Opinion, subjectivity and point of view in language
Research project funded by the Social Sciences and Humanities Research Council of Canada (SSHRC)
Funding period: 2014-2019

PI: Maite Taboada, Simon Fraser University (mtaboada@sfu.ca)

Summary of proposed research

My research will determine how opinion, subjectivity and point of view are expressed in language, and uncover the characteristics of opinionated language. Language is inherently subjective, in that it always expresses opinion and evaluation, but also facts, all from the speaker's or writer's point of view. Although the expression of opinion is one of the most basic functions of language, we do not know enough about how it is regularly conveyed, and how it is interpreted in context. Studies have been undertaken on the difference between objective and subjective language, and on point of view and stance, but no unifying theory on how evaluation and opinion are expressed exists.

   I will begin the study by examining different theoretical approaches, and including them in a thorough study of the expression of opinion and subjectivity, to arrive at a theory that can explain isolated expressions of opinion and how they change in context. The next step in the project will consist of the collection of a large-scale corpus of opinionated language, including newspaper editorials, reviews, blogs and Twitter posts (10 million words).

   The corpus will then be annotated with the main phenomena that contribute to the interpretation of evaluation and opinion. I plan to examine the contribution of three specific phenomena:

- Words and expressions that convey different types of opinion and evaluation (e.g., personal feelings, appreciation for an object, or judgement of a person).
- Operators which may change the polarity or strength of that opinion (e.g., negation or modal verbs such as '*might*')
- Discourse or coherence relations which also affect how evaluation is conveyed (e.g., a condition placed on a very positive evaluation, such as '*Great movie, if you want to be bored*').

   The corpus annotation will allow for a sustained study of all the contextual effects that have a role to play in the interpretation of evaluation and opinion as expressions of subjectivity and point of view. The annotation will also contribute to establishing how opinion is expressed differently across genres, from the formality of newspaper editorials to the more colloquial blog and Twitter posts.

   As the first study to approach evaluation through the annotation of a large corpus, the project will offer lasting contributions to knowledge in linguistics and communication. It will enhance theory that encompasses all aspects of evaluation at all levels of language (from the word and

phrase up to the clause, sentence and discourse level). My investigation will also establish best practices for annotation with multiple annotators. Finally, the project will contribute an annotated corpus, which will be freely available to researchers wishing to test it, or to use it for further research.

   The applications of this work are multifold. From a theoretical point of view, it will help determine how subjectivity is expressed in language, building on research from different disciplines (cognitive, formal and computational linguistics). This study of subjectivity will benefit researchers in those disciplines, but also in other academic fields, such as communication or political science. Potential applications include extracting opinion about products, politicians and current events, which will be of interest to companies in social media analysis, decision-makers, journalists, moderators, and the general public, in particular those who use social media regularly.