Procesamiento del Lenguaje Natural, Revista nº 54, marzo de 2015, pp 77-84 recibido 24-11-14 revisado 27-01-15 aceptado 11-02-15

# Polarity analysis of reviews based on the omission of asymmetric sentences

## *Análisis de la polaridad de comentarios basado en la omisión de oraciones asimétricas*

**John Roberto, Maria Salamó, M. Antònia Martí**
University of Barcelona
Gran Via 585, 08007 Barcelona, Spain
{roberto.john,maria.salamo,amarti}ub.edu

**Resumen:** En este artículo presentamos una aproximación novedosa para el tratamiento de la polaridad en comentarios sobre productos. Nuestro método se centra en identificar y eliminar las oraciones que tienen una polaridad opuesta a la del comentario (oraciones asimétricas) como paso previo a la identificación de los comentarios positivos y negativos. Nuestra hipótesis de partida es que las oraciones asimétricas son morfo-sintácticamente más complejas que las oraciones simétricas (oraciones con la misma polaridad que la del comentario) por lo que es posible mejorar la detección de la polaridad eliminado este tipo de oraciones del texto. Para validar esta hipótesis, hemos medido la complejidad sintáctica de ambos tipos de oraciones en diferentes dominios y hemos contrastado tres configuraciones de datos diferentes basadas en el uso y la omisión de las oraciones asimétricas.
**Palabras clave:** Análisis de la polaridad, minería de opiniones, complejidad sintáctica

**Abstract:** In this paper, we present a novel approach to polarity analysis of product reviews which detects and removes sentences with the opposite polarity to that of the entire document (asymmetric sentences) as a previous step to identify positive and negative reviews. We postulate that asymmetric sentences are morpho-syntactically more complex than symmetric ones (sentences with the same polarity to that of the entire document) and that it is possible to improve the detection of the polarity orientation of reviews by removing asymmetric sentences from the text. To validate this hypothesis, we measured the syntactic complexity of both types of sentences in a multi-domain corpus of product reviews and contrasted three relevant data configurations based on inclusion and omission of asymmetric sentences from the reviews.
**Keywords:** Polarity analysis, opinion mining, syntactic complexity

## 1 Introduction

In recent years, there has been a growing interest in mining opinions from user-generated content on the Web. This interest is motivated in part by an increase in freely available online reviews of products and services.

According to Ricci and Wietsma (2006), a product review can be defined as a subjective piece of text describing the user's experiences, product knowledge and opinions, together with a numerical rating. A common characteristic of a posted review is the presence of an overall opinion polarity, which describes the positive or negative opinion of the author with respect to the evaluated item.

A review, like all other opinionated documents, often consists of some evaluative text units and non-evaluative text units that jointly contribute to the overall polarity of the document. These units have either the same or the opposite polarity as that of the entire review. In this regard, in traditional approaches to polarity analysis, the overall polarity of a text is the average polarity of all its units, mostly words (e.g. adjectives), phrases, and sentences. In contrast to those studies that consider the overall polarity as the result of the average polarity of sentences, in this work we retrieve sentences expressing similar and opposite polarity orientation in relation to the entire review and analyze the differences between them. This paper starts from the premise that both types of sentences use different language constructs becau-

ISSN 1135-5948 

se sentences with the same semantic orientation of the review have less complex structures than sentences with the opposite polarity orientation. On the basis of this assumption, we made the following hypothesis:

*Hypothesis 1*

> *We hypothesize that sentences with the same polarity to that of the entire review are syntactically different from those with the opposite polarity.*

In this paper we call "symmetric sentences" those sentences that have the same polarity as that of the entire review, and "asymmetric sentences" those that do not. Based on this, a second hypothesis can be stated:

*Hypothesis 2*

> *We hypothesize that it is possible to improve the detection of the polarity orientation of reviews by removing from the text the asymmetric sentences.*

With the aim of verifying *hypotheses 1* and *2*, we conducted two experiments with a multi-domain corpus of product reviews in English. These experiments are designed to demonstrate: a) that it is possible to predict accurately when a particular sentence is symmetric or asymmetric, and b) that it is possible to improve the automatic detection of the polarity orientation of reviews by removing asymmetric sentences. The results from both experiments are promising. In particular we show that removing asymmetric sentences improves the performance of the *baseline* to determine the overall polarity of positive and negative reviews.

