Orthographic Social Variation in Online Writing

Jacob Eisenstein

Georgia Institute of Technology

June 6, 2018

Orthographic variation

Orthographic variation is a classic problem for subword models! (Baron & Rayson, 2008)

(1) Original: and drewe vnto hym all ryottours & wylde dysposed persones

(2) Normalization: and drew unto him all rioters and wild disposed persons

Orthographic variation

Orthographic variation is a classic problem for subword models! (Baron & Rayson, 2008)

(1) Original: and drewe vnto hym all ryottours & wylde dysposed persones

(2) Normalization: and drew unto him all rioters and wild disposed persons

But some distinctions make a difference: in contemporary online writing, orthographic variation is recruited to create social meanings.

Orthographic variation today

- How do these variable forms arise?
- What do they mean?
- How do they relate to speech?

Three views of orthographic variation

- in phonological context (Eisenstein, 2015)
- as infectious disease (Goel et al., 2016)
- in stigmatized communities (Stewart et al., 2017)

Three views of orthographic variation

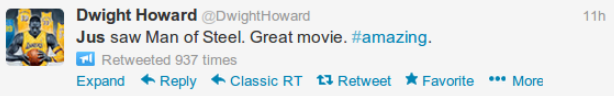
- **in phonological context** (Eisenstein, 2015)
- as infectious disease (Goel et al., 2016)
- in stigmatized communities (Stewart et al., 2017)

Demographic variation in orthography

Some non-standard spellings hint at dialectal pronunciations:

(ing)

(-t,-d)

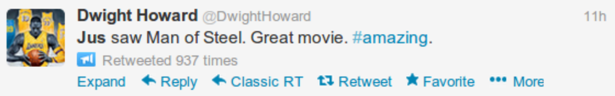


What do these spellings share with the pronunciations that they seem to invoke?

The variable (ing) in speech

- The velar/coronal alternation is ubiquitous, connoting informality throughout world English
- Linguistic constraints and preferences:
  - less likely for nouns than verbs (earing vs cheering);
  - prohibited for monosyllabic words (ring, king)

The variable (td) in speech



- In varieties such as AAE (Green, 2002) and British English (Tagliamonte & Temple, 2005).
- Drosophila of sociophonetics, with several constraints and preferences (Guy, 1991)
- Dispreferred in context preceding a vowel:

  (3) It's just peaches

  (4) It's just apples

Orthographic variables in Twitter

(int)

- 200 most frequent -ing words in 10^5 tweets
- remove three ambiguous examples:
  sing, thing, king

(td)

- Searched non-dictionary words by frequency, identify cases of -td deletion:
  jus(t), nex(t), ain(t), ol(d), tol(d)

G-deletion

	Log odds	%	N
Verb	.227	.200	89,173
Noun	-.013	.083	18,756
Adjective	-.213	.149	4,964
monosyllable	-2.57	.001	108,804
Total		.178	112,893

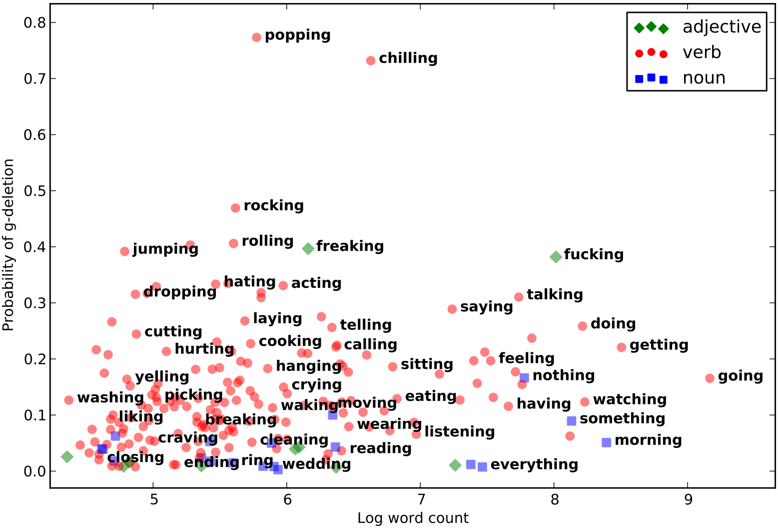
G-deletion

	Log odds	%	N
Verb	.227	.200	89,173
Noun	-.013	.083	18,756
Adjective	-.213	.149	4,964
monosyllable	-2.57	.001	108,804
@-message	.134	.205	36,974
Total		.178	112,893

