## A Value Based Approach to Recipient Sentiment Analysis

Obinna Onyimadu

Abstract

Much of Sentiment Analysis (SA) approaches have focused on classifying the semantic orientation expressed in the text but not the sentiment of the recipient or hearer of the text. For instance, the utterance, “House prices are falling” will most likely be assigned a negative sentiment orientation by contemporary SA models, but humans with diverse interests might assign different sentiments to the utterance. A landlord might assign a negative sentiment to the utterance, while a new buyer will most likely assign a positive sentiment. The question then is how can a sentiment analysis model be trained to predict the sentiment of a recipient/hearer given an utterance.

In this paper, we describe the implementation of a recipient sentiment analysis model that is based on a model of human values. The justification for applying human values is founded on the notion that human values are the primary influencers of behaviour for which sentiment is a type. Therefore, if a persons’ values can be modeled quantitatively, when presented with some text, we should in theory be able to predict the sentiment of the individual. To this end, this paper formalizes human values, modeling it as a generative process. We also describe a unique concept called feature switching, that enables the prediction of recipient sentiment from the value model. Unlike contemporary models of human values and sentiment analysis methodologies, which rely on content analysis or annotation of text, the solution described in this paper is completely devoid of human input thereby making the design process applicable to diverse domains and datasets.

Finally, the model is implemented within the context of politics, basing the value model on the values of major UK Political parties. The recipient sentiment analysis model is shown to have an accuracy of 72.5% and falls within the range of existing models.

## Introduction

Inconspicuous in Sentiment Analysis (SA) state of the art is the ability to predict the sentiment kindled in a hearer or reader when presented a piece of text[[1]](#footnote-1). In other words, what sentiment will a particular hearer or group of hearers express given a piece of text? This question is significantly different from conventional sentiment analysis, which is commonly aimed at predicting the sentiment of the text’s author or the sentiment expressed in the text. Let us consider an utterance: “Company A’s shares will drop next week”. On the surface, this utterance expresses a negative sentiment. However, from a human perspective, the sentiment expressed is open to who the hearer is and the context from which the sentence and its component subjects are perceived by the hearer. Assuming the hearer has a stake in Company A, it is very likely, that such a hearer would express a negative sentiment. Conversely, if the hearer belongs to a rival company, it is not unlikely that such a hearer would express a positive sentiment. This problem is called the author-reader standpoint because the sentiment expressed by the author of a sentence does not necessarily translate to that expressed by the recipient (Liu, 2012). Therefore, there is a gap in current SA research around the prediction of a recipient’s sentiment towards an utterance/sentence, because current SA approaches do not incorporate the subjective human centric aspects which determine the sentiments of individuals. The implication of this is that these SA approaches always return a single sentiment prediction for any sentence. Thus, the purpose of this paper is to develop and implement an innovative approach for predicting recipient sentiment i.e. addressing the author-reader stand point problem.

There have been very few attempts at solving the author-user standpoint problem. However, common approaches have had their foundations in the use of socio-theoretic principles such as Affect Control Theory (Ahothali & Hoey, 2015; Heise, 1987; Mejova, 2012), appraisal theory (Bloom, 2011) and frame semantics (Bhowmick, 2009; Fillmore, 1982). Similarly, the approach taken in this paper also stems from a social concept: human values. We propose that the sentiment of a reader towards an utterance is a form of human behaviour, which in turn is influenced and determined by the values of the reader or hearer (Templeton et al, 2011a, 2011b). In fact, the definition of values as fundamental abstract coordinators of behaviour and guides for preference of one situation over another (Rokeach, 1973), substantiates the link between sentiment as a form of human behaviour and values. It is values that guides preference for one state, substance, entity, concept, idea etc. over another and sentiment in itself is an expression of preference for one state over another. Therefore, to develop a model that predicts reader or hearer sentiment, such a model needs to incorporate the values of the recipient.

One benefit of applying values in recipient sentiment prediction is to enable the estimation of sentiments in objective evaluative sentences (OES). Objective evaluative sentences are sentences which do not contain any explicit sentiments, are not subjective and whose sentiments are implied. Traditional SA methodologies and algorithms have succeeded in identifying these types of sentences but there has not been much success at assigning the polarity. Part of the problem lies in the one-size-fits all nature of SA, as well as the inability of current SA approaches to match the utterance to diverse sentiment holders. Let us consider two examples “*We will exit the EU*”, *“I paid £1000 for the new office software”*. Both these statements are objective and do not carry any explicit sentiments, however both contain implied sentiments. In exiting the ‘EU’, a person with values consistent with remaining in the ‘EU’ will view it as negative and similarly, a person with open-source software values will perceive the latter statement as negative. Traditional sentiment analysis methodologies are unlikely to make these differentiations, and in fact some SA implementations are likely to assign the first sentence a negative polarity because of the presence of the seemingly negative linguistic unit ‘exit’. For such OES, the use of traditional sentiment analysis methods in the determination of the sentiment would result in poor or spurious predictions primarily because of the absence of human centric reference points. This paper proposes that the inclusion of human values in a sentiment model can remedy this problem.

**What are Values?**

Research in the application and formalization of values has been an on-going task in the social sciences and humanities. Expectedly, these fields have evolved a significant proportion of the theory and methodologies adopted today in the classification, formalization and application of values. Several definitions of values exist (Hitlin, 2003; Hitlin and Piliavin, 2004; Rohan, 2000). Values have always been perceived as an abstract concept. In fact, Perry (1926) defined it as a philosophical concept or belief associated closely with virtuous living and morality. Similarly, Williams (1979) expressed values as interests, pleasures, likes, preferences, moral obligations, desires, wants, goals, needs, aversions, attractions and many other kind of selective orientations (Perry, 1926). Rokeach (Rokeach, 1973) attempted to provide a uniform definition and conceptualization of values defining values as “abstract fundamental coordinators of behaviour”. ‘Abstract’ representing an unquantifiable, non-physical entity and ‘coordinators of behaviour’ implying that for any form of behaviour for which sentiment is a type, values represent the primary causal factor. Similarly, Verplanken and Holland (2002) expressed values as “latent variables that have explanatory value for the choices people make”. Schwartz (1996), Feather (1995) and Bardi and Schwartz (2003) reinforce this notion of values as causative to behaviour referring to values as principal determinants of behaviour and attitude.

**How have they been applied?**

Several research works have shown the influence of human values on behaviour. Schwartz (1992) and Schwartz (n.d(a)), showed that on the subject of gay marriage, people with traditional values were more likely to be in opposition or have a negative sentiment. Conversely, people with hedonistic values were most likely to view the subject positively. Bengston et al (2004) showed that on environmental issues, specifically afforestation, the types of values held by individuals had an implication on the policies and goals for public forest management.

