Lexical Features of Sarcasm in Twitter and the Google Books Database

Introduction

- Sarcasm: figurative language where the intended meaning is the opposite of the literal meaning.
- Sarcasm has primarily been studied in terms of pragmatics (e.g., Kreuz, 1996), but lexical features may serve as important discriminating cues.
- An obstacle to studying sarcasm is determining sarcastic intent. One solution is to use corpora such as Twitter (Davidov et al., 2010; González-Ibáñez et al., 2011) and books (Kreuz & Caucci, 2007) where sarcastic intent can be explicitly marked by authors.

Goal: Extend previous work by comparing specific lexical features of sarcastic and non-sarcastic (control) statements in corpora where sarcastic intent is explicitly marked.

Hypothesis: Sarcasm statements will differ from control statements on features that reflect asymmetry of affect, hyperbole, and use of interjections.

Corpora

- **Twitter**
  - Description: Micro-blogging service where users can post short messages (tweets) containing searchable annotations (hashtags)
  - Collected 969 tweets marked with the #sarcasm hashtag
  - Each sarcastic tweet was paired with an earlier tweet from the same user that was not marked as sarcastic

- **Google Books Database**
  - Description: Database of over 15 million scanned books allowing full-text searches and previews
  - Collected 110 quotations marked by the phrase “said sarcastically” (39 from Kreuz & Caucci, 2007)
  - Each sarcastic quotation was paired with an earlier quotation from the same speaker that was not marked as sarcastic

Analyses

- **Linguistic Inquiry and Word Count (LIWC)**
  - Positive emotion, negative emotion, and interjections*

- **Part of Speech (POS) Tagger Word Counts**
  - Stanford POS Tagger v3.0.4
  - Adjectives (JJ), adverbs (RB), adverb + adjective and adjective + adjective co-occurrences, and interjections (UH)

Compared LIWC scores and POS tagger counts of the sarcastic and non-sarcastic (control) statements in each corpus

*This was a custom dictionary of 161 interjections created by the researchers

Results

**LIWC Scores (Table 1)**

- Positive emotion words were more frequent in sarcastic statements across both corpora.
- Negative emotion words were less frequent in sarcastic tweets and equally frequent in sarcastic Google Books quotations.
- Interjections were more frequent in sarcastic statements across both corpora.

**POS Tagger Counts (Table 2)**

- Adjectives and adverb + adjective were more frequent in sarcastic tweets.
- Adverbs were less frequent in sarcastic Google Books quotations.
- Interjections were more frequent in sarcastic statements across both corpora.

Discussion

- Our goal was to use two unique, naturalistic corpora where sarcastic intent was explicitly marked to examine lexical features of sarcastic statements.
- Sarcasm statements in both corpora contained more positive emotion words, but not more negative emotion words. This is consistent with the idea of asymmetry of affect: sarcasm tends to be a positive evaluation of a negative event.
- Interjections were more frequent in sarcastic statements across both corpora, indicating that people may use interjections as lexical cues to sarcasm.
- Results for adjectives and adverbs—potential lexical indicators of hyperbole—were mixed, suggesting that these may not be particularly discriminating features.
- The fact that the emotion word and interjection results converged across two very different corpora suggests that these features are stereotypic of sarcasm and may potentially be useful as discriminating cues in machine learning applications.

References