The rest of this paper is organized as follows: Section 2 looks at the related work on polarity analysis of customer reviews. Next, Section 3 describes syntactic complexity measures. Section 4 explains the experimental analysis (data, tools and results). Finally, we present conclusions in Section 5.

## *2 Related Work*

The traditional state-of-the-art approaches classify polarity of natural language text by analyzing vector representations using, e.g., machine learning (ML) techniques (Pang, Lee, and Vaithyanathan, 2002). ML solutions involve building classifiers from a collection of annotated texts, where each text includes some linguistic-related processing for preparing features such as lemmatization or stemming. Alternative approaches are semantic / lexicon-based (Turney, 2002; Taboada et al., 2011), which renders them robust across domains and texts and enables linguistic analysis at a deeper level. Semantic-based methods involve the use of dictionaries where different kinds of words are tagged with their semantic orientation (SO).

The great majority of works in polarity analysis have mainly focused on analysis of sentences expressing a direct or comparative opinion[1] (Dastjerdi, Ibrahim, and Ghorashi, 2012; Ganapathibhotla and Liu, 2008; Jindal and Liu, 2006). There are few studies analysing how other type of sentences affect the polarity of the entire reviews (cp. Roberto, Salamó, and Martí (2014); Wu and He (2011); Ramanand, Bhavsar, and Pedanekar (2010); Goldberg et al. (2009); and Kim and Hovy (2006)). More specifically, Kim and Hovy (2006) presented a system that automatically extracts the *pros* and *cons* sentences from online reviews. They focused on extracting *pros* and *cons* which include not only sentences that contain opinion-bearing expressions about products and features but also sentences with reasons why an author of a review writes the review.

Goldberg et al. (2009) conducted a novel study on building general "wish detectors"[2] for natural language text, and demonstrated their effectiveness on domains as diverse as consumer product reviews and online political discussions. In the same vein, Ramanand, Bhavsar, and Pedanekar (2010) described rules that can help detect "wishes" from texts such as reviews or customer surveys. Wu and He (2011) analyzed the problem of automatically identifying wishes in product reviews. They built an approach towards such detections, by the use of keyword set constructed by modal words and sequential patterns. Finally, Roberto, Salamó, and Martí (2014) analyzed the role played by narrative sentences in determining the polarity of reviews. Specifically, they applied an algorithm to de-

[1] According to Liu (2010), "direct opinions"give a positive or negative opinion about a object without mentioning any other similar objects and "comparative opinions"declare a preference relation of two or more objects based on some of their shared features.

[2] Wishes are sentences in which authors make suggestions about a product or service or show intentions to purchase a product or service.

tect sentences containing events semantically connected (narrative chains).

## 3 Syntactic Complexity

As we stated in Section 1, we hypothesize that symmetric sentences are syntactically different from asymmetric ones. In this section, we define syntactic complexity and we present a number of different measures of syntactic complexity.

Syntactic complexity refers to "the range of forms that surface in language production and the degree of sophistication of such forms"(Ortega, 2003). Even though there is no single agreed-upon measurement of syntactic complexity, it is mostly a matter of sentence embedding[3] (compare sentences a. and b. in example 1) and non-canonical word order (compare sentences a. and b. in example 2).

(1) a. I eat and you cook.

b. I eat *if you cook.*

(2) a. The student that met the teacher (subject relative clause).

b. The student *that the teacher met* (direct object relative clause).

