Tone Analysis in Spanish Financial Reporting Narratives

Antonio Moreno-Sandoval[†,*], **Ana Gisbert**[‡], **Pablo A. Haya**[*], **Marta Guerrero**[*] and **Helena Montoro**[*,†]

[†]Laboratorio de Lingüística Informática, Autonomous University of Madrid
[‡]Accounting Department, Autonomous University of Madrid
[*]Instituto de Ingeniería del Conocimiento, Autonomous University of Madrid
{antonio.msandoval, ana.gisbert}@uam.es
{pablo.haya, marta.guerrero, helena.montoro}@iic.uam.es

This paper analyzes the tone (including polarity and semantic orientations) in a corpus of financial reports in Spanish. Specifically, we look at the Letter to Shareholders section of the Annual Reports, which focuses on an analysis of the financial performance, corporate strategies and other aspects relevant to investors. We use FinT-esp, a semantic analysis tool developed for Spanish narratives, based on lexicons and phrase-structure information. We divide the corpus in four subgroups based on the net earnings figure as a benchmark, to identify differences in tone between profit firms and loss firms. This paper confirms that Spanish financial narratives suffers a communicative bias towards positive terms (Pollyanna effect). Additionally, we provide a gold standard of financial narratives, based on a random selection of 1% of the sentences of the corpus of Letters. We run a first evaluation of three different sentiment analysis tools (Azure, Watson and FinT-esp) compared to the GS and observe that tone analysis in the financial narratives domain breaks with classical sentiment analysis (based on subjective feelings, value judgments, emotions). Financial narratives tone is linked to measurable facts and figures (financial results) and investors' expectations about the future performance of the firm.

1 Introduction

The last decade presents a unique scenario to extend new techniques in computational linguistics to understand financial narratives. The open access to a wide set of electronic resources of financial texts and the release of additional non-regulated disclosures (i.e. annual reports, earnings press-releases, conferences calls, earnings announcements) creates the perfect scenario to understand how managers make use of the language when communicating with stakeholders. Researchers in accounting and finance need to go beyond the use of manual textual analysis, traditional measures of readability and tone, or "bag-of-words". Computational linguistics and accounting academics must work aligned to advance in domain specific lists and new text mining techniques to understand the semantic orientation of sentences (Malo et al., 2014) and the use of the language to guide users' interpretations of financial texts (Malo et al., 2014; Loughran and McDonald, 2016).

Financial narratives are a central component of the companies' reporting package (Beattie, 2014). However, whereas quantitative disclosures (i.e. Financial Statements) are mostly regulated and subject to periodic controls by auditors and enforcement institutions, other financial and non-financial narratives (i.e. earnings press releases or environmental reports) are unregulated, unaudited and offer a wide degree of discretion to managers. The exponential increase of qualitative disclosures in the last decade has raised a wide debate on whether financial narratives really offer incremental information content on top of the traditionally regulated financial information (Boudt, Thewissen and Torsin, 2018; Plumlee et al., 2015). Managers choose between the use of narratives to increase transparency and reduce information asymmetry or intentionally bias investors' perceptions to obfuscate the reality about firm's performance (Merkl-Davis and Brennan, 2007; Arslan-Ayaydin et al., 2016).

Evidence shows that narratives are indeed value relevant, contribute to the company's reputation (Craig and Brennan, 2012) and investors and analysts decision making process (Boudt et al., 2018; Arslan-Ayayding et al., 2016; Yekini, et al., 2016). Therefore, computational linguistics can play a crucial role in supporting the accounting and finance field to discern about the use and orientation of financial narratives.

The first motivation for this study relies on the idea that company's earnings trend affects the tone of financial narratives. Particularly, for our research question, we investigate whether there are differences in the language orientation (opinion) of the letter to shareholders across benchmark vs. non-benchmark beating companies. We focus the empirical analysis on the letter to shareholders as a document with "*enormous rhetorical importance*" to build credibility and confidence about the company (Hyland, 1998) and influence investors' decisions (Baird and Zelin, 2000; Breton and Taffler, 2001).

