Personality Driven Differences in Paraphrase Preference

Daniel Preoţiuc-Pietro
Bloomberg LP

Joint work with
Jordan Carpenter (Kenan Institute for Ethics, Duke)
Lyle Ungar (Computer & Information Science, UPenn)
while at the University of Pennsylvania

3 August 2017

Penn | World Well-Being Project

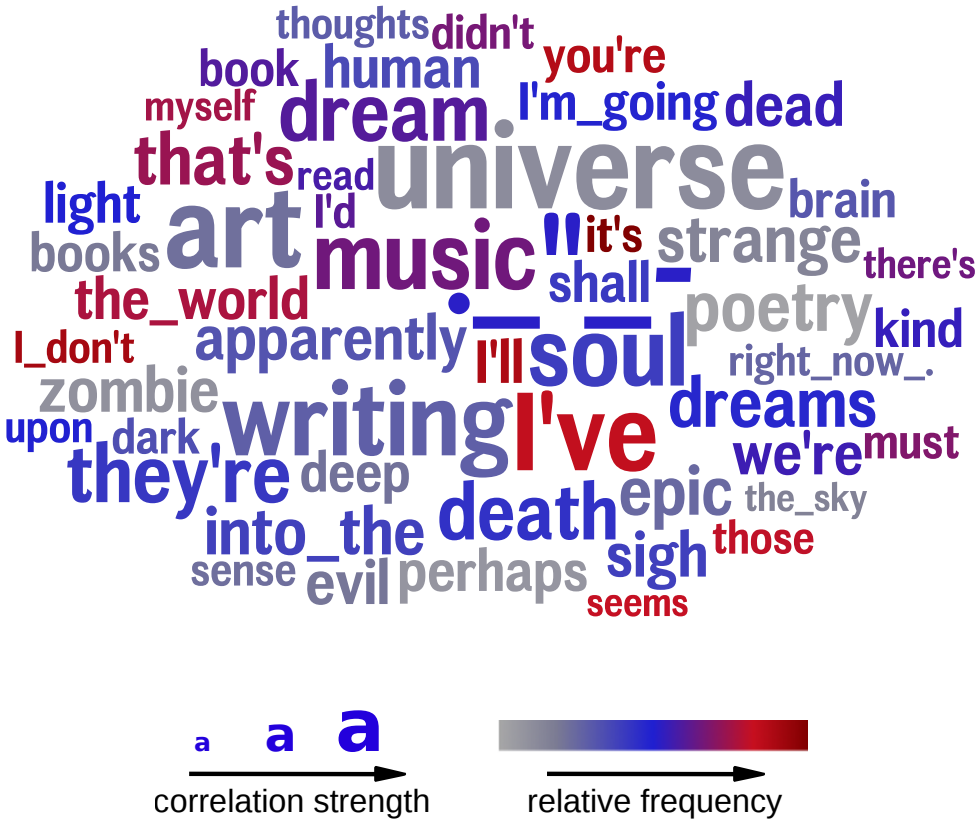
Motivation

User attribute prediction from text is successful:

- Age (Rao et al. 2010 ACL)
- Gender (Burger et al. 2011 EMNLP)
- Location (Eisenstein et al. 2010 EMNLP)
- Personality (Schwartz et al. 2013 PLoS One)
- Impact (Lampos et al. 2014 EACL)
- Political Orientation (Volkova et al. 2014 ACL)
- Mental Illness (Coppersmith et al. 2014 ACL)
- Occupation (Preoţiuc-Pietro et al. 2015 ACL)
- Income (Preoţiuc-Pietro et al. 2015 PLoS One)

However...

Most text prediction methods uncover **topical** differences



Openness to Experience

However...

Most text prediction methods uncover **topical** differences

Extraversion

Stylistic differences

We need to be aware of **style** differences, rather than topical

Not useful for many practical applications that adapt to traits:

- machine translation (Mirkin et al. 2015 EMNLP, Rabinovich et al 2017 EACL)
- agents (e.g. customer service, tutoring)
- controlling for gender or racial bias

Stylistic differences

One type of stylistic difference is **phrase choice** in context.