G-deletion

	Log odds	%	N
Verb	.227	.200	89,173
Noun	-.013	.083	18,756
Adjective	-.213	.149	4,964
monosyllable	-2.57	.001	108,804
@-message	.134	.205	36,974
High Euro-Am county	-.194	.117	28,017
High Afro-Am county	.145	.241	27,022
Total		.178	112,893

("high" / "low" = top/bottom quartile)


adjective
verb
noun
Probability of g-deletion
Log word count
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
5
6
7
8
9
popping
chilling
rocking
rolling
freaking
jumping
fucking
hating
acting
dropping
talking
saying
laying
telling
doing
cutting
cooking
calling
getting
hurting
feeling
hanging
sitting
nothing
going
yelling
crying
washing
picking
eating
watching
waking
moving
having
liking
breaking
something
wearing
listening
craving
cleaning
morning
reading
closing
ending
ring
wedding
everything

-t,-d deletion

	Log odds	%	N
Vowel succeeding context	-.066	.385	9,004
Total		.423	89,174

-t,-d deletion

	Log odds	%	N
Vowel succeeding context	-.066	.385	9,004
@-message	.075	.436	35,240
Total		.423	89,174

-t,-d deletion

	Log odds	%	N
Vowel succeeding context	-.066	.385	9,004
@-message	.075	.436	35,240
High Euro-Am county	-.313	.311	19,992
High Afro-Am county	.065	.508	19,854
High income county	-.107	.388	20,653
Medium income county	.019	.406	43,135
Low income county	.127	.482	25,386
Total		.423	89,174

Discussion

- Orthography recapitulates phonology.
  - g-deletion follows syntactic and phonological patterns of (ing).
  - td-deletion follows phonological patterns of (td).

Discussion

- Orthography recapitulates phonology.
  - g-deletion follows syntactic and phonological patterns of (ing).
  - td-deletion follows phonological patterns of (td).
- Dialect is not destiny.
  - Each variable is correlated with demographic properties of tweet location.
  - But authors modulate usage of these variables depending on the audience (Bell, 1984; Pavalanathan & Eisenstein, 2015).
  - Orthography is a resource that can be deployed to create desired social meanings (Eckert, 2000).

Three views of orthographic variation

- **in phonological context** (Eisenstein, 2015)
- as infectious disease (Goel et al., 2016)
- in stigmatized communities (Stewart et al., 2017)

Three views of orthographic variation

- in phonological context (Eisenstein, 2015)
- **as infectious disease** (Goel et al., 2016)
- in stigmatized communities (Stewart et al., 2017)

Orthography and influence

How does orthographic variation become associated with social meaning?

- It must be a social process.
- For newer variables, this process must have taken place recently.
- Can social metadata explain the fine structure of how this process works?

(Goel et al., 2016)

Language change as epidemiology

Hypothesis: if your "neighbor" adopts a linguistic innovation, you're more likely to adopt it too.

$$r(n) = \frac{\Pr(\text{infection} \mid n \text{ neighbors infected })}{\Pr(\text{infection} \mid 0 \text{ neighbors infected })}$$

Dataset

- Twitter analysis is usually conducted on a **sample** from the streaming API (e.g., Eisenstein et al., 2010; Huang et al., 2016).

Dataset

- Twitter analysis is usually conducted on a **sample** from the streaming API (e.g., Eisenstein et al., 2010; Huang et al., 2016).
- Complete data from all public tweets in the U.S. between 2011-2014:
  - Limited to 20 pre-selected variables undergoing rapid growth in 2013-2014, including LEXICAL, ABBREVIATION, and ORTHOGRAPHIC variables.
  - 4.35 million unique user accounts, with social network ties.
  - Obtained in collaboration with MSR-NYC.

