

# “Haters gonna hate”: challenges for sentiment analysis of Facebook comments in Brazilian Portuguese

**Juliano Desiderato Antonio**  
Universidade Estadual de Maringá  
jdantonio@uem.br

**Ana Carolina Leatte Santin**  
Universidade Estadual de Maringá  
leattesantin@gmail.com

## Abstract

The aim of this paper is to present reflections from Discourse Analysis and from Construction Grammar on the creation of a dictionary for sentiment analysis of Facebook comments. The reflections from Discourse Analysis address problems such as the identification of the semantic orientation of words that present opposite polarities depending on the ideological formation of the speaker. Another reflection from Discourse Analysis regards the fact that the writers of the comments use nouns and noun phrases not only to name some entity, but also to build discourse objects in a way that the label they give to the discourse objects reveals an evaluation. In order to analyze constructions larger than words, such as idioms, we draw on Construction Grammar principles. The investigation of constructions and idioms can provide a better understanding of sentiment in text. The *corpus* consists of comments extracted manually from Facebook public discussion pages related to diverse themes, such as politics, education, religion, music, lifestyle etc.

## 1 Introduction

Facebook is one of the websites with higher data traffic on the internet. On December 2016 it registered 1.86 billion monthly active users (nearly 1 in 4 people worldwide). In Brazil, the ratio is even higher: nearly 55.5% of the population were active Facebook users in November 2016 (Facebook, 2017). Thus, linguists who are interested in investigating language in use have on Facebook an immeasurable research field.

Wilson et al.'s (2012) comprehensive review about scientific research conducted about Face-

book in the social sciences points to the need for social engagement as the main motivation for people to use Facebook. Seidman (2014) investigated the expression of the “true self” on Facebook, which “consists of qualities that an individual currently possesses but does not normally express to others in everyday life”. The need for social engagement and the expression of the true self are encompassed within the metafunctions<sup>1</sup> proposed by Halliday (1985). The ideational metafunction regards self-expression, as it concerns the grammatical features used to construct both the speaker's inner experiences and the experiences with the world around him. On the other hand, the interpersonal metafunction is related to the grammatical resources used by the speaker to interact with his/her interlocutors, assuming social roles and roles concerning the communicative situation (social engagement). Whilst the ideational and the interpersonal are extralinguistic metafunctions, the textual metafunction deals with the presentation of interpersonal and ideational content in the form of information that can be shared by the speaker and his / her interlocutors by means of texts. In other words, texts produced by Facebook users and their linguistic behavior deserve being studied by linguists.

According to Irukieta et al. (2013), Computational Linguistics depends on discourse annotated *corpora* for the creation of automatic applications. The research that resulted in this paper intends to create a dictionary for sentiment analysis by extracting comments from Facebook public pages related to diverse themes, such as politics, education, religion, music, lifestyle etc.

---

<sup>1</sup> “Metafunction refers to the different modes of meaning construed by the grammar.”

However, as we started analyzing the semantic orientation of the comments, we noticed that the same words said by different people had polar opposite semantic orientation, as in examples (1) and (2).

- (1) Só come "pala" da *Direita* quem for analfabeto político e funcional.

Only functional and political illiterate people believe in the lies of the *right wing*.

- (2) ... nós da *Direita* não temos político de estimação.

We, who are *right wing*, do not have pet politicians.

In (1) “right wing” is considered negative by the writer of the comment, as it can be presupposed that assumptions endorsed by right wing are lies. On the other hand, in (2) the writer of the comment, who assumes to have a right wing political orientation, suggests that left wing people have pet politicians, whilst right wing people do not. Therefore, in (2), “right wing” is evaluated positively. That happens because the comments collected were produced by people with different views on the issues discussed in the pages and we decided to draw attention to such problem, as it certainly creates difficulties for sentiment analysis.

Thus, this paper presents some challenges for the creation of a dictionary for sentiments analysis of Facebook comments in Brazilian Portuguese caused by the different positions assumed by the producers of the comments. Furthermore, it is also a goal of the paper to analyze other forms rather than nouns, adjectives, verbs, NPs. The investigation of constructions and idioms can provide a better understanding of sentiment in text.

