

Sentiment Analysis of Online News Text: A Case Study of Appraisal Theory

Christopher Soo-Guan Khoo

Wee Kim Wee School of Communication & Information,
Nanyang Technological University, Singapore

Armineh Nourbakhsh

University of Texas at Austin, Austin, Texas, USA, and

Jin-Cheon Na

Wee Kim Wee School of Communication & Information,
Nanyang Technological University, Singapore

Abstract.

Purpose. Sentiment analysis and emotion processing are attracting increasing interest in many fields. Computer and information scientists are developing automated methods for sentiment analysis of online text. Most of the research have focused on identifying sentiment polarity or orientation—whether a document, usually product or movie review, carries a positive or negative sentiment. It is time for researchers to address more sophisticated kinds of sentiment analysis. This paper evaluates a particular linguistic framework called appraisal theory for adoption in manual as well as automatic sentiment analysis of news text.

Methodology. The appraisal theory is applied to the analysis of a sample of political news articles reporting on Iraq and economic policies of George W. Bush and Mahmoud Ahmadinejad to assess its utility and to identify challenges in adopting this framework.

Findings. The framework was useful in uncovering various aspects of sentiment that should be useful to researchers such as the appraisers and object of appraisal, bias of the appraisers and the author, type of attitude and manner of expressing the sentiment. Problems encountered include difficulty in identifying appraisal phrases and attitude categories because of the subtlety of expression in political news articles, lack of treatment of tense and timeframe, lack of a typology of emotions, and need to identify different types of behaviors (political, verbal and material actions) that reflect sentiment.

Value. The study has identified future directions for research in automated sentiment analysis as well as sentiment analysis of online news text. It has also demonstrated how sentiment analysis of news text can be carried out.

Keywords: Sentiment analysis, appraisal theory, content analysis, political news

1. Introduction

Sentiment analysis and emotion processing is attracting increasing interest in many fields. The large number of online news sources on the Web as well as the explosive growth of social media sites of various kinds means there is a large amount of opinionated text, both formal and user-contributed, on the Web that can be mined for business or research purposes. Computer scientists and computational linguists are developing automated methods for sentiment categorization of text—identifying whether a document carries positive or negative sentiment (Pang & Lee 2008). Automatic sentiment categorization has been applied to product reviews (e.g., Wei & Gulla, 2010), movie reviews (e.g., Thet, Na & Khoo, 2010), discussion of stocks and

financial markets (e.g., Gu et al., 2006), and prediction of election results (e.g., Das, Srihari & Mukund, 2009).

However, the conceptions of sentiment often adopted in these studies are simplistic or ad hoc. Computer and information scientists currently focus on sentiment polarity (also referred to as *orientation* or *valence*)—that is, positive or negative sentiment. Social scientists and communication researchers devise their own frameworks based on their traditions, intuition and research hypotheses. Social psychologists have developed systematic theories of emotion, and linguists have proposed comprehensive frameworks for analyzing sentiment in text, but few computer science and social science researchers have adopted these theories and frameworks in their research. It is true that these frameworks are complex and difficult to apply in automated systems. However, as research in automated sentiment analysis matures, researchers will seek more challenging tasks and attempt to perform more sophisticated sentiment analyses based on systematic and well-grounded sentiment theories and frameworks. Social science and communication researchers should also consider adopting these frameworks as common theoretical platforms for comparing results across different studies.

This paper examines a particular linguistic framework called *appraisal theory* (Martin & White, 2005) in detail. The appraisal theory is applied to the analysis of a sample of political news articles, and evaluated for its utility, ease of use, flexibility, and comprehensiveness. The problems in adopting the framework are identified and possible solutions are proposed. It is hoped that this paper will indicate future directions for research in automated sentiment analysis as well as content analysis of news text.

2. Automatic Sentiment Categorization Research

Research in automatic text classification seeks to develop models (i.e. text classifiers) for assigning category labels to new documents based on a set of training documents. For classification, documents are represented as sets of features representing their content and style, called document vectors. Most studies of automatic text classification have focused on either “topical classification” classifying documents by subject or topic (e.g., *education vs. entertainment*), or “genre classification” classifying documents by document styles (e.g., *fiction vs. non-fiction*).

In recent years, with the tremendous growth of online discussion groups and review sites, researchers are turning their attention to sentiment classification, which predicts an overall sentiment of a review document into positive or negative sentiment. Although machine-learning techniques have long been used in topical text classification with good results, they are less effective when applied to sentiment classification (Pang & Lee 2008). Sentiment classification is a more difficult task compared to traditional topical classification, which classifies articles by comparing individual words (unigrams) in various subject areas. A challenging aspect of sentiment classification that distinguishes it from traditional topic-based classification is that while topics are often identifiable by keywords alone, sentiment can be expressed in a more delicate expression. For example, the sentence “*who would vote for this presidential candidate?*” contains no single word that is obviously negative. Sentiment classification requires more understanding than the usual topic-based classification.

Recently, researchers have gone beyond identifying the overall sentiment of a review document to more in-depth analysis of the different aspects or features of the

object being reviewed. For instance, the reviews of a music album will cover not only the overall sentiment but also the vocals, lyrics, recording quality, creativity, etc. Movie reviews will cover the script, performance of specific actors, cinematography, action sequences, direction, pace, etc. Moreover, researchers have attempted to summarize sentiment from multiple documents. Hu and Liu (2004) summarized customer reviews of electronic products, focusing on the sentiment towards particular product features or aspects (e.g., picture quality and screen size). Thet, Na and Khoo (2010) used clause-level sentiment analysis to extract opinions on different aspects of the movie being reviewed—the cast, director, story and music.

Research in automatic sentiment analysis of text has so far focused on identifying sentiment polarity (positive or negative). Identifying the different types of emotions and finding applications for this more subtle kind of sentiment analysis represent the next frontier in automatic sentiment analysis research. Whitelaw, Garg and Argamon (2005) observed that the current major challenge in sentiment analysis is the accurate identification of full “appraisal expressions” including the appraiser, the appraised and sentiment type. Asher, Benamara and Mathieu (2009) urged researchers to go beyond positive and negative expressions, to identify a wide range of opinion expressions, including motivations, recommendations and speculations, and their discourse relations in the text.