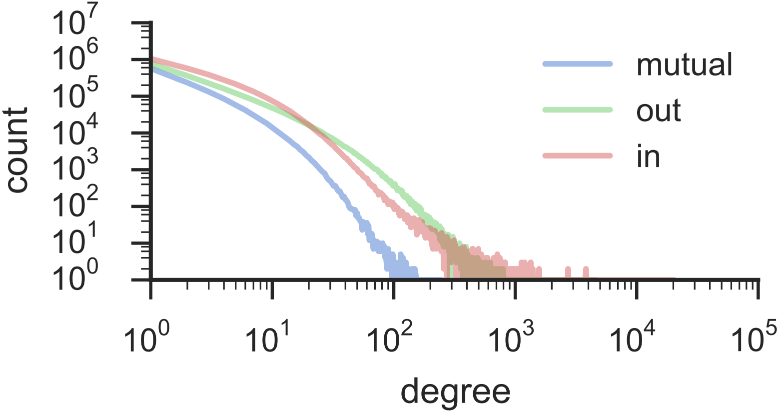
Social network

Two Twitter accounts are socially linked if they have each mentioned each other in a message, e.g.

- User1: @user2 salut
- User2: @user1 what's up?

This *mention network* is more socially meaningful than the *articulated network* of follower-followee links (Huberman et al., 2008).

Social network



The symmetrized ("mutual") mention network yields a more credible degree distribution.

Summary of data

Social network

Bart	Lisa
Bart	Milhouse
Lisa	Homer
Homer	Barney
...	...

Locations

Bart	Los Angeles
Milhouse	Los Angeles
Lisa	Atlanta
Homer	Chicago
...	...

Language

Bart	jawn	Feb 1, 2013, 13:45
Milhouse	jawn	Feb 1, 2013, 13:50
Homer	hella	Feb 1, 2013, 18:15
Bart	lls	Feb 2, 2013, 07:30
Milhouse	lls	Feb 2, 2013, 07:40
...	...	...

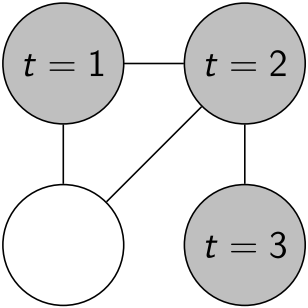
Homophily versus contagion

Hypothesis: if your "neighbor" adopts a linguistic innovation, you're more likely to adopt it too.

- **Contagion**: you caught the disease from your neighbor.
- **Homophily**: you became neighbors because you do both things that make you more likely to catch the disease.

Very hard to distinguish contagion and homophily in general (Shalizi & Thomas, 2011).

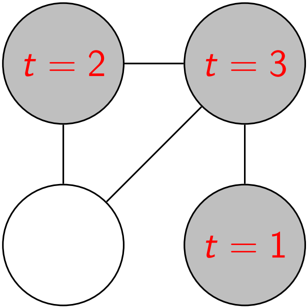
Shuffle test for contagion



- Observed data

$$\Pr(\text{infection} \mid 0 \text{ exposures}) = \frac{1}{4}$$
$$\Pr(\text{infection} \mid \geq 1 \text{ exposures}) = \frac{2}{3}$$

Shuffle test for contagion



- Observed data

$$\Pr(\text{infection} \mid 0 \text{ exposures}) = \frac{1}{4}$$

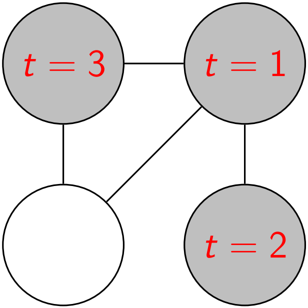
$$\Pr(\text{infection} \mid \geq 1 \text{ exposures}) = \frac{2}{3}$$

- Randomized data

$$\Pr(\text{infection} \mid 0 \text{ exposures}) = \frac{2}{4}$$

$$\Pr(\text{infection} \mid \geq 1 \text{ exposures}) = \frac{1}{2}$$

Shuffle test for contagion



- Observed data

$$\Pr(\text{infection} \mid 0 \text{ exposures}) = \frac{1}{4}$$
$$\Pr(\text{infection} \mid \geq 1 \text{ exposures}) = \frac{2}{3}$$

- Randomized data

$$\Pr(\text{infection} \mid 0 \text{ exposures}) = \frac{2}{4}, \frac{1}{4}, \ldots$$
$$\Pr(\text{infection} \mid \geq 1 \text{ exposures}) = \frac{1}{2}, \frac{2}{3}, \ldots$$

