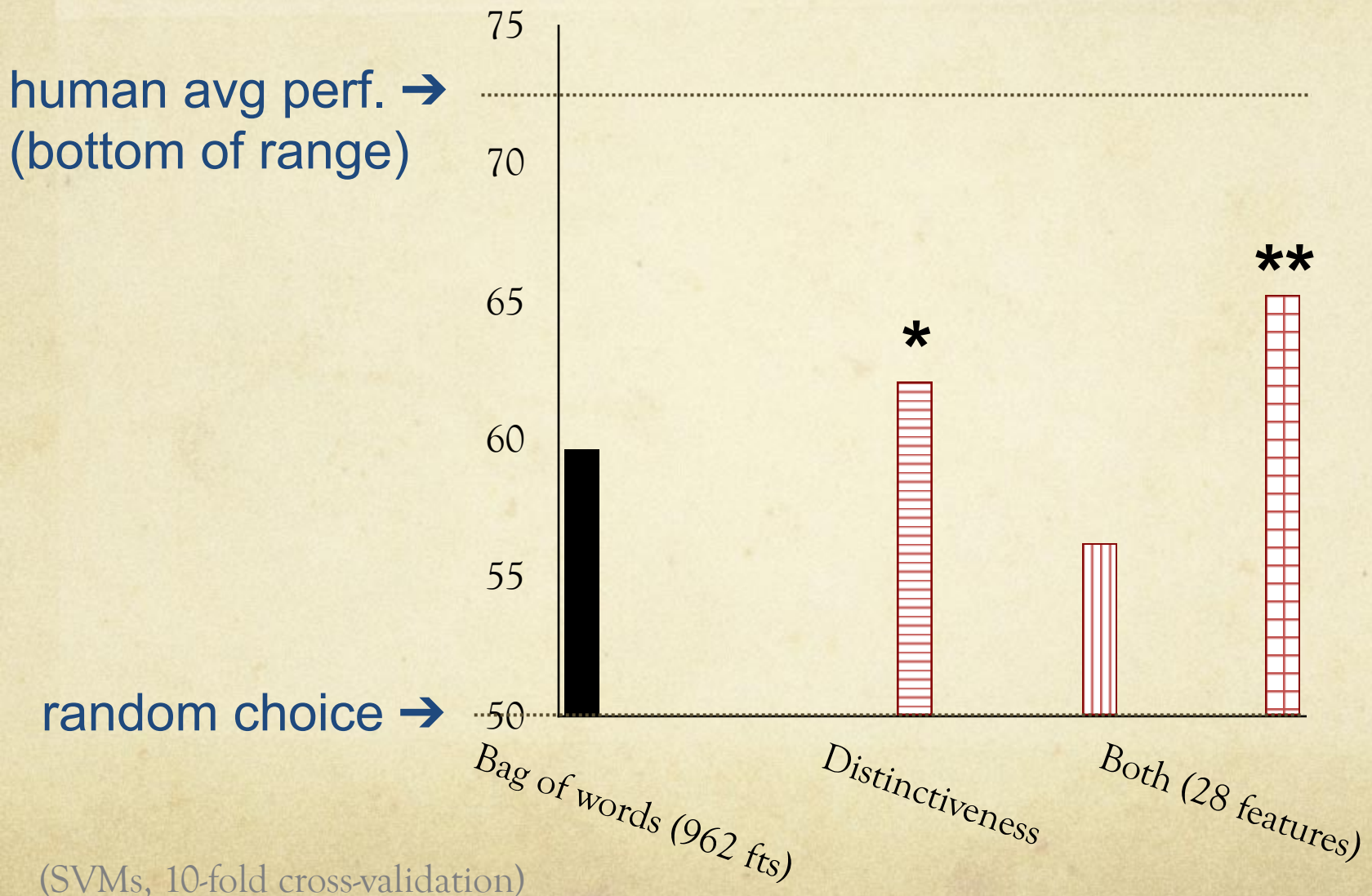
“...aren’t the droids...”

