Corpora and beyond - investigating evaluation in discourse: introduction to the special issue on corpus approaches to evaluation

Gozdz-Roszkowski, Stanislaw; Hunston, Susan

DOI:
10.3366/cor.2016.0089

Citation for published version (Harvard):
Gozdz-Roszkowski, S & Hunston, S 2016, 'Corpora and beyond - investigating evaluation in discourse: introduction to the special issue on corpus approaches to evaluation', *Corpora*, vol. 11, no. 2, pp. 131-141. https://doi.org/10.3366/cor.2016.0089

Corpora and beyond: investigating evaluation in discourse

Introduction to the special issue on corpus approaches to evaluation

Stanisław Goźdź-Roszkowski, Institute of English Studies, University of Łódź

Email: roszkowski@uni.lodz.pl

Susan Hunston, Department of English Language and Applied Linguistics, University of Birmingham

Email: s.e.hunston@bham.ac.uk

Section 1: Evaluative Language

This special issue brings together five papers that take novel approaches to the intersection of discourse analysis, corpus investigation techniques, and evaluative language. In this introductory paper we consider the theoretical bases for these papers and the contribution they make to the study of corpora and of evaluation.

The papers in this issue refer variously to 'evaluation', 'stance(taking)', 'sentiment' and 'appraisal'. These terms originally developed relatively independently, suggesting that several quite disparate research traditions recognised the need to take account of similar ranges of meanings before beginning to explore how their approaches might benefit each other. Thus, while the terms are not synonymous, they share sufficient common ground for us to consider them together before outlining their differences. As is well known, 'evaluation' (as a cover term) has the following distinguishing characteristics:

- It represents kinds of meaning that might be termed 'subjective' or 'attitudinal', and that may be distinguished from the 'objective' or 'factual'.
- It is multi-functional, performing roles in construing relationships between participants in an interaction and in structuring discourse as well as in expressing opinion. The primary function of an evaluative phrase in a text may be, for example, to indicate a change in topic rather than to express an opinion.
- The ways in which evaluative meaning is expressed are notoriously difficult to pin down. Evaluation may be 'inscribed' or explicit, using recognisably evaluative lexis and/or constructions associated with evaluation. This opens the possibility of creating an exhaustive list of evaluative lexis and grammar. Whereas substantial progress has been made in constructing such lists, they are very lengthy and remain open, subject to change as a language changes and to constant expansion as a wider variety of registers is explored.
- As well as being 'inscribed', evaluative meaning might also be 'evoked' or implied. In these cases its presence and classification are open to debate, and researchers may offer a 'reading' of a text rather than an 'analysis'. Its effect is often cumulative and dependent on context. For these reasons, instances of evoked or implicit evaluative meaning are resistant to standard methods of searching corpora for instances of given categories.
- Although it is often argued that the basis of evaluation is the good-bad polarity, a number of parameters of evaluation have been proposed. Crucially, these distinguish both the target of evaluation and the basis of the judgement made. For example,

propositions or statements tend to be evaluated in terms of the degree of certainty or belief to be attached to them, but may also be assessed in terms of how important or how obvious they are. Human behaviour is generally judged along a 'good-bad' scale, but the scale applied may be construed as relatively objective, assessing the success or failure of an endeavour in achieving goals, or as relatively subjective, in terms of its social or ethical acceptability. Philosophers identify 'predicates of personal taste', which express purely personal judgements, such as preferring one food to another, and distinguish these from shared moral judgement, for example (Lasersohn, 2005).

- Concepts of multiglossia and intertextuality are important to evaluative language because evaluation can be attributed as well as averred. This opens the possibility of a complex relationship between 'current speaker' and 'reported speaker' and a consequent manipulation of point of view. Models of evaluative language commonly propose specific categories or taxonomies to account for the value accorded others' discourse.
- Evaluative meaning and its expression in discourse is a key way of both referencing and construing ideologies (both hegemonic and resistant). Implicit evaluation in particular presupposes a shared value-system. For this reason, the study of evaluation is crucial to Critical Discourse Analysis and similar discourse studies.

In short, the study of evaluative language is both problematic (items are difficult to identify) and important (evaluation contributes to the interactive property of language, to the recognition of how a text is organised, and to the connection between discourse and ideology). It is a topic that brings together linguistic concepts, ideas about discourse, and views of society.