3. Sentiment Analysis of News Articles for Research Purposes

Sentiment analysis of news articles have been carried out by social science researchers to investigate research questions related to public opinion and perception, for example the public image and reputation of a specific community, such as a gender, race, social class, and religious minority (Haider-Merkel, Allen & Johansen, 2006; Kerr & Moy, 2002; Len-Ríos, Rodgers & Thorson, 2005). It has also been used to analyze depictions of various politicians in the media and the outcome of such representations on the politicians’ failure or victory in elections, or the public support they receive for their policies (de Vreese & Semetko, 2002; Fransworth & Lichter, 2005). With the growth of online news sources, social science researchers will increasingly employ text mining and natural language processing tools to analyze online text.

Researchers who focus on sentiment in news articles are mainly concerned about the *polarity* of these articles, i.e. whether they convey a positive or negative attitude toward the subject of discussion (Haider-Markel, Allen & Johansen, 2006). Some are also interested in the *attitude* (such as anger, fear, satisfaction, etc.) (Kepplinger, 2002). Uribe and Gunter (2007), for example, studied whether “sensational” news stories were more likely to elicit emotional responses from audiences than other TV news stories.

Benoit and Harthlock (1999) focused on sentiment in political campaign discourse. They identified three major linguistic functions that serve to promote the subject candidate, namely *acclaim*, *attack* and *defence*. Each function can focus on the candidate’s *character* or the candidate’s *policies*. This set of variables, incorporated into a framework called the *functional theory of political campaign discourse*, was used by Cho and Benoit (2006) to study press releases from George W. Bush and John Kerry in the 2004 U.S. presidential election. They suggested that most functions focused on *policy* rather than *character* in campaign news releases. Jamieson and Waldman (1997), de Vreese and Semetko (2002), Fransworth and Lichter (2004 & 2005) analyzed campaign publications and political news releases for their polarity.

Their study focused on a particular political figure such as the President or a candidate.

Most of these studies have used self-designed frameworks with mostly binary or ternary classification (e.g. positive/neutral/negative). A researcher focusing on the polarity of the text would categorize each sentence into *positive/negative*, while another researcher focusing on the type of sentiment might categorize it as *critical/supportive*, and a third one focusing on the way it is expressed might categorize it as *emotional/critical*. This set of categories is usually designed based on the researcher's intuition, and might not be consistent with other similar studies in the field. The frameworks used in different studies are not only incomparable and inconsistent, but are also not grounded in formal linguistic and psychological theories about sentiment and its expression. Adopting a more systematic framework based on linguistic research would facilitate a richer, more detailed and consistent analysis.

4. Linguistic Framework for Sentiment Expression and Appraisal

Two sentiment frameworks have been proposed by computational linguists to annotate texts. Asher, Benamara and Mathieu (2009) used a framework with four top-level categories:

1. Reporting expressions, which conveys the persons commitment to the opinion
2. Judgment expressions, which express evaluations of objects and actions in relation to social and personal norms
3. Advise expressions, which express an opinion on a course of action
4. Sentiment expressions, which express feelings.

Wiebe, Wilson and Cardie's (2005) annotation framework was based on the concept of private states which include opinions, beliefs, thoughts, feelings, emotions, goals, evaluations and judgments. They distinguished between three types of private state expressions:

1. explicit mentions of private states
2. speech events expressing private states
3. expressive subjective elements (i.e. implicit indication of private states).

They planned to expand their attitude types to include subtypes of emotion, warning, stance, uncertainty, condition, cognition, intention and evaluation.

However, the most comprehensive linguistic theory of sentiment that has been employed for sentiment analysis of text is the *appraisal theory* (Martin, 1995). White (2005) characterized the *appraisal* framework as "the language of attitude, arguability and interpersonal positioning." It is a development of one major component of Halliday's (1985) Systemic Functional Grammar (SFG), known as *interpersonal semantics*. SFG is an approach to linguistics that focuses on the function (semantics) of text rather than its form (grammar) (Eggins, 1994).

Martin and White (2005) used the appraisal theory to design a framework for appraisal analysis of text. Their framework defines each chunk of text (e.g. noun phrase, verb phrase or clause) that implies one instance of appraisal as an *appraisal group*. Two major actors are usually involved in each instance of appraisal, namely the *appraiser* and the *object of appraisal*.

The appraisal theory identifies three major aspects to appraisal between these two actors:

1. *Attitude*: Attitude is the essence of emotion the appraiser conveys about the object. Love, anger, fear, jealousy, excitement, hostility, satisfaction, etc. are emotions or attitudes that can be expressed in appraisal groups, such as

“It’s such a wonderful day!” In more formal contexts such as news articles, expressions such as “This policy will not work” are considered appraisal groups conveying attitude.

2. *Graduation*: Graduation refers to the strength or force of emotion and attitude in each appraisal group. Adverbs such as “fairly” or “soulfully” indicate graduation. Words used to express attitude can also contain an indication of graduation, e.g. “dislike” versus “detest”.
3. *Engagement*: Sentiment can be expressed directly or indirectly, or attributed to another source. Engagement is the way speakers or writers express appraisal or *engage* in the argument. The engagement is *monoglossic* if the speaker/writer has directly expressed the appraisal, e.g. “The President has no idea what he is talking about”. It is *heteroglossic* if the speaker/writer has either attributed to another source, or has tried to use other methods to prove the point or make it more credible, e.g. “Democrats criticized his proposal” or “Data from previous studies show there’s little hope that his strategy will work”.

Attitude and engagement can be divided into several subtypes, explained in more detail in the following sections.

