



HAL
open science

Learning Subjectivity Phrases missing from Resources through a Large Set of Semantic Tests

Matthieu Vernier, Laura Monceaux, Béatrice Daille

► **To cite this version:**

Matthieu Vernier, Laura Monceaux, Béatrice Daille. Learning Subjectivity Phrases missing from Resources through a Large Set of Semantic Tests. The 7th International Conference on Language Resources and Evaluation (LREC'10), May 2010, La Valette, Malta. pp.1335–1341. hal-00472168

HAL Id: hal-00472168

<https://hal.science/hal-00472168>

Submitted on 9 Apr 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Learning Subjectivity Phrases missing from Resources through a Large Set of Semantic Tests

Matthieu Vernier, Laura Monceaux, Béatrice Daille

LINA - CNRS UMR 6241 – University of Nantes
2, rue de la Houssinière BP 92208, 44322 NANTES CEDEX 03, France
Matthieu.Vernier@Univ-Nantes.fr, Laura.Monceaux@Univ-Nantes.fr, Béatrice.Daille@Univ-Nantes.fr

Abstract

In recent years, blogs and social networks have particularly boosted interests for opinion mining research. In order to satisfy real-scale applicative needs, a main task is to create or to enhance lexical and semantic resources on evaluative language. Classical resources of the area are mostly built for english, they contain simple opinion word markers and are far to cover the lexical richness of this linguistic phenomenon. We propose a new method, applied on french, to enhance automatically an opinion word lexicon. This learning method relies on linguistic uses of internet users and on semantic tests to infer the degree of subjectivity of many new adjectives, nouns, verbs, noun phrases, verbal phrases which are usually forgotten by other resources.

1. Introduction

Web 2.0 as a free expression area has literally boosted interests for opinion mining research. Through blogs and social networks (Twitter, Facebook), users share their sentiments and give media coverage to their points of view to influence their communities. In computational linguistic, blogs are more often used as a support study for opinion mining (Mishne and Glance, 2006) (Conrad and Schilder, 2007) (Kessler and Nicolov, 2009) but are more complex to process than text reviews according to (Liu, 2009). Like in the recent text mining challenge (DEFT'09), a current problem is to annotate fine-grained subjective segments (Wilson, 2008) instead of classifying text, then to categorize different semantic aspects of these segments and to detect evaluated targets (Stoyanov and Cardie, 2008) (Ruppenhofer and al., 2008).

In this context, *Apopsis* (Vernier and al., 2009b) is a tool for fine-grained subjective segments detection and categorization : axiological polarity, discursive role (assessment, judgement, agreement, disagreement, etc.), enunciative strategy, speaker engagement (does he assume his subjectivity or does he try to hide it ?). This tool is based on a french lexico-semantic resource built manually (982 entries). It was evaluated qualitatively and quantitatively on a corpus of 100 blogs and during the DEFT'09¹ evaluation campaign (Vernier and al., 2009a) where it obtained the best results. These tests show a good precision (from 0.80 to 0.90) but average quality on the quantitative aspect (recall is around 0.50). The resource coverage is the main factor explaining undetected opinions.

In this article, we focus specifically on this issue by presenting a learning method, applied on french, to learn automatically new words and phrases of subjectivity. We attach a particular importance to not bring down the quality of the initial manually-built resource. The learning method relies on document contents indexing by a search engine and results given in response to a large set of queries. The construction of these queries, linguistically motivated,

can infer the subjective degree of many new adjectives, nouns, verbs, noun phrases, verbal phrases. In particular, we argue that this method is able to learn less frequent but meaningful words and phrases of subjectivity which are usually forgotten by other resources and which can be relevant for real applicative tasks.

