

# Discourse Connectors for Latent Subjectivity in Sentiment Analysis

**Rakshit Trivedi**

College of Computing  
Georgia Institute of Technology  
Atlanta, GA 30308, USA  
rtrivedi6@gatech.edu

**Jacob Eisenstein**

School of Interactive Computing  
Georgia Institute of Technology  
Atlanta, GA 30308, USA  
jacobe@gatech.edu

## Abstract

Document-level sentiment analysis can benefit from fine-grained subjectivity, so that sentiment polarity judgments are based on the relevant parts of the document. While fine-grained subjectivity annotations are rarely available, encouraging results have been obtained by modeling subjectivity as a latent variable. However, latent variable models fail to capitalize on our linguistic knowledge about discourse structure. We present a new method for injecting linguistic knowledge into latent variable subjectivity modeling, using discourse connectors. *Connector-augmented transition features* allow the latent variable model to learn the relevance of discourse connectors for subjectivity transitions, without subjectivity annotations. This yields significantly improved performance on document-level sentiment analysis in English and Spanish. We also describe a simple heuristic for automatically identifying connectors when no predefined list is available.

## 1 Introduction

Document-level sentiment analysis can benefit from consideration of discourse structure. Voll and Taboada (2007) show that adjective-based sentiment classification is improved by examining topicality (whether each sentence is central to the overall point); Yessenalina et al. (2010b) show that bag-of-ngrams sentiment classification is improved by examining subjectivity (whether a sentence expresses a subjective opinion or objective fact). However, it is unclear how best to obtain the appropriate discourse analyses. Voll and Taboada (2007) find that

domain-independent discourse parsing (Soricut and Marcu, 2003) offers little improvement for sentiment analysis, so they resort to training a domain-specific model for identifying topic sentences in reviews. But this requires a labeled dataset of topic sentences, imposing a substantial additional cost.

Yessenalina et al. (2010b) treat sentence level subjectivity as a *latent variable*, automatically inducing the “annotator rationale” (Zaidan et al., 2007; Yessenalina et al., 2010a) for each training sentence so as to focus sentiment learning on the subjective parts of the document. This yields significant improvements over bag-of-ngrams supervised sentiment classification. Latent variable subjectivity analysis is attractive because it requires neither subjectivity annotations nor an accurate domain-independent discourse parser. But while the “knowledge-free” nature of this approach is appealing, it is unsatisfying that it fails to exploit decades of research on discourse structure.

In this paper, we explore a lightweight approach to injecting linguistic knowledge into latent variable models of subjectivity. The entry point is a set of *discourse connectors*: words and phrases that signal a shift or continuation in the discourse structure. Such connectors have been the subject of extensive study in the creation of the Penn Discourse Treebank (PDTB: Prasad et al. 2008). The role of discourse connectors in sentiment analysis can be clearly seen in examples, such as “*It’s hard to imagine the studios hiring another manic German maverick to helm a cop thriller. **But** that’s exactly why the movie is unmissable.*” (Huddleston, 2010)

We present a new approach to incorporate discourse connectors in a latent subjectivity model (Yessenalina et al., 2010b). This approach requires no manually-specified information about the meaning of the connectors, just the connectors themselves. Our approach builds on proximity features, which give the latent variable model a way to prefer or disprefer subjectivity and sentiment transitions, usually with the goal of encouraging smoothness across the document. By taking the cross-product of these features with a set of discourse connectors, we obtain a new set of *connector-augmented transition features*, which capture the way discourse connectors are used to indicate subjectivity and sentiment transitions. The model is thus able to learn that subjectivity shifts are likely to be accompanied by connectors such as *however* or *nonetheless*.

We present experiments in both English and Spanish showing that this method of incorporating discourse connectors yields significant improvements in document-level sentiment analysis. In case no list of connectors is available, we describe a simple heuristic for automatically identifying candidate connector words. The automatically identified connectors do not perform as well as the expert-defined lists, but they still outperform a baseline method that ignores discourse connectors (in English). This demonstrates both the robustness of the approach and the value of linguistic knowledge.

