

# **Combining corpus and experimental methods to investigate trust-building strategies in corporate discourse**

Matteo Fuoli (University of Birmingham, UK) and Christopher Hart (Lancaster University, UK)

## **Introduction**

This study combines corpus-based and experimental methods with the aim to (i) identify and (ii) test the effects of several key trust-building strategies used by companies in their public discourse. First, a corpus-based analysis of a specialized corpus of annual and corporate social responsibility (CSR) reports is conducted in order to investigate how companies discursively construct a positive corporate identity in these texts in an attempt to gain the readers' trust. A scenario-based experiment is then carried out to empirically test the persuasiveness of some of the most salient trust-building strategies identified in the corpus analysis. Below we provide details about the two parts of the study, in turn.

## **The corpus-based analysis**

The corpus-based analysis examines how companies use *stance* expressions (e.g. Biber and Finegan, 1989; Biber et al., 1999; Conrad and Biber, 2000) in annual and CSR reports to construct and promote a positive corporate identity and gain the trust of the stakeholder groups that these texts target. The analysis addresses the following research questions:

- Q1: Are there differences in the frequency and type of stance expressions used in annual and CSR reports?
- Q2: How are stance expressions used to discursively construct and promote a trustworthy corporate identity in these two text types?
- Q3: Do companies present themselves differently in annual and CSR reports? And, if so, how do they discursively shape their identity to appeal to and earn the trust of the audiences that these texts target?

The analysis is based on specialized corpus of roughly 2.5 million words, which includes the 2011 annual reports and CSR reports published by a sample of publicly traded multinational corporations belonging to four different industry sectors. Table 1 provides summary information about the corpus.

The analysis is based on a slightly adapted version of the framework for the analysis of stance presented in Biber (2006b: Chapter 5). The framework comprises three main categories of stance devices: attitudinal stance, epistemic stance and modality. Markers of style-of-speaking stance were not considered, as they proved to be extremely infrequent in the corpus. As in Biber (2006b), modality was kept separate from epistemic stance, even though modals and semi-modals may convey epistemic meanings when used in the extrinsic sense, and could therefore, in principle, be assimilated to the category of epistemic stance.

Table 1: Overview of the corpus

Total number of reports		32 (16 annual reports; 16 CSR reports)
Total corpus size		2,433,956
Size of subcorpora	ACR	1,828,029
	CSR	605,927
Mean report size	ACR	114,251.8
	CSR	37,870.4
Sectors and companies	Financial services	JP Morgan Chase
		HSBC
		Wells Fargo
		Bank of America
		Exxon Mobil
	Oil and gas	Royal Dutch Shell
		Chevron
		BP
		Johnson & Johnson
Food processing	Pharmaceuticals	Merck
		GlaxoSmithKline
		Abbott Laboratories
		Nestlé
		Unilever
	Food processing	General Mills
		Kellogg

The procedure for the identification and quantification of stance constructions in the corpus relied on a combination of automatic techniques and manual analysis. First, an extended list of stance markers was created based on the lists provided in Biber and Finegan (1989) and Biber (2006b). By means of a string search Perl script,<sup>1</sup> the stance markers included in the lists were simultaneously and automatically searched for in the corpus. Second, the results of the automatic search were manually inspected in order to identify and remove all false positives. This was accomplished by means of concordance analysis with the aid of the corpus program AntConc (Anthony, 2012). Once concordances for all the stance markers retrieved in step one were produced and all irrelevant entries were removed from them, the frequency of occurrence of each marker in the annual report and CSR report subcorpora was calculated. Frequencies were normalized to allow for a comparison between the subcorpora, which are of different size. Where relevant, pairwise log-likelihood tests (Rayson and Garside, 2000) were performed to determine whether the differences observed in the frequency of stance markers between the annual report and CSR report subcorpora were statistically significant.