... but are built on a more **common** syntactic scaffolding
according to 1-,2-,3-gram part-of-speech language models trained on Brown

“You’re gonna need a bigger boat”

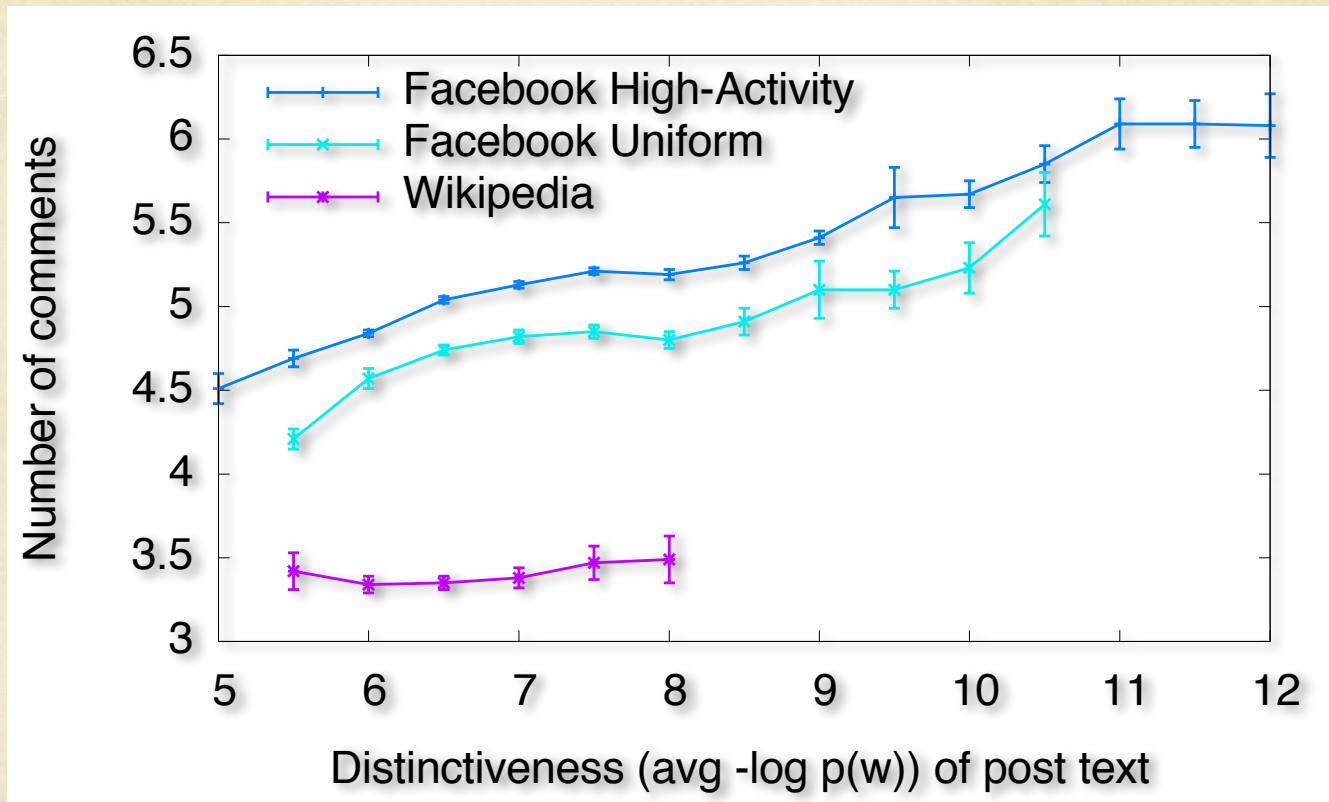
[vs. “You’re gonna need a boat that is bigger”]

Gloss of classification results



Applications to social-media UI

(Backstrom, Kleinberg, Lee, Danescu-Niculescu-Mizil 2013)



More-unusual Facebook posts get more comments

(under certain circumstances), but not so with Wikipedia.