We now turn to the various terms we have used at the beginning of this section. As noted above, **evaluative language** or **evaluation** is sometimes used as a cover term for all the concepts mentioned (see, for example, Hunston and Thompson 2000). As a topic of linguistic research it has its roots in the study of text structure. Labov (1972), in his study of spoken narrative, suggested 'Evaluation' as a tripartite concept: a move delaying the Resolution of the narrative; a series of non-narrative clauses occurring throughout the narrative; and as the indication of the 'point' of the narrative. In later discourse studies, evaluation was proposed as the name of a metalinguistic discourse move that served to indicate the end of a section or phase of either monologic (typically written) or dialogic (typically spoken) discourse (see, for example Hoey 1983; Sinclair 1981). Because evaluation is clearly much more pervasive than this, Sinclair distinguished between evaluation on the autonomous plane that indicated attitude towards entities and evaluation on the interactive plan that served the additional function of structuring the discourse.

The concept of **stancetaking** comes from Conversation Analysis. As du Bois (2007) notes, the use of a term that construes an action (taking stance) rather than an entity emphasises that this is something performed by a discourse participant to achieve various functions in the discourse. The speaker who takes a stance thereby construes a relationship between themselves and an external entity and simultaneously between themselves and their interactant(s). Stance may be taken through a substantive statement ('I like that picture'), or by a more reduced act of agreement ('Yes') or even hesitation ('Well…'). For the study of stancetaking, the act of (dis)alignment between speakers is arguably of greater importance than the kind of stance taken. Alba-Juez and Thompson (2014) argue that stance is a broader term than evaluation because stance can be neutral as opposed to evaluative.

Corpus linguists such as Biber and colleagues also use **stance** as a label for specific words and phrases that inscribe an attitude or stance towards an entity in the text. In work such as Biber et al (1999) the category of stance is subdivided both grammatically (adverbs, verbs, modals, adjectives and nouns) and semantically (epistemic, attitudinal and stylistic). Because each category consists of a defined set of words and phrases, corpus searches for each set can be used to calculate the relative frequency of each grammatical and semantic sub-category, allowing comparisons between registers.

Computational approaches to evaluative language include **sentiment analysis** (see, for example, Turney 2002; Liu 2012; Wiebe, Wilson & Cardie 2005), where the balance and strength of positive and negative attitudes towards a single entity (in, for example, film reviews or political comments) are calculated. Key requirements for successful sentiment analysis include the identification of complete sets of positive and negative lexis and the recognition of contextual elements that will affect either the strength or the polarity of that lexis. For example, 'Only an idiot would find the plot convincing' expresses a negative rather than positive attitude towards a film plot, in spite of the presence of the positive phrase 'find…convincing'. Compiling lists of positive and negative lexis is regarded as essentially a computational task, based on algorithmic learning from small sets of attitudinal lexis compiled by psychologists, for example. Identifying significant contexts requires a more linguistic approach.

Probably the most fully theorised view of evaluative language is taken by Martin and White (2005), who locate this area of meaning within the interpersonal metafunction of Systemic-Functional Linguistics. **Appraisal** is the name given to the whole system, incorporating three sub-systems: attitude, engagement and graduation. The attitude system proposes distinctions between Affect (the expression of feeling), Judgement (evaluation of a person's actions), and Appreciation (evaluation of aesthetic quality). The engagement system encompasses distinctions between the ways that other voices are incorporated in a text, in particular distinguishing between uses of those voices to open up ('expand') the dialogue in a text and those that close it down ('contract'). Appraisal theory draws on the emphasis in SFL on both meaning and choice. That is, although examples of the various categories are given, Martin and White point out that they are distinguishing between meanings rather than forms. For this reason, implicit as well as explicit instantiations of the categories are included in the analysis of any text, and the connection between a particular word or phrase and the category it exemplifies is not determinate. The concept of 'choice' is equally important. The systems express the possibilities that are available when evaluation is to be expressed. For example, a concert might be assessed in personal terms ('We enjoyed the concert') or as a judgement of the orchestra ('The performance was technically accurate but artistically uninspired') or of the music itself ('The music was sublime'). As Martin (2000) demonstrates, the predominance of one choice over the others has ideological consequences.