The results of the corpus-based analysis show that companies profile distinct identities in annual and CSR reports. In annual reports, they use stance resources to portray themselves as unbiased, rational, and competent decision makers. In CSR reports, they present themselves as committed, honest, and caring corporate citizens. These discursive identities are interpreted as strategic self-representations that optimize the

<sup>1</sup> The script, which is called 'filter\_by\_field.pl', is included in the supplementary materials for the following article: M. Baroni and A. Lenci. Distributional Memory: A general framework for corpus-based semantics. 2010. Computational Linguistics 36 (4): 673-721. It can be downloaded, free of charge, from this URL: <http://clic.cimec.unitn.it/dm/materials/scripts.tar.gz> [Last accessed: 20 May 2016]. Instructions on how to run the script are available on demand.

persuasive appeal of the reports by addressing the specific expectations of the target readerships.

## The experiment

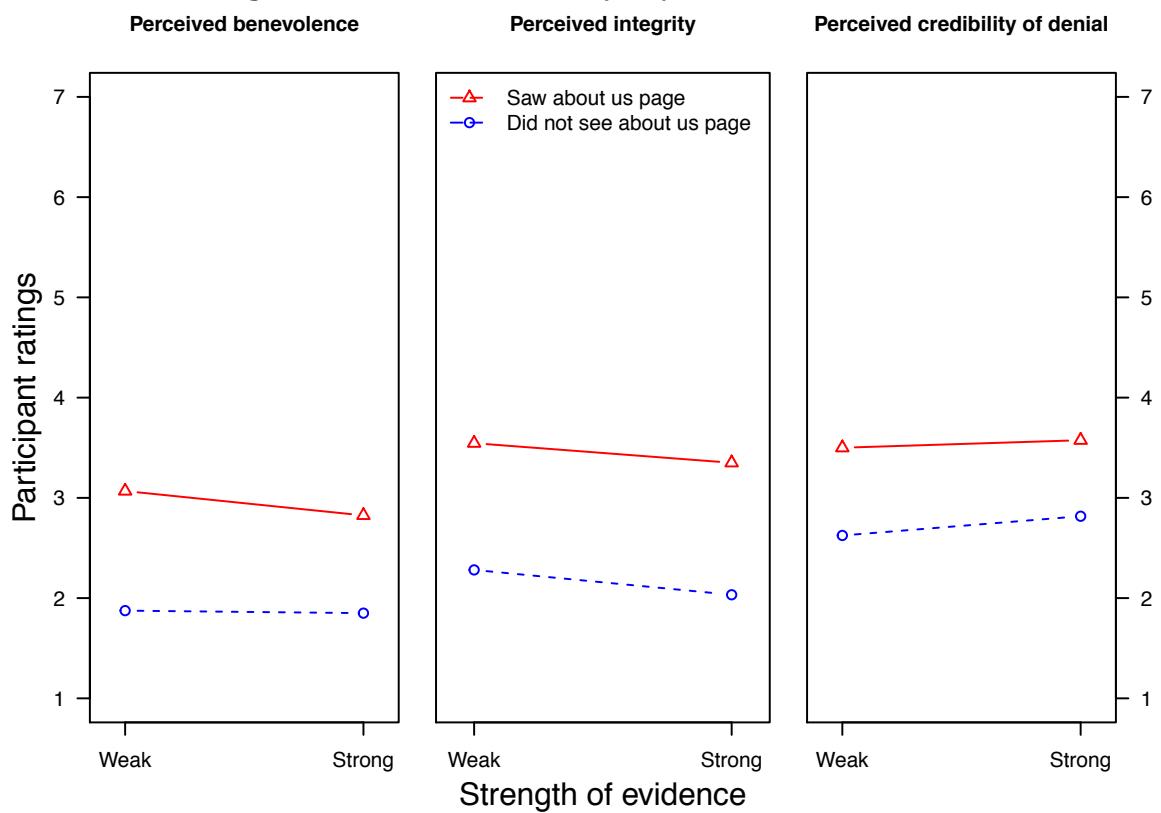
Similar to previous analyses of corporate texts (e.g. Bondi, 2016; Breeze, 2013; Hyland, 1998; Rutherford, 2005), the corpus-based analysis is limited in that it does not provide empirical evidence of the actual persuasiveness of the discourse strategies it brings to light. This limitation is, indeed, common to all qualitative as well as corpus-based discourse analysis research. The experimental part of this study provides the next logical step towards assessing the persuasiveness of corporate narratives. As a follow-up to the corpus-based analysis, we developed and conducted a scenario-based experiment designed to measure the effectiveness of some of the key trust-building discourse strategies that, as the analysis has revealed, are commonly employed by companies in their annual and CSR reports.