Information diffusion

Related work includes: Guerini, Strapparava, Ötzbal, 2011; Petrovic, Osborne, Lavrenko 2011; Artzi, Pantel, Gamon 2012; Tsur, Rappoport 2012; Lakkaraju, McAuley,, Leskovec 2013; Wang, Wen 2015; Li, Ma, Guo, Mei 2017; Vosoughi, Roy, Aral 2018.

The parallel universe

Many Twitter users re-post about the same URL w/in 12 hours, varying their text, with significantly different retweet results



cactus_music @cactus_music · 15 Sep 2012

I know at some point you've have been saved from hunger by our rolling food trucks friends. Let's help support them! bit.ly/P6GYCq



cactus_music @cactus_music · 15 Sep 2012

Food trucks are the epitome of small independently owned LOCAL businesses! Help keep them going! Sign the petition bit.ly/P6GYCq

Try it! <http://chenhaot.com/retweetedmore/quiz>

The parallel universe

Many Twitter users re-post about the same URL w/in 12 hours, varying their text, with significantly different retweet results



cactus_music @cactus_music · 15 Sep 2012

I know at some point you've have been saved from hunger by our rolling food trucks friends. Let's help support them! bit.ly/P6GYCq



cactus_music @cactus_music · 15 Sep 2012

Food trucks are the epitome of small independently owned LOCAL businesses! Help keep them going! Sign the petition bit.ly/P6GYCq



GeorgeMonbiot @GeorgeMonbiot · 4 Oct 2012

read @ameliagentleman's report today, then tell me Tories are no longer the nasty party: guardian.co.uk/society/2012/o...



GeorgeMonbiot @GeorgeMonbiot · 4 Oct 2012

Work capability tests: designed by bastards, performed by idiots. guardian.co.uk/society/2012/o...

Try it! <http://chenhaot.com/retweetedmore/quiz>

Data

From 1.77M URL- and author-controlled tweet pairs using Yang and Leskovec's 236K user-ID list,

... filter out tweets with multiple URLs, users posting about the same URL 5+ times, retweets, non-English tweets, pairs differing only in spacing; cap users at 50 tweets

... slice on follower-count and time-lag to reduce timing effects

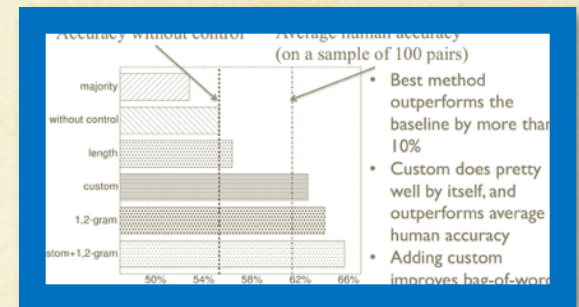
... and filter out the top 50% most similar pairs and the middle 90% most-similarly-retweeted pairs.

Gloss of prediction results

- Task difficulty? 61.3% per-human average accuracy [sample of 100 pairs; 106 judges; 39 judgments/pair], various other **baselines**

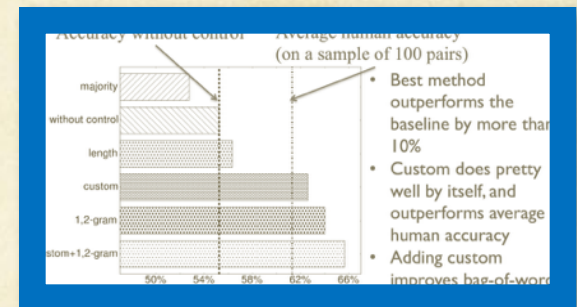
Gloss of prediction results

- Task difficulty? 61.3% per-human average accuracy [sample of 100 pairs; 106 judges; 39 judgments/pair], various other **baselines**
- We find features that yield better cross-validation prediction results, and better results on **truly* held-out data**



