What makes “What's hot” hot?

+Ching-Wan Yip
yipcw@wfu.edu

Written Language in a Digital Era
Dr. Laura Aull

2012.04.30
Agenda

• Google+ What’s Hot
• Characteristics investigated
• Findings
Introduction: Google+

http://plus.google.com/
Newest social networking site - Jun 28, 2011
Currently 170+ million users (April 12, 2012)
vs Facebook 900+ million (Apr 23, 2012)
Introduction: Google+ Stream
Introduction: Google+ Stream
Introduction: Google+ What's hot

Launched October 24, 2011

8 posts in user home stream

cannot be turned off initially
Introduction: #whitespace redesign

also in explore tab after April 2012
UI update
Introduction: volume control

"volume control" available.
See Baron 2008, Vaughan-Nichols 2012
Introduction: how?

- exact algorithm not known
  - not just [...], comments (maximum 500)
Introduction: why? marketing

- Search Engine Optimization (SEO) → social media optimization

- Inc. 500 fast growing corporations
county blogs 2010:50% ↓ 2011:37%
Facebook, Tumblr, Twitter etc.
see Yu 2012
Methods: What makes posts hot?

Research Questions

- Linguistic Features
  - compare with other DMC
- Authors characteristics
Methods: Data Collection from

- 14 days - 2012.03.07-2012.03.21 (continuing)
- use 40 posts each day total 560 posts (collect 50 posts for rejects, max. ~64)
- between 12am-8am
Methods: discards

- duplicate content from different sources
- remove http:// links to youtube and vimeo, more photos from (no longer necessary after recent UI change)
Methods: blank
Methods: embed text

• Most of time not by author, is it content?
Methods: Software

- Laurence Anthony antconc 3.2.4w (Oct 4 2011)
- Cygwin wc utility
- Microsoft Excel 2010
Methods: Author
Results: counts

- word types: 9,026
- word tokens: 46,608
- average “lines” (paragraphy): 6.0
- average words: 80.7
- average chars: 490.0
- median “lines” (paragraphy): 3.0
- median words: 49.0
- median chars: 309.5
## Results: counts

<table>
<thead>
<tr>
<th></th>
<th>G+</th>
<th>TXT (Baron 2008)</th>
<th>IM (Baron 2008)</th>
</tr>
</thead>
<tbody>
<tr>
<td>post</td>
<td>560 (14*40)</td>
<td>191</td>
<td>191</td>
</tr>
<tr>
<td>total words</td>
<td>46,608</td>
<td>1,473</td>
<td>1,146</td>
</tr>
<tr>
<td>average words/tx</td>
<td>80.7</td>
<td>7.7</td>
<td>6.0</td>
</tr>
</tbody>
</table>
Results: features

- emoticons - :D ^_^ etc.
- acroynms - DMCA POV FDA FBI LTE BFF etc.
- abb. - c'mon, specs, admin etc.
- contractions - 'em here/there's, don't etc.
- punctuation - ! ? ... etc.
- ALL CAPS - INGENIOUS SO NOT EVER etc.
## Results: features

<table>
<thead>
<tr>
<th></th>
<th>G+</th>
<th>TXT (Baron 2008)</th>
<th>IM (Baron 2008)</th>
</tr>
</thead>
<tbody>
<tr>
<td>of posts</td>
<td></td>
<td>of words</td>
<td>of words</td>
</tr>
<tr>
<td>emoticons</td>
<td>13%</td>
<td>&lt; 1%</td>
<td>&lt; 1%</td>
</tr>
<tr>
<td>acroynms</td>
<td>23%</td>
<td>&lt; 1%</td>
<td>&lt; 1%</td>
</tr>
<tr>
<td>abb.</td>
<td>8%</td>
<td>3%</td>
<td>0%</td>
</tr>
<tr>
<td>contractions</td>
<td>33%</td>
<td>85%</td>
<td>68%</td>
</tr>
<tr>
<td>apostrophes</td>
<td></td>
<td>32%</td>
<td>94%</td>
</tr>
<tr>
<td>punctuation</td>
<td>58%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALL CAPS</td>
<td>11%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>of sentences</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>overall sentence punctuation</td>
<td>39%</td>
<td>45%</td>
<td></td>
</tr>
<tr>
<td>transmission-final punctuation</td>
<td>29%</td>
<td>35%</td>
<td></td>
</tr>
<tr>
<td>transmission-internal punctuation</td>
<td>54%</td>
<td>78%</td>
<td></td>
</tr>
<tr>
<td>use of required question mark</td>
<td>73%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>use of required period</td>
<td>30%</td>
<td>41%</td>
<td></td>
</tr>
</tbody>
</table>
## Results: authors