Some measures of syntactic complexity are common in first and second language acquisition and development (e.g. Index of Productive Syntax or Developmental Sentence Scoring (Moyle and Long, 2013)). However, the act of giving an opinion is a cognitive activity that does not concern with the language acquisition or development. For this reason, we selected three measures that are not directly linked to language acquisition processes but quantify the demand of cognitive processing of different types of syntactic constructions: Yngve's depth algorithm (Yngve, 1960), Frazier's local nonterminal count (Frazier, 1985), and Pakhomov's length of grammatical dependencies (Pakhomov et al., 2011).

Yngve (1960) assumes that the production of a sentence imposes demands on a limited-capacity working memory. The depth of any word in a sentence represents the number of planned grammatical constituents that have not yet been realized during the production of the sentence. Yngve depth is determined by numbering the branches below each node from right to left in a syntactic tree, starting with zero. The depth of each word was the sum of all the branches connecting the word to the root or top-most node of the sentence. Figure 1 illustrates the calculation of the Yngve depth measure in the sentence "it was still starting but a bit sluggish". In this figure, *Total_Ydepth* is the sum of the depth of each word in the sentence and *mean_Ydepth* is the total divided by the number of words.

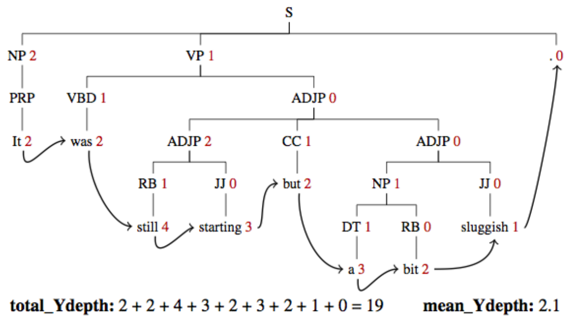



Figura 1: Parse tree fragments with scores for Yngve depth analysis.

Frazier's complexity metric (1985) is based on the idea that syntactic complexity involves the number of non-terminal nodes that the parser must construct when it processes a sentence. The Frazier's approach proceeds in a bottom-up fashion. It traces a path from a word up the tree until reaching either the root of the tree or the lowest node which is not the leftmost child of its parent. Each non-terminal node in the path contributes a score of 1, with 1.5 points for branches from a sentence node (S). Figure 2 illustrates the calculation of the Frazier local non-terminal count measure for the same example sentence "it was still starting but a bit sluggish". *Total_Fdepth* is the sum of the scores for each word in the sentence and *mean_Fdepth* is the total divided by the number of words.

The Pakhomov's scoring method (2011) is inspired in Gibson (1998). It computes the length of grammatical dependencies between lexical items in a sentence based on the Stanford syntactic parser. In the Pakhomov's approach, each dependency relation receives a distance score calculated as the absolute difference between the serial positions of the words that participate in the relation. For example, the distance for the nominal subject relation (*nsubj*) is 4-1 = 3. *Total_SynDepLen* is the sum of all dependencies in the sentence

[3]Embedding refers to the combining of simple sentences into a more complex sentence.

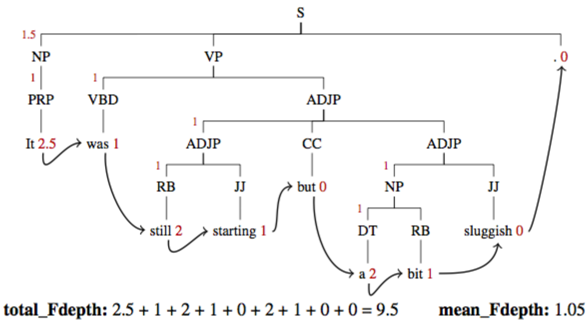
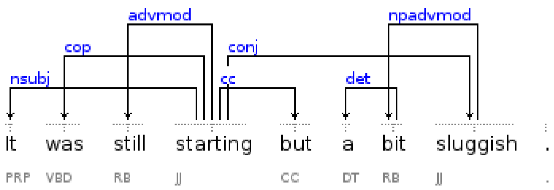