Using the shuffle test

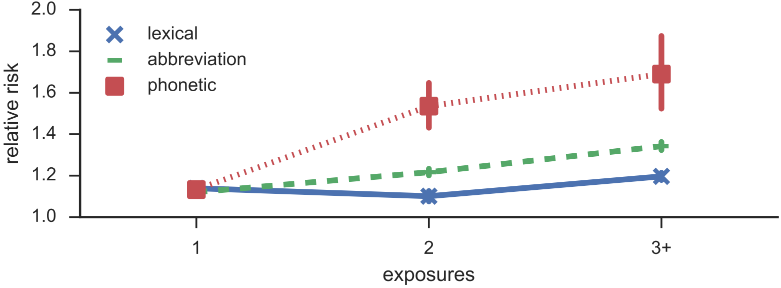
Null hypothesis: relative risk$(n) = 1$.

$$\text{relative-risk}(n) = \frac{r(n)}{\frac{1}{K}\sum_{k=1}^{K} \tilde{r}^{(k)}(n)}$$

If the null hypothesis is rejected, shuffle test provides evidence for contagion under limited conditions:

- network is static;
- homophily effects are static.

Evidence of contagion



- Relative risk > 1: evidence of non-random contagion.
- For phonetic variables, risk increases with multiple exposures.
- This pattern of **complex contagion** is characteristic of the adoption of socially risky innovations.

Three views of orthographic variation

- in phonological context (Eisenstein, 2015)
- **as infectious disease** (Goel et al., 2016)
- in stigmatized communities (Stewart et al., 2017)

Three views of orthographic variation

- in phonological context (Eisenstein, 2015)
- as infectious disease (Goel et al., 2016)
- **in stigmatized communities** (Stewart et al., 2017)

Writing style in stigmatized communities

Subcommunities of Instagram advocate for stigmatized lifestyles, such as eating disorders (tw)

Writing style in stigmatized communities

Subcommunities of Instagram advocate for stigmatized lifestyles, such as eating disorders (tw)

(Chancellor et al., 2016; Stewart et al., 2017) [redacted] Feeling a little better..if i want to reach my goal i have to be patient and work harder. I hope you are doing the same :) #cardio #excercise #loseweight #workout #skinny #thir #thighggap #collarbones

[redacted] #thighgap

thighgap

SOCIETY

Instagram Bans Thinspo Content

Instagram is the latest social media platform to ban thinspiration content. But are these policies effective?

By Heba Hasan @Heba__H | April 26, 2012

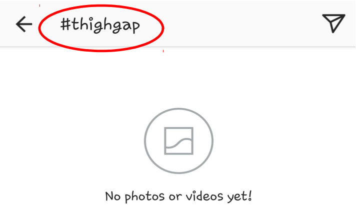
Thinspo content will no longer be welcome on Instagram. Following in the footsteps of Pinterest and Tumblr, Instagram is the latest social media site to ban "thinspiration" photos — images that are meant to provide motivation for those who want to lose weight and which health experts say often contribute to eating disorders.

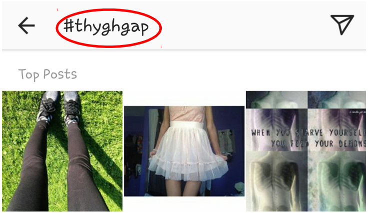
Instagram's new policy doesn't come as a surprise. The app came under scrutiny last week when celebrity and Instagram user Alexa Chung posted a photo of herself and was attacked by users for being too skinny.

Michaela Begsteiger/Getty

Instagram bans thinspo content

Orthographic variation in hashtags





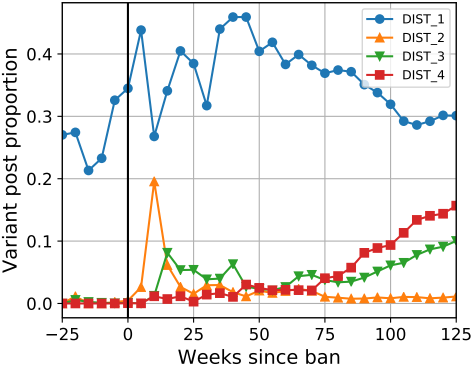
Quantifying the depth of variation

	ED=1	ED=2	ED=3	ED=4
	thyghgap	thyghgapp	thyygap	thyygapp
thighgap	thghgap	thiigap	thighgaappp	thygaps
	thightgap	thightgrap	thightpag	thygsp

Depth

Variation over time

After a crackdown on pro-ED hashtags, variation becomes more frequent and more profound.