Several important works in opinion mining have led to create rich lexical resources manually or semi-automatically : WordNet-Affect (Strapparava and Valitutti, 2004), SentiWordNet (Esuli and Sebastiani, 2006). As noticed by (Banea and al., 2008), these resources are only available for a handful of languages and especially for english. In order to get round the cost of manual creation of such resources, some promising approaches try to determine automatically word's degree of subjectivity (Banea and al., 2008) or word's polarity (Turney, 2002) with the idea that word's polarity can be identified by measuring its co-occurrence with some words whose polarity is known in advance, if a given word occurs with a high probability with a positive (negative) words it can be considered subjective and positive (negative). Results obtained by these methods are interesting but are not able to cover and detect all subjective segments in texts. In particular, these methods are not made for learning infrequent words, subjective phrases or subjective collocations built with objective words : *bol d'oxygène* (\approx a breath of fresh air), *bourreau de travail* (\approx work-a-holic). However, some infrequent words or subjectives phrases are particularly meaningful in appraisal language. The purpose of our work is to take account of these points to enhance the initial french opinion word lexicon.

2. Opinion Word Lexicon

2.1. Linguistic definitions of subjectivity and evaluation

(Lavelle, 1950) defines **evaluation** as *the act of breaking the indifference by which we put things on the same level and we consider all the actions as equivalent*. Every speech act that reveals a break in the indifference results from the evaluative phenomenon. These acts involve complex

¹<http://deft09.limsi.fr/>

semantic, pragmatic or enunciative mechanisms which have been the subject of many research studies (Benveniste, 1974) (Anscombe and Ducrot, 1983). According to (Kerbrat-Orecchioni, 1997), when an enunciative speaker has to select some units from his lexical and syntactic knowledge, he has the choice between two types of formulation :

- **objective** discourse : which strive to hide every marks of the enunciative speaker’s presence,
- **subjective** discourse : in which the enunciative speaker admit his presence explicitly (*Je trouve ça moche* (I find it ugly)) or implicitly (*c’est moche* (This is ugly)).

(Charaudeau, 1992) points out that there are five modalities that allow a speaker to express an evaluation (**opinion, agreement or disagreement, acceptance or refusal, judgement and appreciation**). Each of these modalities reveals a particular attitude of the speaker: his belief which is more or less strong regarding the evaluation he expresses, the experience field from which he takes a stand (ethic, moral, intellectual, aesthetic, etc.), his position in relation to his statement (presence or absence of *I*). These theories are very similar with the *Appraisal Theory* (Martin and White, 2005) for english. According to Charaudeau, these modalities have specific lexical markers and linguistic symbolic structures. (Galatanu, 2000)’s theory on evaluation completes Charaudeau’s as it organizes modalities into hierarchy on a scale of subjectivity. When a speaker structures his statement, he can choose to objectivize or subjectivize his speech by activating some modalities. In the examples (Table 1), the concerned feature *to lie* (modality of judgement (ethic/moral)) is part of a different argumentative strategy. The speaker hides his presence (implicit configuration) and he can sometimes use a modality to modalise another one. Thus, the evaluation *Je n’aime pas qu’il mente* (I don’t like when he lies) will appear more personal (or more subjective) than *Nous condamnons ses mensonges* (We condemn his lies) or *Oui, c’est un menteur* (Yes, he is a liar) even if these phrases use the same evaluative value : *mentir* (to lie).

2.2. Toward an exhaustive resource of subjectivity markers

These linguistic studies show that word subjectivity is particularly context-dependent. An objective word at a semantical level can become subjective at a pragmatical level. Thus, in the following example the adjective *anglais* (english) has in itself a subjective meaning because of the speaker’s enunciation :

- *Il est terriblement anglais (c’est d’ailleurs pour cela que je l’aime autant)*
- He is terribly english (that’s why i like him so much)

Nevertheless, some words or phrases are already subjectives at a semantical level (*mentir* (to lie), *intéressant* (interesting), etc.) or are so much used in a subjective way