In terms of structure, besides the introduction, this paper is divided in 4 more sections. In Section 2 we present a short view of what has been done about sentiment analysis in Linguistics and in NLP. We also introduce in Section 2 some theoretical assumptions from Discourse Analysis in order to face the challenges which are addressed in the paper. A brief review of Constructional Grammar is also presented in Section 2. In Section 3 we present the methodology used in the research and the discussion of the data is provided in

Section 4. The last section of the paper is the Conclusion, followed by the references.

## 2 Theoretical background

In this section we provide a general background about what has been done regarding sentiment analysis and also some contributions from Discourse Analysis which are essential for the challenges discussed in the paper. A brief review of basic assumptions of Constructional Grammar is presented in order to provide a better understanding of the concepts of “construction” and “idiom”.

### 2.1 Sentiment analysis

According to Taboada (2016), “sentiment analysis is a growing field at the intersection of linguistics and computer science that attempts to automatically determine the sentiment contained in text”. Sentiment is conceived as positive or negative evaluation conveyed by linguistic expression, both lexical and grammatical. Beyond defining sentiment analysis, Taboada (2016) also presents a broad view of the contributions of Linguistics to automatic sentiment analysis.

Two main approaches are used for automatic extraction of sentiments: machine learning and lexicon based (Taboada, 2016). We will focus on the latter, as it is the type of method we intend to implement in the subsequent stages of the project we are developing.

Among the lexicon based approaches, Taboada (2016) mentions some dictionaries for sentiment analysis: SentiWordNet (Baccianella et al., 2010) catalogues about 38,000 words regarding their polarity; also based on polarity, Macquarie Semantic Orientation Lexicon (Mohammad et al., 2009) classifies almost 76,000 words; Subjectivity dictionary (Wilson et al., 2009) not only presents the polarity of the words, but also groups them according to their strength (strong positive, weak positive, neutral, weak negative, strong negative); Semantic Orientation Calculator (SO-CAL) (Taboada et al., 2011) stratifies about 5,000 words in a 10-point scale which ranges from -5 to +5.

In Brazilian Portuguese, among many works that deal with sentiment analysis, Sentimeter-Br (Rosa, 2015) is a mechanism for calculating semantic orientation. It is based on a dictionary of words divided according to the area they belong to,

e.g. music, technology, beauty, business. The system also implements a mechanism (Enhanced Sentimeter) which uses the user's profile of a social media to calculate sentiment.

## 2.2 Contributions from Discourse Analysis

Regarding the problem presented in the introduction of the paper with the different evaluations of "right wing" held by the writers of the comments, which have opposite political views, Pêcheux (1975) states that words, expressions, propositions etc do not have a self-contained meaning. On the contrary, their meanings change according to the positions supported by the speakers who use them, i.e. their ideological formations. In (1) "right wing" has a negative evaluation because the comment was written by a person of left wing ideological formation. On the other hand, in (2) "right wing" has a positive evaluation because the writer of the comment belongs to a right wing ideological formation. Thus, ideological formation is an important feature to be taken into account in order to identify sentiment towards propositions.

Another important reflection from Discourse Analysis regards the difference between reference and *référenciation*<sup>2</sup> proposed by Mondada and Dubois (1995). When one makes reference, he / she names in an objective way anything that is in the world (*designatum*, according to Lyons [1977]). On the other hand, in the *référenciation* process, the speaker builds discourse objects in a way that they can be categorized and recategorized. In (3) the writer of the comment, which was taken from a left wing Facebook page, uses the "communist doctrinators" NP in order to show how a group of right wing people refers to teachers. Obviously, it is not only naming. If it were like this, the group which is criticized in the comment would use the noun "teachers". Actually, the "communist doctrinators" NP reveals the treatment of the right wing group towards the discourse object "teachers". External world class "teachers" is not affected by the way the right wing group refers to it, as "communist doctrinators" is a discourse object.

- (3) Por isso mesmo, para isso funcionar, é preciso demonizar a classe dos professores como "douttrinadores comunistas", isto é, duas palavras que a direita adora usar.