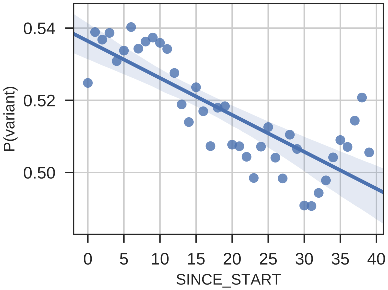


Social factors

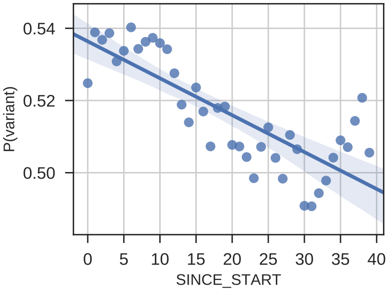
Who drives this ongoing change, and why?

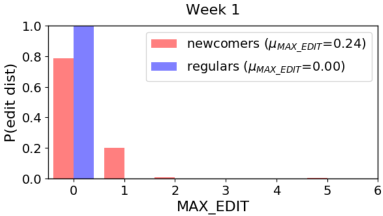
- "Age": length of time that a user account has been posting pro-ED content.
- "Commitment": duration of time that a user account *will post* pro-ED content.

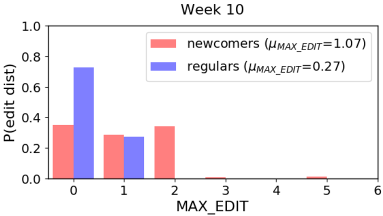
Newcomers use more and deeper variation



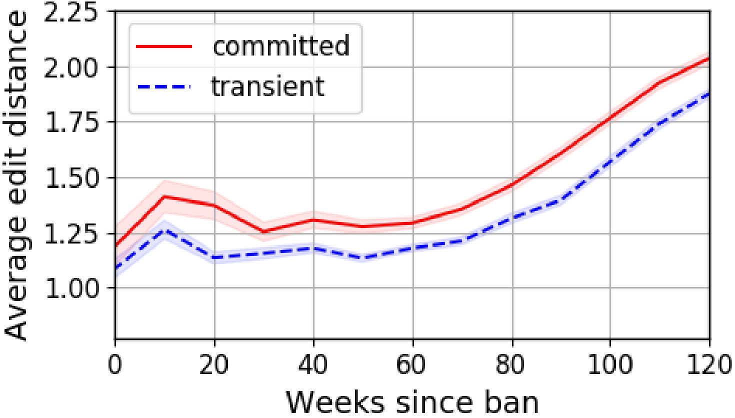
Newcomers use more and deeper variation





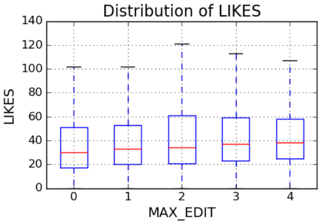


Deeper variation → deeper commitment



Deeper variation → (a few) more likes

	β	*SE*
Dependent variable: LOGLIKES		
USER AGE	-0.0319***	9.03E-4
TOTAL TAGS	0.224***	1.47E-3
VARIANT Y/N?	-1.14E-3	1.98E-3
MAX_POP	-3.89E-3***	7.25E-4
max_edit	0.0130***	3.16E-3



Poisson regression, *** indicates $p < .0001$.

Discussion

Danescu-Niculescu-Mizil et al. (2013): forum members are linguistically **conservative**, retaining the lexical features of their "youth."

Discussion

Danescu-Niculescu-Mizil et al. (2013): forum members are linguistically **conservative**, retaining the lexical features of their "youth."

- We find that old-timers are **regressive**, abandoning their earlier innovative practices.
- Committed newcomers drive variation further.

Discussion

Danescu-Niculescu-Mizil et al. (2013): forum members are linguistically **conservative**, retaining the lexical features of their "youth."