Figura 2: Parse tree fragments with scores for Frazier's node count.

and *mean_SynDepLen* is the total divided by the number of dependencies.




nsubj(starting-4, It-1) =3
cop(starting-4, was-2) =2
advmod(starting-4, still-3) =1
cc(starting-4, but-5) =1
det(bit-7, a-6) =1
npadvmod(sluggish-8, bit-7) =1
conj(starting-4, sluggish-8) =4

**total_SynDepLen:** 3 + 2 + 1 + 1 + 1 + 1 + 4 = 13
**mean_SynDepLen:** 1.8

Figura 3: A graph view of typed dependencies of a sentence as computed by Stanford parser and the calculation of the dependency length for Pakhomov scoring method.

In the experiments we apply the Yngve, Frazier, and Pakhomov's indices of complexity to characterize symmetric and asymmetric sentences.

## 4 *Experiments and results*

This section evaluates the hypotheses stated in Section 1. First, it describes the data and tools used in the experiments. Second, it presents the results obtained in the two experiments addressed to evaluate the hypotheses.

### 4.1 Data and tools

The data used in our experimental analysis is the multi-domain corpus of product reviews by Cruz Mata (2012). Originally, this corpus is a collection of 2,547 reviews extracted from www.ciao.com. This corpus has been chosen in order to analyze symmetric and asymmetric sentences because each review is annotated with the overall polarity and the polarity of its features in every single sentence. A pre-processing of the corpus has been performed automatically in order to remove sentences that do not express any polarity or induce inconsistencies (noise) in the data set. In the latter case, first, we eliminate all text passages that contain multiple sentences with a unique polarity assigned to it. Second, we also subtract those sentence with mixed polarity (e.g. "It is a nice hotel, small but very nice and clean."). Finally, we removed one-word sentences because they are not relevant for our analysis (e.g. "ok", "avoid", "duuuhhh"). Table 1 describes the corpus used in our experiments after removing those sentences.

| Sentence | Cars | Headphones | Hotels | All |
|---|---|---|---|---|
| # *Symm* | 403 | 194 | 334 | 994 |
| # *Asymm* | 403 | 194 | 334 | 994 |
| **Total** | **806** | **388** | **668** | **1988** |

Tabla 1: Number of symmetric (*Symm*) and asymmetric (*Aymm*) sentences in each domain.

As we mentioned in Section 1, sentences with the same polarity as that of the entire review have been referred to as symmetric (*symm*) and those with different polarity as asymmetric (*asymm*).

We used the Computerized Linguistic Analysis System (CLAS) (Pakhomov et al., 2011) for the computation of syntactic complexity measures. CLAS system implements Yngve (1960), Frasier (1985), Gibson (1998), and other computational approaches to establish the syntactic complexity of English sentences. This software uses the Stanford syntactic parser, which provides basic information on the hierarchical constituent structure of the sentence as well as syntactic dependencies between lexical items.

We used the Semantic Orientation CALculator System (SO-CAL) (Taboada et al., 2011) for calculating the polarity orientation of reviews. SO-CAL is a general purpose system that was designed for determining semantic orientation on the level of complete texts. SO-CAL uses manually built dictionaries of words (adjectives, nouns, verbs, and adverbs) annotated with their polarity and strength, and incorporates negation and intensification (e.g., *very*, *slightly*).

## 4.2 Symmetric and asymmetric sentences classification

This experiment attempts to answer the following question: Is it possible to predict accurately when a particular sentence will be symmetric or asymmetric?