To summarise: the study of evaluative language involves the identification of linguistic or formal elements (words, phrases, frames or constructions), a sensitivity towards the discoursal function of individual elements in context, and an awareness of the social significance of individual and cumulative expressions of evaluation. The papers in this special issue address the intersection of these three aspects.

Section 2: Corpora and evaluative language

The study of evaluative language pre-dates the widespread use of language corpora, but the development of corpora has expanded the range of studies of evaluation. As noted above, the ease with which categories of evaluation (stance) can be quantified allows for ready comparison between registers. There is extensive work, for example, on comparisons between academic discourses in terms of the explicit expression of stance (see, for example, Biber, 2006; Hyland, 2005; Jiang and Hyland, 2015). The focus on frames, constructions and patterns that corpus investigations have made possible has enabled sequences associated with evaluation to be identified, along with both typical and creative uses (Hunston, 2011). At the same time, the limitations of corpus searches in identifying and interpreting instances of evaluative language are widely acknowledged. For the most part, corpora can be searched for specific forms only, meaning that only inscribed evaluation can be found. Researchers working with meaning-based rather than form-based categories, such as those used in the appraisal framework, sometimes proceed by annotating a corpus prior to conducting quantitative studies (see, for example, Taboada and Carretero, 2013; Bal, 2014), but such work is laborious, in spite of the assistive software available, and does not really exploit the affordances of corpora. The tendency of corpus work to extract instances of use from their broader context, while highlighting formal patterns, can also have the effect of masking discoursal significance, such as the location of instances with respect to text structure. The challenge for the researcher who wishes to combine the potential offered by corpus investigation techniques with the insights offered by theories of evaluation is to achieve a synergy that is enlightening to both rather than reductive.

The papers in this special issue rise to this challenge. They consider a number of issues: the influence of discourse studies (construed as the study of register, genre and rhetorical structure) on corpus studies of individual stance-indicating items; the contribution of non-corpus-based methods to a corpus study; the contribution of concepts associated with evaluation to the interpretation of other linguistic and social phenomena; the challenge offered by corpus studies to theories of evaluation.

Two themes can be observed: the influence upon corpus studies of concepts taken from discourse studies, including register, rhetorical structure, and newsworthiness; and the influence upon discourse-based concepts, including theories of evaluation, of corpus studies.

Magdalena Szczyrbak in her paper (***Say* and stancetaking in courtroom talk: A corpus-assisted study**") focuses on spoken legal discourse in the UK, and a particular set of features: phrases with the verb SAY in it. Investigating the genre of libel proceedings in a UK court, phrases with SAY are selected as indicative of the alignment function of stance-taking. The paper identifies the frequency of each of these phrases in court proceedings and then it locates each phrase in an extended co-text, identifying a series of distinct pragmatic functions routinely performed by the phrases. The paper describes its approach as 'corpus-assisted', and nicely illustrates the interaction of a corpus approach that prioritises rapid searching for specific forms and quantitative comparisons between corpora, and a discourse approach that examines and interprets the discourse surrounding the target item to identify its function. Those functions cannot be 'read off' from frequency lists or even from limited concordance lines. On the other hand, identification of which phrases are likely to be maximally significant requires the corpus input.

Another aspect of discourse studies is used in a contribution written by Radoslava Trnavac, Debopam Das and Maite Taboada ["**Discourse Relations and Evaluation"**]. In this case the key observation is that the apparent polarity of 'opinion words and phrases' (that is, whether they construe positive or negative judgements) may be altered by the relationship between the clause they occur in and surrounding clauses. For example, if a clause is placed in a concessive relation with another clause, a positive opinion word occurring in the concession clause may be down-toned by the other clause ('I know he's a good actor, but in this film…'). The paper uses Rhetorical Structure Theory (Mann & Thompson, 1988) as a source of categories of clause relations, and the corpus of film reviews is annotated in detail to allocate each clause pair to a category. The number of opinion (or sentiment) words in each clause type is then calculated, along with the proportion of sentiment instances that are altered (downtoned, intensified or reversed) by the clauses around them. Once again the value of incorporating into a corpus study insights from alternative perspectives is demonstrated. It is worth noting, however, that the conclusion of the paper is not that all corpora must be annotated in this way for opinion to be reliably quantified in them, but that sentiment polarity is rarely actually reversed. The discourse-informed study usefully corroborates rather than overturns the less context-sensitive approach.

