Opinion analysis: the effect of negation on polarity and intensity

Lei Zhang, Stéphane Ferrari, Patrice Enjalbert

Opinion analysis: the effect of negation on polarity and intensity

Lei Zhang, Stéphane Ferrari and Patrice Enjalbert
GREYC lab - CNRS UMR 6072
University of Caen Basse-Normandie
`FirstName.LastName@unicaen.fr`

Abstract

This paper presents an ongoing work addressing the problem of opinion analysis. It takes part into a collaborative project with industrial partners, aiming at providing a professional with a help for strategic and technical intelligence. Thus, we focus on local semantic analysis rather than text or sentence classification. The purpose of our task-oriented approach is to characterize the properties of opinion statements in an applicative corpus. Inspired by linguistic models, the method we propose is a compositional one, consisting in detecting and analyzing valence shifters such as negation which contribute to the interpretation of the polarity and the intensity of opinion expressions. We describe our model and its first implementation before discussing the results of a proof-of-concept experiment focusing on adjectival expressions.

1 Introduction

Opinion mining and sentiment analysis have been fields of high interest for the NLP community in recent years. There are many different applications in view, such as strategic intelligence, reputation management as well as automatic gathering of customers opinions and expectations. Our work aims at providing a front-end user, a professional in strategic and technical intelligence, with a help for spotting, analyzing and comparing opinion statements in her corpus. This project, *OntOpiTex*, is a collaboration between institutional research labs and industrial partners.[1] A great part of related works focus on text or sentence classification, proposing methods to spot subjective discourse, and differentiating between positive and negative opinions. Other works studied the constitution of corpus-dependent subjective lexicons. These now well established approaches could have direct applications, but they still need to be completed in order to help a professional user analyze a specific domain. Therefore, the purpose of our project consists in providing a semantic analysis of opinion statements, describing their characteristic properties. In this paper, we focus on two of them, polarity and intensity. We propose a model inspired by linguistic approaches, and present a related experiment.

In the next section, we briefly introduce the generic scope of the *OntOpiTex* project, describing the tasks in view as well as the applicative corpus. In section 2, we present the related works, both in linguistics and in NLP area, paying special attention to the Appraisal theory (Martin and White, 2005), which greatly inspired our approach. In section 4, we describe our method for local semantic analysis of opinion statements and the current implementation of our model. While polarity characteristic is viewed in relation with the 'Graduation' property, the intensity of sentiment statements, is described in terms of 'Force' and 'Focus'. The approach is a compositional one, taking into account different kinds of modifiers and negation. Section 5 details a proof of concept experiment restricted to the class of the adjectives. We also present and discuss the results obtained on the applicative corpus and on a similar corpus. In conclusion, we propose a few perspectives, mainly with regard to the integration

[1] The project and the different partners are anonymous for the submitted version.

of our work in the whole application in view.

2 Scope of the project, task in view

The *OntOpiTex* project is an interdisciplinary one, involving computer scientists and linguists from three diferent laboratories as well as industrial partners, *Noopsis* and *TecKnowMetrix*. Its purpose consists in designing back-end and front-end tools for a professional use in technical strategic intelligence tasks. The whole architecture can be decomposed in two main parts: a set of tools and resources for the automatic analysis and different graphical user interfaces composing a dashboard for the front-end part. In a broad outline, researchers are mainly involved in the linguistic, NLP models and tools, *Noopsis* partner develops the generic architecture and the front-end application while *TecKnowMetrix* provides the applicative frame and the use case for an extrinsic evaluation.

The line of business of *TecKnowMetrix* concerns competitive intelligence in the technology domain. Their experts have to analyze two kinds of texts: patents for dealing with the legal concepts related to intellectual property, and technical journals, in order to analyze furthermore the activity in a specific domain. The *OntOpiTex* project is limited to the latter, where opinion or evaluation statements may have a better chance to occur. The use case in view is the analysis of the competition in the domain of a client, for comparison or position purposes. The applicative corpus is related to the avionic technologies, w.r.t. Boeing and EADS/Airbus companies. It consists in 377 journalistic texts from economics and technical press in French language, representing around 340 000 words.