For this reason, for this to work, it is necessary to demonize the class of the teachers as "communist doctrinators", i.e., two words that right wing loves to use.

In other words, in the view of *référenciation*, nouns and noun phrases do not only name a *designatum*, they also present an evaluation of the discourse object.

## 2.3 Constructions and idioms

Beyond nouns, adjectives, verbs, NPs, other forms should be investigated for a better understanding of sentiment in text. Constructions and idioms are widely used by speakers not only in face to face interactions, but also on social media.

According to Traugott and Trousdale (2013), constructions are form-meaning pairings and include morphemes, words, idioms, and abstract phrasal patterns (Goldberg, 2013; Hoffmann and Trousdale, 2013).

The term "construction grammar" (CG henceforth) refers to a group of distinct frameworks which share some tenets, summarized by Goldberg (2013) as follows:

- i. Constructions are the basic units of grammar;
- ii. Semantic structure is associated directly with syntactic structure without transformations or derivations;
- iii. Constructions form a network in which nodes are related by inheritance links;
- iv. Cross-linguistic variation can be explained in terms of domain-general cognitive processes or by the functions of the constructions involved;
- v. Items and generalizations are part of the knowledge of language (this last tenet is shared by most, but not all approaches).

---

<sup>2</sup> We will use the French word "référenciation" (as Mondada and Dubois 1995) because there is not such word in English.

Formal classes	N
Nouns	136
Adjectives	156
Verbs	117
Constructions	22
Idioms	28

**Table 1:** Quantity of items per class.

Unlike Generative Grammar, CG considers grammar in a holistic way, i.e., no grammatical level is considered core or autonomous; a construction is formed by simultaneous work of phonology, morphosyntax, semantics and pragmatics (Traugott and Trousdale, 2013).

### 3 Methodology

The first step of the Methodology was to collect comments from public Facebook pages which discuss issues such as politics, education, religion, music, lifestyle etc. Nearly 1,000 comments were collected, segmented into EDUs<sup>3</sup> and classified either as subjective (present an evaluation) or objective (do not present an evaluation). The latter were eliminated from the corpus.

The remaining 649 EDUs were classified manually by two annotators as positive, negative or neutral, regarding the evaluation of a discourse object. The words and expressions responsible for the evaluation were extracted manually and divided according to their formal classes.

In Table 1 we present the quantity of words per class. Other features such as intensifiers, adverbs, interjections, signals of irony, laughing etc were also annotated but will not be discussed here as they do not refer directly to the main issue discussed in this paper. It is important to remark that words were counted only once, even if they were used more times in the *corpus*. Thus, the quantity in Table 1 refers to the quantity of words found and not to the amount of times they were used.

---

<sup>3</sup> “minimal building blocks of a discourse tree” (Carlson and Marcu, 2001). In general, EDUs are paratactic or hypotactic clauses, but not restrictive or completive clauses.

## 4 Discussion and analysis

In order to stress the importance of taking constructions into account in sentiment analysis, this Section is divided in two subsections: one for commonly investigated classes and forms such as nouns, adjectives, verbs, NPs, and one for constructions and idioms.

### 4.1 Nouns, adjectives, NPs, verbs

Within the view of *référénciation*, the speaker’s communicative intentions govern his / her linguistic choices (Koch, 2002; 2007), as in example (4).

- (4) Escola é um *depósito de criança*, APENAS, no fundo ninguém tá nem aí pro que é ensinado, só se interessam em ter um lugar pra deixar os “*presentes de Deus*” enquanto estão trabalhando. Por isso que quando tem greve os pais ficam tão irados.

School is a *children warehouse*, ONLY, actually nobody cares about what is taught, they are only interested in having a place to leave “*God’s gifts*” while they are working. That’s why parents get so mad when there is a teacher’s strike.

In example (4) the writer of the comment quotes the opinion of people in general about Brazilian public schools, which are considered “*depósito de criança*” (“children warehouse”), i.e., a place where people leave their kids when they go to work. NP “*depósito de criança*” creates a discourse object which reveals a general conception about the *designatum* school.