The study incorporated a 2x2 between-subjects design. Participants were asked to read a fabricated news story about a fictitious pharmaceutical company who is accused of having paid illegal kickbacks to doctors to help persuade them to prescribe their products over cheaper alternatives. The article reported that the company denied the accusations. Two versions of the article were presented, one in which evidence of the company's guilt was presented as weak, and one in which it was presented as strong. Before reading the article, half of the participants (i.e. the test group) read a fabricated 'About Us' page from the accused company's website. The webpage contained some of the discursive strategies that, based on the corpus-based analysis, we hypothesize perform a trust-building function. Upon reading the text(s), participants completed a questionnaire designed to measure their trust in the accused company.

The results of a pilot study involving 44 individuals show that the discursive strategies contained in the 'About Us' page had a positive effect on participants' trust in the accused company, regardless of the strength of the evidence. As shown in Fig. 1, the participants who were exposed to the 'About Us' text had more positive impressions of the accused company's integrity and benevolence, and perceived the company's denial as more credible compared to the participants who were not exposed to the text. Multiple regression analyses reveal that the differences between the groups are statistically significant.

The results of this study, which are shown in Figure 1 below, provide novel insights into the discursive dynamics of organizational trust, and demonstrate the usefulness of experimental methods as a means for triangulating the findings of qualitative and corpus-based discourse analysis research (Hart, in press).

**Figure 1. Variable means by experimental condition**



## References

- Anthony, L. (2012). Antconc (version 3.2.4) [computer software]. Tokyo, Japan: Waseda University. Available from <http://www.antlab.sci.waseda.ac.jp/> [Last accessed: 20 May 2016]
- Biber, D. (2006b). *University Language: A Corpus-based Study of Spoken and Written Registers*. John Benjamins.
- Biber, D. and E. Finegan. (1989). 'Styles of stance in English: Lexical and grammatical marking of evidentiality and affect'. *Text & Talk* 9/1: 93-124.
- Biber, D., S. Johansson, G. Leech, S. Conrad, and E. Finegan. (1999). *Longman Grammar of Spoken and Written English*. Pearson Education Limited.
- Bondi, M. (2016). 'The future in reports: prediction, commitment and legitimization in CSR'. *Pragmatics and Society* 7/1: 57-81.
- Breeze, R. (2013). *Corporate Discourse*. Bloomsbury.
- Conrad, S. and D. Biber. (2000). 'Adverbial marking of stance in speech and writing' in S. Hunston and G. Thompson (eds): *Evaluation in Text: Authorial Stance and the Construction of Discourse*. Oxford University Press, pp. 56-73.
- Hart, C. (in press). Event-frames affect blame assignment and perception of aggression in discourse on political protests: An experimental case study in critical discourse analysis. *Applied Linguistics*.
- Hyland, K. (1998). 'Exploring corporate rhetoric: Metadiscourse in the CEO's letter'. *Journal of Business Communication* 35/2: 224-244.
- Rutherford, B. A. (2005). 'Genre analysis of corporate annual report narratives a corpus linguistics-based approach'. *Journal of Business Communication* 42/4: 349-378.
- Rayson, P. and Garside, R. 2000. 'Comparing corpora using frequency profiling' in *Proceedings of the workshop on Comparing Corpora*, held in conjunction with the 38th annual meeting of the Association for Computational Linguistics, 1-8 October 2000, Hong Kong, pp. 1-6. Available from <http://ucrel.lancs.ac.uk/lwizard.html> [Last accessed: 20 May 2016]