Example	2nd Modality
<i>Je doute qu’il mente</i> <i>I doubt that he’s lying</i>	Weak Explicit Opinion
<i>Il est évident qu’il ment</i> <i>This is obvious that he’s lying</i>	Strong Implicit Opinion
<i>Oui, c’est un menteur</i> <i>Yes, he is a liar</i>	Agreement
<i>Il ment</i> <i>He is lying</i>	no other modality
<i>Je n’aime pas qu’il mente</i> <i>I don’t like when he lies</i>	Explicit Appreciation
<i>Nous condamnons ses mensonges</i> <i>We condemn his lies</i>	Explicit Judgement

Table 1: Example of evaluative discourse for the same evaluative value : *mentir* (to lie) (1st modality : Implicit Judgement)

at a pragmatical level (*donner de la confiture aux cochons* (to cast pearls before swine), *crier de joie* (to shout for joy), *crier au loup* (to cry wolf)) that it makes sense to add them in a resource for subjectivity processing. Typically, all these pragmatic subjective phrases are not present in the classical lexical resources.

The initial opinion word lexicon that we have built manually, contains 982 lexical entries (most of them are simple words): adjectives (493), verbs (192), nouns (166), etc. It has been built from the annotated evaluative passages of the Blogoscopie corpus. We refer to (Dubreil and al., 2008) for a more accurate description of the annotation methodology. Each lexical entry is described by morphosyntactic and semantic informations according to linguistic theories seen previously and by its context in the corpus. The evaluative term *serious* has the following informations:

- **lemma**: serious
- **grammatical category**: adjective
- **evaluation**: appreciation **polarity**: negative **context**: nothing serious will happen to him
- **evaluation**: judgement **polarity**: negative **context**: raise serious problems
- **evaluation**: judgement **polarity**: positive **context**: He is very serious when he is working

3. Semantic Tests of Subjectivity

Our aim is to enhance the french opinion words lexicon by adding **terms** of subjectivity not present previously. In language, these terms can be words (*néfaste* (harmful), *zizanie* (ill-feeling), *laminer* (to laminate)) or phrases (*rafler la mise* (\approx steal the limelight), *faire un pied de nez* (\approx to thumb one’s nose), *vent de panique* (\approx a wave of panic)). It can be adjectives, nouns (or noun phrases) or verbs (verbal phrases). To achieve this objective, we present the principle of *semantic tests* that underlies the machine learning method.

Some french adjectives (*vrai* (true), *véritable* (real)) or adverbs (*littéralement* (literally, truly), etc) have a particular impact on the enunciation and on subjectivity (Legallois, 2005) (Suhamy, 2006). Thus, it is considered that the word *littéralement* (literally, truly) should not be taken literally, and has instead, by common usage, a feature that reveals the intensive mental representations and the speaker's subjectivity: *le contribuable est littéralement écrasé d'impôts* (\approx the taxpayer is literally crushed by taxation), *il a littéralement déplacé une montagne* (\approx he has literally moved a mountain). We develop this idea to formulate the following hypothesis:

Assumption: A neutral term (adjective, noun or verb) is rarely intensified by an intensity marker.

It makes sense to say:

- Il est *particulièrement* **dynamique**, He is *very* **dynamic**
- C'est *véritablement* **une hérésie**, It's a *true* **heresy**
- Il est *littéralement* **tombe sous le charme**, he *truly* **fall under the spell**

Whereas the following sentences seem semantically badly constructed:

- C'est *terriblement* **scalaire**, It is *terribly* **scalar**
- C'est *littéralement* **un oiseau**, It's *literally* **a bird**
- Il a *littéralement* **mangé au restaurant**, He *truly* **ate at restaurant**

From this principle, we define a set of semantic tests combining :

- an element of an intensity marker list (*Particulièrement, Terriblement, Parfaitement, Véritablement, Littéralement, Réellement, Franchement, Véritable*)
- and a given term with an unknown degree of subjectivity

To perform these tests, we rely on linguistic uses of internet users and their frequencies. The hypothesis is to consider that relevance of an utterance can be established by its number of hits on the web. For example, from Yahoo search engine, the following queries provide an indication on subjectivity degrees of terms in bold:

- *véritablement* **scalaire** (truly scalar) \rightarrow 0 occurrence / **scalaire** (scalar) \rightarrow 650 000 occurrences)
- *littéralement* **mangé au restaurant** (really ate at the restaurant) \rightarrow 0 occ. / **mangé au restaurant** (ate at the restaurant) \rightarrow 25 100 occ.)
- *véritable* **hérésie** (true heresy) \rightarrow 13 occ. / **hérésie** (heresy) \rightarrow 460 000 occ.)
- *littéralement* **soulevé la foule** (really raised the crowd) \rightarrow 15 occ. / **soulevé la foule** (raised the crowd) \rightarrow 9 990 occ.)

hérésie (heresy) and *soulever la foule* (to raise the crowd) are potentially subjectives as these early results.

From a technical standpoint, Yahoo Search BOSS² enables to build and execute automatically these queries toward Yahoo!Search and to get number of hits.

4. Learning Method

4.1. Candidates extraction

For experimental purpose, candidates are extracted automatically by sending queries to Yahoo!Search. Each nouns/noun phrases, verbs/verbal phrases or adjectives following an intensity marker in the Yahoo index is collected. We used eight intensity markers of french : *Particulièrement* (particularly), *Terriblement* (terribly), *Parfaitement* (perfectly), *Véritablement* (really), *Littéralement* (litteraly), *Réellement* (really), *Franchement* (frankly), *Véritable* (real). These adverbs are chose because of their high frequency of occurrence in french. In answer to these eight queries, we consider every *abstracts* given by Yahoo! as a text of our corpus. Then, we use TreeTagger (Schmid, 1994) and a chunking algorithm (Vergne and Giguët, 1998) to extract every nominal phrases and verbal phrases placed just after an intensity marker in this corpus. This process is realised automatically within the UIMA platform (Ferruci and Lally, 2004) and the component *fr.univ.nantes.lina.uima.YahooSearch*³ that we developed to send queries to Yahoo. Thus, approximatively 24,500 different candidates have been collected (9,000 nouns/noun phrases, 6,500 verbs/verbal phrases and 9,000 adjectives).

Adjectives	Nouns	Verbs
aborigène (O)	république (O)	prendre la grosse tête (S)
aboriginal (O)	republic(O)	getting full of yourself (S)
téléphonique (O)	république bananière (S)	tricoter (O)
telephone (O)	banana republic (S)	to knit (O)
néo-nazi (A)	vie de chien (S)	échapper des griffes (S)
nazi (A)	dog's life (S)	to run away from (S)
populiste (S)	vie de famille (S)	voler la vedette (S)
populist (O)	family life (S)	to steal the show (S)
télégénique (S)	souffle de fraîcheur (S)	glandouiller (S)
telegenic (S)	a touch of freshness (S)	to do useless things (S)

Table 2: Examples of candidates for opinion words lexicon extracted by Yahoo! and classified by five human-judges.

4.2. Training data

In order to build a training dataset, five human-judges have manually classified 1,500 of theses candidates : 500 adjectives, 500 nouns or noun phrases, 500 verbs or verbal