4.1 Attitude

Attitude is the main emotion of the appraiser and encompasses three semantic regions (Martin and White, 2005):

1. The emotional region: *Affective* sentences express the appraiser’s emotion. Affect can be an internal mental state (e.g. “I feel *sorry* for people who have lost beloved ones in the incident”), or a behavioral process (e.g. “I couldn’t help *crying* at the sad news”).
2. The ethical region: *Judgmental* sentences involve the appraiser evaluating an intelligent object. The evaluation may concern ethical norms of society (referred to as *social sanctions*) or social norms (*social esteem*). Social sanctions can involve *veracity* (i.e. evaluating someone’s honesty, e.g. “Can we *trust* the President on this?”), or *propriety* (i.e. evaluating someone’s moral behavior, e.g. “A *violent* and *cruel* dictator”). Social esteem may involve assessment of how *normal* someone behaves (e.g. “He is the slightly *off-center* genius ...”), how *capacious* the person is (e.g. “the Iraqi government’s *weakness* in quelling violence”), or how *tenacious* the person is (e.g. “She could work *tirelessly* for a day”).
3. The aesthetic region: *Appreciative* sentences evaluate non-intelligent objects. It is prevalent in political news, since events and policies (both non-intelligent) are two major objects of appraisal in political news. Appreciative appraisal can be expressed as the appraiser’s *reaction* to the object (i.e. whether he or she likes it). This reaction can be about the *impact* of the object on the appraiser (e.g. “an *uninviting* show”), or its *quality* (e.g. “a *lovely* piece”). In addition to reaction, the appraiser might also focus on the *composition* of the object, which may address the *balance* of the object (e.g. “a *consistent* report”) or its *complexity* (e.g. “a *detailed* report”). Finally, the appraiser can talk about the *valuation* of the object, i.e. how worthwhile it is (e.g. “*ineffective* amendments”).

4.2 Engagement

Engagement is the way the speaker/writer engages in the argument being presented to the audience/readers. Heteroglossic engagement includes:

1. *Disclaiming* an argument by either *denying* it using negation, or first presenting it and later *countering* it by using adverbs such as “yet”, “although”, “amazingly”.
2. *Proclaiming* something by either emphasizing it using “in fact”, “of course”, “sure”, etc., or by referring to evidence (e.g. “the incident shows that...”).
3. *Entertaining* alternative arguments by lowering the strength of proclamation using “maybe”, “I think”, “perhaps”, etc.
4. *Attributing* the argument to other sources, by either *acknowledging* the source (e.g. “He *argues* that...”) or *distancing* from it (e.g. “He *claims* that ...”). Attributions can be carried out using direct or indirect quotations (e.g. “He *says* he hates her”), or proclaimed by the speaker/writer (e.g. “*I know* he hates her”).

The speaker/writer is not the only one who can engage in the argument. Other sources to whom the sentiment is attributed have an engagement of their own, but since this is also introduced by the speaker/writer, his or her engagement serves as the major component.

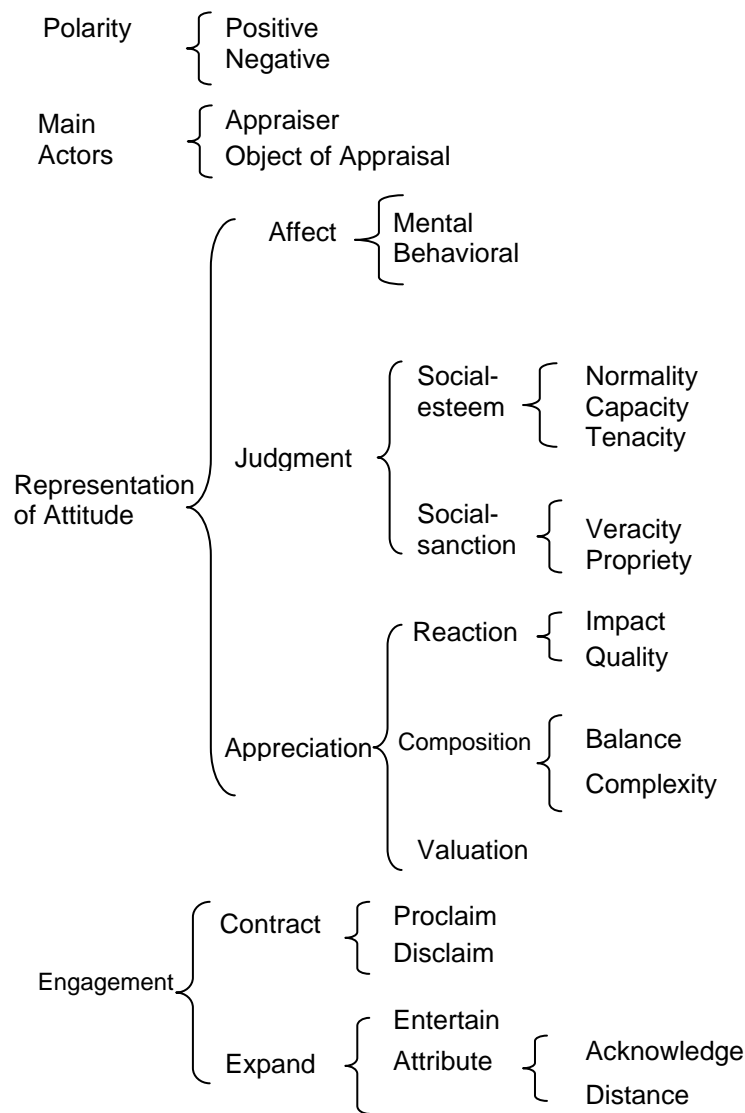
If a statement is expressed directly as an objective fact, suggesting a lack of engagement, this is considered monoglossic.

5. Sentiment Analysis of Political News

We applied Martin and White (2005)’s framework for appraisal analysis on a sample of 30 political news articles, and analyzed them for various aspects of sentiment. The aspect of *graduation* was dropped from the analysis as few social science researchers have shown interest in directly analyzing text sentiment intensity. Rather, this intensity is usually calculated according to the total number of biased sentences that appear in each article (e.g., Cho & Benoit, 2006; Farnsworth & Lichter, 2005). Figure 1 gives a summary of the various aspects of appraisal used in the analysis.