It is often the case that discourse studies challenge corpus findings by showing that generalities about given words or phrases can be disproved by a more detailed study of individual examples. A contribution by Nele Põldvere, Matteo Fuoli and Carita Paradis ["**A study of dialogic expansion and contraction in spoken discourse using corpus and experimental techniques"**] adds a twist to this by showing that investigation of a large amount of data can challenge assertions made based on smaller discourse studies. The assertion in this case is the alignment proposed in appraisal theory (specifically, the engagement system) between certain phrases of the type *I believe* or *I know* and the binary alternatives of dialogic expansion and contraction. The London-Lund Corpus is used to identify large numbers of each of the target phrases, which are then classified as expanding or contracting in function. An interesting feature of this paper is the use of informant intuitions as supplements to interpretation based on linguistic indicators. The authors are then able to specify how a phrase such as *I think* may be modified to give it the function of expanding or contracting the dialogue. Although the paper challenges aspects of the engagement system, ultimately it enhances that system by maintaining the binary distinction proposed and by offering a more detailed specification of what is likely to be perceived by interactants as inviting or disallowing alternative opinions. The use of a large amount of data, along with informant testing, makes such specification possible.

Monika Bednarek's study ["**Investigating evaluation and news values in news items that are shared via social media"**] is a further example of the positive interaction between discourse studies, evaluation and corpora. Like the first contribution to this issue, this paper focuses on a single register, in this case news stories that have been shared through social media. Identifying evaluative language in this study serves the purpose of indicating which news values are being construed, in turn thereby suggesting what makes news items 'shareable'. The paper draws on extensive previous research that links expressions of evaluation to news values, but stresses that the connection can be indicative only. As in all the papers mentioned, there is no one-to-one alignment between form and function, in this case between evaluation and news value. In fact, one of the points made by this paper is that the identification of news values itself is a subjective activity, with values in some texts being 'probably present' rather than definitely so. In this case, evaluative language is a useful pointer towards another kind of entity. Another feature of this paper is that it draws on the UCREL Semantic Analysis System (Archer, Wilson, and Rayson 2002) as an aid to finding news-relevant evaluative features.

In the final paper in this issue, Josef Ruppenhofer and Laura Michaelis ["**Frames, polarity and causation"**], adopt the perspective of Frame Semantics (Fillmore 1985). This proposes that all language is interpreted in terms of schemata or semantic frames (such as the ‘commerce’ frame) that entail the presence of frame elements (such as ‘buyer’, ‘seller’, ‘goods’, ‘money’). Instances of sentences that realise the ‘commerce’ frame can be coded with each of these elements. Within that context, the authors consider the problem of ‘polarity-sensitive items’, that is, items that appear to occur in either positive or negative clauses only. One example given is ‘lift a finger’, which occurs in clauses which are grammatically or semantically negative (‘He never lifted a finger’; ‘They refused to lift a finger to help’) but not in positive equivalents (‘He agreed to lift a finger’). Many PSIs serve to maximise or minimise (‘boost’ or ‘hedge’) large or small scalar values and are associated with what Martin and White might discuss under the heading of Graduation. For example, ‘They refused to lift a finger to help’ is a stronger evaluative statement than ‘They refused to help’. Whereas the function of most PSIs with either positive or negative is predictable from their meaning and polarity (for example, ‘lift a finger’ means a small amount of effort, occurs in negative clauses and has the function of emphasising the smallness of the effort), there are some that break this trend. The task in this study is to explain the behaviour of the ‘rule-breaking’ PSIs. This is done by noting the frame elements realised by such items and in so doing relating the phrases in question to point of view. For example, it is noted that ‘for a song’ and ‘a pittance’ both refer to small amounts of money, but that situations where goods are bought ‘for a song’ take the point of view of the Buyer, who regards paying a small price as a good thing, whereas situations where labour is exchanged ‘for a pittance’ take the point of view of the Seller, who regards receiving a small amount of money as a bad thing. Thus, point of view, evaluative polarity, and frames combine in an explanatory framework to account for a set of linguistically puzzling phrasal items.

Section 3: Evaluation and corpora: the way forward

The five papers in this issue illustrate a few of the range of methods and approaches that come under the heading of evaluative language and corpora. It is worth pausing for a moment to consider the ways in which the papers differ from one another.