As expected, opinion expression is not the main characteristic of such a corpus. However, in our task-oriented application, the professional user knows which targets are of any interest w.r.t. her task. The axiological or evaluative dimension of a statement is mostly a user-centered notion, related to the interpretation process. This is one reason why the resources used for our experiments are currently built to satisfy this constraint, merging generic lexicons (e.g. denoting affect) with very specific ones (e.g. describing technical properties of airplanes). Therefore, local semantic analysis of evaluative statements is proposed to complete any statistical analysis that could be realized on positive and negative tendencies in the corpus. Furthermore, comparing the companies products and activities leads to pay special attention to the intensity of evaluation. For this purpose, we focus on the role of negation and intensity modifiers to detect polarity and intensity variations.

3 Related work

3.1 Opinion mining and sentiment analysis in NLP

Studies in opinion mining and sentiment analysis may be roughly classified in one of the three following tasks: (i) lexicon building, (ii) text or sentence classification and (iii) opinion statements analysis.

(i) The study of Hatzivassiloglou and McKeown (1997) is one of the early works aiming at lexicon building. The authors present a way to determine the orientation of the conjoined adjectives based on constraints on conjunctions. Different approaches use *seeds lists* to initiate a lexicon extraction from corpus, as in (Turney and Littman, 2003). Esuli and Sebastiani (2006) exploited WORDNET to develop the SENTIWORDNET extension, associating two properties (namely subjectivity and semantic polarity) to *Synsets*.

(ii) Text and sentence classification is generally viewed as a binary task: objective *vs* subjective or positive *vs* negative. Most studies are based on data mining and machine learning techniques. Many text genres have been studied. Movie review may be the earliest and most common one (Turney, 2002; Pang and Lee, 2004). Recently, the studies massively focus on customer reviews, weblogs and twitters (Breen, 2012; Singh et al., 2012). Recent approaches such as (Lambov et al., 2010) also propose models for reducing the domain dependence of subjective texts classifiers.

(iii) The last task, opinion statements analysis, consists in determining complementary features. Hu and Liu (2004) study the consumer reviews of product's technical features, pointing the targets of opinion statements as well as their polarities. Our study is included is this third task, based on the *Appraisal* theory (described in 3.3).

3.2 Main approach in linguistics

The French linguistics is currently missing a unified theory for the notion of evaluation. To synthesize, this notion has been considered in studies on subjectivity according to three main points: (i) the study of some modal values (appreciation, evaluation,...); (ii) the analysis and compilation of subjective lexicons and (iii) the study of some related notions in linguistics such as the point of view, the engagement and the endorsement. Different systems of modal values have been proposed by the linguists (Charaudeau (1992), Culioli (1980), Kerbrat-Orecchioni (1999)) with more or less emphasis on those coming under the subjectivity of the speaker (potentially the opinion holder).

The situation in English linguistic is different: opinion has been studied in a more systematic manner. For instance the collective work (Hunston and Francis, 2000) is largely oriented on text analysis. More precise studies (Hunston and Sinclair, 2000) focus on the notion of local grammars of evaluation. Three books propose a complete model of evaluation. The first, Evaluative semantics (Malrieu, 2002) present a model based on a cognitive model of evaluation. The second, Appraisal in media discourse (Bednarek, 2006) present a model based on two main categories: factors and values. The factors are the different appraisal fields (comprehensibility, emotivity, expectedness, importance...). This model is applied in order to classify the evaluations specific of the trustworthy newspapers from the ones specific of the tabloids. The third book, Appraisal in English (Martin and White, 2005) proposes a theory detailed more precisely in next section 3.3.

3.3 Appraisal theory

Appraisal theory is a relatively recent linguistic theory, elaborated for English language.[2] It proposes a complex system for describing the properties of opinion statements according three main aspects: Attitude, Engagement and Graduation (see figure 1).

