



LUND UNIVERSITY

Review of: Geoff Thompson and Laura Alba-Juez (eds.) Evaluation in Context. (Pragmatics & Beyond New Series 242) Amsterdam and Philadelphia: John Benjamins 2014

Fuoli, Matteo

Published in:
Functions of Language

DOI:
[10.1075/fol.22.2.06fuol](https://doi.org/10.1075/fol.22.2.06fuol)

2015

Document Version:
Early version, also known as pre-print

[Link to publication](#)

Citation for published version (APA):
Fuoli, M. (2015). Review of: Geoff Thompson and Laura Alba-Juez (eds.) Evaluation in Context. (Pragmatics & Beyond New Series 242) Amsterdam and Philadelphia: John Benjamins 2014. *Functions of Language*, 22(2), 278-287. <https://doi.org/10.1075/fol.22.2.06fuol>

Total number of authors:
1

General rights

Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Geoff Thompson and Laura Alba-Juez (eds.) *Evaluation in Context*. (Pragmatics & Beyond New Series 242) Amsterdam and Philadelphia: John Benjamins 2014, xi + 418pp. (ISBN 978-90-272-5647-8 (hb)/978-90-272-7072-6 (eb))

Reviewed by Matteo Fuoli (Lund University)

This is a pre-print version of an article that has been published in the journal 'Functions of Language, volume 22, issue 2. For direct quotations and page numbers, please check the published version, which can be downloaded from here: <http://www.jbe-platform.com/content/journals/10.1075/fol.22.2.06fuoli>

1. Summary

Evaluation in Context, edited by Geoff Thompson and Laura Alba-Juez, is intended as the sequel to Susan Hunston's and Geoff Thompson's seminal volume *Evaluation in Text* (Hunston and Thompson, 2000). The book consists of nineteen chapters focusing on various aspects of *evaluation*, i.e. the expression of feelings, attitudes and stances in discourse. It combines articles resulting from a research project funded by the Spanish Ministry of Science and Innovation (FunDETT¹), selected papers from the 'International Conference on the Evaluative Function of Language', held in Madrid in 2011, and several invited contributions. While the book features an international team of scholars, the majority of the chapters are authored or co-authored by researchers affiliated to Spanish universities (at the time of writing).

The considerable length of the volume testifies to the continued and growing interest in the topic of evaluation. It also reflects the complexity of evaluative phenomena in discourse. As Alba-Juez and Thompson explain in the introduction (pp. 6-9), evaluation has multiple 'faces'; it permeates language at all levels and can be expressed, both explicitly and implicitly, through a wide range of linguistic resources. But evaluation also has many 'phases'; it is not a mere textual phenomenon, but rather a discursive process that encompasses the cognitive and intentional states that precede evaluative acts, and the addressee's reception of them. Accordingly, one of the main goals of the volume is to broaden our understanding of

evaluation by exploring features, such as prosody, whose evaluative potential has not so far been systematically studied. But, as the title suggests, the book also emphasizes the importance of *context* in the production and interpretation of evaluative meanings. In this sense, another of its main aims is to examine the “aspects of the discursive context that affect the final evaluative meaning both at the production and reception stage” (p. xi). The articles included in the collection contribute to these goals by tackling the complexity of evaluation from different theoretical perspectives, e.g. *appraisal theory* (Martin and White, 2005), *axiological semantics* (Krzyszowski, 1997) and *politeness theory* (Brown and Levinson, 1987), adopting a range of research methods, including qualitative analysis, corpus-based and experimental techniques, and considering a variety of genres and text types. They are broadly subdivided into theoretically-oriented (chapters 1 to 9) and empirically-oriented studies (chapters 10 to 19). In the following, I shall provide a brief overview of the chapters, before offering a general assessment of the volume.

Chapter 1, by Laura Alba-Juez and Geoff Thompson, introduces the volume by tracing the development of the study of evaluation since the publication of Hunston and Thompson (2000) and by offering an updated definition of this notion. Alba-Juez and Thompson emphasize the pervasive, dynamic and dialogic nature of evaluation in discourse. Evaluation can be realized at all levels of linguistic description, encompassing phonology, morphology, syntax and semantics. It may be communicated both explicitly and implicitly, drawing on the participants’ shared assumptions and values, and contextual cues play an important role in the negotiation of evaluative meanings at all times. The expression of evaluative stances in discourse is also constrained by the speaker’s expectations regarding the potential response from their interlocutor. Evaluation is, therefore, inherently *dialogic* and involves “relational work” (p. 13). The result of these considerations is that evaluation is a highly dynamic phenomenon, which requires special attention to context and defies any static, rigid analytic approach.