**How are they modelled?**

Values are modeled directly from value holders or by analysing recorded communication in textual materials like speeches, debates, testimonies, reports and utterances. In either case, the task involves the detection of value motivations and items, their categorization into inventories or classes and finally aggregation into value orientations. Methods used in accomplishing these aims are centred principally on empirical surveys of value holders (Scott, 1965; Schwartz, 1994; Schwartz, 2004a, 2004b; Schwartz, 2012, McDonald and Gandz, 1991; Braithwaite and Scott, 1991), content analysis of textual content (Cheng and Fleischmann, 2010; Callicott et al, 2000; Ishita et al, 2010) and human/theoretical analysis (Rokeach, 1973). This detection is carried out by researchers and domain/subject experts who using any of the methods, come up with a list of value items which are normally words, expressions or concepts usually nouns, verbs or adjectives that reflect expected desires and actions. The enumerated words could be sourced intuitively, that is the researcher/s uses his/her intuition or experience to itemize a list of expected items. They could also be derived from reviewing literature and conducting surveys on domain experts. For instance, in Schwartz (1994) the goal was to identify a set of basic human values to which 56 basic human value items were identified, Scott (1965), identified 12 value items for the goal of identifying personal traits for ideal relations, Kahle et al (1988), identified 9 value items required to measure consumer attitudes and behaviour and finally, Crace and Brown (1996), in developing values for decision making itemized 14 value items. Items describe abstract values and can be statistically graded e.g. a scale of 1-10 or 1-5. The value items are eventually grouped together into value types by the researchers. A collection of value types form the model of values called a value inventory (VI). The VI is a model that represents a set of value types and their constituent items. It provides explicit categories for the analysis of human values. Several VI’s exist: Schwartz’s (1992, 2012) value conceptualization focused on the notion that human values are based on motivational goals and needs that are basically universal across all cultures and peoples. Hence, the set of values identified by Schwartz are generic and applicable to social issues. Rokeach (1973) conceptualized values from two perspectives, where in one case, the values are perceived as a set of ultimate goals called terminal values, in the second, they are perceived as modes of behaviour. The result of this was a list of 36 value items categorized into terminal and instrumental value types. Bernthal (1962), proposed a hierarchy of values for management decisions based purely on rational reasoning. The inventories contained – the business firm level, economic system level, societal level and individual level. Personal Values Questionnaire (PVQ) (England, 1967) comprised of 66 value items organized into 5 categories. Inventories are thus instruments or models used in determining values.

Nevertheless, VIs have several deficiencies. Value items by nature are an enumerated list, and as such there is no way to determine an exhaustive list of items which capture all possible values especially because values have been shown to vary with time and because contexts evolve (Rokeach, 1973; Gurel-Atay et al, 2010). Value inventories are also subjective and their implementation involves significant human involvement and expense in the form of content analysis or empirical surveys (Cheng and Fleischmann, 2010; Scott, 1965; Takayama et al, 2014). These methodological gaps validate the need for an approach that is flexible, easily applicable to multiple domains and finally one that can be implemented without the express need of human annotations or content analysis. Therefore, the question sought to be answered by this paper is as follows: How can human values be modeled without human input? Secondly, given a sentence with implicit or explicit sentiments, can the value model be applied towards the prediction of the individual’s sentiment polarity? In other words, can the behaviour (sentiment) of the hearer or reader of an utterance or piece of text be predicted from a model of his/her values?

### **Value Models to Sentiment Prediction**

In this paper, we propose …

Structure of Paper

## Related Work

A significant proportion of SA has focused on the sentiment of the author, and as such literature on recipient sentiment prediction is quite limited. Existing research has not solely involved the prediction of the recipient’s sentiment but related areas such as the prediction of the recipient’s emotion or the prediction of the recipient’s stance. The common approaches can be summed up into two distinct categories: Text and Knowledgebase methodologies and socio-theoretic methodologies. We consider these in the next section.

### **Text and knowledgebase Approach**

This approach is based on identifying patterns and features in text, utilizing an external domain knowledgebase or ontology, identifying and applying discourse patterns in sentences and utterances. Tang and Chen (2011) performed emotion detection of writer and recipient emotion in a chat room. Emotion detection involves identifying the emotion expressed in a sentence e.g. anger, disgust, fear, happiness, sadness and surprise (Strapparava & Mihalcea, 2007). Their approach required the identification and annotation of emotions in both the reader and recipient content. Recipients (readers) in the chat network label sentences with optional quantifying emotions like ‘Likes’, ‘Shares’, ‘Gives’, ‘Hates’, ‘Wants’, ‘Wishes’, ‘Needs’, ‘Will’, ‘Hopes’, ‘Asks’ etc. In addition, contributions in the chat room were labelled as positive or negative. As such sentences were labelled for their emotional content and mapped to a semantic orientation, ensuring that both linguistic and human centric features are captured and harnessed in the development of the model. Additional human centric features like the social relations between writers were also captured and fed into the model. The eventual supervised model was shown to have an accuracy in the range of 80.67% and 88.37% for predicting the reader’s emotion. Although this model performs quite well, it required considerable human involvement in the annotation of the content and even in the collection of user centric behaviour. Similarly, Lin et al (2007) and Lin and Chen (2008) adopted a familiar approach in estimating the emotion of readers from a manually tagged Yahoo! Kimo news corpus. A corpus was tagged based on eight emotional classes by humans and linguistic features such as character bigrams, presence of emotional words and content metadata were extracted and applied in their model. They reported an accuracy of 76.88%. Again, in this approach we observe the introduction of human annotations and the division of reader emotion into classes which makes the approach quite rigid since human emotions could belong to more than one class.

Attempts have also been made in predicting the stance of an individual or group. Stance detection is the task of classifying perspectives e.g. for or against something and it has been applied in a variety of sectors and domains ranging from politics (Thomas et al, 2006; Somasundaram and Wiebe, 2009, 2010**)** to online debates on a variety of subjects (Murakami and Raymond, 2010). Besides Sridhar et al (2014), most of the work on stance detection focuses on detecting the stance expressed by the writer (Thomas et al, 2006; Murakami and Raymond, 2010; Somasundaran and Wiebe, 2009, 2010). Sridhar et al (2014) implemented a stance detection approach using both linguistic and the structural arrangement of the debates in online posts as features in classifying the stance on gun control and gay marriage. Linguistic features such as the length of a speech, word counts, discourse cues and punctuation count are applied. The most unique feature applied was the incorporation of author information. However, this implementation was dependent on hand annotated stances for each sentence in the training set. That is each sentence in the training set contained a marker saying if it was pro or anti a subject. They obtain an average F1 score of 74% for the positive class.