Splendid
Excellent
Remarkable

Openness

Magnificent
Fabulous
Tremendous

Extraversion

Data

We study the Big Five personality traits:

- 115,312 Facebook users
- Personality scores obtained through the MyPersonality app (Kosinski et al, 2013)
- For each trait, take top and bottom 20% of users

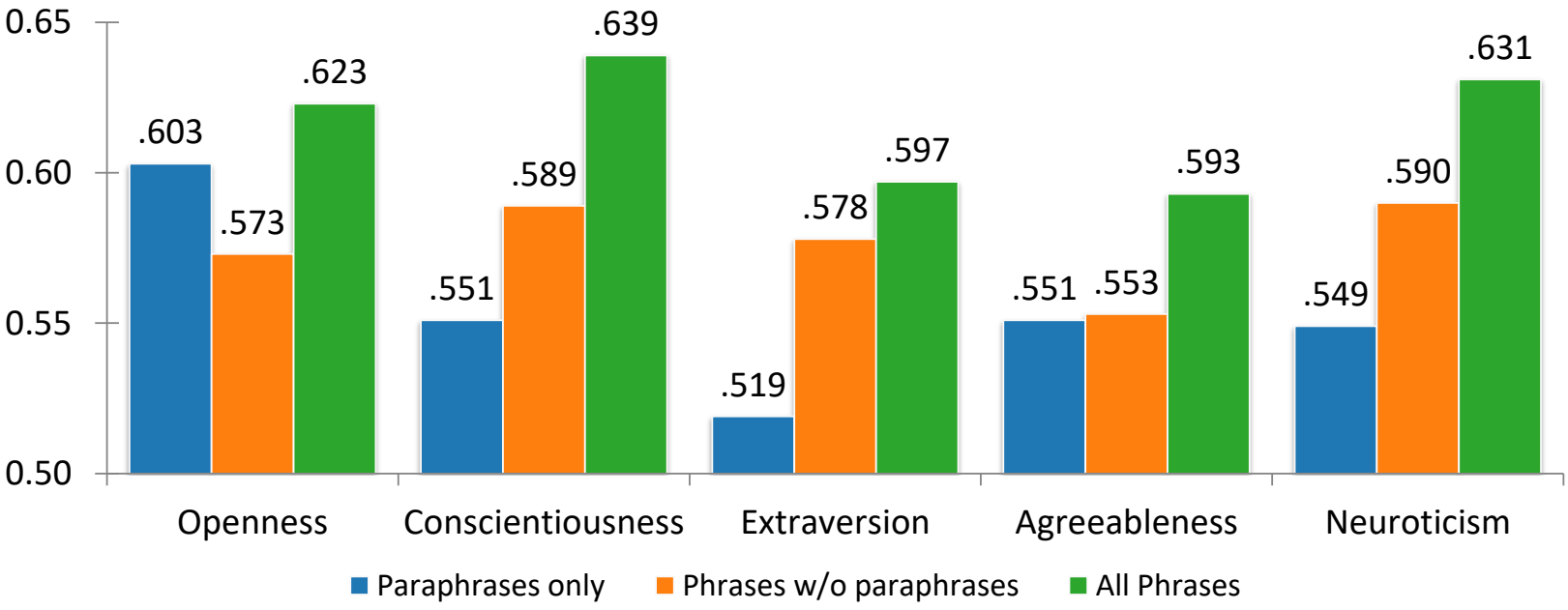
Paraphrasing

Paraphrases – alternative ways to convey the same information

Paraphrase Database (PPDB) 2.0 (Pavlick et al. 2015 ACL):

- annotated with type and confidence (filter 'equivalent' paraphrases with >.2 confidence)
- >6M automatically derived paraphrase pairs
- we use only 1–3 grams
- difference in a pair more than just change of stopwords or root form of word

Prediction



Accuracy, Naive Bayes, 90-10 training-testing, balanced data

Quantifying Preference

Straightforward measure:

$$\text{Extraversion}(w) = \log\left(\frac{\text{Extravert}(w)}{\text{Introvert}(w)}\right) \tag{1}$$

Within a paraphrase pair (w_1, w_2), the difference $\text{Extraversion}(w_1) - \text{Extraversion}(w_2)$ is the stylistic distance.