To answer this question, we analyzed the syntactic complexity of opinionated sentences from reviews using the Computerized Linguistic Analysis System (Pakhomov et al., 2011). The fifteen scores obtained with this tool (listed in Table 2) were used as attributes to train and test different classifiers in Weka (Witten et al., 1999).

| N. | Attributes |
|---|---|
| 1 | Mean of Frazer depth scores on individual tokens (*mean_Fdepth*). |
| 2 | Sum of Frazer depth scores on individual tokens (*total_Fdepth*). |
| 3 | Mean of Yngve depth scores on individual tokens (*mean_Ydepth*). |
| 4 | Sum of Yngve depth scores on individual tokens (*total_Ydepth*). |
| 5 | Mean of syntactic dependency lengths in the dependency parse (*mean_SynDepLen*). |
| 6 | Sum of syntactic dependency lengths in the dependency parse (*total_SynDepLen*). |
| 7 | Number of “S”nodes in the parse tree. |
| 8 | Raw count of nouns. |
| 9 | Raw count of adjectives. |
| 10 | Raw count of adverbs. |
| 11 | Raw count of verbs. |
| 12 | Raw count of determiners. |
| 13 | Raw count of conjunctions. |
| 14 | Raw count of prepositions. |
| 15 | Raw count of proper nouns. |

Tabla 2: List of the fifteen scores/attributes generated by the Computerized Linguistic Analysis System (CLAS).

With the aim of determining the consistency of the scores obtained automatically by CLAS, we randomly selected 30 sentences from the corpus (10 for each domain) that were labeled and scored by a trained linguist. To compare automatic and human scores, we have used a Kappa statistic approach. The average Kappa score was 0.76, showing an acceptable degree of agreement for the task.

Additionally, we performed a linear transformation on the original data to scale the value of all features in the range [0..1] using the R package “ppls”(Krämer and Sugiyama, 2011). For classification, a 10-fold cross validation methodology was performed from which we report *accuracies*.

Table 3 shows the accuracies obtained for each one of the classifiers analyzed. First column contains the list of classification algorithms that have been tested. Subsequent columns list the distribution of accuracies per domains (cars, headphones, hotels) and all domains as a whole. Finally, the last row contains the average accuracies obtained for each domain.

| Algorithm | Cars | Headphones | Hotels | All |
|---|---|---|---|---|
| *BayesNet* | 70.3 | 70.1 | 67.2 | 73.1 |
| *LWL* | 65.0 | 71.9 | 66.5 | 76.3 |
| *DTNB* | **72.3** | 70.6 | 69.5 | 75.8 |
| *Decis.Table* | 71.5 | 72.7 | 71.0 | 76.0 |
| *JRip* | 70.3 | 70.6 | 68.7 | 76.3 |
| *Ridor* | 71.7 | 70.4 | 69.9 | 76.4 |
| *ADTree* | 70.6 | **73.7** | 71.1 | 76.0 |
| *BFTree* | 71.7 | 71.1 | 70.4 | 76.4 |
| *LADTree* | 71.3 | 72.9 | 71.7 | 75.5 |
| *REPTree* | 70.7 | 71.6 | 69.5 | 75.5 |
| *SimpleCart* | 70.7 | 71.6 | **72.2** | 76.3 |
| *Average* | *70.6* | *71.6* | *69.8* | *75.8* |

Tabla 3: Percentage of symmetric and asymmetric sentences correctly classified in each domain.

The findings of this study reveal that a good accuracy can be obtained using syntactic complexity for determining symmetric and asymmetric sentences. Note that on average all the results are around 70 %. In particular, a 70.6 % in the cars domain, a 71.6 % in the headphones domain, and 69.8 % in the hotels domain. In the cars domain, the best classifier achieves an accuracy of 72.3 % for distinguishing symmetric from asymmetric sentences. The best accuracy estimated using the same syntactic complexity measures is 73.7 % and 72.2 % for headphone and hotel domains, respectively. The best results are achieved bringing all domains: all classification accuracies are above 73 % and the general average is 75.8 %.