In applying this approach, a group of recurrent limitations are observed. The involvement of humans in annotating, tagging and classifying a corpus for ground truth, makes it expensive and time consuming. In addition, dependence on manually constructed knowledgebase or affect lexicon results in a model that is rigid, domain dependent and incapable of catering to diverse contexts. Finally, in basing the approach on patterns and linguistic features, human centric features such as behaviour, relationship between the subjects are not emphasized, although Sridhar et al (2014) and Tang and Chen (2011) incorporate such features. In the next section, we describe socio-theoretic approaches which tend to focus more on incorporating human centric features.

### **Socio-Theoretic Approaches**

These approaches tend to incorporate social theories which model or describe abstract human behaviour in sentiment or emotion detection. Examples of social theories include Affect Control Theory, Frame theory and Appraisal theory. They are typically applied in stance detection, emotion classification, identification of implicit objective sentences.

In this section, we discuss Affect Control Theory (ACT), a sociology theory that has been applied in SA and in particular recipient sentiment prediction.

ACT is a social psychological theory of human interaction (Heise, 2007). It suggests that “certain cultural norms dictate the affective meanings of words that people in a culture with a common language share”. It computes this affective meaning of an event or concept - events or concepts are expressed as words - in a multi-dimensional semantic space (Robinson and Smith-Lovin, 2006; Mejova, 2012) that consists of Evaluation, Potency and Activity (EPA).

In ACT, empirical equations are derived for a wide-ranging set of situations associated with an event. The affective sentiment of cultures is derived or measured using a survey technique called semantic differential derived by Osgood et al (1957), the basis of which is not so different from the empirical surveys applied in modeling values. Basically, individuals with knowledge of a culture rate concepts on a numerical scale with opposing adjectives at each end. In fact, a database of concepts expressed as words and their average EPA ratings derived from survey participants who are knowledgeable about their culture has been collected in Heise (2010).For instance, in the example given by Ahothali and Joey (2015), the culturally shared EPA for the concept ‘mother’ in Ontario Canada is given as [2.74, 2.04, 0.67] which is interpreted as quite good, quite powerful and slightly active. Whereas in the same place, the concept ‘daughter’ has an EPA of [2.18, -0.01, 1.92], which is interpreted as quite good, less powerful and more active than mother. These values are derived from ACT equations. ACT lexicons have been compiled for several countries and cultures including USA, Canada, Germany, China and Northern Ireland (Robinson and Smith-Lovin, 2006; Mejova, 2012). Additionally, ACT lexicons have also been developed for groups within societies such as religious groups (Smith-Lovin and Douglas, 1992), state troopers (Heise, 1979) and internet users (King, 2001).

Mejova (2012), showed that sentiment orientation classifiers which make use of ACT lexicons outperforms traditional SA classifiers. Mejova showed that using three variations of ACT compared to a sentiment analysis algorithm, the accuracy of the system was between 71.9% and 80.3%. The accuracy of the positive class polarity was between 64.2% and 85.7%, while the accuracy of the negative polarity class was between 77.5% and 78.1%. However, it was indicated in the experiments of Mejova (2012) that a major flaw in the ACT approach is that the ACT lexicon is limited and so does not necessarily account for all possible words that can be used to describe a situation. However, Ahothali and Joey (2015) implemented an approach for increasing the dataset or vocabulary of ACT words. Unlike the work of Mejova (2012) which focused on the sentiment of the reader, Ahothali and Joey (2015) applied ACT in analysing reader sentiment towards factual objective content. They computed reader sentiment using ACT equations and evaluated their approach against traditional SA approaches on news headlines. This resulted in a precision of between 68% and 82%. Like Mejova (2012), they also showed better performance compared to traditional SA methods.

A unique benefit of ACT is that due to the lexicons and equations obtained for each culture, using the approach in Ahothali and Joey (2015) or Mejova (2012), it is possible to predict the sentiment of a recipient in cultures for which there exists a lexicon. Nevertheless, these lexicons are limited and do not encompass all cultures or situations. More so, the lexicon is generated by empirical surveys, which involve considerable human effort and time. Therefore, while ACT clearly incorporates human centric features, there is still a gap in the research methodologies for an approach that is independent of human annotations or input and one that is not dependent on a knowledgebase of human values.

Other socio-theoretic approaches used in SA have focused on the sentiment of the writer. One such theory involves frames. Frames “capture the background knowledge that competent speakers use when producing and understanding utterances” (Ruppenhofer, 2013). The fundamental idea behind frames is that people understand the meaning of a word based on the frames they evoke and that these frames are “story fragments which serve to connect a group of words to a bundle of meanings” (Ruppenhofer et al, 2016). Ruppenhofer et al (2016) illustrates with an example, where the term avenger evokes the Revenge frame, which describes a complex series of events and the group of participants involved in the event. The knowledgebase of frames is collated from human annotations of sentences, involving the identification of possible frames expressed in the sentence and the participants. As a resource, frames capture the contextual implication of words and participants involved in the discourse and so this makes it ideal for SA related tasks. Frames have been applied in aspects of sentiment analysis including the identification of multiple opinions, identification of opinion source and opinion target (Ruppenhofer, 2013). More so, the work of Bhomwick et al (2009) which classifies the emotion of readers from sentences is the only work that was identified in this research that applies frames in recipient emotion. The emotion of readers was categorised into four classes - disgust, fear, happiness and sadness. They showed that the inclusion of word frames as feature vectors performed better than the use of just words and their POS. The overall F1 score of their approach was 82.1%. Nevertheless, the use of frames highlights the gap in the research in that the inclusion of frames in the SA methodology requires the annotation or labelling of sentences into emotion classes as was the case in Bhomwick et al (2009). By having a fixed set of emotion classes, the application is already restricted to those classes and thus unable to account for variations in emotion or even sentences that portray multiple emotions.

It is evident that both text/knowledgebase and Socio-theoretic approaches are hampered by the challenge associated with the inclusion of human involvement in their implementation. We also observe that both approaches emphasize different feature sets: The former emphasizing social and human-centric features while the latter emphasizes textual patterns and linguistic clues. The goal then is to implement a solution that addresses the issues associated with human involvement i.e. eliminate or vastly reduce the need for human annotations, tags or empirical surveys, while still incorporating human-centric social features and linguistic clues and patterns embedded in the text. In the next section, we introduce values, explore the relationship between values and sentiment, review approaches and problems associated with modeling values after which we describe the implementation of a value model.