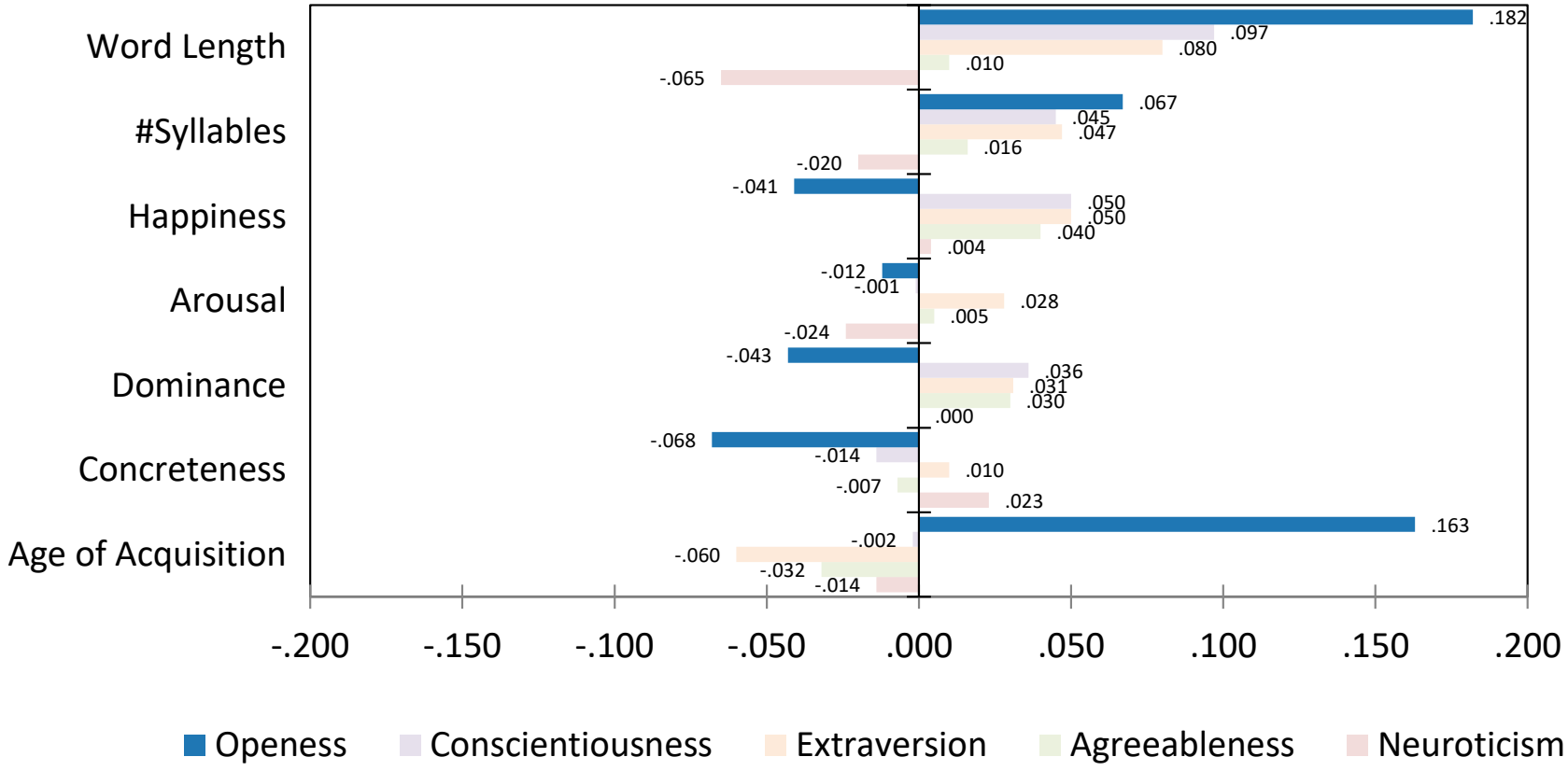
Used previously to study paraphrase preference across age, gender and occupational class (Preoţiuc-Pietro, Xu & Ungar, AAAI 2016).

Linguistic Theories

Study which attributes of words in a pair are preferred by one group:

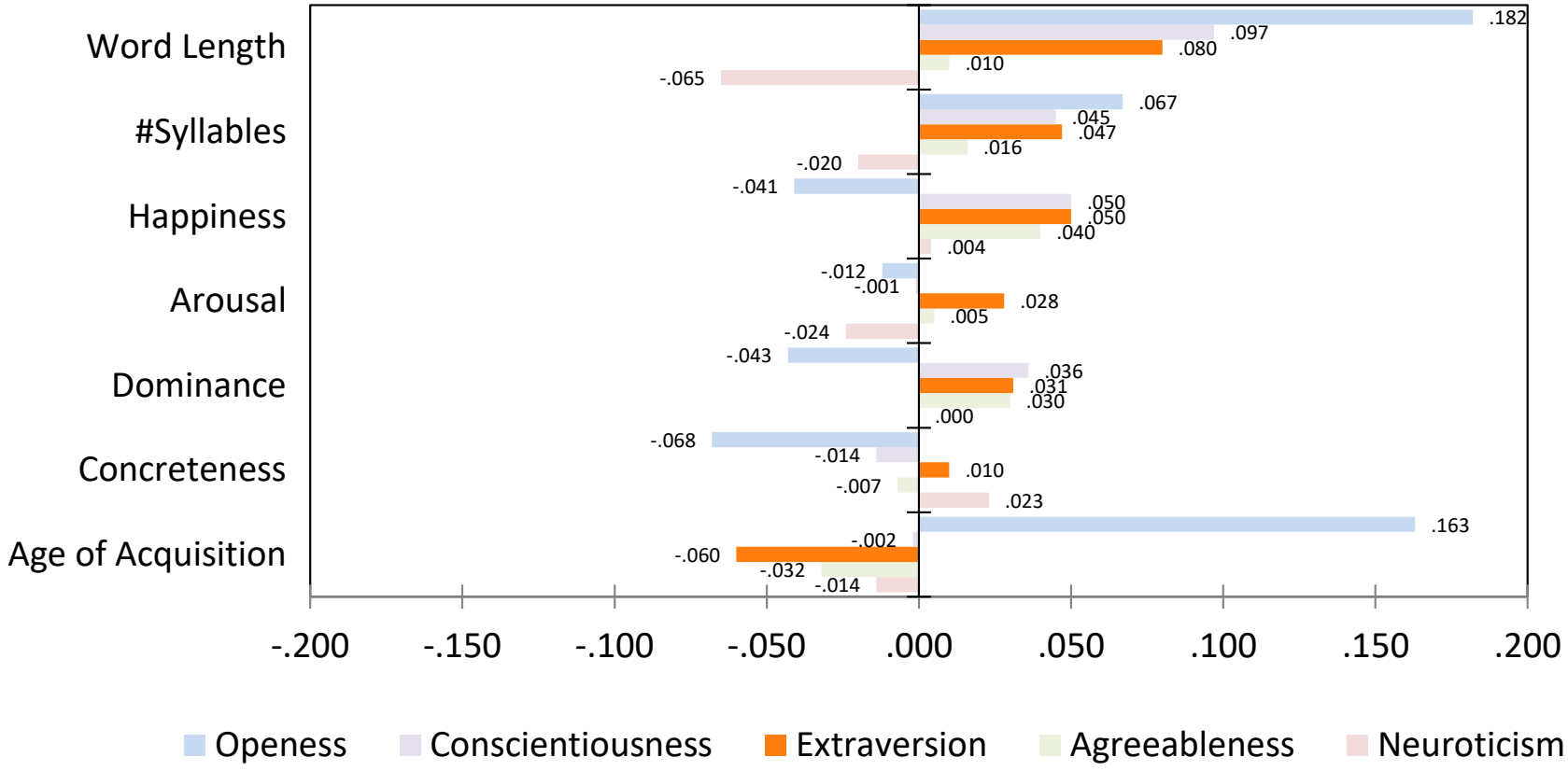
- Word Length in Characters
- Word Length in Syllables
  Simple proxies for word complexity
- Affective Norms: Valence, Arousal, Dominance
  14k rated words
  Valence: suicide (0.15) → bacon (0.70) → laughter (1)
- Concreteness
  40k rated words: spirituality (1) → morning (3.44) → tiger (5)
- Age of Acquisition
  30k rated words: great (5.05) → splendid (7.22) → tremendous (10.63)
- More in the paper ...

Linguistic Theories



Correlation coefficients between paraphrase pair preference and user group usage.

Linguistic Theories



Correlation coefficients between paraphrase pair preference and user group usage.

Take Aways

- Stylistic difference between user groups have important applicability
- Paraphrase choice contains valuable information
- Shed light on psycholinguistic theories
- Potential way to generate text perceived to be from a different user trait

See our EMNLP 2017 paper (Preoţiuc-Pietro, Guntuku, Ungar - Controlling Human Perception of Basic User Traits)

Thank you!

Questions?