Managers' choice to bias or enrich financial information depends on a set of incentives. Previous literature document significant capital market rewards (penalties) for benchmark (non-benchmark) beating companies (Graham et al. 2005). We consider these capital markets' rewards and potential penalties a clear incentive to manage upwards the tone of narrative disclosures, avoid negative messages and therefore, affect investors' perceptions about the performance of the firm. Previous literature in the US document that managers structure their narratives to manage investors' perceptions about the company performance (Alee and Deangelis, 2015). Li (2008) finds that firms with lower reported earnings have less readable annual reports (10-K) and more recently, Iatridis (2016), Davis and Tama-Sweet (2012) or Feldman et al. (2010), finds that benchmark beating, and high-growth firms tend to use less pessimistic language.

For a final sample of 76 companies listed in the Spanish Stock Exchange, we apply NLP techniques for tone analysis, and we measure the degree of accuracy of the use of these techniques in the domain of financial narratives.

The performance of current sentiment analysis (SA) systems seems less accurate when used in the financial domain compared to other narrative contexts as social media messages. We posit that the underperformance of the different tools is linked to the specific language complexity of financial narratives due to its impact on users' decisions that may affect the company's market value.

In spite of the caveats and limitations, this study is one of the first attempts to identify automatically the tone and semantic orientations of financial narratives in the Spanish language.

2 Characteristics of the corpus

The potential sample consists of 125 companies listed in the Madrid Stock Exchange. For each company, we accessed the corporate website in order to retrieve all the publicly available Annual Reports for the four-year period 2014-2017. However, the Spanish accounting regulation does not require the preparation of this document and therefore it is not available for all companies. We finally retrieved the Annual Reports dataset files in PDF format for a final sample of 76 reports.

Annual Reports have not a standardized format across companies, its content and structure vary significantly and therefore, they are rarely comparable documents. One of the few comparable sections across companies is the *Letter to shareholders*.

Due to the relevance of the "letter to shareholders", we focus the analysis on this specific and relevant section of the Annual Report. The letter to shareholders it is not subject to accounting regulation and it offers managers with a great opportunity to use their writing style to change investors' perceptions about the past, present and future performance of the company (Hooghiemstra, 2010). Previous literature documents that investors decisions are clearly influenced by the information presented in the letter to shareholders (Baird and Zelin, 2000; Breton and Taffler, 2001).

In order to identify differences in the language style across benchmark vs. non-benchmark beating companies, we group companies in groups based on the company's financial performance. For this purpose, we download financial data from ORBIS, a Bureau Van Dijk database with financial information for over 300 million companies across the Globe. ORBIS is key source of financial data for professional and academic use. More specifically, we download the net income figure (NI_{it}) for each sample company across the time-period 2013-2017 to classify firms in the following four groups as follows:

- **Group 1**: Companies reporting positive earnings (profits) ($NI_{it} > 0$) and improving past performance. That is, increasing earnings compared to the preceding year [$(NI_{it} - NI_{i,t-1})/ |NI_{it}| > 0$].

- **Group 2:** Companies reporting positive earnings (profits) ($NI_{it} > 0$) and declining past performance. That is, decreasing earnings compared to the preceding year $[(NI_{it} - NI_{i,t-1})/ |NI_{it}| < 0]$.
- **Group 3**: Companies reporting negative earnings (losses) ($NI_{it} < 0$) but improving past performance. That is, decreasing the amount of losses from the preceding year $[(NI_{it} - NI_{i,t-1})/ |NI_{it}|) > 0]$.
- **Group 4:** Companies reporting negative earnings (losses) ($NI_{it} < 0$) and declining past performance. That is, increasing losses compared to the preceding year $[(NI_{it} - NI_{i,t-1} / |NI_{it}|) < 0]$.

The initial corpus of the Letter to shareholders was composed of a total of 385 text, 462,189 words, 16,800 sentences, and 8,682 paragraphs (Moreno et al., 2019). However, we excluded from the final corpus those letters for companies with missing net income data in the ORBIS database (7 documents).

For the normalization of the corpus, each letter is in a separate file -encoded in UTF-8-, one sentence per line and double carriage return separating each paragraph.

The final 378 texts are distributed across the four groups as follows (Table 1):

	Num. of texts
Profits / Δ Earnings	258
Profits / ∇ Earnings	68
Losses / Δ Earnings	20
Losses / ∇ Earnings	32
Total	378

Table 1: Company classification and number of texts

3 Applying an opinion and semantic tool

Robo-readers (Malo et al., 2014) can extract opinion and semantic orientations from reports to identify how financial sentiments relate to future company performance. In this article, we apply a sentiment analysis engine to analyze the tone of a corpus of Spanish financial narratives. More specifically, instead of using informal texts in social media (see Section 4.2), we have focused the analysis on the sentiment and opinion in domain specific texts, Letter to shareholders.