## Model Design

The basis of the value model described in this paper, stems from the notion that although values are abstract and unseen, they are implicitly observed in human behaviour and the utterances they make. According to Entman (1993), a speaker conveying a message about an entity or subject will “select aspects of a perceived reality, and make it more salient in a communicating text, in such a way as to promote’’ a particular view point. Therefore, the verbal descriptions or linguistic units used in describing an event portray the underlying values held by the speaker towards the event or subject matter. Based on this observation, the requisite question would be what set of linguistic units, features or identifiers embedded within a sentence represent expressed values.

To answer this, we undertake a process called value decomposition which entails the identification of the constituent parameters which make up values from observed text. The necessity for value decomposition arises from the requirement to not introduce human input in the value model process. Values decomposition provides a theoretical framework for the extraction of value items and their categories. To accomplish this, we identify clues, and recurring patterns in the definition and characteristics of values and use these to define the constituent features of values. Consider the following values definitions.

Values are -

“The criteria through which people use to evaluate actions, people and events.” (Schwartz, 2006)

“Abstract coordinators of behaviour.” (Rokeach, 1973)

“Latent variables that have explanatory value for the choices people make.” (Verplanken and Holland, 2002).

“The belief that a specific mode of conduct or end state is personally or socially preferable to an opposite mode of conduct or end state.” (Rokeach, 1973)

“Determining factors for choices and a guide for determining what is desirable.” (Guth and Taguiri, 1965; Kluckhorn, 1951).

“Concepts that point out why a behaviour is acceptable or which state or behaviour is most acceptable from a set of options.” (Hutcheon, 1972).

“What is important to an individual” (Friedman et al, 2006).

“Principles encompassing abstract goals in life and modes of conduct that an individual or a collective considers preferable across contexts and situations” (Braithwaite and Blamey, 1988).

From these definitions, when values are portrayed as beliefs, we observe that the definitions refer to the ‘end state’ or ‘mode of conduct’. As concepts, it refers to ‘determining what is desirable’ which reflects a question of choice i.e. selecting from a choice of states, features or possible events etc. While as a motivation, reference is made to ‘what is important’ or the ‘motivation behind an action’. We observe from this that every value as a belief, concept or motivation is directed at an object or entity, which could be abstract or real and would have a set of states[[2]](#footnote-2), properties or features. The value could also address a state or aspect of the event/entity.

Therefore, a value is made up of the following constituent parameters:

* Value Holder () – Values are developed and applied by value holders. Value holders could be individuals or groups such as societies, clubs, political parties.
* Subject of the Value () – According to Schwartz (2006) values are criteria through which people use to evaluate actions, people and events. Therefore, a value must refer to a subject. The subject of the value is the object, entity, event, person, item, place that is referred to in the expression of the value. The subject of the value could be a real or abstract entity. Linguistically, they are typically nouns or Noun Phrases (NP).
* State () - In the definitions, above, phrases such as ‘end state’, ‘mode of conduct’, ‘what is desirable’ reflect a question of choice and preference. A person’s value for a subject will be his preference for a state of the subject; where state refers to a position, marked feature or property of the subject that is preferred amongst a set of features. For instance, on the subject ‘*house prices’*, states could include words and phrases such as ‘*high*’, ‘*low*’, ‘*expensive*’, ‘*exorbitant*’, ‘*affordable*’, ‘*stable*’ etc. (Use of these words and phrases are a form of linguistic grading). The value of value holders would be a preference for one of the states. Linguistic elements used to express states are adjectives or adverbs.
* Action () – For any subject, the preference of the value holder also extends to the preferred action or activity to be carried out or performed by or on the subject. Reference to phrases like ‘motivation’, ‘specific mode of conduct’, ‘specific mode of action’ and ‘end state’ suggest that the essence of the value also refers to an action, a conduct or an activity to be performed. In fact, the preferred state action refers to the ‘action, activity or process’ undertaken on the subject or by the subject that is most preferred by the value holder. For instance, in the sentence, “*We have a plan to destroy terrorist groups across the region*”, the preferred action on the subject ‘*terrorist group’*, is ‘*the plan to destroy*’. Using the ‘*house prices*’ analogy, actions could include phrases such as ‘*increased*’, ‘*reduce*’, ‘*rocketed*’, ‘*risen*’, ‘*plummeted*’, ‘*rebounded*’, ‘*raised*’. Actions are typically verbs or verb phrases.
* Context () – Context refers to existential factors that are usually outside the control of the value holder. These factors include elements such as time and place, background knowledge of the speaker or hearer, the expectations of people, the location and the nature of the subject matter. This list of existential factors is not exhaustive. They could also be static - factors that are fixed and unchanging such as date of birth, address etc. or dynamic – factors that are constantly changing e.g. temperature, preference, desires, social environment etc. (Henricksen et al, 2002). The value conceptualization of Braithwaite and Blamey (1988) which states that values are ‘principles encompassing abstract goals in life and modes of conduct that an individual or a collective considers preferable across contexts and situations’ reinforces the notion that the applicability of a value is dependent on the context.

In summary, these definitions show that values are made up of five parameters: the value holder (), the subject of the value (), the preferred state (), preferred action () and the context (). Together, these parameters are called Value Components (VC) and represent a formalism for translating abstract values as sentences. Thus, given a large corpus of utterances by a value holder, the value model would be a function (), that takes a text and maps it to a value representation:

To identify VCs embedded in utterances, three assumptions are made:

* The granularity of an utterance[[3]](#footnote-3) expressing a value is a sentence called a Value Laden Sentences (VLS). These are statements that impart a personal value that may not be true in the strictest sense but are based on personal opinions or values. They reflect the bias of an author or the speaker while also reflecting the priorities and ideas of the speaker.
* Each VLS can be expressed as a sequence of words.
* Consequently, a VLS is a sequence of words 1...n(where is an integer and , which express the preferred action a or preferred state s of a subject . Subjects, actions and states could be multiple words or phrases which function as a single unit. For example, the sentence *“We will reduce taxes by 2 percent in our first year in Government”* contains the subject *‘taxes’* and a multi-word action *‘will reduce’*. In the sentence *“EU Taxes will be reduced by 2 percent to take it to an all-time low”* contains a multi-word subject *‘EU Taxes’*, an action *‘will be reduced’* and a preferred state *‘all time low’*.