## 2 Model

Given accurate labels of the subjectivity of each sentence, a document-level sentiment analyzer could safely ignore the sentences marked as non-subjective.<sup>1</sup> This would be beneficial for training as well as prediction, because the learning algorithm would not be confused by sentences that contradict the document label. But in general we cannot rely on having access to sentence-level subjectivity annotations. Instead, we treat subjectivity as a *latent variable*, and ask the learner to impute its value. Given document-level sentiment annotations and an initial

<sup>1</sup>Discourse parsing often focuses on sub-sentence *elementary discourse units* (Mann and Thompson, 1988). For simplicity, we consider units at the sentence level only, and leave finer-grained analysis for future work.

model, the learner can mark as non-subjective those sentences whose analysis disagrees with the document label.

More formally, each document has a label  $y \in \{-1, 1\}$ , a set of sentences  $\mathbf{x}$ , and a set of per-sentence subjectivity judgments  $\mathbf{h} \in \{0, 1\}^S$ , where  $S$  is the number of sentences. We compute a set of features on these variables, and score each instance by a weighted combination of the features,  $\mathbf{w}^\top \mathbf{f}(y, \mathbf{x}, \mathbf{h})$ . At prediction time, we seek a label  $y$  which achieves a high score given the observed  $\mathbf{x}$  and the ideal  $\mathbf{h}$ .

$$\hat{y} = \arg \max_y \left( \max_{\mathbf{h}} \mathbf{w}^\top \mathbf{f}(y, \mathbf{x}, \mathbf{h}) \right). \quad (1)$$

At training time, we seek weights  $w$  which achieve a high score given all training examples  $\{\mathbf{x}, y\}_t$ ,

$$\hat{w} = \arg \max_w \sum_t \max_{\mathbf{h}} \mathbf{w}^\top \mathbf{f}(y_t, \mathbf{x}_t, \mathbf{h}). \quad (2)$$

We can decompose the feature vector into two parts: polarity features  $\mathbf{f}_{\text{pol}}(y, \mathbf{x}, \mathbf{h})$ , and subjectivity features  $\mathbf{f}_{\text{subj}}(\mathbf{x}, \mathbf{h})$ . The basic feature set decomposes across sentences, though the polarity features involve the document-level polarity. For sentence  $i$ , we have  $\mathbf{f}_{\text{pol}}(y, \mathbf{x}_i, h_i) = y h_i \mathbf{x}_i$ : the bag-of-words features for sentence  $i$  are multiplied by the document polarity  $y \in \{-1, 1\}$  and the sentence subjectivity  $h_i \in \{0, 1\}$ . The weights  $\mathbf{w}_{\text{pol}}$  capture the sentiment polarity of each possible word. As for the subjectivity features, we simply have  $\mathbf{f}_{\text{subj}}(\mathbf{x}_i, h_i) = h_i \mathbf{x}_i$ . The weights  $\mathbf{w}_{\text{subj}}$  capture the subjectivity of each word, with large values indicate positive subjectivity.

However, these features do not capture transitions between the subjectivity and sentiment of adjacent sentences. For this reason, Yessenalina et al. (2010b) introduce an additional set of *proximity* features,  $\mathbf{f}_{\text{prox}}(h_i, h_{i-1})$ , which are parametrized by the subjectivity of both the current sentence  $i$  and the previous sentence  $i - 1$ . The effect of these features will be to learn a preference for consistency in the subjectivity of adjacent sentences.