- We find that old-timers are **regressive**, abandoning their earlier innovative practices.
- Committed newcomers drive variation further.
- Orthography complements the lexical features typically used in computational sociolinguistics, measuring **depth** of variation away from standard.

Three views of orthographic variation

- in phonological context (Eisenstein, 2015)
- as infectious disease (Goel et al., 2016)
- in stigmatized communities (Stewart et al., 2017)

Last thoughts

- Much of the work in SCLEM is about **generalization** over orthographic differences (e.g., Pinter et al., 2017).
- But orthography also creates socially meaningful **distinctions**.
- Human readers generalize about propositional content, while distinguishing social meanings.
- Can we learn representations of subword structure that do this too?

Acknowledgments

- **Students**: Stevie Chancellor, Rahul Goel, Umashanthi Pavalanathan, Yuval Pinter, Sandeep Soni, Ian Stewart, Yi Yang.
- **Collaborators**: Munmun De Choudhury, Fernando Diaz, Hanna Wallach
- **Sponsors**: NSF, AFOSR, NIH, DTRA, Google

References I

Baron, A. & Rayson, P. (2008). Vard2: A tool for dealing with spelling variation in historical corpora. In *Postgraduate conference in corpus linguistics*.

Bell, A. (1984). Language style as audience design. *Language in Society*, *13*(2), 145–204.

Chancellor, S., Lin, Z. J., Goodman, E. L., Zerwas, S., & De Choudhury, M. (2016). Quantifying and predicting mental illness severity in online pro-eating disorder communities. In *Proceedings of Computer-Supported Cooperative Work (CSCW)*.

Danescu-Niculescu-Mizil, C., West, R., Jurafsky, D., Leskovec, J., & Potts, C. (2013). No country for old members: User lifecycle and linguistic change in online communities. In *Proceedings of the Conference on World-Wide Web (WWW)*, (pp. 307–318).

Eckert, P. (2000). *Linguistic variation as social practice*. Blackwell.

Eisenstein, J. (2015). Systematic patterning in phonologically-motivated orthographic variation. *Journal of Sociolinguistics*, *19*, 161–188.

Eisenstein, J., O'Connor, B., Smith, N. A., & Xing, E. P. (2010). A latent variable model for geographic lexical variation. In *Proceedings of Empirical Methods for Natural Language Processing (EMNLP)*, (pp. 1277–1287).

Goel, R., Soni, S., Goyal, N., Paparrizos, J., Wallach, H., Diaz, F., & Eisenstein, J. (2016). The social dynamics of language change in online networks. In *The International Conference on Social Informatics (SocInfo)*.

Green, L. J. (2002). *African American English: A Linguistic Introduction*. Cambridge, U.K.: Cambridge University Press.

Guy, G. R. (1991). Contextual conditioning in variable lexical phonology. *Language Variation and Change*, *3*, 223–239.

Huang, Y., Guo, D., Kasakoff, A., & Grieve, J. (2016). Understanding us regional linguistic variation with twitter data analysis. *Computers, Environment and Urban Systems*, *59*, 244–255.

Huberman, B., Romero, D. M., & Wu, F. (2008). Social networks that matter: Twitter under the microscope. *First Monday*, *14*(1).

Pavalanathan, U. & Eisenstein, J. (2015). Audience-modulated variation in online social media. *American Speech*, *90*(2).

References II

Pinter, Y., Guthrie, R., & Eisenstein, J. (2017). Mimicking word embeddings using subword RNNs. In *Proceedings of Empirical Methods for Natural Language Processing (EMNLP)*.

Shalizi, C. R. & Thomas, A. C. (2011). Homophily and contagion are generically confounded in observational social network studies. *Sociological methods & research*, *40*(2), 211–239.

Stewart, I., Chancellor, S., Choudhury, M. D., & Eisenstein, J. (2017). #anorexia, #anarexia, #anarexyia: Characterizing online community practices with orthographic variation. In *Proceedings of IEEE Big Data: Workshop on Natural Language Processing for Social Media (SocialNLP)*.

Tagliamonte, S. A. & Temple, R. (2005). New perspectives on an ol' variable: (t,d) in British English. *Language Variation and Change*, *17*, 281–302.