Additionally, we apply four well known selection methods to pick up the five most informative attributes that are used to classify symmetric and asymmetric sentences, as shown in Table 4. In general, we have found that the attributes based on the use of syntactic complexity measures ($total\_SynDepLen$, $total\_Ydepth$, $total\_Fdepth$, $mean\_Fdepth$, $mean\_Ydepth$, and $mean\_SynDepLen$) are among the five most discriminative attributes.

| CARS | | | |
|---|---|---|---|
| **Chi-squared** | **Gain Ratio** | **Info. Gain** | **ReliefF** |
| total_SynDepLen | total_Ydepth | total_SynDepLen | total_SynDepLen |
| total_Ydepth | total_SynDepLen | total_Ydepth | total_Ydepth |
| total_Fdepth | det_count | total_Fdepth | total_Fdepth |
| mean_Ydepth | total_Fdepth | mean_Ydepth | mean_Ydepth |
| mean_SynDepLen | mean_Ydepth | mean_SynDepLen | mean_SynDepLen |
| HEADPHONES | | | |
| **Chi-squared** | **Gain Ratio** | **Info. Gain** | **ReliefF** |
| total_Ydepth | total_Ydepth | total_Ydepth | total_Fdepth |
| total_SynDepLen | total_SynDepLen | total_SynDepLen | total_Ydepth |
| total_Fdepth | mean_Ydepth | total_Fdepth | total_SynDepLen |
| mean_Ydepth | total_Fdepth | mean_Ydepth | adj_count |
| num_clauses | num_clauses | num_clauses | conj_count |
| HOTELS | | | |
| **Chi-squared** | **Gain Ratio** | **Info. Gain** | **ReliefF** |
| total_Ydepth | total_SynDepLen | total_Ydepth | total_Fdepth |
| total_SynDepLen | total_Ydepth | total_SynDepLen | num_clauses |
| noun_count | mean_Ydepth | mean_Ydepth | verb_count |
| mean_Ydepth | noun_count | noun_count | mean_Ydepth |
| total_Fdepth | total_Fdepth | total_Fdepth | total_Ydepth |

Tabla 4: The most relevant features retained by the attribute selection methods for each domain.

In summary, the most discriminative features are the ones based on the syntactic complexity measures and the accuracies obtained using these features support the hypothesis that it is possible to predict accurately when a particular sentence express the same (symmetric) or the opposite (asymmetric) polarity to that of the entire review.

### 4.3 Polarity classification

This experiment attempts to answer the following question: Is it possible to improve the accuracy of polarity classifiers by removing asymmetric sentences from reviews?

To answer this question, we calculated the overall polarity of reviews using the Semantic Orientation CALculator System (Taboada et al., 2011). We contrasted three relevant data configurations based on the extraction of different types of sentences from the reviews. These configurations are:

1. **Gold standard:** the polarity analysis was performed using only symmetric sentences based on a hypothetical prediction accuracy of 100 % for the detection of this type of sentences. The input to the SO-CAL system is formed by all the sentences labeled with the same polarity to that of the entire review.
2. **Baseline:** the polarity analysis was performed in standard fashion, that is, by using the entire review. The input to the SO-CAL system is formed by all the sentences from the review.
3. **Approach:** the polarity analysis was performed removing some of the asymmetric sentences from reviews based on the factual categorization accuracies obtained in experiment one (see Table 3). The input to the SO-CAL system is formed by all the sentences from the review except the 70 % of asymmetric sentences for cars, the 71 % of asymmetric sentences for headphones, and the 69 % of asymmetric sentences for hotels.

The results of this experiment are summarized in Table 5. The performance of the *baseline* is consistent with other published studies (Taboada, 2011). The so-called *gold standard* configuration improves from 88 % to 93.2 % the performance of the *baseline* for positive reviews and from 76.8 % to 84.5 % the performance of the negative reviews. Recall that the *gold standard* configuration is based on hypothetical accuracies for symmetric and asymmetric sentences categorization.