By augmenting the transition features with the text  $\mathbf{x}_i$ , we allow this preference for consistency to be modulated by discourse connectors. We design the *transition* feature vector  $\mathbf{f}_{\text{trans}}(\mathbf{x}_i, h_i, h_{i-1})$

to contain two elements for every discourse connector, one for  $h_i = h_{i-1}$ , and one for  $h_i \neq h_{i-1}$ . For example, the feature  $\langle \textit{moreover}, \text{CONTINUE} \rangle$  fires when sentence  $i$  starts with *moreover* and  $h_{i-1} = h_{i,i}$ . We would expect to learn a positive weight for this feature, and negative weights for features such as  $\langle \textit{moreover}, \text{SHIFT} \rangle$  and  $\langle \textit{however}, \text{CONTINUE} \rangle$ .

### 3 Experiments

To evaluate the utility of adding discourse connectors to latent subjectivity sentiment analysis, we compare several models on movie review datasets in English and Spanish.

#### 3.1 Data

We use two movie review datasets:

- 50,000 English-language movie reviews (Maas et al., 2011). Each review has a rating from 1-10; we marked ratings of 5 or greater as positive. Half the dataset is used for test and half for training. Parameter tuning is performed by cross-validation.
- 5,000 Spanish-language movie reviews (Cruz et al., 2008). Each review has a rating from 1-5; we marked 3-5 as positive. We randomly created a 60/20/20 split for training, validation, and test.

#### 3.2 Connectors

We first consider single-word discourse connectors: in English, we use a list of all 57 one-word connectors from the Penn Discourse Tree Bank (Prasad et al., 2008); in Spanish, we selected 25 one-word connectors from a Spanish language education website.<sup>2</sup> We also consider multi-word connectors. Using the same sources, this expands the English set to 93 connectors, and Spanish set to 80 connectors.

In case no list of discourse connectors is available, we propose a simple technique for automatically identifying potential connectors. We use a  $\chi^2$  test to select words which are especially likely to initiate sentences. The top  $K$  words (with the lowest  $p$  values) were added as potential connectors, where  $K$  is equal to the number of “true” connectors provided by the gold-standard resource.

<sup>2</sup>[russell.famaf.unc.edu.ar/~laura/shallowdisc4summ/discmar/](http://russell.famaf.unc.edu.ar/~laura/shallowdisc4summ/discmar/)

Finally, we consider a model with connector-augmented transition features for all words in the vocabulary. Thus, there are four connector sets:

- **true-unigram-connectors:** unigram connectors from the Penn Discourse Treebank and the Spanish language education website
- **true-multiword-connectors:** unigram and multiword connectors from these same resources
- **auto-unigram-connectors:** automatically-selected connectors using the  $\chi^2$  test
- **all-unigram-connectors:** all words are potential connectors

#### 3.3 Systems

The connector-augmented transition features are incorporated into a latent variable support vector machine (SVM). We also consider two baselines:

- **no-connectors:** the same latent variable SVM, but without the connector features. This is identical to the prior work of Yessenalina et al. (2010b).
- **SVM:** a standard SVM binary classifier

The latent variable models require an initial guess for the subjectivity of each sentence. Yessenalina et al. (2010b) compare several initializations and find the best results using OpinionFinder (Wilson et al., 2005). For the Spanish data, we performed initial subjectivity analysis by matching against a publicly-available full-strength Spanish lexicon set (Rosas et al., 2012).

#### 3.4 Implementation details

Both our implementation and the baselines are built on the latent structural SVM (Yu and Joachims, 2009; <http://www.cs.cornell.edu/~cnyu/latentssvm/>), which is in turn built on the SVM-Light distribution (<http://svmlight.joachims.org/>). The regularization parameter  $C$  was chosen by cross-validation.