We use a lexicon and rule-based sentiment engine instead of an ML classifier, with a general polarity lexicon and a phrased-structure grammar. The domain-independent lexicon is made up of about 8,000 single word entries and more than 20,000 multiword expressions. The grammar is a modified version of the Spanish FreeLing (Padró and Stanilovsky, 2012). The grammar is used to identify semantic groupings at a phrase-structure level and to project polarity information up to the upper level. The label (Positive 100 to 1; Neutral 0; Negative -1 to -100) assigned to each sentence is the result of the projection of the different phrase units in the construction of the parsing tree (similar to the one described by Malo et al., 2014).

3.1 Preparing the corpus

The corpus consists of 16,800 sentences, but it includes lines for the names of Presidents/CEOs, their positions and section titles. To remove this irrelevant content for sentiment analysis, we delete all sentences with less than 4 words. The final corpus contains 14,812 sentences to run the opinion engine.

3.2 Output

We separate the final corpus into the four categories shown in Table 1. The opinion sentiment engine provides a numeric value for each sentence between -1 and 1, where 0 is the neutral value. Overall, the results clearly show that a positive opinion prevails in all categories (see Figure 1).

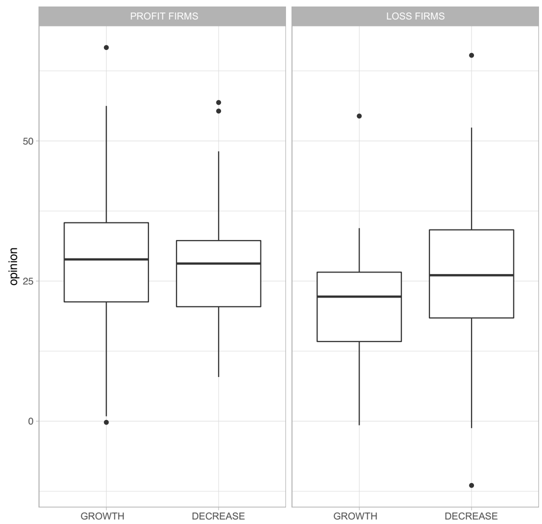

Figure 1: Opinion for the four categories

These results are consistent with expectations and previous literature that suggests the tendency of managers to present the analysis of the company's financial results from the best possible perspective. Figure 1 shows two remarkable results: (1) companies with losses and decreases in performance (group 4) have a higher positive tone compared to companies with losses but increases in performance (group 3); (2) Group 4 companies have a tone similar to group 1. That is, the worst performing companies maintain a positive tone similar to the best performing companies in the narratives of the letter to shareholders.

4 Results evaluation

In order to measure the performance of the opinion sentiment tool, we compiled a "Gold Standard" (FinT-esp GS) annotated by human experts.

4.1 Building an evaluation Gold Standard

The objective of the GS is to assess the accuracy of competing tools and observe the polarity distribution in the financial narratives domain.

Sample selection: We randomly selected 1% (148 sentences) of the sentences of the final corpus (14,812 sentences) as a significant representation of the complete dataset. Annotators tag a total of 130 sentences from profit companies and 18 sentences for companies with loss.

Annotation instructions: Annotators are informed about the requirement to assess the tone of the sentence from investor's perspective. This implies that the "referee" for disagreements between the two annotators must have financial knowledge:

- **Neutral**: statements without positive or negative judgements about the information (i.e. without adjectives and adverbs, such as "better", "increasingly", "significant", "unfortunately" etc.) Example: "*Nos dirigimos, un año más, a ustedes para informales sobre los resultados del ejercicio 2016 cuyas cuentas se someten a su aprobación*" (Trans. 'Once again, we are writing to inform you of the results for the fiscal year 2016, the annual accounts of which are submitted for your approval'). Additionally, sentences are considered neutral if includes the same amount of both positive and negative statements that compensate with each other and therefore, the tone of the message is neutralized. Conversely, if the number of positive statements predominates the sentence is considered as "positive", "negative" otherwise. That is, when the number of negative statements predominates.
- **Positive**: "good news" messages based on real economic facts. Example: "*En Abril del 2017 tenía el placer de comunicarles un inicio de acuerdo con el fondo de pensiones APG para la creación de una Socimi especializada en activos residenciales*." (Trans. 'In April 2017 I had the pleasure to inform you about the beginning of an agreement with the APG pension fund for the creation of a Real Estate Investment Trust specialized in residential assets').
- **Negative**: "bad news" messages or "positive" expressions that mask losses or decreases in earnings. Example: "*Esta presentación se produjo como consecuencia de la demora sufrida dentro del proceso negociador con el pool bancario en referencia a la reestructuración de la deuda*." (Trans. 'This presentation occurred as a consequence of the delay in the negotiation process with the banking pool regarding debt restructuring')