²<http://developer.yahoo.com/search/boss/>

³Tutorial and sources available here: <http://www.uima-fr.org>

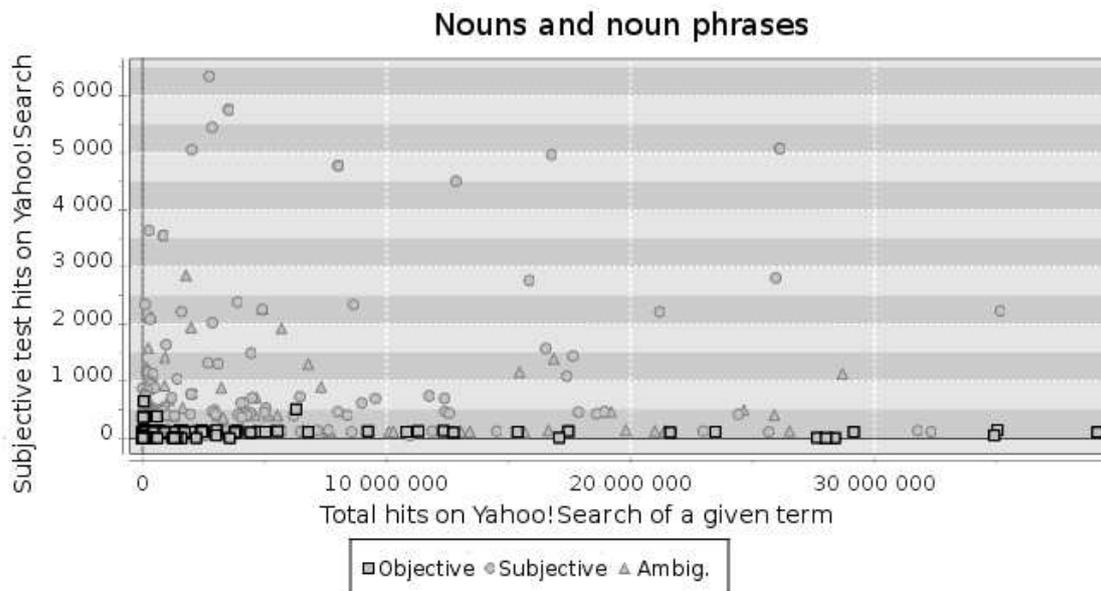


Figure 1: Nouns and noun phrases distributions (with human-judge categorizations) along two axes : number of hits of a given noun (X) and number of hits of the given noun with an intensity marker (Y) in Yahoo!Search Index.

phrases. For each term, human-judges have to decide if the candidate is : *subjective* (S), *objective* (O) or *ambiguous without context* (A). The agreement (0.70) is measured with Fleiss Kappa (Fleiss, 1971). Table 2 shows examples of words classified by human-judges.

In the training dataset, several attributes are automatically added to each candidate : we measure pointwise mutual information between each intensity marker (x) and each candidate (y) considering number of hits given by Yahoo!Search.

$$SI(x, y) = \log \frac{hit(x, y)}{hit(x)hit(y)}$$

As an example, the adjective *anglais* (english) can be subjective : the value of $hit(x, anglais)$ is more than 500, but $hit(anglais)$ is more than 300.000.000, so this adjective might not be a good candidate for the lexical resource.

4.3. Supervised classification of new terms

In a two-dimensional representation, figure 1 shows distributions of nouns and noun phrases classified by human-judges along two axes :

- **Y** : number of hits of an intensity marker followed by a given term (semantic test of subjectivity) (ex : *littéralement pété les plombs* (1 330 hits) (literally *flip my lid* (70 hits))
- **X** : number of hits only for the given term (ex : *péter les plombs* (78 700)) (*flip my lid* (23,700 hits))

Two-dimensional representations of verbs/verbal phrases and adjectives are equivalent to figure 1. From the training dataset and for each grammatical category, we trained a Support Vector Machine (Joachims, 1997) classifier to search for an optimal hyperplan. Then, we apply the

classification function to separate subjective and objective candidates in the initial list of 24,500 terms.

5. Evaluation

type	number	examples (random selection)
Adject.	596	larmoyant (whining), exorbitant (exorbitant), opiniâtre (\approx obstinate, bulldog), lunatique (moody person), incestueux (incestuous), cocace (comical), famélique (scrawny), infantile (childish), subversif (subversive)
Nouns, noun phrases	1,390	régal (delight), fléau (plague), plébiscite (plebiscite), camouflet (poking), marée humaine (\approx human tide), descente aux enfers (descent into hell), gain de temps (time-savings), cacophonie (cacophony), bouffée d'air frais (breath of fresh air), capharnaüm (\approx shambles, souk)
Verbs, verbal phrases	488	jouer un rôle décisif (\approx to play a decisive role), faire basculer le match (\approx to change the momentum of a game), subjuguier (to subjugate), voler la vedette (\approx to steal the show), toucher le fond (to plumb the depths), ovationner (\approx to greet somebody with wild applause)

Table 3: Examples of terms added to lexicon (random selection).