### 4 Results

Table 1 shows the sentiment analysis accuracy with each system and feature set. The best overall results in both language are given by the models with

	system	English	Spanish
true-multiword-connectors		91.25	<b>79.80</b>
true-unigram-connectors		<b>91.36</b>	77.50
auto-connectors		90.22	76.90
all-unigram-connectors		87.60	74.30
No-connectors		88.21	76.42
SVM		84.79	69.44

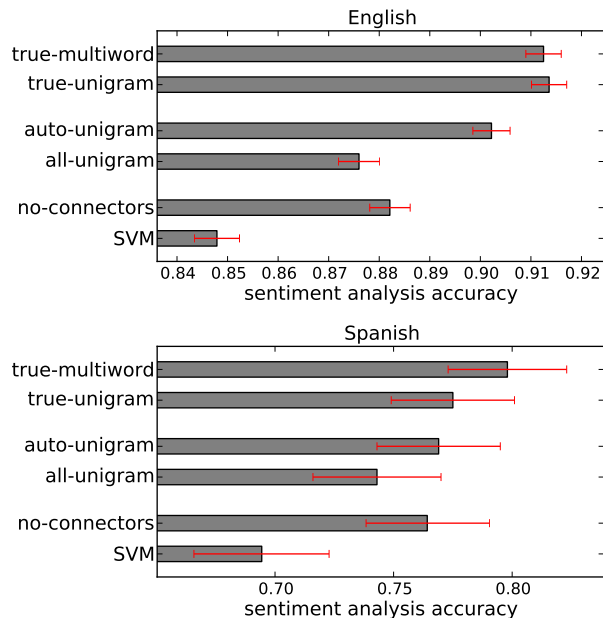


Figure 1: Document-level sentiment analysis accuracy. The 95% confidence intervals are estimated from the cumulative density function of the binomial distribution.

connector-augmented transition features. In English, the multiword and unigram connectors perform equally well, and significantly outperform all alternatives at  $p < .05$ . The connector-based features reduce the error rate of the latent subjectivity SVM by 25%. In Spanish, the picture is less clear because the smaller test set yields larger confidence intervals, so that only the comparison with the SVM classifier is significant at  $p < .05$ . Nonetheless, the connector-augmented transition features give the best accuracy, with an especially large improvement obtained by the multiword connectors.

Next, we investigated the quality of the automatically-induced discourse connectors. The  $\chi^2$  heuristic for selecting candidate connectors gave results that were significantly better than the no-connector baseline in English, though the

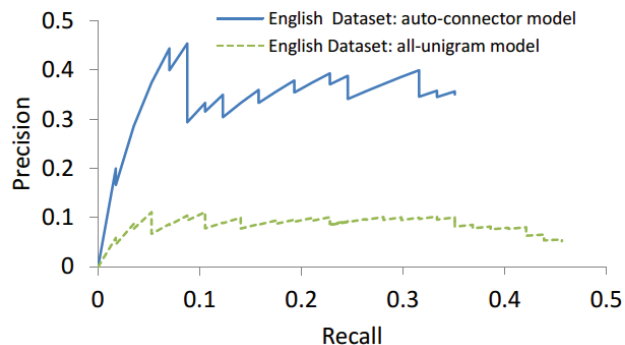


Figure 2: Precision-Recall curve for top-K discovered connectors when compared with PDTB connector set

difference in Spanish was minimal. However, when every word is included as a potential connectors, the performance suffers, dropping below the accuracy of the no-connector baseline. This shows that the improvement in accuracy offered by the connector features is not simply due to the increased flexibility of the model, but depends on identifying a small set of likely discourse connectors.

For a qualitative evaluation, we ranked all English-language unigram connectors by their feature weights, and list the top ten for each subjectivity transition:

- SHIFT: *however; though; but; if; unlike; although; while; overall; nevertheless; still*
- CONTINUATION: *as; there; now; even; in; after; once; almost; because; so*

Overall these word lists cohere with our intuitions, particularly the words associated with SHIFT transitions: *however, but, and nevertheless*. As one of the reviewers noted, some of the words associated with CONTINUATION transitions are better seen as discourse **cues** rather than connectors, such as *now*. Other words seem to connect two **subsequent** clauses, e.g., *if Nicholas Cage had played every role, the film might have reached its potential*. Incorporating such connectors must be left for future work.