Based on these assumptions a structure for value laden sentences emerges. A VLS as a sequence of words made by a value holder under a particular context . Since the objective of the VLS is to express the preferred state or action on a given subject matter/s, the sequence of words which constitute the VLS must include at least one value subject ( and express at least one action and/or preferred state . For example, the sentence *“We will reduce taxes by 2 percent in our first year in Government”* contains one main subject *‘taxes’*, an action *‘reduce’* and a preferred state *‘by 2 percent’*. These core words are priority words called content words. Therefore, the VCs in the utterance are content words.

Assuming the value holder, - the speaker - is a known entity, the remaining words in the sentence belong to a second category of words called function or helper words whose aim is to connect the actions, states and subjects in a manner that conveys the intended meaning. Function words are a class of words in English grammar that are contrary to content words in that they bear no semantic relevance (Fries, 1952). They are classified as closed vocabulary words because it is quite uncommon to create new ones. They are generally restricted to 9 grammatical classes as shown in table 4.

Table 4: Function word classes and examples

|  |  |
| --- | --- |
| **Word Class** | **Example** |
| Auxiliary verbs | Am, are, be, is |
| Conjunction | Or, and, but, while |
| Determiner | A, the |
| Exclamation | Yes, No |
| Interjection/Disfluencies | Uh, em, huh, duh |
| Modals | Could, would |
| Particles | No, not, then, if, thus |
| Preposition | Of, in, at, between |
| Negation | Not, never, no |

Let actions, subject and preferred states – the lexical components of the value – as , , and the function words as , a VLS could take any of the following sample formats (The ‘’ and ‘’ signal to mark the beginning and end of the sequence) -

Sentence 1 –

Sentence 2 –

Sentence 3 –

Consider the sentence, “*We will implement more apprenticeship programmes for young people in the winter”.*

Assuming the value holder is known, the sequence of words expressed above can be represented as the sequence

, where,

*‘We’* =

*‘will implement’*=

*‘more’* =

*‘apprenticeship programmes for young people in the winter’* =

The subject *‘apprenticeship programmes for young people in the winter’* is a long phrase consisting of nested related subjects. So, to make processing and analysis easier, it is split further to show each of the related subjects and the function words linking them. Thus, the concept *‘apprenticeship programmes for young people in the winter’* is expressed as the sequence:

, where,

*‘apprenticeship programmes’* =

*‘for’* =

*‘young people’* =

*‘in’* =

*‘the’* =

*‘winter’* =

The full sentence *“We will implement more apprenticeship programmes for young people in the winter”* can thus be expressed as the sequence

〈*START*, *F1*, *λA*,*λS*, *λ* Ө*1*, *F2*, *λ Ө2*, *F3*, *F4*, *λ* Ө*3*, *STOP*〉

Following the above examples, a formal structure of VLSs emerges and it consists of a sequence of strings composed of one or more subject expressions drawn from a countably infinite vocabulary of subjects, one or more semantically relevant actions or states which are also drawn from a countably infinite vocabulary of actions and states and at least one function word drawn from a finite set of function words. We propose that the vocabulary of subjects is countably infinite because subjects could be about literally anything. Finally, all words and expressions drawn belong to a universal set of expressions.

Thus, the generation of a VLS is defined by the following parameters:

= A countably infinite set of Subjects, Actions, State, Function words and Value holder entity names and expressions (Vocabulary of vocabularies)

*λ* Ө = A countably infinite set of subject

*λS =* A countably infinite set of States

*λA* = A countably infinite set of Actions

= A finite set of function expressions and words

START = Marker representing the start of a sentence

STOP = Marker representing the end of a sentence

Where, (*λ* Ө, *λS, λA*, F) Є ∑ and {START STOP} Є ∑

Following the derivation of the formal structure, how then does a value holder generate a sentence?

We elucidate this by considering the example of a machine programmed with pro open-source software values. Imagine that this machine needs to generate a VLS that responds adequately to a statement or question e.g., a statement that challenges the ‘benefit of open source software as compared to proprietary software’ e.g. *“Should companies developing high precision software use proprietary software or open source software”*. The goal of the machine is to satisfy three conditions –

* It must generate a response that is grammatically correct.
* It must generate a sentence that fits the requisite value.
* It must generate a sentence with actions and states that are semantically relevant.

Assume the machine has access to the sets *λ* Ө, *λS, λA*, , {}.

To generate the VLS, the machine could commence with the base subject expressions *‘proprietary software’* and *‘open source software’* drawn from the vocabulary *λ*Ө. Since the resulting sentence is a sequence of words and expressions, and assuming the machine prefers the expression ‘*open source software’* to precede *‘proprietary software’* in the resulting sequence, the goal of the machine is to generate a set of words to fill the empty slots as seen in figure 5.

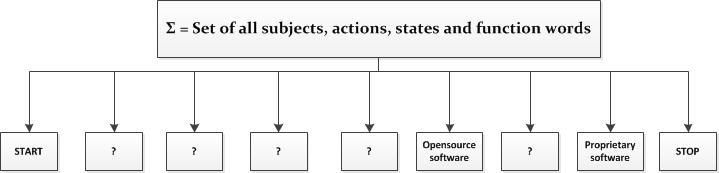


Figure 5: Illustration of Sentence generation by a machine and showing empty word slots

We know already that slots containing actions and states must be filled with expressions that are in the same semantic field as the subjects. So, the machine generates each word in the sequence by picking semantically relevant words from each vocabulary in as seen in figure 6.

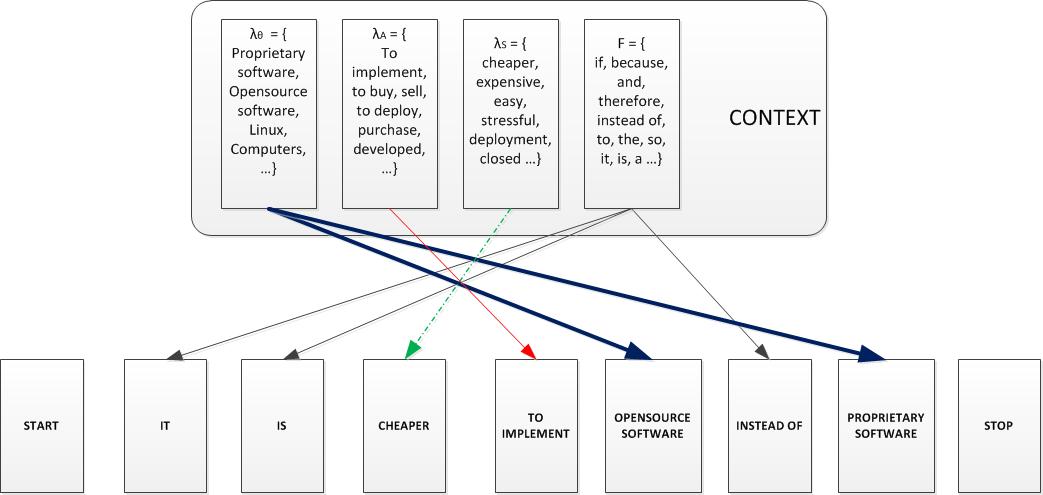


Figure 6: Illustration of Sentence generation by a machine with filled word slots