Attitude itself divides into three sub-systems: Affect, Judgment (human behavior) and Appreciation (objects and products). Attitudinal meaning can be clearly conveyed by individual words – for example, 'angry', 'brave' and 'beautiful'. But most of the times, it is not individual words but words combinations which convey Attitude – for example, 'his election [is] an affront to the democratic principle'. Therefore, though individual words may be 'attitudinal', it is better to see Attitude as a feature or property of complete statements and of stretches of language which present a complete proposition or proposal. In the Appraisal theory, Polarity is part of the Attitude system.

Engagement concerns the intersubjective dimension, linked to linguistic marks which explicitly position texts proposals and propositions inter-subjectively. For example: modals of probability – 'perhaps', 'I think...', 'surely'; expectation – 'predictably', 'of course', etc.

Graduation involves two dimensions : 'Force' (variable scaling of intensity) and 'Focus' (sharpening or blurring of category boundaries). It can apply to both Attitude (graduating the opinion itself) and Engagement (graduating the endorsement). For example: force – 'very', 'completely', 'rather'; focus – 'a kind of', 'effectively', 'a true friend'.

In (Whitelaw et al., 2005), Appraisal groups used for sentiment analysis are described in terms of Attitude, Polarity, Force and Focus. The authors use the example of *not very happy* to illustrate the role of negation and intensity modifiers in such groups. This example is discussed further in the next section.

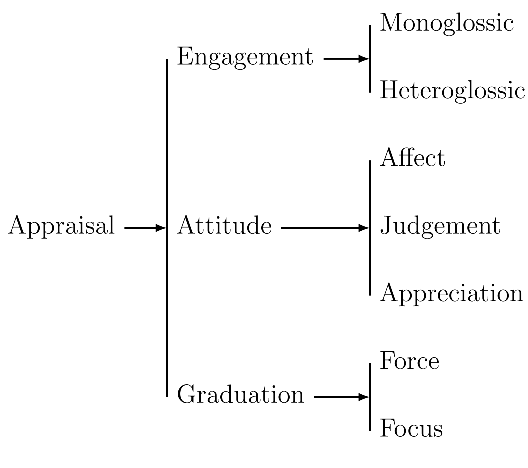

Figure 1: the Appraisal framework (simplified)

[2]See http://www.grammatics.com/appraisal/ for an outline.

3.4 Negation and graduation

Valence shifters play a crucial role in sentiment analysis. Among different aspects of valence shifters, three types are frequently considered: negation, intensifiers and diminishers.[3]

Negatives are the most obvious shifters that affect polarity and force/focus. How 'not' can flip the valence of a term has been discussed in several works: (Pang et al., 2002), (Hu and Liu, 2004), etc. Wiegand et al. (2010) propose an interesting survey on the role of negation in sentiment analysis. According to the authors, negation is highly relevant for sentiment analysis. However, negation is a complicated phenomenon, and despite the existence of several approaches to negation modeling, current models are still incomplete (Wilson et al., 2005).

Intensifiers and diminishers are also important valence shifters. They can belong to all open lexical classes. The calculation of polarity modified by negatives or intensifiers/diminishers has often been done separately. Whitelaw et al. (2005) chose to reverse the force and the polarity in the context of conjoined negation and intensifier: 'very happy' (polarity:*positive*, force:*high*) $\rightarrow$ '*not* very happy' (polarity:*negative*, force:*low*). In our opinion, for French language at least, this result depends the force level (low, high, extreme), as presented in the following section 4.

4 Model for graduation analysis

Under Graduation, our concerns cover two dimensions: Force and Focus.

Force Force has to do with the intensity of a word or expression – *assez* "rather", *très* "very", etc. It includes quantity and proximity modifiers – *un peu* "a few", *beaucoup* "many"; *quasiment* "almost", *presque* "nearly", etc. It should also be noted that this principle of force grading operates intrinsically across values of attitude in the sense that each particular attitudinal meaning represents a particular point along the scale of low to high intensity. For example, some adjectives represent the highest scaling such as *extraordinaire* "extraordinary", *magnifique* "brilliant", *énorme* "huge", etc. In addition, some prefixes can be interpreted as 'very': super-, hyper-, extra-, etc.