Finally, in learning weights for each connector feature, our model can be seen as *discovering* discourse connectors. We compare the highly weighted discovered connectors from the all-unigram and auto-unigram settings with the one-word connectors from the Penn Discourse Tree Bank. The results

of this comparison are shown in Figure 2, which traces a precision-recall curve by taking the top  $K$  connectors for various values of  $K$ . The auto-unigram model is able to identify many true connectors from the Penn Discourse Treebank, while the all-unigram model achieves low precision. This graph helps to explain the large performance gap between the auto-unigram and all-unigram feature sets; the all-unigram set includes too many weak features, and the learning algorithm is not able to distinguish the true discourse connectors. The Spanish discourse connectors identified by this approach were extremely poor, possibly because so many more of the Spanish connectors include multiple words.

## 5 Related Work

Polanyi and Zaenen (2006) noted the importance of accounting for *valence shifters* in sentiment analysis, identifying relevant connectors at the sentence and discourse levels. They propose a heuristic approach to use shifters to modify the contributions of sentiment words. There have been several subsequent efforts to model within-sentence valence shifts, including compositional grammar (Moilanen and Pulman, 2007), matrix-vector products across the sentence (Yessenalina and Cardie, 2011), and methods that reason about polarity shifters within the parse tree (Socher et al., 2012; Sayeed et al., 2012). The value of discourse structure towards predicting opinion polarity has also demonstrated in the context of multi-party dialogues (Somasundaran et al., 2009). Our approach functions at the discourse level within single-author documents, so it is complementary to this prior work.

Voll and Taboada (2007) investigate various techniques for focusing sentiment analysis on sentences that are central to the main topic. They obtain negative results with the general-purpose SPADE discourse parser (Soricut and Marcu, 2003), but find that training a decision tree classifier to identify topic-central sentences yields positive results. Wiebe (1994) argues that in coherent narratives, objectivity and subjectivity are usually consistent between adjacent sentences, an insight exploited by Pang and Lee (2004) in a supervised system for subjectivity analysis. Later work employed struc-

tured graphical models to model the flow of subjectivity and sentiment over the course of the document (Mao and Lebanon, 2006; McDonald et al., 2007). All of these approaches depend on labeled training examples of subjective and objective sentences, but Yessenalina et al. (2010b) show that subjectivity can be modeled as a *latent variable*, using a latent variable version of the structured support vector machine (Yu and Joachims, 2009).

Our work can be seen as a combination of the machine learning approach of Yessenalina et al. (2010b) with the insight of Polanyi and Zaenen (2006) that connectors play a key role in transitions between subjectivity and sentiment. Eisenstein and Barzilay (2008) incorporated discourse connectors into an unsupervised model of topic segmentation, but this work only considered the role of such markers to differentiate adjoining segments of text, and not to identify their roles with respect to one another. That work was also not capable of learning from supervised annotations in a downstream task. In contrast, our approach uses document-level sentiment annotations to learn about the role of discourse connectors in sentence-level subjectivity.

## 6 Conclusion

Latent variable machine learning is a powerful tool for inducing linguistic structure directly from data. However, adding a small amount of linguistic knowledge can substantially improve performance. We have presented a simple technique for combining a latent variable support vector machine with a list of discourse connectors, by creating an augmented feature set that combines the connectors with pairwise subjectivity transition features. This improves accuracy, even with a noisy list of connectors that has been identified automatically. Possible directions for future work include richer representations of discourse structure, and the combination of discourse-level and sentence-level valence and subjectivity shifters.

## Acknowledgments

Thanks to the anonymous reviewers for their helpful feedback. This work was supported by a Google Faculty Research Award.