After the *START* sign, the machine selects the words *‘it’* and *‘is’* from the vocabulary of function words. To convey the notion that open-source software is preferable to proprietary, it associates the expressions *‘cheaper’* and *‘to implement’*, which are drawn from the state and action vocabulary. The functional expression *‘instead of’* plays the role of a comparator, bridging the two subject expressions. Once the machine completes the sentence generation sequence it returns the *‘STOP’* sign to mark the end of the sentence. Thus, the final sentence sequence becomes,

〈*START, It, is, cheaper, to implement, open-source software, instead of, proprietary software, STOP*〉.

This sentence is grammatically correct and satisfies the values of the machine because the actions and states associated with the subjects are semantically and contextually relevant.

However, it is possible for the sentence to be grammatically correct but have actions that fail to satisfy semantic correctness. For instance, an element in the set of actions could be the expressions *‘to marry’*, which when substituted for the action *‘to implement’* forms a grammatically correct sentence that makes no sense semantically – *“It is cheaper to marry open-source software instead of proprietary software”*. This reemphasizes the importance of the semantic relationship between subjects, actions and states, and that each subject in the vocabulary can be mapped only to a subset of semantically relevant states and actions.

In addition, the context and value holder further constrains the choice of action and state.

Therefore, in generating each word for a sentence, the machine determines the likelihood of each word/event by considering several factors including the nature of the subjects, the grammatical structure of the sequence of events etc. Mathematically, we can say that each word event generated has certain probabilities associated with it, and the probability of each word or expression is consistent with the value, grammatical correctness and context. Following this illustration, the value model development is about developing a function capable of estimating these probabilities so that for any subject under any context and for a particular value holder, a sentence can be generated. In conclusion, the process of generating sentences from values is a generative process, where words as events are generated from a vocabulary, and the probability of each event occurring in the sequence is a function of the grammatical relationship between the events, the context, the level of semantic relevance between the events and the event generator that is the value holder. In the next section, we describe how the generation of VLSs can be mathematically represented.

## Representation of Value Models

Having shown that the generation of VLSs is a generative process, our goal is to learn a value model from all the VLSs made by a value holder (). This model must be able to estimate the probability of a sequence of words in a VLS. To accomplish this task, we assume that there exists a large corpus of VLSs made by , and that each sentence in the corpus represents a distribution of possible utterances that can be made by the value holder.

The estimation of these distributions can be broken down into two related tasks.

1. Using the example of the generative machine mentioned earlier (figures 5 and 6), how can the machine correctly generate each word in the sequence so that it forms a correct sentence.
2. How do we estimate the probability of a sentence?

Since each sentence in the corpus represents a sequence of random variables, , 〉 where each random variable can take any value in a vocabulary of possible words, the probability of any such sequence of random variables can be expressed as:

where and is an element of the vocabulary for .

Based on the chain rule of probabilities, (1) above can be expressed as,

Thus answering the second question. As for the first question, we can reformulate it as a word prediction task. Assume that our machine in figures 5 and 6 was asked to generate a word to complete the sentence

‘*We will reduce ­­\_\_\_’*.

Assuming the machine had the option of three words, ‘*taxes*’, ‘*fishes*’ and ‘*riches*’. The machine could reference the corpus of sentences to find how frequently the sentences “*We will reduce taxes”*, “*We will reduce fishes*” and “*We will reduce riches*” occur. In other words, the machine makes an estimation of the likelihood of a word by looking at some reference history. This analogy can be represented mathematically as the conditional probability of a word given its history (**)**, that is .

These estimations falls under a category of statistical models called Language Models (LMs). LMs belong to a class of models called generative models that have been applied to a wide variety of applications such as hand writing recognition (Russell and Norvig, 2002), spelling correction (Kukich, 1992), text prediction and machine translation.

In LMs, a simplifying assumption based on the Markov assumption which allows us estimate the probability of each event by conditioning only on words in its immediate past is applied to equation 3. The result of this simplification means that equation 3 becomes:

)

Where, the probability is conditioned on the previous word (bigram LM), or

Where the probability is conditioned on the previous two words (trigram LM).

Although the trigram language model has been shown to produce good model estimates, the above model does not fully capture all the properties of the VLS. For instance, the above model only captures local dependencies and does not incorporate the semantic aspects and contexts of each word. In addition, the relationship between the aspects (subject, action, state) of VLSs are not captured at all as each word is only dependent on a word or words seen only in a short window span (one or two words for the bigram and trigram case respectively).

Consider the sentence sequence *“UKIP will implement more apprenticeship programmes for young people in the winter”*, the action *‘will implement’* is related to the value holder *‘UKIP’* and the subjects *‘apprenticeship programmes’*, *‘young people’*. Both subjects *‘apprenticeship programmes’*, *‘young people’* are related subjects and share a semantic relationship demonstrated through the state word *‘more’*. These relationships are not captured by the trigram model. Other types of language models like skip N-gram[[4]](#footnote-4) - where the context skips over some words so that the probability estimate becomes and variable length N-gram which support conditioning on additional contextual information might aid in addressing the problem of long distance dependencies and local context but fail to capture the semantic relationships between the value components of the sentence (Ney et al, 1994; Kneser, 1996).

To modify the LM for VLSs, the LM must capture the syntactic and semantic relationships between actions, states and subjects. It must also be tailored to the context and value holder. For now, we assume that the value holder () and context () are known entities. We can incorporate and into the model by also conditioning the probability of each word in addition to the history on and . The estimation equation becomes,

,

where is a history in the trigram

or bigram case for a VLS made by a value holder under a context .

Equation (8), must also capture the relationships between the value components that make up the sentence i.e. the subject expressions (), Action expressions , State expressions () and function words (). Using the earlier example in Figure 5 and 6, for the machine to generate an action expression or state for the subject, it must generate expressions that are semantically relevant to the subject, in other words, it must select actions or states that are in the same semantic field as the subject. Since these words (actions, states, subjects) represent the main substance of the VLS, they have priority status as the most important events in the sequence. Therefore, we assume that for the VLS to make sense, each Action, State or subject occurring in the sentence is additionally dependent on any other and , that is in direct semantic relationship with it. As for the function words, since their primary function is to connect priority expressions, they are estimated from their history alone. Thus, the value language model takes the form,

Equation (9) captures a richer informational and semantic context and the net effect is a high order LM. In summary, unlike regular LMs, the probability of each word in the VLS is if or if . In the next section, we describe how the value model is applied in estimating recipient sentiment.