[3] In our work, Graduation covers diminishers and intensifiers.

In our system, we consider five different values of Force: **low** – *un peu* "a little", **moderate** – *moyennement* "fairly", **standard**, **high** – *très* "very" and **extreme** – *extrêmement* "extremely". An adjective without intrinsic intensity label is assigned the 'standard' value by default.

Figure 2 describes how negation acts on force and polarity of a word or expression. Our model of negation is inspired by French linguists (Charaudeau (1992), Muller (1991), etc.). We chose to present this model with the opposite pair *bon* "good" $\leftrightarrow$ *mauvais* "bad", because the antonymy relation characterizes most gradable words, placing them on a scale of values from a negative extremity to a positive one.

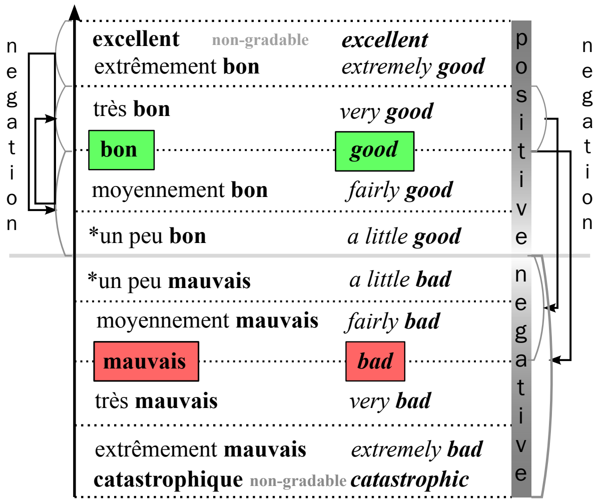

Figure 2: Our model of negation and graduation

The rules for negation and positive polarity are the following:

- when acting on 'extreme' values, force is lowered and polarity is preserved (not extremely good ≡ a little good);
- when acting on 'high' values, force is lowered and polarity is reversed (due to euphemism in discourse – not very good ≡ a little bad);
- when acting on 'moderate' or 'low' values, force is raised and polarity is preserved (not a little/fairly good ≡ very good);
- when acting on 'standard' values, polarity is reversed but force stays uncertain.

The first rule concerning 'extreme' values was established through an experiment:[4] two experts

[4] Detailed in another submitted paper.

annotated 125 sentences containing axiological adjectives in a negative context. The disagreement (4 %) over the polarity is mainly due to some adjectives with extreme value in negative context such as "n'est pas catastrophique" *is not catastrophic* (probably in relation to the effect of rhetoric).

These rules also apply on negative polarity with one exception: Muller (1991) points out that when negation acts on 'standard' negative values, the degree of value can be anything but 'standard' in the positive pole. More precisely, *pas mauvais* "not bad" means whether *moyennement bon* "fairly good") or *très bon* "very good" but not just *bon* "good".

For applicative purposes, all values have to be instantiated in the implementation. Therefore, the last rules are not exactly respected, a 'standard' value being currently attributed to uncertain cases (in these particular cases, standard is viewed as the mean of the possible values): not good ≡ bad, not bad ≡ good.

Focus In the Appraisal theory, Focus is used to intensify not gradable categories: modifiers such as 'true', 'pure' or 'sort of' sharpen or soften the belonging to a category, therefore intensifying the associated axiological value, if any.

In our model, we also consider focus modifiers with the two possible values 'sharpening' or 'softening'. The following rules are used in order to combine them with the other modifiers:

- when acting with a non-standard Force, a sharpening focus modifier is interpreted as a force intensifier, and a softening focus modifier as a force diminisher, the previous rules for Force are applied;
- when acting with a standard Force, Focus is reversed by negation: SHARPEN (central) ↔ SOFTEN (peripherical).