*Approach* is the second best configuration but, in contrast to the *gold standard*, it is based on the factual data gathered from the first experiment. Under this configuration, the positive polarity obtains an average accuracy of

| Configurations | Gold | | Baseline | | Approach | |
|---|---|---|---|---|---|---|
| Review Polarity | + | − | + | − | + | − |
| Cars | 93.4 | 85 | 88.1 | 77.3 | **89.4** | **83.2** |
| Headphones | 89.2 | 80.6 | 81 | 74.3 | **83.3** | **78.7** |
| Hotels | 97 | 88 | 95 | 78.8 | **96.7** | **81.9** |
| Averages | 93.2 | 84.5 | 88 | 76.8 | **89.8** | **81.2** |

Tabla 5: Polarity analysis of product reviews (based on SO-CAL system) under four different configurations. The reported values are classification accuracies, that is, the percentage of correct choices.

89.8 %. This is a worthy improvement over the 88 % that results when all sentences are used (*baseline*). Nevertheless, the most significant increment is observed in the case of negative reviews: *approach* configuration improves from 76.8 % to 81.2 % the average accuracy of the *baseline* for negative reviews. This huge improvement is shared by every domain.

In summary, these results show that the polarity analysis of reviews improves by removing their asymmetric sentences, as shown in the *Averages* at the bottom of Table 5.

## 5 Conclusions

In this paper we analyze the function of symmetric and asymmetric sentences (opinionated sentences expressing similar and opposite polarity orientation in relation to the entire document) with the aim of improving the polarity detection of reviews. For this purpose we have performed two tasks.

The first task consists of the evaluation of the usefulness of different syntactic complexity measures to characterize both symmetric and asymmetric sentences. To this end, we have trained a cascade of classifiers using the Weka Environment. Our experiments show that syntactic complexity is an effective way to characterize symmetric and asymmetric sentences and it is possible to detect accurately when a particular sentence is symmetric or asymmetric.

The second task consists of the classification of the reviews as being positive or negative. To this end, we contrasted three relevant data configurations based on the removal of different types of sentences from the reviews. The experimental results indicate that removing asymmetric sentences increases the performance on the determination of the overall polarity of reviews. There is a noticeable improvement in the case of the negative reviews.

## 6 Acknowledgments

This work was supported by projects SGR-2014-623, TIN2012-38603-C02-02 and TIN2012- 38584-C06-01, as well as a FI grant (2010FI_B 00521).

## References

Cruz Mata, Fermín L. 2012. *Extracción de opiniones sobre características: un enfoque práctico adaptable al dominio*. Colección de monografías de la Sociedad Española para el Procesamiento del Lenguaje Natural. Sociedad Española para el Procesamiento del Lenguaje Natural.

Dastjerdi, Niloufar Salehi, Roliana Ibrahim, and Seyed Hamid Ghorashi. 2012. Product feature extraction using natural language processing techniques. *Journal of Computing*, 4(7):39–43.

Frazier, Lyn, 1985. *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives*, chapter Syntactic complexity, pages 129–189. Cambridge University Press, Cambridge, UK.

Ganapathibhotla, Murthy and Bing Liu. 2008. Mining opinions in comparative sentences. In *Proc. of the 22Nd International Conference on Computational Linguistics*, volume 1 of *COLING '08*, pages 241–248, Stroudsburg, PA, USA. Association for Computational Linguistics.

Gibson. 1998. Linguistic complexity: locality of syntactic dependencies. *Cognition*, 68(1):1–76.

Goldberg, Andrew B., Nathanael Fillmore, David Andrzejewski, Zhiting Xu, Bryan Gibson, and Xiaojin Zhu. 2009. May all your wishes come true: a study of wishes and how to recognize them. In *Proc. of Human Language Technologies: The 2009*

*Annual Conference of the North American Chapter of the Association for Computational Linguistics*, NAACL '09, pages 263–271, Stroudsburg, PA, USA. Association for Computational Linguistics.