Chapter 2, by Ángel Felices-Lago, provides a general overview of the main contributions to axiological semantics stemming from the linguistic traditions of structural linguistics and transformational-generative grammar. Felices-Lago seeks to show that, in spite of the widely-held assumption that the Saussurean and Chomskyan schools would have neglected the value-laden components of language, several scholars within these traditions realized their centrality, indirectly anticipating many ideas now at the heart of modern functionalist approaches.

Chapter 3, by Geoff Thompson, and chapter 4, by Mary Macken-Horarik and Anne Isaac, address several theoretical and methodological questions that have emerged from research based on appraisal theory, an increasingly influential model of evaluation developed within the systemic functional linguistic tradition (see, e.g. Martin, 2000; Martin and White, 2005). Both chapters offer suggestions on how to refine the model, tackling well-known and yet unresolved issues. In chapter 3, Thompson considers three problems that might arise when applying the model to the analysis of texts. The first issue concerns whether descriptions of third person's emotions should be categorized in the same way as feelings expressed by the speaker or explicitly attributed to the addressee. The second issue regards the often-fuzzy distinction between the categories of *judgement*, which encompasses normative assessments of human behaviour, and *appreciation*, which subsumes aesthetic evaluations of human artefacts and natural phenomena. Finally, the author examines what he calls the "Russian doll syndrome" (p. 59), which arises when an evaluative expression instantiating one category indirectly realizes other categories in a recursive manner. Throughout the chapter, Thompson offers concrete suggestions for addressing these issues that should, according to the author, help make analyses based on the appraisal model more consistent and replicable.

In chapter 4, Macken-Horarik and Isaac present three challenges that emerge when applying the appraisal model to the analysis of narrative (and indeed, all kinds of) texts. First, they examine some of the complexities involved in accounting for implicit ('invoked' in

appraisal terms) evaluation. Second, they observe that the evaluative meaning of lexical items is often shaped by the ‘global’ stance that is constructed in text through the interplay of evaluative resources, raising additional analysis issues. Third, they discuss the complex relationship between texts and the cultural context in which they operate, and the consequent clash that often arises between the general-purpose architecture of the appraisal model and the “contextual specificity of evaluation” (p. 70). In response to these challenges, Macken-Horarik and Isaac propose a series of theoretical and methodological principles that should facilitate the application of the model.

In chapter 5 the focus turns from written to spoken discourse, and to the relation between evaluation and another important interpersonal discourse function, i.e. verbal irony. Laura Alba-Juez and Salvatore Attardo present the findings of a survey conducted with native speakers of English and Spanish with the aim to determine whether the contrast between positive and negative evaluation is a defining component of verbal irony. The results of the study show that evaluation is indeed a key feature of this trope, but the stance conveyed by ironic utterances may lie on a cline from ‘good’ to ‘bad’, and irony may also be used to amuse the receiver, rather than to express approval or disapproval of some person, entity or situation (‘neutral irony’).

One of the areas where evaluative aspects of language have received a great deal of attention is that of computational linguistics, in particular within the branch of Natural Language Processing (NLP). Chapter 6, by Ángel Felices-Lago and María Enriqueta Cortés-de-los-Ríos, describes how axiological aspects of word meaning are encoded in FunGramKB, a multi-purpose lexico-conceptual resource for NLP systems. The article illustrates how axiological parameters are incorporated in the computational representation of the meaning of words, and examines the distribution of value-laden verbs in FunGramKB.

In chapter 7, Istvan Kecskes investigates the evaluative functions of *situation-bound utterances*, i.e. highly conventionalized expressions whose use is tied to specific

communicative situations (e.g. *help yourself, welcome aboard*), from an intercultural communication perspective. The author argues that, where SBUs are used to express interpersonal attitudes, intercultural differences might cause their evaluative force to be lost or undesired evaluative effects to emerge. Throughout the chapter, Kecskes discusses the role of context and common ground in determining this type of misunderstanding.