Although NP “*presentes de Deus*” (“God’s gifts”) carries nouns of positive semantic orientation, it is written between “quotation marks”, in a sarcastic way, which results in a negative evaluation of that discourse object. In other words, the writer of the comment means that there are moments when children are a nuisance to parents, who are not interested in their education, but only in having a place to leave them while they are at work.

Depreciatory collective nouns reveal the speaker’s negative evaluation of a discourse object, especially when the referent is human (Neves, 2000). It is the case of the noun “*bando*” (which could be

roughly translated into English as “gang”). In example (5), scoping adjective “demente” (“demented, in English), it evaluates negatively a group of people. In example (6), it is used to convey negative evaluation of a group of “machos” (“chauvinists”, in English), which are also strongly qualified in a negative manner as “asquerosos” (“loathful).

- (5) É só o reflexo de como viramos um *bando de dementes*.

It is only the reflex of how we became a *gang of demented*.

- (6) *Bando de macho asqueroso!*

*Gang of loathful chauvinists!*

However, negative characteristics may be assumed by a group and transformed into positive evaluation. That’s what happens with the expression “gang of crazy people”<sup>4</sup>, in example (7), used by the supporters of Brazilian football team Sport Club Corinthians as a motivation yell.

- (7) Aqui tem um *bando de loucos*, loucos por ti, Corinthians.

Here there is a *gang of crazy people*, crazy for you, Corinthians.

The comparison between examples (5) and (7) shows similarities – both NPs have the noun “gang” scoping an adjective from insanity semantic field – and also differences – in example (5) the evaluation is negative, while, in example (7), the evaluation is positive. To explain the differences, we have to cite Pêcheux (1975) again, to whom the meanings of words change according to the positions supported by the speakers who use them.

Nouns may lose their referential function in order to express quality (Neves, 2000) and, in such use, they can convey negative or positive evaluation. In example (8), noun “massa” (“mass”) is not used to refer to “matter with no definite shape”, but to qualify noun “página” (“page”) as “cool”.

- (8) Aparece uma página que parece ser *massa*.

A page that seems to be *cool* shows up.

The same happens to noun “show” in example (9). Instead of naming a spectacle, it classifies the “debate” as “amazing, spectacular”.

- (9) O debate foi *show*.

The debate was *amazing*.

In example (10) noun “shit” qualifies noun “time” (“team”) in an extremely pejorative way. In examples (11) and (12), besides the the change of position in the NP, the negative evaluation assigned to noun “filme” (“film”) remains the same<sup>5</sup>. The possibility of such positional change is a particular characteristic of the grammar of Brazilian Portuguese. As it can be noticed, the translation into English is not even possible.

- (10) Que time *merda*.

\*What a *shit* team.

- (11) *Merda* de filme.

\**Shit* of film. (The appropriate translation would be “shitty film”.)

- (12) Filme *de merda*.

\*Film of *shit*. (The appropriate translation would be “shitty film”.)

Adjectives have been widely investigated in sentiment analysis due to their nature. According to Taboada (2016), “adjectives convey much of the subjective content in a text”. However, not all adjectives can be used to evaluate. Classifier adjectives only subcategorize the nouns that they modify (Neves, 2000). As a result, they are not suitable for subjective evaluation, as in example (13), in which adjective “sexual” only specifies the type of

<sup>4</sup> In Brazilian Portuguese, the NP “bando de loucos” (“gang of crazy people”) does not have the noun “people”, as the adjective “crazy” can be used as a noun. Roughly, the literal translation would be “gang of crazies”.

<sup>5</sup> In examples (10), (11) and (12), the correct translation would be “shitty”. As in Brazilian Portuguese the noun “merda” (“shit”) functions as an adjective, the same construction in English is ungrammatical.

“option” (“opção”) mentioned by the writer of the comment.

- (13) O que que a gente tem a ver com a opção *sexual* do outro?

What do we have to do with other people’s sexual options?

On the other hand, qualifier adjectives, in a predication process, attribute properties to the nouns they modify. In example (14), the adjective assigns the quality “feliz” (“happy”) to anyone who fulfills the conditions presented in the subject clause.