<table>
<thead>
<tr>
<th>user ID</th>
<th># of posts</th>
<th>category</th>
<th>in # of circles</th>
<th># of twitter followers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mike Shaw</td>
<td>12</td>
<td>photographer</td>
<td>1,076,596</td>
<td>N/A</td>
</tr>
<tr>
<td>Trey Ratcliff</td>
<td>10</td>
<td>photographer</td>
<td>1,866,504</td>
<td>32,380</td>
</tr>
<tr>
<td>Mike Elgan</td>
<td>9</td>
<td>executive</td>
<td>969,232</td>
<td>18,895</td>
</tr>
<tr>
<td>Alex Koloskov</td>
<td>8</td>
<td>photographer</td>
<td>1,060,812</td>
<td>6,012</td>
</tr>
<tr>
<td>Cadbury UK</td>
<td>8</td>
<td>company</td>
<td>853,744</td>
<td>11,700</td>
</tr>
<tr>
<td>Tom Anderson</td>
<td>8</td>
<td>executive</td>
<td>1,834,245</td>
<td>192,393</td>
</tr>
<tr>
<td>Amanda Blain</td>
<td>7</td>
<td>executive</td>
<td>971,718</td>
<td>15,846</td>
</tr>
<tr>
<td>Felicia Day</td>
<td>6</td>
<td>celeb</td>
<td>1,910,120</td>
<td>1,896,041</td>
</tr>
<tr>
<td>IGN</td>
<td>6</td>
<td>game</td>
<td>380,590</td>
<td>409,456</td>
</tr>
<tr>
<td>The Hunger Games</td>
<td>6</td>
<td>entertainment</td>
<td>967,682</td>
<td>434,632</td>
</tr>
<tr>
<td>BBC News</td>
<td>5</td>
<td>news</td>
<td>384,418</td>
<td>653,497</td>
</tr>
<tr>
<td>Hugh Jackman</td>
<td>5</td>
<td>celeb</td>
<td>1,938,792</td>
<td>1,235,457</td>
</tr>
<tr>
<td>Jay Patel</td>
<td>5</td>
<td>photographer</td>
<td>1,104,512</td>
<td>N/A</td>
</tr>
<tr>
<td>Jim Goldstein</td>
<td>5</td>
<td>photographer</td>
<td>1,083,729</td>
<td>10,944</td>
</tr>
<tr>
<td>Mashable</td>
<td>5</td>
<td>website</td>
<td>353,093</td>
<td>108,346</td>
</tr>
<tr>
<td>The Next Web</td>
<td>5</td>
<td>website</td>
<td>506,584</td>
<td>695,842</td>
</tr>
<tr>
<td>Victor Bezukov</td>
<td>5</td>
<td>photographer</td>
<td>1,111,814</td>
<td>2,352</td>
</tr>
<tr>
<td>Angry Birds</td>
<td>4</td>
<td>game</td>
<td>993,226</td>
<td>415,885</td>
</tr>
</tbody>
</table>
Discussion - situation

• public posts, target audience
• "on stage"
Discussion - real authors?

- male/female/company?
- demographics?
Discussion – thousand words

• heavy on media sharing – images, videos
• One liners - context
Discussion - emoticons

• emoticons as graphics elements
## Discussion – 2008+

- **textism more common?**

<table>
<thead>
<tr>
<th></th>
<th>G+ of posts</th>
<th>TXT (Baron 2008) of words</th>
<th>IM (Baron 2008) of words</th>
</tr>
</thead>
<tbody>
<tr>
<td>emoticons</td>
<td>13%</td>
<td>&lt; 1%</td>
<td>&lt; 1%</td>
</tr>
<tr>
<td>acroynms</td>
<td>23%</td>
<td>&lt; 1%</td>
<td>&lt; 1%</td>
</tr>
<tr>
<td>abb.</td>
<td>8%</td>
<td>3%</td>
<td>0%</td>
</tr>
<tr>
<td>contractions</td>
<td>33%</td>
<td>85%</td>
<td>68%</td>
</tr>
<tr>
<td>apostrophes</td>
<td></td>
<td>32%</td>
<td>94%</td>
</tr>
<tr>
<td>punctuation</td>
<td>58%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALL CAPS</td>
<td>11%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Discussion - further analysis

- Tagliamont and Dennis (2008) break out of characteristic IM forms - frequency and proportion of total word count, distribution of first-person singular pronouns, laughter variants etc.

- Rosen (2010) linguistic textisms vs contextual textisms, quantitative statistics - std. dev., p-value etc.
Discussion - future research

- compare with Facebook, LinkedIn, twitter for “professional”, post length limit effects
- almost no foreign language posts (at least not in the one we can access), could content be different in other locales?
Discussion - future research

- including other users comments
- trending topics
References

- Wikipedia Google+
- Vasile, Gabriel Google+ number of users
  https://plus.google.com/10639347869568433143/posts/QkhjpiEEOFm
- Google+ Update: See What Posts are Hot on Google+
  http://www.youtube.com/watch?v=xcgCtYOlxhk
- Volume Control
  Baron, Naomi S. 2008 Always On - Chapter 3: Controlling the Volume - Everyone a Language Czar
  Vaughan-Nichols, Steven J. Google+’s best feature: The power to shut fools up 2012/04/19
  http://www.zdnet.com/blog/networking/google-s-best-feature-the-power-to-shut-fools-up/2247
- Yu, Roger More companies quit blogging, go with Facebook instead 2012/04/21
  http://www.usatoday.com/tech/news/story/2012-04-19/corporate-blogging/54419982/1
- Antconc
  http://www.antlab.sci.waseda.ac.jp/software.html#antconc
- Cygwin
  http://www.cygwin.com/
- Baron, Naomi S. 2008 Always On: Language in an Online and Mobile World - Chapter 7, Another Pilot Study - location 2975, 42% Kindle format
- Baron, Naomi S. 2008 Always On: Language in an Online and Mobile World - Chapter 5 - Sleeping ... or am i :) away message in IM location 1445, 20% Kindle format