Thesis Summary: Emotional Language in Persuasive Communication

Felicia Oberarzbacher

Student-Id.: 174809

Tutor: Dr. Raffaella Bernardi
CIMeC, University of Trento

Co-Tutor: Dr. Carlo Strapparava
HLT-NLP, FBK-irst

I. Motivation & Objective

This thesis is concerned with the relation between emotion-evoking words and audience reactions in persuasive communication. It merges the research on computational emotional analysis in *natural language processing (NLP)* and the automatic analysis and detection of persuasive attempts in political speeches, namely in CORPS (CORpus of Political Speeches). The aim of this study is to gain insight into the impact of emotional words on audience reactions and to provide indications for further research by showing the distribution of emotion-evoking words in political speeches, by giving an overview of emotional lexical resources and by setting a baseline for distinguishing audience reactions based on emotions.

As persuasion is widely used in politics, advertising and in everyday human interactions, it is useful to understand the factors leading to successful persuasion. Although persuasion is also possible using logical argumentation only, in practice persuasive communication usually takes advantage of emotions and attempts to evoke them in the audience. Therefore, it is interesting to explore the relation between conveyed emotions and audience reactions.

Previous work investigating CORPS focused on the polarity (valence) before audience reactions, i.e. *positive*, *negative*. However, psychological research on emotions and persuasion indicates that distinct emotions have a different impact on the persuasiveness of an argument, calling for a more fine-grained, beyond-sentiment approach which is taking into account specific emotions like *anger*, *fear*, *joy* or *trust*.

II. Method & Course of Action

A *computational, knowledge-based approach* has been taken to detect emotional words before audience reactions using NRC Word-Emotion Association Lexicon (EmoLex). Several programs have been built in Python to preprocess the corpus and the lexicons, to extract windows of sentences before audience reactions, to count the emotional words, to look them up in emotional lexicons and to implement a formula to retrieve the *persuasive impact* words, i.e. the most distinctive words and their associated emotion before each audience reaction. In this thesis, the absolute frequencies of words associated with different emotions before audience reactions, the emotional words in the whole corpus and in text passages not followed by an

audience reaction have been compared, as well as frequencies normalized by the number of words in the sentence window. Three different emotional lexical resources have been tested for building frequency lists of emotional words, and the most suitable one has been chosen for the rest of the analysis.

The assumption has been made that before audience reactions the distribution of emotional words was different from the distribution in text passages not followed by audience reactions. Given that there is no clear boundary between text passages preceding an audience reaction and text passages that are not followed by an audience reaction, audience reaction windows of different sizes have been extracted.

Furthermore, the emotional frequencies of CORPS have been confronted with the emotional frequencies of non-persuasive texts, sampled from the BNC (British National Corpus), in order to observe differences in their distribution of emotional words. Besides, for different audience reactions, statistics of the words with the highest *persuasive impact* with their associated emotion have been computed, which have been extracted using a method introduced by Guerini et al., 2008.[1]

Finally, a machine learning experiment with Support Vector Machines (SVM) trying to distinguish between a pair of audience reactions has been conducted.

III. Results & Conclusions

Several existing lexical resources for emotion detection have been thoroughly examined, so that a suitable one, NRC Word-Emotion Association Lexicon, has been chosen for the task. Quantitative analyses have been made that illustrate the high frequency of the positive emotions *trust* and *joy* in the corpus. They also show that negative emotions like *anger* are present before negative audience reactions as well as before positive ones, but being especially frequent before the audience reaction *booing*. Although there are variations in the distribution of emotional words depending on the audience reaction, there are emotional words that are frequent in general, e.g. *trust* is consistently the prevalent emotion, independently of the reaction. In comparison to BNC the rate of emotional words in CORPS is higher, especially for positive emotions like *joy* and *trust*.

Finally, a machine learning experiment demonstrated that, using only the emotional words preceding an audience reaction, it is possible to discriminate that reaction with above-chance accuracy. Good results can be obtained in particular with *booing* and *laughter*, probably indicating that these reactions are connected to a distinctive “emotional signature", i.e. there is a strong presence of emotional words before these reactions, and the distribution of the different emotions is characteristic of these two categories.

[1]Guerini, M., Strapparava, C., and Stock, O. (2008). CORPS: A corpus of tagged political speeches for persuasive communication processing. Journal of Information Technology & Politics, 5(1):1932