

# Anonymous Free Speech

Carl Vogel

Computational Linguistics Group  
Centre for Computing and Language Studies  
School of Computer Science and Statistics  
Trinity College Dublin  
Dublin, Ireland  
DIMPLE – SLANDAIL

May 31, 2014

The Plan:

- 1 What modes of communication are prerequisite to trust?
- 2 Is anonymity inherently asocial?
- 3 How do disgust, offence and politeness relate?
- 4 What is the role of anonymity in offence management?

# Conclusions

- Anonymity in communication has some positive consequences.
- Fully anonymous messages may be achieved in a medium that subverts both identification of individuals and development of public authorial personae.
- Trust may associate to fully anonymous messages.

# Modes of communication

- Nonymity
- Persononymity
- Anonymity

# Demands on communications during disasters

- Correctness
- Completeness
- Trustworthiness
  - vested interest
  - accountability

- Evaluated as attached to personae
- Politeness matters
  - epistemic hedges
  - indirectness
- Friendliness matters (emoticons)

(Vogel & Mamani Sanchez, 2012; Mamani Sanchez & Vogel, 2012, 2013)

# Anonymity: social, asocial or antisocial?

- Negatives

- Greater aggression in race conflict  
(Donnerstein, Donnerstein, Simon, & Ditrachs, 1972)
- Greater aggression in road rage  
(Ellison-Potter, Bell, & Deffenbacher, 2001)
- Diminished (25%) generocity  
(Alpizar, Carlsson, & Johansson-Stenman, 2008)
- Diminished rule maintenance  
(Nogami & Takai, 2008)
- Diminished positive feedback in performance evaluations  
(Stone, Spool, & Rabinowitz, 1977)

# Anonymity: social, asocial or antisocial?

- Positives
  - No effect on aggression where aggression was established (Diener, 1976)
  - Less pronounced economic punishment behaviors (Connolly, Jessup, & Valacich, 1990)
  - Greater conformance to group norms (Postmes, Spears, Sakhel, & de Groot, 2001; Smith, Terry, & Hogg, 2007)
  - Greater medical self reporting (Durant, Carey, & Schroder, 2002)
  - Most prolific idea generation condition (Connolly et al., 1990)

- Face-work: (Goffman, 1967; Brown & Levinson, 1987)
  - Power
  - Distance
  - Rank of imposition
- Relational work:  
(Watts, 2003; Bousfield & Locher, 2008; Culpeper, 2008)
- Offence/Disgust Management: (Vogel, 2014a, 2014b)
  - Politeness: minimize disgust
  - Impoliteness: maximize disgust



# The power of disgust avoidance

Fig. 1: A superlatively upsetting image

This is a picture of the  
most disgusting thing  
you can imagine.

# Disgust (Kolnai, 1929a, 1929b; Kelly, 2011)

- The disgust response
  - Universal oral/nasal nausea-symptathetic response
  - Involuntary
  - Learned instantly
  - Irreversible
  - Contagious
- Disgust triggers
  - toxins
  - decay & fecundity
  - contagion yields culture-specificity

# Disgust, Politeness and Anonymity

- Expect politeness:  
 $offensiveness_{addresser}(Addresser) > offensiveness_{addresser}(Addressee)$
- Expect impoliteness:  
 $offensiveness_{addresser}(Addresser) < offensiveness_{addresser}(Addressee)$
- Expect politic behavior:  
 $offensiveness_{addresser}(Addresser) = offensiveness_{addresser}(Addressee)$
- With possibility of anonymity, disgust not attached to individuals, nor personae, expect politic behavior
- Therefore, provide for communication of ideas through a system that thwarts even personae development: FreeSpeech

## Closing remarks

- FreeSpeech is a platform that supports anonymous posting of ideas.
- Posts are anonymous – a history of posting by an individual cannot be developed unless a poster signs messages.
- Anonymous posting of messages does not lead inevitably to defamation or offensiveness.
- Anonymity in communication has some positive consequences
- Fully anonymous messages may be achieved in a medium that subverts both identification of individuals and development of public authorial personae
- Trust may associate to fully anonymous messages
- (Used as a sandbox for text classification and to be used to study posting behaviors.)

# Closing remarks

- FreeSpeech is a platform that supports anonymous posting of ideas.
- Posts are anonymous – a history of posting by an individual cannot be developed unless a poster signs messages.
- Anonymous posting of messages does not lead inevitably to defamation or offensiveness.
- Anonymity in communication has some positive consequences
- Fully anonymous messages may be achieved in a medium that subverts both identification of individuals and development of public authorial personae
- Trust may associate to fully anonymous messages
- (Used as a sandbox for text classification and to be used to study posting behaviors.)
- **Thank you for your time and attention!**

# Closing remarks

- FreeSpeech is a platform that supports anonymous posting of ideas.
- Posts are anonymous – a history of posting by an individual cannot be developed unless a poster signs messages.
- Anonymous posting of messages does not lead inevitably to defamation or offensiveness.
- Anonymity in communication has some positive consequences
- Fully anonymous messages may be achieved in a medium that subverts both identification of individuals and development of public authorial personae
- Trust may associate to fully anonymous messages
- (Used as a sandbox for text classification and to be used to study posting behaviors.)
- **Thank you for your time and attention!**
- **Thanks to the EC FP7 (Grant Agreement Number: 607691)!**