For example, in *les résultats ne sont pas vraiment très bons* "the results are not really very good", 'vraiment' is computed as a force modifier:
'bon':(force:standard..polarity:pos)
→'très bon':(force:high..polarity:pos)
→'vraiment très bon':(force:extreme..polarity:pos)
→'pas vraiment très bon':(force:low..polarity:pos)
≡ "a little positive"

In *les résultats ne sont pas vraiment bons* "the results are not really good", the same word is computed as a focus modifier:
'vraiment bon':(focus:sharpen..polarity:pos)
→'pas vraiment bon':(focus:soften..polarity:pos)
≡ "hardly positive"

In *les résultats ne sont vraiment pas bons* "the results are really not good", the order of the modifiers changes the final interpretation:
'pas bon':(focus:standard..polarity:neg)
→'vraiment pas bon':(focus:sharpen..polarity:neg)
≡ "truly negative"

5 Implementation and first results

5.1 Implementation: broad outline through an example

The previous model is currently implemented in an IDE, using *Noopsis* plugin to the Eclipse environment. Different built-in modules help preparing texts to analyses: extraction of textual parts from the XML source texts (XSLT module), tokenizers for words and sentences (mostly regular expressions), POS tagging. Different lexicons can be designed via this environment (XML format, associating feature sets to forms or lemma entries, including multi-words expressions) and projected on the texts. A chunking module has also been developed for the purpose of the project, producing simple chunks as well as 'groups' of chunks (ideally syntactic groups). The entries for our core analyzer can therefore be described as hierarchic structured texts (texts ⊃ paragraphs ⊃ sentences ⊃ groups ⊃ chunks ⊃ words), with a feature set associated to each unit.

The core analysis, mainly implemented in prolog, consists in the three following steps:

1. projecting resources on the texts, both lexicons for opinion words or expressions and lexicons for negation and graduation modifiers;
2. filtering opinion sentences (presence/absence of an opinion word);
3. analyzing filtered sentences.

The latter step produces a feature set for each opinion word (i.e. issued from the corresponding resource), activating the opinion analyzer. If word is ambiguous, e.g. potentially both modifier and axiological, this activation depends on the local context. For instance, in *une belle réussite* “a great success”, though potentially axiological, the adjective ‘belle’ is here considered as a modifier, intensifying the noun ‘réussite’, no opinion analysis will be activated for the adjective itself.

The opinion analyzer is a set of (prolog) rules, adapted to the POS tag of the initiator. When activated, it creates a new feature set added at the sentence level. Three main features are created: **init**, describing the word or expression which activated the analysis, **tgt**, describing the part of the sentence which contains information related to the target of the opinion, and **graduation**, indicating the polarity, force and focus. Lets consider the following example:

> *≪ Trouver les ressources nécessaires pour l’A320 NEO n’a pas vraiment été chose facile ≫, a déclaré Tom Enders dans un communiqué.*
>
> “Finding the necessary funds for A320 NEO was not really an easy thing”, stated Tom Enders in a press release.

Figure 3 and following show different partial views of the feature sets created. The output is serialized in an XML format where features are represented by XML elements and features values are their textual content.

```
– <annotation type="sentence">
  – <text>
      « Trouver les ressources nécessaires pour l'A320 NEO n'a pas
      vraiment été chose facile », a déclaré Tom Enders dans un
      communiqué.
    </text>
  – <features>
      <subjective>true</subjective>
    – <opil>
      + <init></init>
      + <tgt></tgt>
      – <grad>
          <polarity>positive</polarity>
          <force>STD</force>
          <focus>SOFTEN</focus>
        + <from></from>
        </grad>
      </opil>
    </features>
  </annotation>
```

Figure 3: Example of feature set

The **init** feature (*initiator*) contains information about the lexicon entry. Here, the adjective *facile* “easy” initiated the analysis. The lexicon gives information only about its polarity (‘positive’) and force (‘standard’) when used in an axiological way. For this specific example, the initiator is not known as a potential modifier, so the opinion analyzer is automatically activated.