Gloss of prediction results

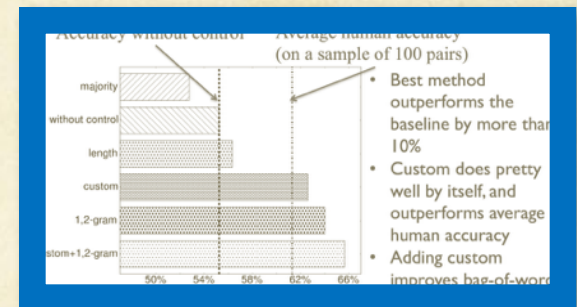
- Task difficulty? 61.3% per-human average accuracy [sample of 100 pairs; 106 judges; 39 judgments/pair], various other **baselines**
- We find features that yield better cross-validation prediction results, and better results on **truly* held-out data**



*We ran only one experiment on it, and that was right before submission

Gloss of prediction results

- Task difficulty? 61.3% per-human average accuracy [sample of 100 pairs; 106 judges; 39 judgments/pair], various other **baselines**
- We find features that yield better cross-validation prediction results, and better results on **truly* held-out data**



*We ran only one experiment on it, and that was right before submission

Try our prediction tool for your own tweets:
<https://chenhaot.com/retweetedmore/>

Teaser results

On average,

preferred versions are more similar to Twitter's overall language (unigram or bigram language model)

and more similar to the user's prior language.

(unigram language model)

Table 3: Explicit requests for sharing (where only occurrences POS-tagged as verbs count, according to the Gimpel et al. (2011) tagger).

	effective?	author-preferred?
rt	↑↑↑↑ *	YES (59%)
retweet	↑↑↑↑ *	YES (56%)
spread	↑↑↑↑ *	—
please	↑↑↑↑ *	—
pls	↑↑↑↑ *	—
plz	↑↑	—

Table 4: Informativeness.

	effective?	author-preferred?
length (chars)	↑↑↑↑ *	YES (54%)
verb	↑↑↑↑ *	YES (56%)
noun	↑↑↑↑ *	YES (51%)
adjective	↑↑↑↑ *	YES (55%)
adverb	↑↑↑↑ *	NO (45%)
proper noun	↑↑↑↑ *	NO (48%)
number	↑↑↑↑ *	—
hashtag	↑↑↑↑ *	—
@-mention	↑↑↑↑ *	—

Table 6: LM-based resemblance to headlines.

	effective?	author-preferred?
headline unigram	↑↑	YES (53%)
headline bigram	↑↑↑↑ *	YES (52%)

Table 7: Retweet score.

	effective?	author-preferred?
rt score	↑↑↑↑ *	NO (49%)
verb rt score	↑↑↑↑ *	YES (50%)
noun rt score	↑↑↑↑ *	YES (51%)
adjective rt score	↑↑↑↑ *	NO (48%)
adverb rt score	↑↑↑↑ *	—
proper noun rt score	↑	—

Table 8: Sentiment (contrast is measured by presence of both positive and negative sentiments).

	effective?	author-preferred?
positive	↑↑↑ *	—
negative	↑↑↑ *	—
contrast	↑↑↑ *	—

Table 9: Pronouns.

	effective?	author-preferred?
1st person singular	—	YES (51%)
1st person plural	—	YES (52%)
2nd person	—	YES (57%)
3rd person singular	↑↑	YES (55%)
3rd person plural	↑	YES (58%)

Table 10: Generality.

	effective?	author-preferred?
indefinite articles (a,an)	↑↑↑ *	—
definite articles (the)	↑↑↑ *	—

Table 11: Readability.

	effective?	author-preferred?
reading ease	↑↑	YES (52%)
negative grade level	↑↑	YES (52%)



Persuasion

Other *ACL work includes: Guerini, Strapparava, Stock 2010; Mochales, Moens, 2011; Althoff, Danescu-Niculescu-Mizil, Jurafsky 2014; Wei, Liu, Li 2016; Persing, Ng 2017; Simpson and Gurevych 2018; Xiao 2018; Anselma, Mazzei (this INLG)

Audience effects: Ding and Pan, 2016; Lukin, Anand, Walker, Whittaker 2017; Durmus, Cardie 2018

Persuasion attempts in ChangeMyView

Original post

CMV: the Tontine should be legalized and made a common retirement strategy.