The method described above enables to extract 2,474 new terms (Tab.3) and to add them to the initial french opinion words lexicon (initially 982 lexical entry) with a metadata to inform that each new term have been added automatically. At this point, we don't address the problem of word's polarity categorization. We plan to also use pointwise mutual information as described in several

works (Turney, 2002) (Bestgen and al., 2004) for this purpose. Nevertheless, we argue that the step of candidates extraction based on semantic tests is important to reduce noise and improve resource coverage.

5.1. Lexicon enhancement evaluation without textual context

Lexicon enhancement evaluation would require a standard resource on which to compare, but our work is precisely motivated by the lack of such resource. In consideration of this aspect, we made a first validation of lexicon enhancement considering the list of 1,500 terms classified by human-judges as a reference. We used a ten cross-validation method during the learning phase to measure the precision and the recall (table 4).

<i>Axiology</i>	Precision	Recall
Objective	75,49% (687/910)	94,62% (687/726)
Subjective	77,28% (456/590)	61,81% (356/576)
Ambiguous	0% (0/0)	0% (0/198)

Table 4: Objective terms and subjective terms classification results in comparison with 1,500 terms classified by human-judges.

Results shows that even if recall measure of subjective terms is quite low (Figure 1 shows that lots of subjective nouns are approximately in the same area of objective terms), this method enables to extract 456 subjective candidates with an interesting precision. Nevertheless, we underline that contextual evaluation is a predominant aspect, in particular in opinion mining and subjectivity research domain. This bring us to suggest a different evaluation protocol to observe lexical enhancement impact on an real applicative task. Our evaluation protocol differs from (Turney, 2002) who evaluate his method by comparing results to General Inquirer resource. This protocol has the tendency to consider always correct some *subjective* terms which can be used in an objective context.

5.2. Lexicon enhancement evaluation with textual context

For this second evaluation protocol, we extract 5,000 posts from french blog platform Over-blog without any constraints on themes or on post sizes. Then, we use the tool Apopsis to annotate fine-grained subjective appraisal segments. This tool relies on approximately 2,000 grammar rules and on pattern recognition method for evaluative segment detection in texts (Vernier and al., 2009b).

At the end of the natural language process, two files (CSV) are generated to list : on the one hand, subjective appraisal segments detected with initial opinion words lexicon and on the other hand, segments detected with enhanced lexicon part. Two human-judges estimate the precision of the enhanced part. Observed agreement between human-judges is 0.76. Results are sum up in table 5.

Table 5 shows that mistakes and disagreements between humans-judges are particularly concerned by nouns and

Words/Phrases	Total	Verbs	Nouns	Adject.
TOTAL	17,669	2,132	11,235	4,250
not correct	3,632	330	3,024	230
correct	13,450	1 793	7,657	4,000
can't be evaluated	587	9	606	20
PRECISION	78,7%	84,5%	72,0%	94,6%

Table 5: Precision of fine-grained subjective segments detection with subjective words and phrases learnt automatically.

noun phrases. This grammatical category is inclined to activate different cultural stereotypes between human-judges. Thus, many examples containing the following expressions lead to disagreement: *crise économique* (economical crisis), *politique écologique* (ecological politic), *terrorisme* (terrorism) ou *pandémie* (pandemic). For examples :

- *La pandémie de grippe, réelle ou inventée, permet de mettre en scene le final[...]* (the flu pandemic, real or invented, enables to put a spotlight on[...])
- *Le mot pandémie est d'actualité, nous l'entendons même depuis des mois.* (The word pandemic is buzzing, we hear about it for months)

As noticed by DEFT'09 program committee, the lack of standard corpus to evaluate fine-grained subjectivity detection is still not resolved for french. Recall measure can't be evaluated, it would require an exhaustive manual annotation by human-judges. Nevertheless, we estimate the lexicon enhancement on quantitative aspect by comparing number of subjective segments annotated with initial lexicon (68 536) and by subjective segments annotated correctly with enhanced lexicon (+13 450) : +15,6%.