The **tgt** feature is an intermediary result, spotting the words related to the opinion target, to be combined with a domain ontology and an anaphora resolver (out of the scope of the present paper). It also indicates which rule has been applied for analyzing the initiator. In figure 4, an adjective activated the *AdjInGN* rule, which looks for a name in the same chunk to spot the target, and which also explores the context in order to find potential negation and modifiers before the chunk.

```
– <tgt>
    <rule>AdjInGN1</rule>
    <t>chose</t>
    <lemma>chose</lemma>
    <sg>[ chose facile ]</sg>
  </tgt>
```

Figure 4: Information about the target

Each rule has its specific exploration process, a systematic local exploration (inside the chunk), and a contextual exploration which depends on the rule. For adjectives, we currently use 4 specific rules and one default (X_{init}: element (here only adjectives) activating the analysis; X_{focus}, X_{force}, X_{neg}: focus, force or negation modifiers $[G]_{context}$: statement explored outside the chunk boundaries):

- AdjInGN: an epithet, inside a nominal phrase – the context is explored to identify a possible attributive verb before. For example: It is [really$_{focus}$ not$_{neg}$]$_{context}$ an [important$_{init}$ contract$_{target}$]$_{NP}$.
- GNGVAttrGAdj: attribute of a subject, with an attributive verb already identified. For example: [The model]$_{target}$ is [particularly$_{focus}$ innovative$_{init}$]$_{AP}$.
- GNGAdjAppo: affixed adjective phrase – no context exploration.
- GAdjAppoGN: affixed adjective phrase before the qualified nominal phrase – no context exploration. For example: [Very$_{force}$ ergonomic$_{init}$ and comfortable$_{init}$]$_{AP}$, [the new model]$_{target}$...

- GAdj: default, no target, no context exploration.

For each rule, the embedding group is explored to build a list of close modifiers (including negation). The context exploration process may produce a second list of modifiers (including negation) found in the verbal group and appended to the previous one.

The **graduation** processed is based on the initiator description and the (possibly empty) list of modifiers. Figure 5 shows the result for the previous example.

```
- <grad>
    <polarity>positive</polarity>
    <force>STD</force>
    <focus>SOFTEN</focus>
  - <from>
    - <mod_int>
        <lemma>pas</lemma>
        <dir>none</dir>
        <force>false</force>
        <focus>false</focus>
        <neg>true</neg>
      </mod_int>
      <polarity>positive</polarity>
      <force>STD</force>
      <focus>SHARPEN</focus>
    - <from>
      - <mod_int>
          <lemma>vraiment</lemma>
          <dir>SHARPEN</dir>
          <force>false</force>
          <focus>true</focus>
          <neg>false</neg>
        </mod_int>
        <polarity>positive</polarity>
        <force>STD</force>
        <focus>STD</focus>
      </from>
    </from>
  </grad>
```

Figure 5: Graduation feature with embedded 'from' features

A **from** feature helps understanding the applied modifications (and checking the modifiers list): *pas* (not) applied to *vraiment* (really) applied to the initial adjective.

In the project, this feature is produced to be used by a following module dedicated to discourse level, in order to allow revisions. When two statements are incoherent, negation interpretation may be revised (out of the scope of the present paper). Indeed, the modifiers are combined as proposed in our model, but we had to choose which force value to associate with negation rather than presenting the final user with a possible choice between multiple values (the two cases of negation acting on standard values, with positive or negative polarity).

Next section presents an experiment applying this implementation for adjectives only.

5.2 First results

The applicative corpus is constituted of 377 texts, mainly economics articles from the French newspaper *Les Échos*. Its size corresponds to the mean size of corpus *TecKnowMetrix* currently deals with to handle other real cases. Texts are all about at least one of the two avionic companies Boeing and EADS/Airbus.

As already observed in many previous works, adjectives are the most frequent forms used for producing opinion statements. Due to the small size of our applicative corpus, this first experiment has been limited to this category. The lexicon was built by experts observing the occurences of adjectives used in the corpus(Enjalbert et al., 2012). It consists in 283 adjectives, some related to the generic categories of Attitude proposed in Appraisal, some specific to the application domain.