[Reference URL omitted] Basically, today we have a huge problem with retirement [...+73 words]

A tontine for retirement looks like [...+56 words] The yearly sum is divided evenly for all the surviving participants [...+25 words]. The key advantages as I see it are:

*We don't need actuaries [...+29 words...]

*Management fees can be quite low [...+22 words]

* [Another reason]

* [Another reason]

But CMV. Are there major risks I am not foreseeing? [+2 more questions]



A tontine is a pretty crappy retirement vehicle for most people. It pays out the least when you need the most, and the most when you need the least.

People's income needs in retirement generally fall as they age. [...+35 words]

[URL]



Very interesting. I'll give a Δ because I didn't have any idea that was true and changes my idea of how the tontine should work. That said, I don't think it's unsolvable [...+44 words]



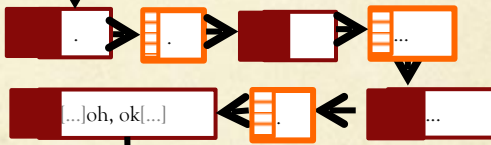
[DeltaBot] Confirmed: 1 delta awarded to [red]

The Social Security system is basically one giant Tontine [...+13 words] So it's already legal.

There are some key differences though. First, Social Security is defined by the government [...+36 words]

I'd imagine the tontine as a secondary system to social security though, one that is optional for people to do, not mandatory like social security. [...+11 words]

And a tontine would be defined by your bank [...+79 words]



Good points. I still am not sold [... > 100 words]

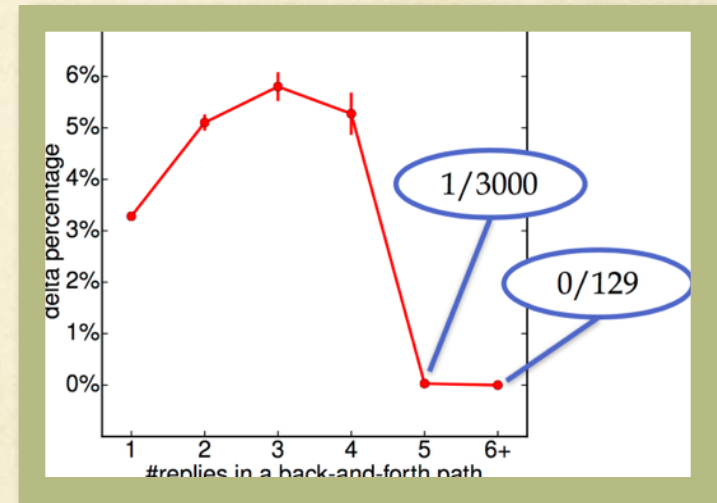
Then your back to needing actuaries, to predict [...+11 words]

Depends how exact you need to be [...+33 words]

Teaser results

Data: 20K argument trees, 1.1M contributions, 80K unique participants. Truly held-out results.

On average, persuasive “versions” are more dissimilar in content words to the original argument, but more similar in non-content words.



Too much back ‘n’ forth

Many other findings plus caveats; see paper



All papers and datasets are on my homepage:
<http://www.cs.cornell.edu/home/llee>

Looking forward

Even deeper interplay between
natural language processing
and

how people use and are affected by language
is a huge opportunity for all concerned.

I think this is the beginning
of a beautiful friendship.



I think this is the beginning
of a beautiful friendship.

Thanks!