Jindal, Nitin and Bing Liu. 2006. Mining comparative sentences and relations. In *Proc. of the 21st National Conference on Artificial Intelligence - Volume 2*, AAAI'06, pages 1331–1336. AAAI Press.

Kim, Soo-Min and Eduard Hovy. 2006. Extracting opinions, opinion holders, and topics expressed in online news media text. In *Proc. of the Workshop on Sentiment and Subjectivity in Text*, SST '06, pages 1–8, Stroudsburg, PA, USA. Association for Computational Linguistics.

Krämer, Nicole and Masashi Sugiyama. 2011. The degrees of freedom of partial least squares regression. *Journal of the American Statistical Association*, 106(494):697–705.

Liu, Bing, 2010. *Handbook of Natural Language Processing*, chapter Sentiment Analysis and Subjectivity, pages 627–666. CRC Press, Connecticut, USA.

Moyle, Maura and Steven Long, 2013. *Encyclopedia of Autism Spectrum Disorders*, chapter Index of Productive Syntax (IPSyn), pages 1566–1568. Springer.

Ortega, Lourdes. 2003. Syntactic complexity measures and their relationship to l2 proficiency: A research synthesis of college-level l2 writing. *Applied Linguistics*, 4(24):492–518.

Pakhomov, Serguei, Dustin Chacon, Mark Wicklund, and Jeanette Gundel. 2011. Computerized assessment of syntactic complexity in alzheimer's disease: a case study of iris murdoch's writing. *Behavior Research Methods*, 43(1):136–144.

Pang, Bo, Lillian Lee, and Shivakumar Vaithyanathan. 2002. Thumbs up?: sentiment classification using machine learning techniques. In *EMNLP '02: Proc. of the ACL-02 conference on Empirical methods in natural language processing*, pages 79–86, Morristown, NJ, USA. Association for Computational Linguistics.

Ramanand, J., Krishna Bhavsar, and Niranjan Pedanekar. 2010. Wishful thinking: finding suggestions and 'buy' wishes from product reviews. In *Proc. of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, CAAGET '10, pages 54–61, Stroudsburg, PA, USA. Association for Computational Linguistics.

Ricci and Wietsma. 2006. Product reviews in travel decision making. In *Information and Communication Technologies in Tourism 2006*, pages 296–307.

Roberto, John, Maria Salamó, and M. Antònia Martí. 2014. The function of narrative chains in the polarity classification of reviews. *Procesamiento del Lenguaje Natural*, 52:69–76.

Taboada, Maite. 2011. Stages in an online review genre. *Text and Talk. An Interdisciplinary Journal of Language, Discourse & Communication Studies*, 31(2):247–269.

Taboada, Maite, Julian Brooke, Milan Tofiloski, Kimberly Voll, and Manfred Stede. 2011. Lexicon-based methods for sentiment analysis. *Computational Linguistics*, 37(2):267–307, June.

Turney, Peter. 2002. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In *Proc. of the 40th Annual Meeting of the Association for Computational Linguistics (ACL-02)*, pages 417–424, Philadelphia, Pennsylvania.

Witten, Ian, Eibe Frank, Len Trigg, Mark Hall, Geoffrey Holmes, and Sally Cunningham. 1999. *Weka: Practical Machine Learning Tools and Techniques with Java Implementations. (Working paper 99/11).* Hamilton, New Zealand: University of Waikato, Department of Computer Science.

Wu, Xing and Zhongshi He. 2011. Identifying wish sentence in product reviews. *Journal of Computational Information Systems*, 7(5):1607–1613.

Yngve, Victor. 1960. A model and an hypothesis for language structure. *Proc. of the American Philosophical Society*, 104(5):444–466.