## References

- Fermin L. Cruz, Jose A. Troyano, Fernando Enriquez, and Javier Ortega. 2008. Clasificación de documentos basada en la opinión: experimentos con un corpus de críticas de cine en español. *Procesamiento de Lenguaje Natural*, 41.
- Jacob Eisenstein and Regina Barzilay. 2008. Bayesian unsupervised topic segmentation. In *Proceedings of EMNLP*.
- Tom Huddleston. 2010. Review of *The Bad Lieutenant: Port of Call New Orleans*. *Time Out*, May 18.
- Andrew L. Maas, Raymond E. Daly, Peter T. Pham, Dan Huang, Andrew Y. Ng, and Christopher Potts. 2011. Learning word vectors for sentiment analysis. In *Proceedings of ACL*.
- William C Mann and Sandra A Thompson. 1988. Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3).
- Yi Mao and Guy Lebanon. 2006. Isotonic conditional random fields and local sentiment flow. In B. Schölkopf, J. Platt, and T. Hoffman, editors, *Advances in Neural Information Processing Systems 19*.
- Ryan McDonald, Kerry Hannan, Tyler Neylon, Mike Wells, and Jeff Reynar. 2007. Structured models for fine-to-coarse sentiment analysis. In *Proceedings of ACL*.
- Karo Moilanen and Stephen Pulman. 2007. Sentiment composition. In *Proceedings of RANLP*.
- Bo Pang and Lillian Lee. 2004. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In *Proceedings of ACL*.
- Livia Polanyi and Annie Zaenen. 2006. Contextual valence shifters. *Computing attitude and affect in text: Theory and applications*.
- Rashmi Prasad, Nikhil Dinesh, Alan Lee, Eleni Miltasakaki, Livio Robaldo, Aravind Joshi, and Bonnie Webber. 2008. The penn discourse treebank 2.0. In *Proceedings of LREC*.
- Veronica Perez Rosas, Carmen Banea, and Rada Mihalcea. 2012. Learning sentiment lexicons in spanish. In *Proceedings of LREC*.
- Asad B. Sayeed, Jordan Boyd-Graber, Bryan Rusk, and Amy Weinberg. 2012. Grammatical structures for word-level sentiment detection. In *Proceedings of NAACL*.
- Richard Socher, Brody Huval, Christopher D. Manning, and Andrew Y. Ng. 2012. Semantic compositionality through recursive matrix-vector spaces. In *Proceedings of EMNLP-CoNLL*.
- Swapna Somasundaran, Galileo Namata, Janyce Wiebe, and Lise Getoor. 2009. Supervised and unsupervised methods in employing discourse relations for improving opinion polarity classification. In *Proceedings of EMNLP*.
- Radu Soricut and Daniel Marcu. 2003. Sentence level discourse parsing using syntactic and lexical information. In *Proceedings of NAACL*.
- Kimberly Voll and Maite Taboada. 2007. Not all words are created equal: Extracting semantic orientation as a function of adjective relevance. In *Proceedings of Australian Conference on Artificial Intelligence*.
- Janyce M. Wiebe. 1994. Tracking point of view in narrative. *Computational Linguistics*, 20(2).
- Theresa Wilson, Paul Hoffmann, Swapna Somasundaran, Jason Kessler, Janyce Wiebe, Yejin Choi, Claire Cardie, Ellen Riloff, and Siddharth Patwardhan. 2005. Opinionfinder: A system for subjectivity analysis. In *Proceedings of HLT-EMNLP: Interactive Demonstrations*.
- Ainur Yessenalina and Claire Cardie. 2011. Compositional matrix-space models for sentiment analysis. In *Proceedings of EMNLP*.
- Ainur Yessenalina, Yejin Choi, and Claire Cardie. 2010a. Automatically generating annotator rationales to improve sentiment classification. In *Proceedings of ACL: Short Papers*.
- Ainur Yessenalina, Yisong Yue, and Claire Cardie. 2010b. Multi-Level structured models for Document-Level sentiment classification. In *Proceedings of EMNLP*.
- Chun-Nam John Yu and Thorsten Joachims. 2009. Learning structural svms with latent variables. In *Proceedings of ICML*.
- Omar F. Zaidan, Jason Eisner, and Christine Piatko. 2007. Using "annotator rationales" to improve machine learning for text categorization. In *Proceedings of HLT-NAACL*.