Beyond Sentiment: Social Psychological Analysis of Political Facebook Comments in Hungary

Márton Miháltz[1], Tamás Váradi[1], István Csertő[2], Éva Fülöp[2], Tibor Pólya[2], Pál Kővágó[3]

[1]Research Institute for Linguistics, Hungarian Academy of Sciences
[2]Institute of Cognitive Neuroscience And Psychology, Hungarian Academy of Sciences
[3]Institute of Psychology, University of Pécs

Motivation

- **Social Media content analysis to assess political attitudes**
- **Current approaches mainly focus on sentiment analysis only** (*Ceron et al. 2014; Chen et al. 2010; Costa et al. 2015; Hammer et al. 2014; O'Connor et al. 2010; Sobkowicz, Kaschesky and Bouchard 2012; Tumasjan et al. 2010*)
- **Introduce new, social psychology-motivated measures**
- **Use Facebook: dominant SM in Hungary**
  4.25M users, 59.2% penetration (of people with internet access)
  Analyze public comments for public posts on political pages
- **Examine correlation with traditional opinion poll results**

Data

3 major election events in Hungary in 2014
(Hungarian Parliament, European Parliament and municipal elections)
1341 Facebook pages:

- Members of parliament
- Election candidates
- Political organizations and subbranches
- Official and fan pages

Total 1.9M public comments from 226K users for 141K posts

- 46M tokens (running words)
- Download via Facebook Graph API, weekly
- October 2013-September 2014

Processing pipeline

- Tokenization, segmentation
- PoS-tagging, lemmatization
- Named entity recognition: political actors (persons, organizations + party affiliations)
- Sentiment and social psychology analysis (FSAs)

NLP for Social Media

Social media text is different

- Typos, spelling errors
- non-standard punctuation
- Slang words
- Emoticons, creative use of characters
- Unaccented Hungarian text

Investigation corpus

- 1.2M comments, 29M words
- Processing with vanilla NLP tools
- Unknown tokens f>=15: 14,000 types
  Manual analysis: common, regular problem types, lists of unknown frequent and important words

2-fold approach:

- Normalize input to standard language
  Pre- & postprocessing, normalization lists
- Adapt tools to SM language
  Add unknown words to lemmatizer's lexicon

Analysis

Sentiment + 4 new indicators
Building on *narrative psychology* and *social psychology* research
Custom lexicons and grammars
Compiled into *NooJ* finite state automatons

Sentiment

- 500 positive, 420 negative entries (content words, multiwords, emoticons)
- negation rules
- Score: (n_positive - n_negative) / n_tokens

Agency & Communion

- 2 dimensions in social value judgements (*Abele and Wojciszke 2007; Abele et al. 2008*)
- **Agency**: describes an individual in terms of the efficiency of their behavior oriented to their personal goals: *motivation, competence, control*
- **Communion**: describes the moral and emotional aspects of an individual's relations to other group members, individuals or groups: *cooperation, social benefit, honesty, self-sacrifice, affection, friendship, respect, love etc.*
- Both with positive-negative values
- Lexicon: 650 words and multiwords
- Scores: (n_positive - n_negative) / n_tokens

Optimism-pessimism

- Time of events plays a role in individual thinking (*Habermas et al. 2008; Kunda 1999*)
  - dominated by **past**: they view the world unchangeable
  - dominated by **present**: importance of realistically attainable goals
  - **future**-dominated thinking: sees open possibilities
- Based on PoS and morphology annotations + time expressions
- 2 measures for degree of optimism:
  - present_verbs / past_verbs
  - future_verbs / (present_verbs + past_verbs)

Individualism-collectivism

- Individualism: importance of the category of the self when thinking about the world
  - Individualistic: focus on individuals' actions
  - Collectivist: focus on actions of groups
- Correlation between usage/omission of personal pronouns (**pronoun drop**) and levels of individualism in societies (*Kashima and Kashima 1999*)
- Extended to measure individualism/collectivism in groups
  - Frequent use of personal pronouns: high level of individualism
  - Pronoun drop (verbs and nouns with personal inflections): lower level
- Measure: pers_pronouns / (verbs_with_inflection + nouns_with_inflection)

Development of lexicons

- Corpus: 176K comments, 5.45M words
- 3500 most frequent words (f>=100): coding for categories by 6 annotators

Evaluation

Gold Standard:

- 1008 comments from all political parties with same distribution as in complete corpus (*FIDESZ-KDNP 25.2%, EGYÜTT-2014 19.3%, JOBBIK 19.2%, MSZP 16.6%, DK 12.5%, PM 4.2%, LMP 2.9%*)
- 3 human annotators

Annotation Type	Precision	Recall	F1
Named entities	98.38	57.14	72.29
Sentiment: positive	82.56	74.50	77.38
Sentiment: negative	67.03	53.68	59.62
Agency: positive	70.59	69.43	52.83
Agency: negative	65.79	25.51	36.76
Communion: positive	65.75	38.40	48.48
Communion: negative	96.39	41.45	57.97
Individualism: pers. pron.	35.20	65.63	45.82
Individualism: inflections	77.27	94.74	85.12
Optimism: past	78.90	93.97	85.78
Optimism: present	31.40	92.54	46.88
Optimism: future	32.80	67.03	44.04

Comparison with Poll Data

How well new measures indicate changes in political attitude during Hungarian parliamentary elections in April 2014?

- Traditional public opinion poll data from *Tárki*
- Facebook data: 1.9M comments 6+6 months before/after elections, scores aggregated monthly for each party's pages

Individualism, Optimism

- **Individualism correlates with party popularity over 12 months** (r=.22, p=.052)
  Higher individualism - higher responsibility for party choices - higher party popularity
- **Individualism increased after elections**
  Decline of significance of cooperation and unity
- **Optimism increased after elections only for winning parties**
  Different experiences of success and failure

Agency & Communion

- **Expected: negative correlation between both positive agency & negative communion and party popularity**
  - *Intergroup bias: overrate in-group & devalue out-group in intergroup competence or conflict*
  - *Judge in-group through agency, out-group through communion*
- 6 m. before elections, correlation w/ party popularity:
  - **Negative correlation for positive agency** (r=-.429, p=.05)
  - **Negative correlation for agency score** (r=-.677, p=.05)
- 6 m. after elections, correlation w/ party popularity:
  - **Negative correlation for negative communion** (r=-.574, p=.01)
  - **Negative correlation for communion score** (r=-454, p=.05)
- Averages for all parties:
  - Pos. agency > neg. agency for 12 m. (p=.001)
  - Neg. communion > pos. communion for 12 m. (p=.001)
  - Pos. agency decreased after elec. (p=.01)

LaTeX TikZposter