First, the papers differ in the selection of items under investigation and the reasons for that selection. In three of the papers, a small set of items is pre-selected for study. In “*Say* and stancetaking in courtroom talk: A corpus-assisted study”, for example, it is a single verb SAY, selected because observation of interactions has suggested that phrases including this verb play a significant role in the management of interaction. In “A study of dialogic expansion and contraction in spoken discourse using corpus and experimental techniques” it is a small set of phrases, selected because the theory of engagement suggests they reliably map on to the categories in that taxonomy. In “Frames, polarity and causation” it is a number of phrases selected because they exemplify a category of item. In each case, the set of items is suggested by research undertaken with corpus linguistics, and corpus studies allow a more detailed investigation of the items. In the other two papers there are no pre-selected items. Both “Discourse Relations and Evaluation” and “Investigating evaluation and news values in news items that are shared via social media” start with the analysis of texts, using RST in one case and the concept of news values in the other, and discover the evaluative words or phrases as the study unfolds.

Then, the papers differ in the type of corpus they use and the reason for choosing or compiling that corpus. The papers vary from corpora carefully designed to represent specific genres or registers (**Magdalena Szczyrbak** and **Monika Bednarek**) to corpora that are large enough and general enough to have sufficient examples of the target items but where the balance and representativeness of the corpus is less important (**Josef Ruppenhofer** and **Laura Michaelis**).

The papers differ also in the methods of corpus investigation used and the extent of the integration between corpus methods and other approaches. Ruppenhofer and Michaelis use basic search techniques to identify the target items, then interpreting the output with relation to Frame Semantics. At the opposite end of the scale, Bednarek first annotates the corpus using a system designed to annotate corpora automatically, while Trnavac, Das and Taboada first annotate the corpus with a model designed for individual texts. Szczyrbak as well as Põldvere, Fuoli and Paradis start with phrases identified through basic searches but then expand the co-text so that extended stretches of discourse can be examined. In fact, the method employed by Szczyrbak is more akin to conversation analysis than to traditional corpus linguistics.

Fourthly, the papers are not identical in the approach to evaluation they draw on. Szczyrbak takes a stance approach while Trnavac, Das and Taboada contribute to sentiment analysis while using some concepts from appraisal theory. Põldvere, Fuoli and Paradis both use and challenge engagement. Bednarek uses a parameter-based model that is based on theories of news values rather than on evaluation per se, while for Ruppenhofer and Michaelis evaluation might be called a by-product of the selection of PSIs for study.

Finally, it is interesting to note that in each of the papers the concept of evaluation itself makes a different contribution to each of the papers in this issue. As noted above, Ruppenhofer and Michaelis give important information about a set of items that have an evaluative resonance, but evaluation does not lie at the heart of the paper. In Bednarek's study, evaluative meaning is used as a means to identify news values, but it is the identification and quantification of those values themselves that is important. Three concepts of evaluation – engagement, sentiment, and stance – are central to the purpose of the other three papers. Szczyrbak demonstrates how phrases associated with stance contribute to the achievement of social and interactional organisation. Both Trnavac, Das and Taboada and Põldvere, Fuoli and Paradis show how evaluative meaning is determined not by single words or phrases but by other aspects of the linguistic environment, thereby challenging existing orthodoxies.

In short, taking each of these aspects of the papers, no two of them are identical in approach. At the very least, then, a considerable amount of diversity in this field of study is demonstrated. But more may be said. These papers are not alone in mixing methodologies - it might be said that a trademark of current corpus linguistics is that it supports and is supported by other approaches to language. Increasingly, corpus investigation tools are being used by researchers who would not describe themselves as corpus linguists, or even linguists. It is interesting, however, that the focus on the evaluative area of meaning seems to invite extensive variation in methodology and indeed innovation in methodology. There seems to be a recognition that: the role of evaluation in texts and registers needs to be a focus of attention in corpus studies; that attention may be focused on the 'supporting roles' of evaluation as much as on the evaluation itself; the study of such an amorphous concept requires some ingenuity in combining methods of investigation. It is tempting, perhaps, to imagine that, in the study of evaluative language, in the end 'Truth Will Out' (Gilbert & Mulkay 1984: 90ff)), that is, that if enough studies are done a single model and a single methodology will emerge as 'the best'. The papers in this issue suggest that such as eventuality is not, in fact, to be desired. Apart from further methodological innovation, future research into evaluation is likely to explore new territories by focusing on specialized (and/or spoken) genres or registers and by testing the validity and usefulness of existing theoretical frameworks in such new contexts. This trend is already present in this issue. Szczyrbak investigates stance-taking in a spoken legal genre and Põldvere, Fuoli and Paradis apply the category of engagement to spoken interaction. Both specialised genres, especially in legal contexts, and spoken registers, remain under-researched as regards the phenomenon of evaluation.