The sample articles were taken from the New York Times, the Times of London, and the Sydney Morning Herald in a six-month long period from October 2006 to March 2007. Fifteen articles discussed Mahmoud Ahmadinejad, the President of Iran, and fifteen discussed George W. Bush, the former U.S. President. All articles focused on the Presidents’ economic or Iraq policies. They all contained the keywords *Bush* or *Ahmadinejad*, along with *Iraq* or *economic* in their titles or leading paragraphs. The two political figures and two topics were chosen because we expected that sentiment analysis of news articles was likely to be used to study people’s sentiment towards government policies and towards political figures, especially approaching an election.

The Iraq war was a recent controversial issue at the time of data collection and attracted intensive coverage from all three newspapers. The nature of discourse was argumentative and attitudinal, and news articles on the issue were good candidates for sentiment analysis. The economy on the other hand had more objective reportage and allowed for analyzing the critical tone.



Note: *Graduation* has been left out due to its infrequent use in political news analyses

Figure 1. Martin and White's (2005) framework for appraisal analysis

The articles were analyzed for appraisal groups—chunks of text (usually phrase or clause) that contains one instance of appraisal or sentiment expression. Totally 967 appraisal groups were collected from the sample, among which 645 groups discussed or engaged George W. Bush, and 419 groups discussed or engaged Mahmoud Ahmadinejad. There were totally 97 overlapping groups, meaning they engaged both Presidents (in 45 groups the appraiser was George W. Bush and the object of appraisal was Mahmoud Ahmadinejad, in 52 appraisal groups the reverse occurred).

All appraisal groups were divided into categories that showed which role each President had played in them. Consequently three categories were identified—groups that engaged either President as the appraiser, groups that engaged either President as the object of appraisal, and groups that involved either President appraising his own policies. All groups were then analyzed for polarity (or orientation), type of sentiment (i.e. *attitude*), actors involved in the process of appraisal (i.e. appraiser and the object of appraisal), and the way sentiment is expressed (i.e. *engagement*).

Table 1 – Sample sentiment analysis of a news excerpt using appraisal theory

Appraisal group	Appraiser	Object of Appraisal	Engagement
The Bush administration reveled in the economy's continued expansion [<i>affect: behavioral: +</i>] ¹	the Bush administration	economy's expansion	attribute: acknowledge
In a buoyant [<i>appreciation: reaction: quality: +</i>] annual report,	author	Bush's economic advisors' report	monoglossia
<i>President Bush's Council of Economic Advisers said the economy was poised to keep growing</i> [<i>appreciation: reaction: quality: +</i>] at a solid 3 percent a year	Bush's Council of Economic Advisers	economy	attribute: acknowledge (through indirect quotation)
But the report set up a clash [<i>appreciation: reaction: quality: -</i>] with Democrats	author	Bush's economic advisors' report	disclaim: counter

¹ Note: explanations on the type and polarity of attitude are embedded within appraisal groups (+ stands for positive and - stands for negative)

Table 1 shows a sample analysis for the following excerpt that discusses President George W. Bush's economic policies:

“The Bush administration reveled in the economy's continued expansion on Monday [...] In a buoyant annual report, President Bush's Council of Economic Advisers said the economy was poised to keep growing at a solid 3 percent a year [...] But the report set up a clash with Democrats and even some Republicans in Congress [...]”

As Table 1 indicates, the statement *“The Bush administration reveled in the economy's continued expansion”* expresses a *behavioral affect* attitude. This indicates that the administration has manifested sentiment through a behavioral surge (*“reveling”*). The sentiment is *positive* (as indicated by the plus sign). It is experienced by the Bush administration (as indicated in the *appraiser* column), and is about the expansion of economy (indicated in the *object of appraisal* column). The sentiment has been *attributed* to the Bush administration by the author in an affirmative way (by *acknowledging* the feeling rather than rejecting or ridiculing it).

The author then talks about *“a buoyant annual report”* indicating the author's sentiment is *positive* and is directed toward the report (*object of appraisal*). This is an instance of *monoglossia*, when sentiment about the report is directly expressed to readers.

In the third statement, the author *attributes* the sentiment to Bush's Council of Economic Advisors using indirect quotation. The council is acclaiming the *quality* of the U.S. economy and hence are *positive* about it.

But the last sentence reveals another side to the story. The author now discusses the deficiencies of the report that led to *“a clash with Democrats”*. The author here is expressing *negative* attitude about the report. He does it by *disclaiming* the positive arguments that were previously presented in the article (*“but”* is used here to express disclamation).

The same analysis was performed on the rest of the 967 appraisal groups, and the results of the analysis are presented in the next section.

6. Results

The analysis starts by investigating *polarity*, which is useful for determining the overall sentiment orientation or bias of a document or person towards a particular object. Table 2 presents four cross-tabulation tables of polarity (positive/negative) versus actors (appraiser/object of appraisal/both) for Bush and Ahmadinejad, for the two topics of economy and Iraq war. For example, the first cross-tabulation (upper left quadrant) shows the number of instances of positive and negative sentiment expressed when Bush was the appraiser, object of appraisal or both (i.e. Bush appraising his own policy). A Chi-square test of independence was carried out on each cross tabulation, and the relationship between polarity and actor was significant at the 0.001 level for appraisal groups involving Bush and economy, Bush and Iraq war, and Ahmadinejad and economy. The relationship was significant at the 0.01 level for Ahmadinejad and Iraq war. The numbers in bold print in Table 2 are frequencies that are substantially higher than expected.

On the topic of *economy*, both Bush and Ahmadinejad had a positive bias toward their own policies, and a negative bias toward others. There were 41 instances of Bush as appraiser on an issue of *economy*, and 26 (63%) of these were negative. There were 19 instances of Bush as both appraiser and appraisee (i.e. appraising himself), and 18 (95%) of these were positive. In contrast, when they were the object of appraisal, the sentiments expressed were largely negative—69% negative sentiment for Bush and 92% negative sentiment for Ahmadinejad.