- (14) *Feliz* é aquele que encontra a felicidade nos pequenos gestos.

*Happy* is the one who finds happiness in the small gestures.

Regarding syntactic use, adjectives can be predicative, as in example (15), in which “feio” (“ugly”) is the nucleus of the VP, or adnominals, as in example (16), in which “louca” (“crazy”), “desequilibrada” (“unbalanced”) and “insuportável” (“unbearable”) modify noun “gente” (“people”) within the NP.

- (15) ... se vier me perguntar se tu é feio...

... if you ask us if you are ugly...

- (16) ... gente muito louca, desequilibrada e insuportável...

... very crazy, unbalanced and unbearable people...

The semantic orientation of the NP is usually given by the adjective. In example (17), adjective “maravilhoso” (“wonderful”) is responsible for the positive semantic orientation of the NP, whilst in example (18) adjective “fake” is responsible for the negative semantic orientation of the NP.

- (17) Bar maravilhoso

Wonderful pub

- (18) Sorriso falso

## Fake smile

Verbs occupy the central role of a predication (Ilari and Basso, 2008). Thus, the correct identification of the semantic orientation of a clause depends to a great extent on the verb.

Some verbs convey negative (“odeio” – “hate”) or positive (“adoooooro” – “loooove”; “prefiro” – “prefer”) evaluation by their basic meaning, as in examples (19) and (20).

- (19) Pão com ovo. *Adoooooro*.

Bread with fried egg. I *loooove* it.

- (20) Eu não uso isso. Eu *odeio*. *Prefiro* meus tênis.

I don’t wear this. I *hate* it. I *prefer* my tennis shoes.

On the other hand, evaluation conveyed by other verbs can only be identified by the analysis of their arguments. In example (21) and (22) verb “merecer” (“to deserve”) can point to a negative or to a positive evaluation, depending on the semantic orientation of its second argument (A2). In (21) semantic orientation is positive (“compliments”), while in (22) it is negative (“jail”).

- (21) Acompanho e continuarei acompanhando seu trabalho, que merece, sim, *elogios* em muitos pontos.

I follow and I will keep on following your work, which deserves, yes, *compliments* in many aspects.

- (22) Ela merece *cadeia*.

She deserves *jail*.

However, there are constructions in which objects cannot be considered arguments of the verb. Neves (2002), following Ashby and Bentivoglio (1993), calls them “constructions with support verbs”. In such constructions, the object NP forms a predicate with the verb and thus the construction must be analyzed as a whole. In Brazilian Portuguese, the most productive verbs in such constructions are “dar” (“to give”) and “fazer” (“to do”, “to make”). In example (23), constructions with sup-

port verb “dar” have opposite polarities, and the semantic orientations of the comment must be determined by discourse structure (Taboada, 2016). Conjunction “e” (“and”) is usually associated with the idea of additive parataxis. However, it has other uses in Brazilian Portuguese, such as signaling contrast (Camacho, 1999), which is the case in example (23).

(23) Tinha tudo pra *dar errado e deu certo*.

It had everything to *go wrong* and it *went right*.

Neves (2002) presents some reasons why speakers use support verbs, such as the obtainment of more communicative adequacy, more semantic precision and more syntactic versatility. In example (24), the construction with support verb provides more syntactic versatility by the determination of reflexive possession.

(24) Os pais nunca se dão conta disso não, eles acham q o professor não *faz mais do que a obrigação deles...* e se quiser ganhar mais trabalhe mais

Parents do not realize this, they think that teachers don't *do more than their obligation...* and if teachers want to earn more, they have to work more.

## 4.2 Constructions and idioms

The construction presented in example (25) is used to convey positive evaluation. The noun “tatuadores” (“tattoo artists”) can be replaced by a variety of nouns which nominate, among other possibilities: *i*) professions such as “teachers”, “policemen”, “doctors” etc; *ii*) human referents such as “mothers”, “students”, “children” etc; *iii*) animate referents, such as “dogs”, “cats”, “dolphins” etc; *iv*) inanimate referents, such as “cars”, “softwares”, “cell phones” etc; *v*) places such as “schools”, “churches”, “malls” etc.