The relationship between phonological patterns and evaluation has so far received scant attention in the literature. Chapters 8 and 9 focus on this level of analysis, investigating how certain prosodic features may be used to express evaluative meanings in spoken discourse, with the aid of experimental techniques. In chapter 8, Victoria Escandell-Vidal, Victoria Marrero Aguiar and Pilar Pérez Ocón examine the role of prosodic lengthening (e.g. *Juan odia el bré:col*) as a means of marking information structure and communicating a particular evaluative stance to the addressee. The results of their perceptual experiment show that native speakers of Spanish consistently recognize prosodic lengthening as a means for emphasizing the truth-value of an utterance (*verum focus*) and communicating certainty in response to disagreement. In addition, this device is used to convey an attitude of insistence and impatience on the part of the speaker.

Chapter 9, by Eva Estebas-Vilaplana, investigates evaluation in spoken discourse from the hearer's perspective. It reports on an experiment designed to test the effect of pitch range variability, i.e. whether sentences are uttered with a low or high amount of pitch movement, on the hearer's assessment of the speaker's attitude in English and Spanish. The results indicate that pitch range variability has an impact on hearers' evaluation of the speaker's stance in both languages, but also reveal clear cross-linguistic differences. Utterances produced with a high pitch range tend to be assessed as natural and polite by the English speakers, but as unexpected and overexcited by the Spanish speakers. Conversely, low pitch variability utterances are interpreted as polite in Spanish but as unexpected and rude in English. These findings clearly show that intonation can be exploited as an "off-record strategy" (p. 184) to

express interpersonal attitudes, and provide further evidence for the idea that evaluation pervades all aspects of language, including phonological processes.

The empirically-oriented section of the book begins with chapter 10, by Monika Bednarek, which presents a corpus-based analysis of evaluation in a small collection of *DVD blurbs*, i.e. short promotional texts printed on the back of DVD box sets of television series. The chapter provides a detailed descriptive account of evaluative language in the genre of DVD blurbs and highlights similarities between this discourse type, the closely related genre of book blurbs, and advertising discourse.

In chapter 11, Marta Carretero and Maite Taboada combine the appraisal framework with corpus-based methods to explore cross-linguistic differences in the use of expressions of *graduation*, i.e. the linguistic devices that function to boost or down-tone assessments in discourse (e.g. *extremely*, *somewhat*), in a specialized corpus of English and Spanish book and movie reviews. The authors manually annotate the corpus with the aid of the UAM corpus tool (O'Donnell, 2008) in order to generate quantitative data on the use of this type of expressions in the corpus. Overall, the results show striking cross-linguistic similarities, but several differences are also observed.

One of the domains where evaluation has been more extensively investigated is that of academic discourse (e.g. Biber, 2006; Bondi and Mauranen, 2003; Hunston, 1994). Chapter 12, by Stefania Degaetano-Ortlieb and Elke Teich, contributes to this line of research by presenting a corpus-based analysis of expressions of epistemic and attitudinal stance in scientific research articles across nine disciplines, spanning both traditional and interdisciplinary fields. The results of the analysis provide some evidence for discipline-specific tendencies in the type of evaluative meanings that are more frequently expressed.

Chapter 13, by Laura Hidalgo-Downing, examines patterns of co-occurrence or 'synergies' of negative and modal markers (e.g. *I don't believe*) in Charles Darwin's *The Origin of Species*, with the aim to shed light on the role played by evaluative resources in the

author's presentation of his novel and potentially controversial ideas. The results of her study show that negative-modal synergies are a distinctive feature of Darwin's authorial stance style in *The Origin of Species* and perform an important role in the discursive negotiation of his new ideas. The frequent use of these expressions reflects a tension between certainty and caution in the scientist's writing, and reveals the inherently dialogic nature of scientific progress.

Chapter 14, by Marianna Ryshina-Pankova, contributes further insights into how evaluative resources are used to introduce new knowledge in academic discourse. The chapter focuses on argumentation in a corpus of blogs written by a group of university students as part of their academic writing training. Drawing on the appraisal framework, it investigates the strategies employed by students to construct their arguments and negotiate the validity of their ideas. The findings of the study have implications for the teaching of academic writing. They show the usefulness of blogs as a pedagogical tool and suggest that a linguistic approach to argumentation may help scaffold learners' argumentative thinking and writing skills.