On the issue of Iraq war, both Bush and Ahmadinejad had been expressing predominantly negative sentiments. Bush's appraisal of himself was more balanced, with about equal numbers of positive and negative sentiments. What is interesting is that, even though Bush's appraisal of issues related to the Iraq war was mostly negative, his was significantly more positive than other people's appraisal of him. Bush had 22% positive sentiment as appraiser and 48% positive sentiment when he appraised himself, compared to 15% positive sentiment when others appraised him. It is curious that Ahmadinejad exhibited the same pattern: he had a higher positive sentiment as appraiser (36%) than when he was the object of appraisal (16%).

Table 2 – Sentiment polarity of phrases involving Bush and Ahmadinejad as Appraiser and Object of Appraisal

Stance		Economy		Iraq war	
		Pos	Neg	Pos	Neg
Bush	Appraiser	15 (37%)	26 (63%)	29 (22%)	106 (79%)
	Object of Appraisal	47 (31%)	104 (69%)	40 (15%)	230 (85%)
	Both	18 (95%)	1 (5%)	14 (48%)	15 (52%)
	Total	80 (38%)	131 (62%)	83 (19%)	351 (81%)
Ahmadinejad	Appraiser	7 (13%)	48 (87%)	43 (36%)	78 (65%)
	Object of Appraisal	11 (8%)	124 (92%)	14 (16%)	75 (84%)
	Both	12 (75%)	4 (25%)	3 (100%)	0 (0%)
	Total	30 (15%)	176 (85%)	60 (28%)	153 (72%)

Table 3 – Most frequent appraisers of Bush

Appraiser	Economy		Iraq war		Total	Percentage
	Pos	Neg	Pos	Neg		
Author	27	47	7	81	162	38.5
Ahmadinejad			8	25	33	7.8
Democrats		14		8	22	5.2
Public/Americans	3	4		10	17	4
Congress		5		5	10	2.4
Iran			1	8	9	2.1
Unknown		1		8	9	2.1
Syria			3	5	8	1.9
Charles E. Schumer	2	5			7	1.7
Republicans	1		1	4	6	1.4

Note: Coverage of appraisal groups in this table is 67%.

Table 4 –Most frequent appraisers of Ahmadinejad

Appraiser	Economy		Iraq war		Total	Percentage
	Pos	Neg	Pos	Neg		
author	2	47	1	21	71	31.7
America		4		12	16	7.1
Bush				9	9	4
United Nations	1	6		2	9	4
Vali Nasr	1	6			7	3.1
American government	2	1		2	5	2.2
Ayatollah Khamenei	2	3			5	2.2
Iranian voters		5			5	2.2
Nuri Kamal al-Maliki			4	1	5	2.2

Note: Coverage of appraisal groups in this table is 58.7%.

We can go on to analyze what various people think of Bush and Ahmadinejad, and their policies. Table 3 and 4 lists the most frequent appraisers of Bush and Ahmadinejad. Predictably, the author is the most frequent appraiser, comprising 39% of the sample. The appraisers mostly have negative sentiment toward Bush and Ahmadinejad. This type of analysis can be useful in identifying the allies and opponents of each President, and the policies these people support or reject. Table 4 suggests that Ahmadinejad had the support of Nuri Kamal al-Maliki on issues of the Iraq War.

Going on to analyze the types of sentiment expressed towards Bush and Ahmadinejad, Table 5 and 6 lists the types of attitudes (affect, judgement or appreciation) attributed to the major appraisers of Bush and Ahmadinejad.

Table 5 shows that the most prominent attitude type for the authors is *appreciation* and *judgement*, rather than *affect* (emotion). They mainly focus on critiquing non-intelligent objects (i.e. policies and plans). Occasionally, they have also been judgmental of the President's character. Democrats, Ahmadinejad, Congress, Iran, Americans, and Syria appeared to be more affective or emotional.

The same pattern emerges for authors who have appraised Ahmadinejad (Table 6). *Affect* is again the major attitude of other appraisers.

Comparing *economy* and the *Iraq war*, we find that economic issues attract more objective evaluation (appreciation) than the Iraq war which attracts more emotion (affect).

Table 5 - Attitudes of the main appraisers of Bush

Appraiser	Attitude	Economy		Iraq war		Total	Percentage
		Pos	Neg	Pos	Neg		
author	appreciation: reaction: quality	13	11	4	39	67	15.9
author	appreciation: composition: balance	11	12	1	8	32	7.6
author	appreciation: valuation	2	6	1	7	16	3.8
author	judgment: social-sanction: veracity		4		11	15	3.6
Democrats	affect: behavioral		7		4	11	2.6
author	judgment: social-sanction: propriety		5		5	10	2.4
author	judgment: social-esteem: capacity		1	1	7	9	2.1
Ahmadinejad	affect: mental			5	2	7	1.7
Congress	affect: behavioral		4		3	7	1.7
Iran	affect: behavioral			1	6	7	1.7
Americans	affect: mental		2		4	6	1.4
author	appreciation: reaction: impact		4		2	6	1.4
Ahmadinejad	affect: behavioral			1	4	5	1.2
Ahmadinejad	appreciation: reaction: quality				5	5	1.2
Syria	affect: behavioral			3	2	5	1.2
a freshman Democrat	appreciation: composition: balance		4			4	0.9

Note: Coverage of appraisal groups in this table is 50.4%.

Analyzing the subcategories of affect, judgement and appreciation can yield additional insights, though it will be more difficult to draw quantitative conclusions as the sample size decreases. For example, the various categories of *appreciation* can clarify the appraiser's stance. Appreciation on *quality*, for example, is more indicative of whether or not the appraiser likes the object (e.g. President's policy) and is satisfied/dissatisfied with it, and hence more critical/supportive of the President. In contrast, most appreciations of the *composition* of the object are focused on external issues and situations that Bush finds himself in. The analysis is focused on the *balance* of the situation, and blame is not directed at the President but the situation.