2323 opinion statements have been processed on the whole corpus. Only 1 feature set was created in 1755 sentences, 2 feature sets in 225, 3 feature sets in 34 and 4 feature sets in 4. In other words, 87% of the subjective sentences contains one opinion statement only (involving an opinion adjective). Table 1 shows raw results: for 2323

type of unit	*# occ.*
analyzed sentences	2018
created feature sets	2323
feature sets with modifier	365 (15.7%)
negation in feature set	85 (3.7%)
feature sets with >1 modifier	6 (0.3%)

Table 1: Raw results

analysis, a modifier list was built 365 times, 85 times involving a negation. Only 6 occurrences of more than one modifier were found. The example described in the previous section is one of them, showing a combination between negation and a focus modifier. The other ones are combinations between a negation and a force modifier, like the following:

≪ *Il y a encore moins d'un an, Airbus*

n'était pas très favorable au GTF, reconnaît David Hess.≫

"Less than one year ago, Airbus was not very disposed to GTF, admits David Hess."

Our tool produced the following feature set for the previous: (polarity:negative..force:LOW..focus:STD), taking into account *pas*, *très* acting on *favorable*. Due to the size of this applicative corpus, statistics are not relevant.

We built a similar corpus from the French newspaper 'Le Monde', extracting all articles (6227 articles) about EADS/Airbus or Boeing from years 1987 to 2006 in order to evaluate our approach.

Recall could not be evaluated, because of the rarity of the studied phenomenon: 6 examples found in a 340 000 words corpus. It must also be noted that the lexicon used in our experiment is not exactly designed for the second corpus. Articles in specialized press are written for professional readers, while a daily paper like 'Le Monde' adresses a larger readership. Different opinion adjectives should also be added before evaluating recall in this context.

In order to evaluate precision, we make the assumption that the processing of intensity and polarity does not depend on the context. We observed the opinion statements involving the same adjectives in the new context, focusing on the most elaborated expressions. In a sample of examples combining negation and modifiers (65 segments: sentences or paragraphs), 83.9% are correctly analyzed, w.r.t our model. The main errors are due to tagging or chunking problems (6.5%), difficulty to take punctuation into account when in the context exploration process (4.8%), temporal aspects combining with axiological expression (3.2%) (for exemple, *jamais aussi actif* "never so active"), as well as insufficient resources (*rien de* "nothing", *ni* ... *ni* "neither ... nor") for the remaining.

The adjective roles (epithet, attribute, affixed) allowed rather simple and limited context exploration strategies. We currently are generalizing the model to nouns, verbs and adverbs categories, using the same rules for Graduation computing. Context exploration strategies are not as easy to design as for adjectives. This work is realized in collaboration with the linguist partners who identified relevant patterns, which may be more specific of the genre of the applicative corpus.

6 Conclusion

In this paper, we addressed the problem of opinion analysis. We first described the application in view, a help for strategic and technical intelligence, and the related use-case corpus: articles related to EADS/Airbus and Boeing companies issued from specialized press. The constraints of our task led us to focus on local semantic analysis of opinion statements, with special attention on their polarity and intensity for comparison purposes.

Relying on the Appraisal theory (Martin and White, 2005) and other linguistic studies, we proposed a model for computing the values of Force and Focus, two Graduation characteristics reflecting intensity, w.r.t. negation and other modifiers. The implementation of this model is a specific module in the whole project application. We detailed its behavior regarding adjectives analysis. The results on a small applicative corpus are not relevant, because the most elaborated rules are barely activated: only 0.3% of the outputs combine at least two modifiers. An experiment on a similar corpus has therefore been realized in order to evaluate the accuracy of these rules. With a current precision of 83.9%, we consider the module not accurate enough for professional purposes.[5]

However, this first experiment allowed us to identify the remaining problems. Tagging errors and wrong contextual analysis are the most frequent errors encountered. Our further works will focus on the integration of our module in the whole application. In this scope, we ought to take advantage of other analysis, which may correct some of the current errors: domain ontology and terminology as well as domain specific patterns established by linguistic partners should improve the tagging and context exploration around opinion statements. We plan an extrinsic evaluation after integration, with returns from the final user on the accuracy of the whole application.