6. Discussion & Conclusion

From a quantitative point of view, french opinion words lexicon raised from 982 to 3,456 entries (+252%). A first comment concern subjective segments detection which has improved of *only* 20% in comparison. However, this improvement is far from being not significant for the following reason : compared with terms from the initial french opinion word lexicon (*beau* (*beautiful*), *inquiétude* (*worry*), *aimer* (*to love*)), enhanced terms (*blasphématoire* (*blasphemous*), *la politique de l'autruche* (\approx *ostrich policy* : *to bury one's head in the sand*)), *faire tordre de rire* (\approx *to convulse*) have a lower frequency - explaining why they are often forgotten by manually-built resources - but are meaningful for real applicative tasks.

New adjectives and verbs/verbal phrases enables to annotate subjective segments with good accuracy. Ambiguity and mistake sources come from new nouns and noun phrases. Nouns are more polysemic in french, like *farce* or *daube* which have real subjective uses (*c'est une farce cette assemblée de politiciens* (these politicians are a joke)) but also a culinary meaning. As noticed by (Kerbrat-Orecchioni, 1997), subjective words and phrases are not stable in language (terms like *collaboration*

or *to collaborate* do not represent the same cultural stereotype nowadays and in the context of the second world war). Our method is based on linguistic uses of internet users and in this way can follow the evolution of some cultural stereotypes : the most admitted ones (*pantouflard* (\approx stay-at-home), *négaționniste* (holocaust denier), *escroquerie* (swindling)), but also the most recent (thus *écologie*, *écologique* (*ecology*), *pollution* (*pollution*)) are automatically classified as subjective because of the intensity expressed around these concepts at present).

The french opinion words lexicon will be made publicly available at the following address: <http://www.blogoscopie.org>