Similarly, different types *judgment* carry different implications. For judgments focusing on *veracity*, more criticism is directed at the President than for judgments on *propriety* (compare the sentences "he is a *liar*" and "he is *weak*"). Judgments focusing on *social sanction* show a bias toward moral issues, while judgments focusing on *social esteem* show a bias toward culture-of-honor type values.

Table 6 - Attitude of the main appraisers of Mahmoud Ahmadinejad

Appraiser	Attitude	Economy		Iraq war		Total	Percentage
		Pos	Neg	Pos	Neg		
author	appreciation: reaction: quality	2	15	1	6	24	10.7
author	appreciation: composition: balance		14		2	16	7.1
America	affect: behavioral		3		12	15	6.7
author	judgment: social-sanction: propriety		4		9	13	5.8
United Nations	affect: behavioral		6		1	7	3.1
author	appreciation: reaction: impact		4		2	6	2.7
author	judgment: social-sanction: veracity		4		2	6	2.7
Vali Nasr	appreciation: composition: balance	1	4			5	2.2
author	judgment: social-esteem: capacity		4			4	1.8
Bush	affect: behavioral				4	4	1.8
Ayatollah Khamenei	affect: behavioral	1	2			3	1.3
Bush	judgment: social-sanction: propriety				3	3	1.3
Iranian voters	affect: behavioral		3			3	1.3
Abbas Milani	appreciation: reaction: quality		2			2	0.9
American government	appreciation: valuation	2				2	0.9

Note: Coverage of appraisal groups in this table is 50.3%.

In analyzing appraisers other than the author of the article, it should be kept in mind that all the statements go through the author's filter before appearing in the article. The author has substantial control over the choice of content, words and context, and will shape the material to support his or her argument. Therefore, sentiment of people other than the author should be analyzed as a sentiment that the author has attributed to them. For example, if a non-author appraiser engages in more emotional appraisal than the author, it might be because the author has presented them as more emotional or irrational.

One way of analyzing these implications is by doing *engagement* analysis. Table 7 shows the results of engagement analysis of the major appraisers of Bush. Predictably, authors have directly expressed appraisal on most occasions. Sometimes they have used *proclamation* (i.e. provided evidence for their argument or have stressed its correctness by using "in fact", "of course", etc.). Appraisal has mostly been *attributed* to other appraisers using direct or indirect quotation. However, there are instances where the author has attributed appraisal to other sources without providing any quotation or evidence (e.g. "Democrats believe that..."). External sources appear to be documented more often (e.g. government of other countries, foreign politicians,

etc.), than internal sources such as Democrats and Americans, which tend to be attributed appraisal without documentation, evidence or quotation.

Table 7 - Engagement of the main appraisers of George W. Bush

Appraiser	Engagement	Economy		Iraq war		Total	Percentage
		Pos	Neg	Pos	Neg		
author	monoglossia	8	29	5	49	91	21.6
author	contract: proclaim	17	12		28	57	13.5
Ahmadinejad	expand: attribute (through direct quotation)			4	17	21	5
Democrats	expand: attribute (through indirect quotation)		9		3	12	2.8
Democrats	expand: attribute		4		4	8	1.9
Iran	expand: attribute (through indirect quotation)			1	7	8	1.9
Ahmadinejad	expand: attribute (through indirect quotation)			2	5	7	1.7
Americans	expand: attribute		2		5	7	1.7

Note: Coverage of appraisal groups in this table is 50.1%.

Table 8 - Main issues for which Bush was appraised

Object of Appraisal	Economy		Iraq war		Total	Percentage
	Pos	Neg	Pos	Neg		
Bush's character	4	24	7	39	74	17.6
Bush's surge plan			9	47	56	13.3
America		2	2	18	22	5.2
Bush's invasion of Iraq			4	10	14	3.3
American troops				13	13	3.1
Bush's economic policy	6	6			12	2.8
Bush's economic advisors' report	1	10			11	2.6
America's economic situation	6	4			10	2.4
America's economic situation	6	4			10	2.4
Bush's performance in Iraq				9	9	2.1
Republicans	3	3		3	9	2.1

Note: Coverage of appraisal groups in this table is 56.9%.

Table 9 - Main issues for which Mahmoud Ahmadinejad was appraised

Object of Appraisal	Economy		Iraq war		Total	Percentage
	Pos	Neg	Pos	Neg		
Ahmadinejad's character	4	28	2	15	49	21.9
Iran	1	2	6	25	34	15.2
Iran's nuclear program	1	23		8	32	14.3
Ahmadinejad's situation		10		1	11	4.9
Ahmadinejad's speeches		5			5	2.2
Iran's nuclear ambitions		2		3	5	2.2

Note: Coverage of appraisal groups in this table is 60.7%.

Table 10 - Attitude and object of the main appraisers of George W. Bush

Appraiser	Attitude	Object of Appraisal	Economy		Iraq war		Total	Percentage
			Pos	Neg	Pos	Neg		
author	appreciation: reaction: quality	America's economic situation	5	4			9	2.1
author	appreciation: composition: balance	Bush's situation		3		5	8	1.9
author	appreciation: reaction: quality	Bush's Iraq surge plan			1	7	8	1.9
author	judgment: social- sanction: veracity	character				8	8	1.9
author	appreciation: composition: balance	economic situation	5	2			7	1.7
author	judgment: social- esteem: capacity	character		1	1	5	7	1.7
author	judgment: social- sanction: propriety	character		4		3	7	1.7
author	appreciation: composition: balance	Bush's economic advisors' policy	5	1			6	1.4
author	appreciation: reaction: quality	Bush administration's performance in Iraq				5	5	1.1
author	appreciation: reaction: quality	Bush's invasion of Iraq				5	5	1.1

Note: Coverage of appraisal groups in this table is 16.6%.

More detailed analysis can focus on what objects and issues related to the economy and Iraq war were being appraised (Tables 8 and 9), and details of the appraiser -> attitude -> object relation (Table 10).