[5] If not totally bad, this precision means more than 1 error every 10 results.

References

M. Bednarek. 2006. *Evaluation in media discourse: analysis of a newspaper corpus*. Continuum Intl Pub Group.

J.O. Breen. 2012. Mining twitter for airline consumer sentiment. *Practical Text Mining and Statistical Analysis for Non-structured Text Data Applications*, page 133.

P. Charaudeau. 1992. *Grammaire du sens et de l'expression*. Hachette París.

A. Culioli. 1980. Valeurs aspectuelles et opérations énonciatives: l'aoristique. *La notion d'aspect*, pages 181–193.

P. Enjalbert, L. Zhang, and S. Ferrari. 2012. Opinion mining in an informative corpus: Building lexicons. In *PATHOS 2012*.

A. Esuli and F. Sebastiani. 2006. Sentiwordnet: A publicly available lexical resource for opinion mining. In *LREC'06: Proceedings of the 5th Conference on Language Resources and Evaluation*, pages 417–422.

V. Hatzivassiloglou and K.R. McKeown. 1997. Predicting the semantic orientation of adjectives. In *Proceedings of the eighth conference on European chapter of the Association for Computational Linguistics*, pages 174–181, Morristown, NJ, USA. Association for Computational Linguistics.

M.Q. Hu and B. Liu. 2004. Mining and summarizing customer reviews. In *KDD '04: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177.

S. Hunston and G. Francis. 2000. *Pattern grammar: A corpus-driven approach to the lexical grammar of English*, volume 4. John Benjamins Publishing Company.

S. Hunston and J. Sinclair. 2000. A local grammar of evaluation. *Evaluation in Text: Authorial stance and the construction of discourse*, pages 74–101.

C. Kerbrat-Orecchioni. 1999. *L'Énonciation. De la subjectivité dans le langage.* Armand Colin, Paris.

D. Lambov, G. Dias, and J.V. Graça. 2010. Multi-view learning for text subjectivity classification. In *Proceedings of the Workshop on Computational Approaches to Subjectivity and Sentiment Analysis of the 19th European Conference on Artificial Intelligence (ECAI 2010)*, pages 30–35.

J.P. Malrieu. 2002. *Evaluative semantics: Cognition, language and ideology*, volume 3. Routledge.

J.R. Martin and P.R.R. White. 2005. *The Language of Evaluation: Appraisal in English*. Palgrave.

C. Muller. 1991. *La négation en français*. Librairie Droz.

B. Pang and L. Lee. 2004. A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts. In *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, pages 271–278.

B. Pang, L. Lee, and S. Vaithyanathan. 2002. Thumbs up? sentiment classification using machine learning techniques. In *EMNLP'02: conference on Empirical methods in natural language processing*, pages 79–86.

V.K. Singh, M. Mukherjee, G.K. Mehta, S. Garg, and N. Tiwari. 2012. Opinion mining from weblogs and its relevance for socio-political research. *Advances in Computer Science and Information Technology. Computer Science and Engineering, Part II*.

P.D. Turney and M.L. Littman. 2003. Measuring praise and criticism: Inference of semantic orientation from association. *ACM Trans. Inf. Syst.*, 21(4):315–346.

P. Turney. 2002. Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. In *Proceedings of the 40th Annual Meeting of the ACL (ACL'02)*, pages 417–424. Philadelphia, Pennsylvania, USA, ACL.

C. Whitelaw, N. Garg, and S. Argamon. 2005. Using appraisal groups for sentiment analysis. In *CIKM '05: Proceedings of the 14th ACM international conference on Information and knowledge management*, pages 625–631, New York, NY, USA. ACM.

M. Wiegand, A. Balahur, B. Roth, D. Klakow, and A. Montoyo. 2010. A survey on the role of negation in sentiment analysis. In *Proceedings of the Workshop on Negation and Speculation in Natural Language Processing*, pages 60–68. Association for Computational Linguistics.

T. Wilson, J. Wiebe, and P. Hoffmann. 2005. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pages 347–354. Association for Computational Linguistics.