# References I

- Alpizar, F., Carlsson, F., & Johansson-Stenman, O. (2008). Anonymity, reciprocity, and conformity: evidence from voluntary contributions to a national park in costa rica. *Journal of Public Economics*, 92(56), 1047 – 1060.
- Bousfield, D. & Locher, M. (Eds.). (2008). *Impoliteness in Language: Studies on its Interplay with Power in Theory and Practice*. Berlin: Mouton de Gruyter.
- Brown, P. & Levinson, S. (1987). *Politeness: Some Universals in Language Usage*. Cambridge University Press.
- Connolly, T., Jessup, L. M., & Valacich, J. S. (1990). Effects of anonymity and evaluative tone on idea generation in computer-mediated groups. *Management Science*, 36(6), 689–703.
- Culpeper, J. (2008). Reflections on impoliteness, relational work and power. In Bousfield, D. & Locher, M. (Eds.), *Impoliteness in Language: Studies on its Interplay with Power in Theory and Practice*, pp. 17–44. Berlin: Mouton de Gruyter.
- Diener, E. (1976). Effects of prior destructive behavior, anonymity, and group presence on deindividuation and aggression.. *Journal of Personality and Social Psychology*, 33(5), 497 – 507.
- Donnerstein, E., Donnerstein, M., Simon, S., & Ditrichs, R. (1972). Variables in interracial aggression: anonymity, expected retaliation, and a riot.. *Journal of Personality and Social Psychology*, 22(2), 236 – 245.
- Durant, L. E., Carey, M. P., & Schroder, K. E. (2002). Effects of anonymity, gender, and erotophilia on the quality of data obtained from self-reports of socially sensitive behaviors. *Journal of Behavioral Medicine*, 25(5), 439–467.
- Ellison-Potter, P., Bell, P., & Deffenbacher, J. (2001). The effects of trait driving anger, anonymity, and aggressive stimuli on aggressive driving behavior. *Journal of Applied Social Psychology*, 31(2), 431–443.
- Goffman, E. (1967). On face-work. In *Interaction Ritual: Essays in Face-to-Face Behavior*, pp. 5–45. New Brunswick, NJ: Transaction Publishers. Reprinted 2008.
- Kelly, D. (2011). *Yuck! The Nature and Moral Significance of Disgust*. MIT.
- Kolnai, A. (2004 (1929)a). Disgust. In Smith, B. & Korsmeyer, C. (Eds.), *On Disgust*, pp. 29–91. Open Court Press: Chicago. Translated by Elisabeth Gombrich, Elizabeth Kolnai and Barry Smith.
- Kolnai, A. (2004 (1929)b). The standard modes of aversion: fear, disgust and hatred. In Smith, B. & Korsmeyer, C. (Eds.), *On Disgust*, pp. 93–108. Open Court Press: Chicago. Translated by Elisabeth Gombrich, Elizabeth Kolnai and Barry Smith.

# References II

- Mamani Sanchez, L. & Vogel, C. (2012). Emoticons signal expertise in technical web fora. In Esposito, A., Squartini, S., & Palm, G. (Eds.), *Computational Intelligence in Emotional or Affective Systems*. WIRN2012: Special Session Post-Proceedings (in press). Springer LNCS.
- Mamani Sanchez, L. & Vogel, C. (2013). Imho: an exploratory study of hedging in web forums. In *Proceedings of the SIGDIAL 2013 Conference*, pp. 309–313 Metz, France. Association for Computational Linguistics.
- Nogami, T. & Takai, J. (2008). Effects of anonymity on antisocial behavior committed by individuals. *Psychological Reports*, 102, 119–130.
- Postmes, T., Spears, R., Sakhel, K., & de Groot, D. (2001). Social influence in computer-mediated communication: the effects of anonymity on group behavior. *Personality and Social Psychology Bulletin*, 27(10), 1243–1254.
- Smith, J. R., Terry, D. J., & Hogg, M. A. (2007). Social identity and the attitudebehaviour relationship: effects of anonymity and accountability. *European Journal of Social Psychology*, 37(2), 239–257.
- Stone, E. F., Spool, M. D., & Rabinowitz, S. (1977). Effects of anonymity and retaliatory potential on student evaluations of faculty performance. *Research in Higher Education*, 6(4), 313–325.
- Vogel, C. (2014a). Denoting offence. *Cognitive Computation*. to appear.
- Vogel, C. (2014b). Intended and unintended offence. In Bassis, S., Esposito, A., & Morabito, F. C. (Eds.), *Recent Advances of Neural Network Models and Applications*, Smart Innovation, Systems and Technologies Volume 26, pp. 395–404. Springer International Publishing.
- Vogel, C. & Mamani Sanchez, L. (2012). Epistemic signals and emoticons affect kudos. In Baranyi, P. (Ed.), *3rd IEEE Conference on Cognitive Infocommunications*, pp. 517–522. [Best Paper Award]; Track VI: *Communicative Social Signals: Computational and Behavioural Aspects of Human-Human and Human-Machine Interaction*. Anna Esposito, Klara Vicsi, Alessandro Vinciarelli (Sub-Eds.).
- Watts, R. J. (2003). *Politeness*. Cambridge University Press.