While the study of evaluation in written discourse has produced a large body of scholarship, research on evaluation in multimodal discourse is still in its infancy (e.g. Bednarek and Caple, 2012). Chapter 15, by Ruth Breeze, contributes to filling this gap by exploring how evaluative meanings are co-constructed across verbal and visual modes in online news discourse. The author presents a case study analysis of a corpus of articles dealing with the controversy surrounding the wearing of religious items by Muslims, Sikhs and Christians that made headlines in the British news media in early June 2010. The analysis focuses on how the different religious groups are represented and evaluated through textual and visual resources. The results seem to reveal a conflict between the neutral tone of the articles, and the more overtly evaluative character of the headlines and images that accompany them.

Chapter 16, by Elena Martínez Caro, adds to the discussion of the evaluative functions of

news discourse by presenting an analysis of evaluative language patterns in a specialized corpus of articles from two weekly news magazines, i.e. *Time* and *The Economist*. Drawing on Hunston and Sinclair's (2000) notion of 'local grammar of evaluation', the analysis sets out to identify common lexico-grammatical patterns used to express evaluation in these texts. The result is an extremely detailed description of evaluative phraseology in the corpus.

Similarly to Martínez Caro, in chapter 17 Donna R. Miller and Jane H. Johnson take a phraseological approach to the analysis of evaluation, focusing on the institutional discourse of congressional debates. The study combines quantitative corpus-based methods and qualitative analysis grounded in the appraisal framework. The authors investigate the distribution and functions of the phrase "*it * time to/for/that*" in a specialized corpus of speeches concerning the Iraq war that were given at the U.S. House of Representatives in the wake of the controversy over Saddam Hussein's alleged stockpile of weapons of mass destruction (WMDs). The results show that this phrase was typically used by the Opposition party to confront and criticize the ruling party. Some evidence of gender differences is also found.

One of the domains where the centrality of evaluation in discourse was first recognized is that of narrative research (e.g. Labov, 1972). Evaluation is the key discursive resource through which narrators "show how they intend the narrative to be understood and what the point is" (Cortazzi and Jin, 2000: 103). In Chapter 18, Manuela Romano analyses evaluation in a corpus of unplanned and highly emotional narratives delivered by anonymous speakers during a Spanish late-night radio program. The author describes some of the most productive strategies narrators use to express evaluation, and how evaluative elements are distributed in their accounts.

The last chapter of the book, by Carmen Santamaría-García, explores evaluation in social networking websites, where the expression of attitudes and preferences is one of the users' main communicative concerns. The study combines politeness theory and appraisal theory to

investigate how evaluation is used in this context to manage interpersonal relationships and build communities of shared values. The data consists of a tiny corpus of Facebook posts (1414 words), which is annotated using the UAM corpus tool (O'Donnell, 2008) based on Brown and Levinson's (1987) typology of positive and negative politeness strategies and Martin and White's (2005) appraisal framework. The results of the analysis indicate that expressions of *affect*, i.e. relating to emotions, represent the most frequent type of appraisal used in the corpus. These expressions are often integrated into positive politeness strategies for the purpose of building common ground, conveying cooperation and fulfilling addressee's wants.

2. Assessment

Evaluation in Context provides a comprehensive overview of emerging research on evaluation and contributes novel perspectives on this phenomenon, shedding light on aspects that had not been systematically analysed before. Despite its usefulness, however, the book has several limitations. First, the range of phenomena considered under the umbrella term 'evaluation' is so broad that it is often difficult to know where the boundaries of this concept lie. Some of the contributions conflate evaluation and closely related, but not identical notions, such as politeness and axiology. In chapter 7, for example, politeness appears to be treated as a sub-type of evaluation, as the author writes that "SBUs rarely have an evaluative function except from politeness and impoliteness" (p. 149). The distinction between evaluation and politeness is blurred in chapter 9 too, where informants are asked to rate utterances as either "natural, polite or nice" or "rude, impolite or over-excited" (p. 188), and one of the conclusions is that "pitch range differences should be taken into account to explain cross-cultural differences in the perception of politeness" (p. 191). The terms *axiology* and *axiological semantics*, which are used in chapters 2 and 6, are not clearly defined and the difference between axiology and evaluation is not discussed anywhere in the book. Many readers who, like the present reviewer, are more familiar with functional approaches to

evaluation could have benefitted from a clear illustration of this concept and area of research.