Table 8 and 9 indicate that Bush's character, surge plan, invasion of Iraq and economic adviser's report received very critical appraisal. So did Ahmadinejad's character and nuclear program.

Table 10 shows that objects such as situations and problems have been appraised for their *composition*, while more subjective issues such as the Presidents' behaviors, or their policies and plans have been appraised for their *quality*.

It should be noted that all the above analyses can be carried out by considering more variables such as region, time, newspaper, etc. The analysis can be very useful for a thorough investigation of the bias, style, and authoritativeness of news publications.

7. Challenges

This section discusses the challenges encountered in the study.

7.1 *Identification of appraisal groups*

An appraisal group was earlier defined as a chunk of text that conveys an instance of appraisal. As appraisal is multi-dimensional encompassing polarity, attitude, actors involved, graduation and engagement, it is sometimes difficult to identify the text span containing all these components. Instances of appraisal can be subtle, ambiguous and overlapping, making it difficult to tease out an instance. Appraisal can also be multi-layered. A statement of an actor's sentiment toward an object can be overlaid with other people's sentiment of the situation, the author's sentiment, and the sentiment that the reader is expected to have. Pronouns, anaphors and ellipses complicate the analysis.

A possible solution is to ignore less important components of appraisal when identifying appraisal groups. In our analysis, we focus on *polarity* and *attitude* in identifying appraisal groups. Other components of appraisal are taken from the bigger context when needed.

7.2 *Mental versus behavioral affect*

As indicated earlier, *affect* (emotion) can be an internal mental state or a behavioral process. The distinction is not always clear. In this study, we include as *behavioral affect* political actions and verbal actions (e.g. *rejected* Bush's proposals, *opposed* his plans, etc.), as well as material actions (e.g. attack, war).

7.3 *Ambiguity of attitude*

It is sometimes difficult to identify the type of attitude expressed, especially in a political context, where appraisers tend to disguise the attitudinal nature of their argument using more objective language. Appraisal theory is based on a functional approach to linguistics and hence provides little grammatical and textual clues to categorization. Martin and White (2005) discussed the problems that ambiguity can cause and proposed guidelines for easier classification of clauses.

7.4 *Tense*

Some researchers are interested in analyzing changes in media bias toward a specific object (e.g. politician, country, party, etc.) over time (Jhally, 1997). This requires a longitudinal analysis of publications over a period of time and comparing their stance. The tense of a sentiment is tricky to analyze. It has to be determined by the semantic function of the appraisal group rather than the grammatical tense of verbs. For example, in the clause "Republicans rejected the proposal", the grammatical tense is past tense, but the author is referring to the Republicans' current objection to the proposal. The same holds for future-tense sentiment, where the author is usually evaluating a current issue (e.g. "The bill will surely face resistance in Congress").

Some appraisal groups refer to an old state-of-affairs between actors. For example, the phrase "America's *old* foe" has been used to convey that the recently reconciled countries U.S and Vietnam had previously been enemies. In our analysis, there were about 6% of appraisal groups that expressed no-longer valid sentiment.

7.5 Types of Affect

Appraisal theory does not specify the types of affect (emotions) to identify in the text. It is not known whether the basic types of emotions developed by social psychologists would be useful for analyzing news text. Asher, Benamara and Mathieu (2009) used the following sentiment categories based on French word classes (verbs and nouns) derived by Mathieu (2005) using lexical-semantic principles:

1. anger/calm down
2. astonishment
3. love, fascinate
4. hate, disappoint
5. fear
6. offense
7. sadness/joy
8. bore/entertain.

Further research is needed to identify the main types of emotions found in news text.

8. Conclusion and Discussion

Martin and White (2005)'s framework for appraisal analysis was adapted and applied to a sample corpus of political news articles that discussed two Presidents on two major issues, economy and Iraq war.

The framework proved useful in uncovering various aspects of sentiment that should be useful to researchers, such as the appraisers and object of appraisal, bias of the appraisers and the author, type of attitude and manner of engaging in the sentiment expression. Generally, the framework was capable of answering the following questions:

1. Who were the major appraisers of each President?
2. What was each appraiser's prominent polarity and attitude type?
3. How did each appraiser express appraisal?
4. What was each appraiser more concerned with? What did the appraiser evaluate?
5. What were the major targets of appraisal for each President?
6. What was each President's sentiment polarity and attitude type toward each object?
7. How did each President express his appraisal?
8. What was the bias of each President when talking about himself/his own policies? How did each President appraise himself/his own policies?
9. What was each President's attitude type about himself/his own policies?
10. What were the major issues each President appraised when talking about himself/his own policies?

Each of these questions can be answered considering the time, newspaper, and region the article is published in. All aspects of the framework are independent from each other and can be analyzed separately. Thus, researchers can customize the framework based on their own needs. For example, researchers focusing on media bias can focus on *polarity*, *attitude*, and specific subcategories of *engagement* (e.g. *acknowledgement* versus *distance*), while researchers who are interested in style analysis, can concentrate more on *engagement* than on other aspects.

The framework suffers from a few problems, including difficulty in identifying appraisal groups and attitude categories because of the subtlety of expression in

political news articles, lack of treatment of tense and timeframe, lack of a typology of emotions, and need to identify different types of behavioral affect (political, verbal and material actions that reflect sentiment).

Aside from illustrating how appraisal theory can be applied in content analysis of news text, the analyses in this paper also suggest future directions in automatic sentiment analysis research. As pointed out earlier, automatic sentiment analysis research has so far focused on identifying sentiment polarity (positive or negative). It is time for researchers to address more challenging tasks of automatically identifying appraisal groups in text, the actors (appraiser and object of appraisal) involved, and the type of attitude and engagement expressed. Further research can grapple with the issues of time, identification of different types of sentiment-biased behaviors (political, verbal and material actions), and analysis of different perspectives (author's perspective, reader's or specified actor's).

Since sentiment frameworks are complicated and time consuming to apply manually to annotate texts, and moreover it is difficult to achieve high inter-coder agreement (Read, Hope & Carroll, 2007), automatic or semi-automatic tools for carrying out different types of analysis specified in the framework will be a boon to linguistics, communication and social science researchers.

