Language as Influence(d): Power and memorability

Lillian Lee
Dept. of Computer Science, Cornell University
http://www.cs.cornell.edu/home/llee

Removed: image licensed for powerpoint presentation only; preview can be viewed here: https://cartoonbank.licensestream.com/LicenseStream/Store/Content?Filter.SearchString=23622&search=
O brave new world, that has such people in't

Lots of on-line conversations
(Facebook, Slashdot,...; YouTube comments...; Yelp reviews..., ...)

= Many systems with humans and language as key components

= Fantastic opportunities for NLP + the social sciences
to improve such systems and create new ones
Preview of case study 1

Who’s in charge?

Did we have similar language in Cherokee?

Didn’t we have similar language in Cherokee Nation?

Who’s in charge?
Preview of case study 2

Does phrasing affect memorability?  
Or is it all language-external factors?

You can put *lipstick on a pig*, but it’s still a pig.

*Palin’s lipstick remark*  
*Can wording also have an effect?*
Aside: on presentation style

Your goal is not to convince your audience 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 believe that you are correct.

-- paraphrase of Stuart Shieber
Echoes of Power:

Language effects & power differences in social interaction

Cristian Danescu-Niculescu-Mizil, Lillian Lee, Bo Pang & Jon Kleinberg
WWW 2012
Example motivating applications

Proactive task-discussion management
  • Steering conversations back on track

Conversation facilitation systems
  • Defusing imminent flame-wars
  • Measuring engagement in MOOC settings
  • e-govt participation initiatives

[Farina, Newhart, Cardie and Cosley '11]
Language reveals power: “easy” cases

Your Honor, I agree.

I’d love to get your thoughts on this when you are available.

Thank you.

Let’s discuss later.


What about general (domain-independent) signals?
Who has the (conversational) lead?

Communicative behaviors are “patterned and coordinated, like a dance” [Niederhoffer and Pennebaker 2002]

Look for asymmetric adaptation of linguistic style
Defining linguistic style coordination

Direct repetition: under speaker’s control, could just be choice of topic.

Function-class matching: unconscious & frequent
[Niederhoffer and Pennebaker 2002]

At least you were outside.

It doesn’t make much difference...

Doesn’t really matter...

It’s not important...
Measuring immediate influence

How much does speaker $x_1$ immediately trigger $x_2$’s use of function-word class $c$?

= how much does $x_2$ *coordinate* to $x_1$ on $c$?

[Danescu-Niculescu-Mizil, Dumais, Gamon WWW 2011]

\[ \Pr (x_2 \text{ uses } c \mid x_1 \text{ uses } c, x_2 \text{ immediately replies}) \]

= \[ \Pr (x_2 \text{ uses } c \mid x_2 \text{ immediately replies to } x_1) \]

(reported as % (multiplied by 100) in the following)
Status in US Supreme Court transcripts
50K exchanges + metadata (download link on my homepage)

low status to high status
high status to low status
Status change in Wikipedia

240K exchanges + metadata (download from my webpage)
Dependence in Supreme Court transcripts

Power differences can arise from dependence [Emerson 1962]

“needed to change” J’s mind

J “had your opinion”
Evidence of domain independence

SVM classification with various features

- bag of words (20,000 features)
- coordination (9 features)
- "stylistic" (18 features)
You Had Me at Hello: How phrasing affects memorability

Joint with Cristian Danescu-Niculescu-Mizil, Justin Cheng & Jon Kleinberg
ACL 2012

You can put lipstick on a pig, but it's still a pig…
Does phrasing affect memorability?

Motivations:
• Ranking what material to promote
• Understanding influence, information diffusion, and persuasion.

[Much related work in many fields; see paper for refs. Our direct inspiration: Leskovec, Backstrom, & Kleinberg '09; Meme modification: Simmons, Adamic & Adar '11]

• What kind of data?
• What counts as “memorable?”
• How to tease apart language from context? (the lipstick-on-a-pig problem)
Movie quotes: massively viral

"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: 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.
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]
Obi-Wan: Move along.
Stormtrooper: [ditto]

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.
Data

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

Available at http://www.mpi-sws.org/~cristian/memorability.html
Pilot study

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

http://www.mpi-sws.org/~cristian/memorability.html

<table>
<thead>
<tr>
<th>Movie</th>
<th>First Quote</th>
<th>Second Quote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jackie Brown</td>
<td>Half a million dollars will always be missed.</td>
<td>I know the type, trust me on this.</td>
</tr>
<tr>
<td>Star Trek: Nemesis</td>
<td>I think it’s time to try some unsafe velocities.</td>
<td>No cold feet, or any other parts of our anatomy.</td>
</tr>
<tr>
<td>Ordinary People</td>
<td>A little advice about feelings kiddo; don’t expect it always to tickle.</td>
<td>I mean there’s someone besides your mother you’ve got to forgive.</td>
</tr>
</tbody>
</table>

50% impossible

72-78% interesting task

100% trivial

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

... 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”]
On average, memorable quotes (significantly)…

… are more portable/general:

- more indefinite articles

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

- fewer 3rd person pronouns (“he” vs “we”),
- fewer past tense verbs and more present tense verbs
Improved classification accuracy (SVMs, 10-fold cross-validation)

human avg perf. (bottom of range)

random choice

Bag of words (962 fts)
Distinctiveness
Generality
Both (28 features)
Ad slogans are closer to memorable quotes in lexical and syntactic choice:

Memorable-quote LMs predict slogans better than non-memorable-quote LMs

... and in generality:

% 3rd pers. pronoun:
% indefinite articles:
% past tense:
More-unusual Facebook posts get more comments (under certain circumstances), but not so with Wikipedia.
Aside: re-entry prediction:
A new "interestingness" characterization

(1) Length/volume

(2) Re-entry
Conclusions (I)

I had not yet come to the conviction: It is not what you say that matters but the manner in which you say it; there lies the secret of the ages.

--- William Carlos Williams, preface to Selected Essays

Social power relationships affect language even at extremely fine-grained levels.

Phrasing does indeed have some effect on memorability and engagement.
Improved computational techniques for socio-computational systems is an important broader direction.

I view deeper interplay between constructing NLP-enabled systems and understanding how people use & are affected by language as a huge opportunity for all concerned.
I think this is the beginning of a beautiful friendship.

Thanks!