Several chapters adopt a rather vague notion of evaluation. In chapter 15, for example, the author annotates and quantifies ‘evaluative elements’ in her corpus of news texts, but she does not discuss the criteria used to identify them, nor does she give any concrete linguistic example from the corpus or describe the coding procedure in any detail. The evaluative function of various examples given in chapter 16, e.g. “drought and famine persist” (p. 328), is not obvious, and, indeed, the nineteen patterns surveyed by the author seem to perform diverse functions in discourse. Thus, in general the chapters do not appear to conform to a single, coherent definition of evaluation. Hunston (2011: chapter 2) provides an excellent synthesis of current approaches to the analysis of evaluation in discourse and highlights both areas of overlap and differences between them. The “points of consensus” (ibid.: 12ff) that she identifies could have been a useful guideline for delimiting the notion of evaluation and achieving a more coherent and principled account of this phenomenon. Quite surprisingly, however, Hunston’s book is not even mentioned in the introduction, where the editors present their updated definition.

One of the main strengths of the volume is that it includes numerous empirical studies of evaluation in different discursive contexts. However, some of these studies are rather weak in terms of methodological rigor. Some of them are based on manual corpus annotation. Manual annotation can be a useful method for analysing evaluation, as it facilitates comprehensive and detailed analyses that would not be possible with purely automatic techniques (Fuoli and Hommerberg, *forthcoming*). But manual annotation is also a complex and subjective task, and poses numerous challenges to achieving reliable and replicable analyses. A common measure for assessing the reliability of annotations is *inter-coder agreement* (e.g. Artstein and Poesio, 2008; Fuoli, 2012; Fuoli and Hommerberg, *forthcoming*; Read and Carroll, 2010; Taboada and Carretero, 2012). If independent annotators can be shown to consistently make equivalent coding choices, we can assume that the annotation guidelines are sufficiently well defined

and, accordingly, that the results are reliable and potentially replicable (Artstein and Poesio, 2008). Out of four studies that use some form of manual annotation, only one, i.e. chapter 14, reports inter-coder agreement results. In addition, several studies include quantitative data, but do not perform any statistical test to validate their findings.

Indeed, a number of crucial and yet unresolved methodological questions concerning the analysis of evaluation are, unfortunately, not directly addressed in the volume. For instance, how can we reliably identify evaluative expressions in discourse? How should we deal with the problem of subjectivity in the analysis process, in particular when quantitative data are produced from manual analysis? How should we account for our decisions so that our analyses are explicit, transparent and maximally replicable? Chapters 3, 4 and 11 do offer some valuable suggestions for making analyses more systematic, but fail to consider these questions. For this reason, I believe that *Evaluation in Context* has missed an opportunity to promote a much-needed discussion of some of the most vexing methodological problems that arise when analysing evaluation in discourse.

The volume covers a wide range of methods and approaches to the analysis of evaluation, adding fresh insights into this phenomenon and promoting dialogue between different theoretical and empirical perspectives. The inclusion of a chapter (chapter 6) dealing with evaluation from a computational linguistics point of view is particularly welcome, given the recent dramatic surge of interest in evaluative aspects of language that this field has experienced. However, one of the most vibrant strands of research in this context, namely *sentiment analysis* (for an overview, see Pang and Lee 2008 and Hunston 2011: 53-55) is merely mentioned in passing in chapter 12. This is a particularly unfortunate omission, as there are many areas of overlap between work in this area and more linguistic-oriented approaches to evaluation. For example, the appraisal model has gained increasing recognition as a useful theoretical basis for improving sentiment analysis systems (e.g. Argamon et al., 2009; Read and Carroll, 2010; Taboada and Grieve, 2004; Whitelaw et al., 2005).

Finally, overall the book reads well, but the contributors' writing style is quite uneven. Several empirically-oriented chapters, e.g. 10 and 16, contain an overwhelming wealth of detail and an excessive number of examples, to the point that it becomes difficult to understand what the main findings and the central theses are. Chapters 11 and 19 include very large data tables that could have been replaced by graphs for improved readability and clarity. Overall the book is well produced, even though there are several typographical errors and inaccuracies.

Despite these limitations, the book makes several valuable theoretical and methodological contributions. It clearly shows that evaluation permeates language at all levels and demonstrates this empirically by investigating aspects that have so far received little attention, such as the evaluative potential of certain phonological patterns. The book does not only greatly expand our understanding of evaluation at different levels of linguistic description, it also brings about an important conceptual shift by viewing evaluation as a dynamic interpersonal and dialogic process, rather than as a mere textual phenomenon. In this sense, the inclusion of the addressees' perspective and the investigation of their interpretation of evaluative acts carried out in chapters 5, 8 and 9 represent milestone advancements and open up many exciting new avenues for research. More generally, the wide variety of theoretical frameworks and methods represented in the book make it a valuable resource and a rich source of inspiration for scholars interested in exploring the multiple faces and phases of evaluation in discourse.