References

- Asher, N., Benamara, F. and Mathieu, Y.Y. (2009), "Appraisal of opinion expressions in discourse", *Linguisticae Investigationes*, Vol. 32 No. 2, pp. 279-292
- Benoit, W. L. and Harthlock, A. (1999), "Functions of the great debates: acclaims, attacks and defence in 1960 presidential debates". *Communication Monographs*, Vol. 66, pp. 341-357.
- Cho, S. and Benoit, W. (2006), "2004 presidential campaign messages: A functional analysis of press releases from President Bush and Senator Kerry", *Public Relations Review*, Vol. 32 No. 1, pp. 47-52.
- Das, P., Srihari, R. and Mukund, S. (2009), "Discovering voter preferences in blogs using mixtures of topic models", in *Proceedings of The Third Workshop on Analytics for Noisy Unstructured Text Data (AND '09)*, ACM, New York, pp. 85-92.
- de Vreese, C. H., & Semetko, H. A. (2002), "Cynical and engaged: strategic campaign coverage, public opinion, and mobilization in a referendum", *Communication Research*, Vol. 29 No. 6, pp. 615-641.
- Eggs, S. (1994), *An Introduction to Systemic Functional Linguistics*, Pinter, London.
- Farnsworth, S. J. and Lichter, S. R. (2004), "Source material-New presidents and network news: Covering the first year in office of Ronald Reagan, Bill Clinton, and George W. Bush", *Presidential Studies Quarterly*, Vol. 34 No. 3, pp. 674-690.
- Farnsworth, S. J. and Lichter, S. R. (2005), "Reporting on two presidencies: News coverage of George W. Bush's first year in office", *Congress and the Presidency: A Journal of Capital Studies*, Vol. 32 No. 2, pp. 91-109.
- Gu, B., Konana, P., Liu, A., Rajagopalan, B. and Ghosh, J. (2006), *Predictive Value of Stock Message Board Sentiments*, McCombs Research Paper No. IROM-11-06, available at: <http://ssrn.com/abstract=946498> (accessed 1 July 2011).
- Haider-Markel, D. P., Allen, M. D. and Johansen, M. (2006), "Understanding variations in media coverage of U.S. Supreme Court decisions: Comparing media outlets in their coverage of Lawrence v. Texas", *The Harvard International Journal of Press/Politics*, Vol. 11 No. 2, pp. 64-85.
- Halliday, M. A. K. (1985), *An Introduction to Functional Grammar*, Edward Arnold, London.

- Hu, M., and Liu, B. (2004), "Mining and summarizing customer reviews", in *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 168-177.
- Jamieson, K. H. and Waldman, P. (1997), "Mapping campaign discourse: An introduction", *American Behavioral Scientist*, Vol. 40 No. 8, pp. 1133-1138.
- Jhally, S. (1997), *The Myth of Liberal Media: the Propaganda Model of News*, Media Education Foundation, North Hampton, MA.
- Kepplinger, H. M. (2002), "Mediatization of politics: Theory and data", *Journal of Communication*, Vol. 55 No. 4, pp. 972-986.
- Kerr, P. A. and Moy, P. (2002), "Newspaper coverage of fundamentalist Christians, 1980-2000", *Journalism & Mass Communication Quarterly*, Vol. 79 No. 1, pp. 54-72.
- Len-Ríos, M. E., Rodgers, S. and Thorson, E. (2005), "Representation of women in news and photos: Comparing content to perceptions", *Journal of Communication*, Vol. 55 No. 1, pp. 152-168.
- Martin, J. R. (1995), "Interpersonal meaning, persuasion, and public discourse: Packing semiotic punch", *Australian Journal of Linguistics*, Vol. 15, pp. 3-67.
- Martin, J. R. and White, P.R.R. (2005), *The Language of Evaluation: Appraisal in English*, Palgrave Macmillan, London.
- Mathieu, Y.Y. (2005), "Computational semantic lexicon of French verbs of emotion", in J.G. Shanahan, Y. Qu and J. Wiebe (Eds.), *Computing Attitude and Affect in Text*, Springer, Dordrecht.
- Pang, B. and Lee, L. (2008), "Opinion mining and sentiment analysis", *Foundations and Trends in Information Retrieval*, Vol. 2 No. 1-2, pp. 1-135.
- Read J., Hope D. and Carroll J. (2007), "Annotating expressions of appraisal in English", in *Proceedings of the Linguistic Annotation Workshop, ACL 2007*, pp. 93-100, available at: <http://portal.acm.org/citation.cfm?id=1642074> (accessed 12 June 2011).
- Thet, T. T., Na, J.-C. and Khoo, C.S.G. (2010), "Aspect-based sentiment analysis of movie reviews on discussion boards", *Journal of Information Science*, Vol. 36 No. 6, pp. 823-848.
- Uribe, R. and Gunter, B. (2007), "Are 'sensational' news stories more likely to trigger viewers' emotions than non-sensational news stories?: A content analysis of British TV news", *European Journal of Communication*, Vol. 22 No. 2, pp. 207-228.
- Wei, W. and Gulla, J. A. (2010), "Sentiment learning on product reviews via sentiment ontology tree", in *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL '10)*, Association for Computational Linguistics, Stroudsburg, PA, pp. 404-413.
- White, P.R.R. (2005), *Appraisal Website Homepage*. Retrieved 1 January 2010 from <http://www.grammatics.com/Appraisal/>
- Whitelaw, C., Garg, N. and Argamon, S. (2005), "Using appraisal groups for sentiment analysis", in *Proceedings of the 14th ACM international conference on Information and knowledge management (CIKM '05)*, pp. 625-631, available at: <http://portal.acm.org/citation.cfm?id=1099714> (accessed 12 June 2011).
- Wiebe, J., Wilson, T. and Cardie, C. (2005), "Annotating expressions of opinions and emotions in language", *Language Resources and Evaluation*, Vol. 1 No. 2, pp. 165-210.