## Applying Value Model to Sentiment Prediction

To apply the value model towards predicting the sentiment of a recipient, we assume that the utterance for which we intend to predict the recipient's sentiment satisfies the formal structure of VLSs.

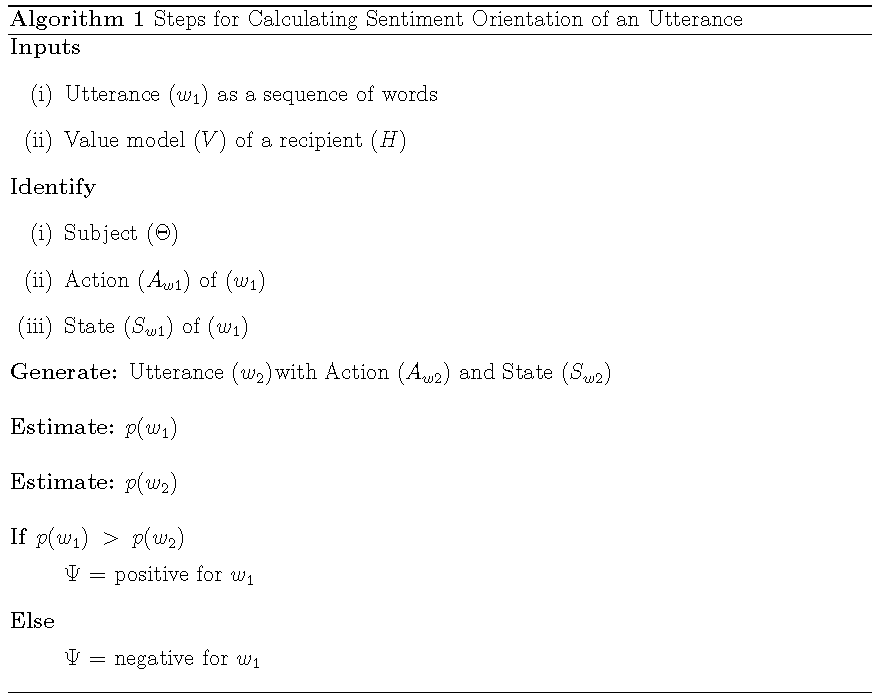
We know from value theory that for the utterance to evoke either negative or positive sentiment, the state or action on the sentence’s subject must be in line with the recipient’s preferred state or action or vice versa. Assuming we identify the states or actions in the sentence, the sentiment of the recipient can be manually determined by comparing the states and actions in the utterance against previous statements or utterance made by the recipient i.e. compare the expressed state or action against a history of the recipient’s utterance (This is synonymous to a language model). In this regard, we are also asking ‘what is the likelihood of the recipient generating an utterance with the states or actions expressed in the utterance?’. Therefore, assuming all utterances are based on people’s values, the behaviour of the recipient to a new utterance () can be predicted as a measure of how likely it is for the recipient to make the utterance .

However, this estimate does not tell us anything about the sentiment of the recipient i.e. if the utterance is positive or negative. To accomplish some sort of measurable reference for the estimate, we introduce a second assumption. Since we know the state () and action () of the subject in and we have a universal set of expressions **∑,** and a value model of the recipient, we can generate a second utterance such that the action ) or state () is completely opposite in sentiment to the state or action in the original sentence . This state/action of the new sentence must be drawn from a finite subset of **,** containingsemantically relevant state/actions associated with the subject and the newly generated sentence must also be syntactically correct. By expressing contrary action or states we attempt to generate a syntactically correct and relevant sentence that connotes sentiments opposite to the original sentence . To this new sentence , we can estimate the likelihood of it being made by the via the value language model. This process results in two probability estimates () where utterances and portray opposite sentiments.

Therefore, if , then we can infer that the recipient is more likely to make the statement and so, it means he is more likely to prefer the state and actions expressed in which will most likely result in positive behaviour or sentiment. Conversely, if , the recipient is less likely to make the statement and we conclude that his sentiment is likely to be negative, because he is less likely to utter the original statement and more likely to utter or make the new sentence with the opposite sentiment. In other words, the recipient is more likely to prefer the state or action expressed in the new sentence .

Based on this, the sentiment () is a measure of the difference between the probability estimate of the actual utterance () and the probability estimate of a new utterance () depicting a state or action that is opposite in sentiment to the initial utterance.

Another key aspect of this approach to recipient sentiment prediction is the ability to detect the intensity of the sentiment as a measure of the difference between and . The margin of the difference between the probability likelihoods can allow us infer sentiment orientations like extremely negative, negative, positive, extremely positive. The steps in algorithm 1 provide a stepwise guide to predicting the sentiment of an utterance.



## Implementation

In implementing the value model, a corpus of spoken or written content is required as the source of values. The domain of interest is politics, and data is sourced from speeches, policy manifestos and debates in compiling training and test corpus. Specifically, the implementation focuses on two timely topics in UK politics and they are *‘Immigration’* and the ‘*European Union (EU)’*. Our goal is to build a model of values for these two subjects such that the model is representative of three major UK political parties (value holders). Afterwards, we apply these models towards predicting the sentiment of the value holders on sentences. The political parties are The Conservative party, Labour party and Liberal Democrats (LD).

For the sake of this implementation, the VC, context (C) which is ambiguous and able to take an innumerable number of forms (Henricksen et al, 2002) is the subjects *‘EU’* and ‘*Immigration*’. We can view these domain subjects or contexts as conditions or scenarios. For example, the United Kingdom Independence Party (UKIP) manifesto, suggests that in the context of the EU, they are opposed to unlimited relationship and uncontrolled migration with and from the EU but open to migration and more relationship with the Commonwealth. Whereas, in the context of Immigration, they advocate limited migration into the UK from anywhere (Europe and outside Europe). So, in one topical subject context, they support migration and in another they are opposed to it.

Although the implementation focuses on value holders from the three parties mentioned, we also downloaded documents associated with UKIP. This UKIP corpus is used in evaluating the model of the three parties. The idea behind this evaluation is that UKIP views on the EU and Immigration are very well established and not as unclear as the three main parties and as such we should be able to compare similarities in the value orientation and sentiment of each of the three political parties in relation to UKIP policies and sentiment. Finally, for this implementation, we introduce an expression ‘context-party’ pair, which refers to the corpus of training data for a party or value holder under a context. For example, the expression ‘EU-Labour’ pair refers to data by a Labour value holder under the EU context.