(25) *Por mais tatuadores como esse...*

*For more tattoo artists like this...*

In example (26), the construction conveys negative evaluation. In the first six clauses, the writer of

the comment criticizes characteristics and actions of a politician. In the seventh clause, he / she presents the action which he / she considers to be the worst of all. In terms of argumentation, this construction saves the strongest argument for last and can be represented as *Não basta PREDICATION, tem que PREDICATION*.

(26) Não basta ser golpista, não basta ser corrupto, não basta comprar a grande maioria dos parlamentares, não basta criar 14 mil cargos desnecessários, não basta conceder aumento desproporcional ao STF, não basta extinguir os ministérios da cultura e da previdência social: ele tem que aniquilar os direitos sociais e trabalhistas!

It's not enough to be a state stroker, it's not enough to be corrupt, it's not enough to buy the majority of the parliamentarians, it's not enough to create 14 thousand unnecessary public job roles, it's not enough to grant a disproportionate pay rise to STF, it's not enough to wipe out the ministries of culture and social care: he has to annihilate the social and labor rights!

In idioms, words do not have independent meanings and the idiosyncrasy (meaning cannot be predicted from form) must be stored in the speakers' long-term memory (Jackendoff, 2013). In example (27), idiom “chutar cachorro morto” (literal translation: “to kick a dead dog”) in Brazilian Portuguese means being aggressive or even doing harm to somebody who is not able to defend him / herself and is insignificant to society. The idiom is used by the writer of the comment to convey a view of society towards teachers which is growing common: teachers are not important.

(27) Mexer com professor é como *chutar cachorro morto*, ninguém liga mais.

Messing with teachers is like *kicking a dead dog*, nobody cares.

In example (28), idiom “beijinho no ombro”, literally translated as “little kiss over the shoulder” has become one of the most widely used idioms in Brazil after it was used as the chorus and the title of a song by a popular Brazilian funk singer. A

quick search on Google, for instance, leads to more than two million results. As it happens with idiom “chutar cachorro morto”, its meaning cannot be predicted from the form, as it expresses a gesture of superiority over envious people, negative people who only criticize, haters etc. In the context of the example, the writer of the comment uses the idiom “beijinho no ombro” in order to show that he / she does not care about peoples’ critics towards the Facebook page she follows. She sends to those people a “beijinho no ombro” to show that she is superior to their negative and critical behaviour.

(28) Não curto página de escola, a única com a qual me identifiquei foi essa por justamente mostrar meus medos, ansios enquanto professora. Se não estão satisfeitas, só lamento. *Beijinho no ombro.*

I don’t follow pages of schools, the only one I identified myself with was this one exactly because it shows my fears, my yearnings as a teacher. If they are not happy with it, I’m sorry. *Little kiss over the shoulder.*

## 5 Conclusion and future work

This paper aimed at discussing some challenges found for the creation of a sentiment analysis dictionary for Facebook comments in Brazilian Portuguese.

The analysis of the *corpus* showed that the same words spoken by different people may have polar opposite semantic orientations. We also noticed that the writers of the comments use nouns and noun phrases not only to name some entity, but also to build discourse objects in a way that the label they give to the discourse objects reveals an evaluation. We propound reflections about such problems within the Discourse Analysis framework, mainly Pêcheux (1975) and Mondada and Dubois (1995).

Besides taking into account reflections from Discourse Analysis, another suggestion of the paper is to use assumptions from Construction Grammar to analyze constructions and idioms rather than only nouns, adjectives, NPs, verbs etc. The investigation of constructions and idioms can provide a better understanding of sentiment in text.

In future works, we intend to expand the dictionary and create a test *corpus* in order to try to create algorithms for automatic evaluation of sentiment of Facebook comments in Brazilian Portuguese.