References

Argamon, Shlomo, Kenneth Bloom, Andrea Esuli & Fabrizio Sebastiani. 2009. Automatically determining attitude type and force for sentiment analysis. In Zygmunt Vetulani & Hans Uszkoreit (eds.), *Human Language Technology. Challenges of the Information Society*, 218-231. Springer.

- Artstein, Ron & Massimo Poesio. 2008. Inter-coder agreement for computational linguistics. *Computational Linguistics*, 34(4):555-596.
- Bednarek, Monika & Helen Caple. 2012. *News Discourse*. London & New York: Continuum International Publishing Group Ltd.
- Biber, Douglas. 2006. Stance in spoken and written university registers. *Journal of English for Academic Purposes*, 5(2):97-116.
- Bondi, Marina & Anna Mauranen. 2003. Evaluation in academic discourse. *Special issue of Journal of English for Academic Purposes*, 2(4).
- Brown, Penelope & Stephen Levinson. 1987. *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.
- Cortazzi, Martin & Lixian Jin. 2000. Evaluating evaluation in narrative. In Susan Hunston & Geoff Thompson (eds.), *Evaluation in text: Authorial stance and the construction of discourse*, 102-120. Oxford: Oxford University Press.
- Fuoli, Matteo. 2012. Assessing social responsibility: A quantitative analysis of appraisal in BP's and IKEA's social reports. *Discourse & Communication*, 6(1):55-81.
- Fuoli, Matteo & Charlotte Hommerberg. Forthcoming. Optimizing transparency, reliability and replicability: Annotation principles and inter-coder agreement in the quantification of evaluative expressions.
- Hunston, Susan. 1994. Evaluation and organization in a sample of written academic discourse. In Malcolm Coulthard (ed.), *Advances in written text analysis*, 191-218. London: Routledge.
- Hunston, Susan. 2011. *Corpus Approaches to Evaluation: Phraseology and Evaluative Language*. New York and Oxon: Taylor & Francis.
- Hunston, Susan & John Sinclair. 2000. A local grammar of evaluation. In Susan Hunston &

- Geoff Thompson (eds.), *Evaluation in text: Authorial stance and the construction of discourse*, 74-101. Oxford: Oxford University Press.
- Hunston, Susan & Geoff Thompson. 2000. *Evaluation in text: Authorial stance and the construction of discourse*. Oxford: Oxford University Press.
- Krzeszowski, Tomasz P. 1997. *Angels and Devils in Hell: Elements of Axiology in Semantics*. Warsaw: Energeia.
- Labov, William. 1972. The transformation of experience in narrative syntax. In William Labov (ed.), *Language in the Inner City: Studies in the Black English Vernacular*, 354-396. Philadelphia: University of Pennsylvania Press.
- Martin, James R. 2000. Beyond exchange: appraisal systems in English. In Susan Hunston & Geoff Thompson (eds.), *Evaluation in text: Authorial stance and the construction of discourse*, 142-175. Oxford: Oxford University Press.
- Martin, James R. & Peter R. R. White. 2005. *The language of evaluation: Appraisal in English*. London & New York: Palgrave Macmillan.
- O'Donnell, Mick. 2008. Demonstration of the UAM CorpusTool for text and image annotation. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics on Human Language Technologies: Demo Session*, 13-16. Association for Computational Linguistics.
- Pang, Bo & Lillian Lee. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2):1-135
- Read, Jonathon & John Carroll. 2010. Annotating expressions of appraisal in English. *Language Resources and Evaluation*, 46(3):421-447.
- Taboada, Maite & Marta Carretero. 2012. Contrastive analyses of evaluation in text: Key issues in the design of an annotation system for attitude applicable to consumer

reviews in English and Spanish. *Linguistics and the Human Sciences*, 6(1-3):275-295.

Taboada, Maite & Jack Grieve. 2004. Analyzing appraisal automatically. In *Proceedings of AAAI Spring Symposium on Exploring Attitude and Affect in Text*, Stanford University, CA, pp. 158-161. AAAI Press.

Whitelaw, Casey, Navendu Garg & Shlomo Argamon. 2005. Using appraisal groups for sentiment analysis. In *Proceedings of the 14th ACM international conference on Information and knowledge management*, 625-631. ACM.