References

Alba-Juez, L. and Thompson, G. 2014. 'The many faces and phases of evaluation' in G. Thompson and L. Alba Juez (eds.) *Evaluation in Context*, pp. 3-23. Amsterdam, Philadelphia: John Benjamins.

Archer, D., A. Wilson, and P. Rayson. 2002. 'Introduction to the USAS Category System'. Retrieved from http://ucrel.lancs.ac.uk/usas/usas guide.pdf

Bal, K. B. 2014. 'Analyzing Opinions and Argumentation in News Editorials and Op-Eds', *International Journal of Advanced Computer Science and Applications, Special Issue on Natural Language Processing,* pp. 22-29.

Biber, D. 2006. *University Language. A Corpus-based Study of Spoken and Written Registers.* Amsterdam, Philadelphia: John Benjamins.

Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. 1999. *Longman Grammar of Spoken and Written English*. Amsterdam, Philadelphia: John Benjamins.

du Bois, J.W. 2007. 'The stance triangle' in R. Englebretson (ed) *Stancetaking in Discourse: Subjectivity, Evaluation, Interaction,* pp. 139–182. Amsterdam, Philadelphia: John Benjamins.

Fillmore, C. J. 1985. 'Frames and the semantics of understanding', *Quaderni di Semantica* 6, pp. 222-254.

Gilbert G.N. and Mulkay M. 1984. *Opening Pandora's box: a sociological analysis of scientists' discourse.* Cambridge: Cambridge University Press.

Hoey, M. 1983. *On the surface of discourse.* London: Allen and Unwin.

Hunston, S. 2011. *Corpus Approaches to Evaluation: Phraseology and Evaluative Language*. New York/Oxon: Routledge.

Hunston and Thompson (eds.) 2000. *Evaluation in Text: authorial stance and the construction of discourse.* Oxford: Oxford University Press.

Hyland, K. 2005. 'Stance and Engagement: A Model of Interaction in Academic Discourse', *Discourse Studies* 7/2, pp. 173 – 192.

Jiang, F. and Hyland, K. 2015. 'The fact that': Stance nouns in disciplinary writing' Discourse Studies 17, pp. 529 - 550.

Labov, W. 1972. *Language in the Inner City*. Philadelphia: University of Pennsylvania.

Lasersohn, P. 2005. 'Context dependence, disagreement, and predicates of personal taste. *Linguistics and Philosophy* 28, pp. 643-686.

Liu, B. 2012. *Sentiment analysis and opinion mining*. San Rafael, CA: Morgan & Claypool Publishers.

Mann, W. C., & Thompson, S. A. 1988. 'Rhetorical Structure Theory: Toward a functional theory of text organization', *Text*, 8(3), pp. 243-281.

Martin, J.R. 2000. 'Beyond Exchange: Appraisal Systems in English' in Hunston and Thompson (eds) *Evaluation in Text*, pp. 142 – 175. Oxford: Oxford University Press.

Martin, J.R. and White, P. 2005. *The Language of Evaluation. Appraisal in English.* Basingstoke: Palgrave Macmillan.

Sinclair, J. 1981. 'Planes of discourse' in S.N.A. Rizvi (ed) *The Two-Fold Voice: Essays in Honour of Ramesh Mohan*. Salzburg: University of Salzburg, pp. 70-89.

Turney P. 2002. 'Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. *Proceedings of 40th Meeting of the Association for Computational Linguistics*, pp. 417 – 424. Philadelphia, PA.

Taboada, M. and Carretero, M. 2013 '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' in Taboada, M., Doval Suárez and Álvarez (eds) *Contrastive Discourse Analysis*, pp. 262 – 281. Sheffield: Equinox.

Wiebe, J., Wilson, T and Cardie C. 2005. 'Annotating expressions of opinions and emotions in language', *Language Resources and Evaluation* 39 (2), pp. 165-210.