## References

- W. J. Ashby and P. Bentivoglio. 1993. Information flow in spoken French and Spanish: a comparative study.
- Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. 2010. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC’10)*, 0(January):2200–2204.
- Roberto Gomes Camacho. 1999. Estruturas coordenadas. In Maria Helena Moura Neves, editor, *Gramática do português falado - volume 7*, pages 351–405. Humanitas/FFLCH/USP, S. Paulo.
- Lynn Carlson and Daniel Marcu. 2001. Discourse Tagging Reference Manual. Technical report, University of Southern California, Los Angeles.
- Facebook. *Company Info*. Available at <www.facebook.com>. Accessed in <07/04/2017>.
- Adele E. Goldberg. 2013. Constructionist Approaches. In Elizabeth Closs Traugott and Graeme Trousdale, editors, *The Oxford Handbook of Construction Grammar*, pages 15–31. Oxford University Press, Oxford.
- Michael Alexander Kirkwood Halliday. 1985. *An introduction to functional Grammar*. Edward Arnold, Baltimore.
- Thomas Hoffmann and Graeme Trousdale. 2013. *The Oxford Handbook of Construction Grammar*. Oxford University Press, Oxford.
- Rodolfo Ilari and Renato M. Basso. 2008. O verbo. In Rodolfo Ilari and Maria Helena Moura Neves, editors, *Gramática do português culto falado no Brasil - volume 2*, pages 163–365. Ed. da Unicamp, Campinas.
- Mikel Iruskietia, Arantza Diaz de Ilarraza, and Mikel Lersundi. 2013. Establishing criteria for RST-based discourse segmentation and annotation for texts in Basque. *Corpus Linguistics and Linguistic Theory*:1–32.
- Ray Jackendoff. 2013. Constructions in the parallel architecture. In Thomas Hoffmann and Graeme Trousdale, editors, *The Oxford Handbook of Construction Grammar*, pages 70–92. Oxford University Press, Oxford.
- Ingedore G. V. Koch. 2002. *Desvendando os segredos do texto*. Cortez Editora, S. Paulo.
- Ingedore G. V. Koch. 2007. *O texto e a construção dos sentidos*. Contexto, S. Paulo.



- John Lyons. 1977. *Semantics*. volume 2. Cambridge University Press, Cambridge.
- Saif Mohammad, Cody Dunne, and Bonnie Dorr. 2009. Generating high-coverage semantic orientation lexicons from overtly marked words and a thesaurus. *EMNLP '09 Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, 2(August):599–608.
- Lorenza Mondada and Danièle Dubois. 1995. Construction des objets de discours et catégorisation: una approche des processus de référénciation. *Tranel*, 23:273–302.
- Maria Helena Moura Neves. 2000. *Gramática de Usos do Português*. Ed. da Unesp, São Paulo.
- Maria Helena Moura Neves. 2002. Estudo das construções com verbo-suporte em português. In Ingedore G. V. Koch, editor, *Gramática do português falado - volume 6*, pages 209–238. Editora da Unicamp, Campinas, 2. ed. edition.
- Michel Pécheux. 1975. *Les vérités de la Palice*. Maspero, Paris.
- Renata Lopes Rosa. 2015. *Análise de sentimentos e afetividade extraídos das redes sociais*. Ph.D. thesis, USP.
- Gwendolyn Seidman. 2014. Expressing the “true Self” on Facebook. *Computers in Human Behavior*, 31(1):367–372.
- Maite Taboada. 2016. Sentiment Analysis: An Overview from Linguistics. *Annual Review of Linguistics*, 2(1):325–347.
- Maite Taboada, Julian Brooke, Milan Tofiloski, Kimberly Voll, and Manfred Stede. 2011. Lexicon-Based Methods for Sentiment Analysis. *Computational Linguistics*, 37(2):267–307.
- Elizabeth Closs Traugott and Graeme Trousdale. 2013. *Constructionalization and constructional changes*. Oxford University Press, Oxford.
- R. E. Wilson, S. D. Gosling, and L. T. Graham. 2012. A Review of Facebook Research in the Social Sciences. *Perspectives on Psychological Science*, 7(3):203–220.
- Theresa A. Wilson, Janyce Wiebe, and Paul Hoffmann. 2009. Recognizing Contextual Polarity: an exploration of features for phrase-level sentiment analysis. *Computational Linguistics*, 35(3):399–433.