

Analysing Text Virality in Social Media

George Markopoulos
Department of Linguistics
School of Philosophy

Topics

- Language and Communication in Social Media
- Text Virality and Discourse Analysis
- Sociolinguistics of Virality
- Machine Learning and Virality

Internet Linguistics

- Netspeak: a radically new linguistic medium (Crystal 2001)
- Huge growth of the range and variety of language.
- Interactivity and personal creativity
- Computer-mediated communication and information exchange
- Social media: verbal, nonverbal and audiovisual cues, Internet memes

Internet English

- Changes in spelling and grammar
- Widespread use of memes, acronyms and abbreviations
- Emotive punctuation
- Case folding (e.g. use of capitals for emphasis)
- Blends (*netizen, cybercide, bugzilla*, etc.)
- Unconventional use of language in digital communication

Virality and Content

- Relationship between virality and the content of messages
- Strong connection between virality of content and its nature (Strapparava et al. 2011)
- Affect and emotion are crucial factors for successful viral diffusion
- Positive sentiment encourages retweeting
- Negative sentiment propagates better news related content (Hansen et al. 2011)

Discourse Analysis of Virality

- *Expressive* and *persuasive* discourse functions influence social media virality (Kinneavy 1969)
- Virality benefits various political practices:
 - Decision-making
 - Volunteer activation
 - Petition signing
 - Crowdfunding

Discourse Analysis of Virality: Linguistic Means

Linguistic means that foster virality:

- Choice of informal register as a result of meticulous audience design - accommodation to the target group's needs
- Conversationalization of public discourse and construction of a public colloquial language
- Selection of pronouns creates a sharp distinction between in-groups and out-groups: "us" vs "them", collective "we", preference for "you"
- Modality as a linguistic device indicates the degree to which an observation is possible, probable, certain etc.

Discourse Analysis of Virality: Linguistic Means (*cont.*)

Linguistic means that foster virality:

- Use of storytelling and conversational narratives (especially personal stories and life anecdotes)
- Humour and jokes creating positive atmosphere (closeness and intimacy) or targeting the opponent (gossip, rumor spreading)
- Metaphor and other rhetorical figures (such as the rule of three as a writing principle, parallelism, repetition etc.)
- Use of ethically-loaded appeals as persuasion mechanism

Language and Gender in Discourse

- Women and men express emotions differently and thus use language differently
- Female writing style: more contextual, personal and emotional
- Male writing style: more impersonal, formal and judgmental
- Gender-biased variety of text genres ranging from online personal diaries to fiction and non-fiction novels
- The gender information could prompt a political candidate to redefine the targets of political campaigns, to shape and maintain his/her public image

Pitfalls

Obstructions of the message's diffusion:

- *Homophily principle* (McPherson et al 2001), i.e. the tendency to follow others with the same number of followers as yourself
- Creation of *echo chambers* as a direct result of homophily (Colleoni 2014)
- *Power law* distribution (e.g. Wikipedia: in average 5% of people contribute to 80% of its content)
- Adorno's notion of *pseudoindividualization* (= false intimacy): fine line between self-promotion and self-absorption/narcissism

Sociolinguistics of Virality: The emergence of "cool"

- A new sociolinguistic realism: the non-linearity of meaning in the case of "cool" (Blommaert 2015)
- Meaning as an outcome based on different facets of a sociolinguistic system in which whole communities appear as actors, rather than individual users.
- Another example: The very notion of sexuality overaccumulates and virally proliferates in the form of linguistic and visual signification, e.g. "travel porn", "food porn"

Machine Learning and Virality

- In the age of Big Data and AI, machine generated content could also become viral
- Unsupervised Deep Learning language models achieve human level performance in natural language generation and understanding
- Fake news and the rise of fact-checking algorithms indicate symptoms of shifting paradigms in post-truth societies
- Dramatic change in information dissemination and information consumption

References

- J. Blommaert. *Meaning as a nonlinear effect: The birth of cool*. Theory in Applied Linguistics Research, 28, 2015.
- E. Colleoni. *Echo Chamber or Public Sphere? Predicting Political Orientation and Measuring Political Homophily in Twitter Using Big Data*. Journal of Communication, 64, 2014.
- D. Crystal. *Language and the Internet*. Cambridge University Press, 2001.
- L. Hansen, A. Arvidsson, F. Nielsen, E. Colleoni, and M. Etter. *Good Friends, Bad News - Affect and Virality in Twitter*. Communications in Computer and Information Science, 185, 2011.
- J. Kinneavy. *The Basic Aims of Discourse*. College Composition and Communication, 20, 1969.
- V. Miller. *Phatic culture and the status quo: Reconsidering the purpose of social media activism*. Convergence, 23, 2015.
- C. Strapparava, M. Guerini, and G. Özbal. *Persuasive Language and Virality in Social Networks*. In Proc. 4th International Conference on Affective Computing and Intelligent Interaction (ACII), 2011.