Therefore, implementation data consists of two categories. The first category consists of relevant policy documents, manifestos and reports related to the two subjects (Immigration and EU). The second was drawn entirely from Parliamentary debate transcripts - Hansard - covering the periods between 2010 and 2015.

### **Data Preparation**

Data collected for the four political party is grouped by the subject. For each context-party pair, e.g `EU-Labour', `Immigration-Labour', the data was prepared. The first part of data preparation is Content renaming. It is not uncommon for relevant VCs observed in training to be non-existent or have a low count in the test set. To account for such low frequency or uncommon VCs, the objective of content renaming is to identify and rename rare or unseen semantically relevant lexical units. These units are typically VCs such as names, titles, locations and product types etc.The renamed linguistic units include:

* Names -- This includes the names of persons. In the political corpus, person names are quite prevalent. People are identified via their titles e.g. `The Honourable MP', pronouns e.g `he', or by their names, e.g. `Theresa May'. Using a compiled gazetteer and ontology of all UK MP names as well as a named entity transducer, all identified names are mapped to the pseudo-word `PERSONNAME'. Similarly, abbreviations e.g. `NATO' and `TTIP', Locations e.g. `Manchester' and alphanumeric named entities which are commonly products or even places e.g. 'A40' are respectively mapped to the pseudo-words `LOCATIONNAME', `ABBREVIATIONNAME' and `ALPHANUMERICNAME'.
* **Dates, Numbers and Currency -** All date expressions are assigned the pseudo-word ‘FIGUREDATE’. These include expressions such as ‘*20th century’*, ‘*12th of August*’, ‘*July*’, ‘*12-11-2001*’. As for numbers and currencies, this category includes numbers expressed either as words or Arabic numerals. Since numeric elements can range from , we assign a single generic pseudo-word for all numbers, currencies and percentages that are less than to be ‘FIGURENEGATIVENUMBER’ for numeric expressions, ‘FIGURENEGATIVEPERCENT’ for percentages and ‘FIGURENEGATIVEMONEY’ for currencies. For numeric expressions, greater than we use a naming convention that combines the expression ‘FIGURE’ followed by the number of digits expressed in words and the category ‘PERCENT’, ‘MONEY’ or ‘NUMBER’.

Since the unit of analysis is a sentence, each document is converted corpus to a corpus of sentences so that Our training and test set will be made up of a collection of sentences. Table 19 shows the original number of documents prior to pre-processing while table 20 shows the total number of sentences in the corpus associated with each party and domain after data preparation. Table 21 portrays the total number of tokens (\emph{N}) in each training corpus. The comparative difference in corpus and token size between Conservative/Labour and the Liberal Democrats is accounted for by the comparative size of the parties and the number of MPs. Since Labour and Conservative party have more MPs, it is expected that they would make more contributions, thus a larger data set.

Table 19: Number of documents extracted for domains

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **YEAR** | **Conservative** | **Labour** | **LD** | **UKIP** |
| **2010** | 25556 | 37108 | 1001 |  |
| **2011** | 28003 | 19011 | 3103 |  |
| **2012** | 30965 | 17882 | 2782 |  |
| **2013** | 32517 | 18429 | 2840 |  |
| **2014** | 30151 | 16407 | 2298 |  |
| **2015** | 29532 | 14182 | 1048 | 22 |
| **Total** | **176724** | **123019** | **13072** | **22** |

Table 20: Number of sentences post-data preparation

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Data** | **Cons. EU** | **Cons. Immigration** | **Lab. EU** | **Lab. Immigration** | **LD EU** | **LD Immigration** |
| **Train** | 238135 | 194838 | 165769 | 135628 | 20132 | 16470 |
| **Test** | 85050 | 69585 | 59203 | 48439 | 7190 | 5883 |
| **Dev** | 17008 | 13918 | 11840 | 9688 | 1439 | 1177 |
| **Total** | **340193** | **278341** | **236812** | **193755** | **28761** | **23530** |

***Table 21: Total Number of Tokens across training corpus***

|  |  |  |
| --- | --- | --- |
| **Party** | **EU** | **Immigration** |
| **Conservative**  **Labour**  **Liberal Democrats** | 2974827  2488436  385171 | 2577704  2355246  356076 |

### **VC Identification and Implementation**

Having collected the corpus of sentences, we commence with the identification of VCs. The value holders $H$ are known, as the author of each sentence or document was mapped to his/her party. As such $H$ are collectively viewed as the parties themselves. The contexts $C$ was described earlier to be the subjects of interest i.e. `Immigration' and `EU', which leaves the action $A$, state $S$ and subject $\theta$.

A n{\"a}ive approach to identifying content and function words would have entailed identifying all function words in the corpus so that the remainder would be content words. Conversely, we could identify all content words i.e. NPs, Nouns, verbs, adjectives and adverbs in the corpus so that what we are left with are the function words. However, simply identifying these elements was not informative enough for the model implementation since it did not reflect the relationship or dependency between any of the linguistic units nor portray the primacy of the content words either as independent units or when compared to other units.

Having provided the reasons for using political data, we proceed to discussing the implementation setup.

The first section describes the implementation of a set of processes designed to prepare the data for building the model.

The second part addresses the extraction of value parameters from sentences. Chapter 5, showed how syntactic clues aid the identification of value parameters and highlighted the importance of capturing and introducing the relations between value parameters in the model. This section, describes how these relationships are captured and represented by introducing the concept of Dependency Grammars (DG) - a grammar formalism that plays a major role in how subjects, actions and states are identified and the relationships between them. Applying a well-established theory like DGs enables the automation of the value component extraction process that is independent of the subject domain or document type.

The third part addresses the implementation of the value model and its adaptation for sentiment analysis. This section is divided into two: The first part focuses on the implementation of the value language model, illustrating how the rich context of the model is implemented combining language models and a maximum entropy (maxent) classifier. The second part demonstrates the implementation of sentiment prediction via the LM and maxent classifiers through a method we’ve termed feature switching. Figure 9, acts as an illustrative stepwise guide for the implementation beginning with the document preparation and terminating with the prediction of recipient sentiment.

1. Piece of text describes all forms of written and spoken text including utterances, documents, phrase, sentences. [↑](#footnote-ref-1)
2. State refers to a mode which the object can exist in [↑](#footnote-ref-2)
3. Utterance and sentence are used interchangeably. Utterances are vocalized sentences. [↑](#footnote-ref-3)
4. Language models are also called N-grams [↑](#footnote-ref-4)