7. References

- J.C. Anscombe and O. Ducrot. 1983. *L'argumentation dans la langue*. Pierre Mardag.
- Carmen Banea, Rada Mihalcea, and Janyce Wiebe. 2008. A bootstrapping method for building subjectivity lexicons for languages with scarce resources. In *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco, may. European Language Resources Association (ELRA). <http://www.lrec-conf.org/proceedings/lrec2008/>.
- Emile Benveniste. 1974. *Problèmes de linguistique générale II*. Gallimard edition.
- Yves Bestgen, Cédric Fairon, and Laurent Kerves. 2004. Un baromètre affectif effectif: Corpus de référence et méthode pour déterminer la valence affective de phrases. In *Journées internationales d'analyse statistique des données textuelles (JADT)*, pages 182–191.
- Patrick Charaudeau. 1992. *Grammaire du sens et de l'expression*. Hachette Education, COMMUNICATION, PARA UNIVERSITAIRE.
- Jack G. Conrad and Frank Schilder. 2007. Opinion mining in legal blogs. In *ICAIL '07: Proceedings of the 11th international conference on Artificial intelligence and law*, pages 231–236, New York, NY, USA. ACM.
- Estelle Dubreil, Matthieu Vernier, Laura Monceaux, and Béatrice Daille. 2008. Annotating opinion - evaluation of blogs. In *Workshop on LREC 2008 Conference, Sentiment Analysis: Metaphor, Ontology and Terminology (EMOT-08)*.
- Andrea Esuli and Fabrizio Sebastiani. 2006. SentiWordNet: A publicly available lexical resource for opinion mining. In *Proceedings of Language Resources and Evaluation (LREC)*.
- D Ferruci and A Lally. 2004. Uima : an architectural approach to unstructured information processing in the corporate research environment. In *Natural Language Engineering*, 10(3-4), pages 327–348.
- J. L. Fleiss. 1971. Measuring nominal scale agreement among many raters. In *psychological Bulletin*, Vol. 76, No.5, pages 378–382.
- Olga Galatanu. 2000. Signification, sens, formation. In *Education et Formation, Biennales de l'éducation, (sous la direction de Jean-Marie Barbier, d'Olga Galatanu)*, Paris. PUF.
- T. Joachims. 1997. Text categorization with support vector machines: Learning with many relevant features. Rapport Interne Ls8-Report 23, Universität Dortmund.
- C. Kerbrat-Orecchioni. 1997. *L'Énonciation, de la subjectivité dans le langage*. Colin (réédition 2002).
- Jason S. Kessler and Nicolas Nicolov. 2009. Targeting sentiment expressions through supervised ranking of linguistic configurations. In *3rd Int'l AAAI Conference on Weblogs and Social Media (ICWSM 2009)*.
- Louis Lavelle. 1950. *Traité des valeurs*, volume tome 1. PUF.
- Dominique Legallois. 2005. Pour une définition énonciative de l'enclosure vrai. In *Les marqueurs linguistiques de la présence de l'auteur*. David Banks, L'Harmattan.
- Bing Liu. 2009. Sentiment analysis. In *Invited talk at the 5th Annual Text Analytics Summit*.
- James Martin and Peter White. 2005. *The Language of Evaluation, Appraisal in English*. Palgrave Macmillan, London, New York.
- Gilad Mishne and Natalie Glance. 2006. Predicting movie sales from blogger sentiment. In *AAAI Symposium on Computational Approaches to Analysing Weblogs (AAAI-CAAW)*, pages 155–158.
- Josef Ruppenhofer, Swapna Somasundaran, and Janyce Wiebe. 2008. Finding the sources and targets of subjective expressions. In *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco, may. European Language Resources Association (ELRA). <http://www.lrec-conf.org/proceedings/lrec2008/>.
- Helmut Schmid. 1994. Probabilistic part-of-speech tagging using decision trees. In *Proceedings of the International Conference on New Methods in Language Processing*, pages 44–49.
- Veselin Stoyanov and Claire Cardie. 2008. Annotating topics of opinions. In *Proceedings of the Sixth International Language Resources and Evaluation (LREC'08)*, Marrakech, Morocco, may. European Language Resources Association (ELRA). <http://www.lrec-conf.org/proceedings/lrec2008/>.
- Carlos Strapparava and Alessandro Valitutti. 2004. Wordnet-affect: an affective extension of wordnet. In *Proceedings of LREC*, volume 4, pages 1083–1086.
- Henri Suhamy. 2006. Métaphore et dualité. In *Bulletin de la Société de stylistique anglaise (ISSN 0240-4273)*, numéro 28, pages 9–24.
- Peter Turney. 2002. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In *Proceedings of the Association for Computational Linguistics (ACL)*, pages 417–424.
- Jacques Vergne and Emmanuel Giguët. 1998. Regards théoriques sur le tagging. In *Actes de Traitement Automatique des Langues Naturelles (TALN'98)*, pages 22–31.
- Matthieu Vernier, Laura Monceaux, and Béatrice Daille. 2009a. Deft'09 : détection de la subjectivité et catégorisation de textes subjectifs par une approche

mixte symbolique et statistique. In *Actes de l'atelier de clôture du 5ème Défi Fouille de Textes (DEFT'09)*, pages 101–112, Paris, France, June.

Matthieu Vernier, Laura Monceaux, Béatrice Daille, and Estelle Dubreil. 2009b. Catégorisation des évaluations dans un corpus de blogs multi-domaine. In *Numéro spécial de la revue RNTI (Revue des Nouvelles Technologies de l'Information) - fouille de données d'opinion*.

Theresa Ann Wilson. 2008. *Fine-grained Subjectivity and Sentiment Analysis: Recognizing the Intensity, Polarity, and Attitudes of Private States*. Ph.D. thesis, University of Pittsburgh.