Annotation guides have been created before the manual annotation from a sample of 1% of the dataset different from the one used in the GS.

Percent overall agreement = 80.41%
Free-marginal kappa = 0.71
Fixed-marginal kappa = 0.62

Table 2: Inter-annotators agreement

Annotators and "referee": Two expert linguists have tagged all 148 sentences independently. In addition, a financial expert has reviewed all the tone assessments and has decided the correct one in case of discrepancy between annotators. Only in few cases, based on her knowledge of the domain, the "referee" has corrected the annotations shared among linguists.

Table 2 shows the inter-annotator agreement for the 148 cases and 3 categories. Noteworthy is the fact that the annotators agreed more in the 130 sentences from the profit companies than in the 18 sentences from loss companies: 82.41% vs. 66.67%. The results are indicative of the difficulty of analyzing the tone of the narratives of companies with financial problems.

Following Fleiss's rule of thumb, kappa values from 0.40 to 0.75 are considered "*intermediate – good*." Therefore, the values obtained for the annotation procedure are quite satisfactory.

Finally, the financial expert's revision of the linguists' annotation makes the GS highly reliable for assessing the accuracy of the opinion analysis tools.

4.2 Semantic tone of the GS

Considering the GS a representative sample of the financial reporting domain, we focus on the distribution of polarity values. Results in Figure 2 show that the positive tone (70%) prevails over the others, with very few negative messages (8%).

These results contrast with the distribution of negative vs. positive tone in other highly studied domain: social media. Taking as a reference the TASS competition[1] developed between 2012 and 2017 for Sentiment Analysis in Spanish datasets, the InterTass2017 reports the following distribution for a Twitter GS with 1,625 tuits: 13% (neutral), 47% (+) and 40% (-) (see Figure 3). (Martínez-Cámara et al., 2017).

In Twitter, negative messages are close in number to positive messages. This great difference in polarity distribution forces sentiment systems to make a strong adaptation. The next section explains differences in performance across the three sentiment analysis tools: Watson, Azure, and our FinT-esp. Most tools are usually applied in sentiment analysis in social media (i.e. Twitter). Therefore, we could use to the values of the last competition InterTASS2017 as a reference for the state-of-the-art in Spanish (see Table 3)[2].

Best systems	M-F1	Acc.
ELiRF-UPV-run1	0.493	0.607
RETUYT-svm	0.471	0.596
ELiRF-UPV-run3	0.466	0.597

Table 3: Best systems in InterTASS 2017

5. Performance comparison of three sentiment tools

We have chosen two professional applications to evaluate their performance in an unusual domain.

Microsoft Azure Sentiment Analysis is included in the Text Analytics API service. It is based on machine learning algorithms and does not require training data. Azure uses neural network technology and word embeddings. The evaluation of each sentence has been done from the demo page[3] copying the results into a spreadsheet.

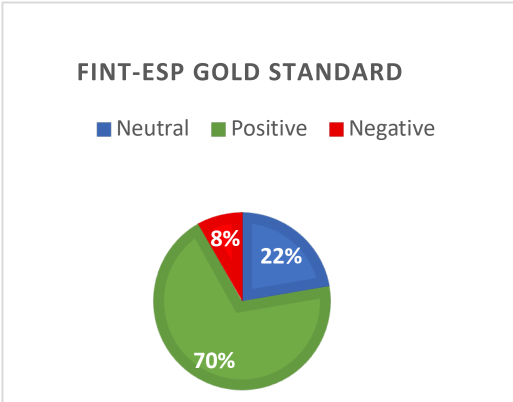

Figure 2: Polarity distribution in the Financial narratives

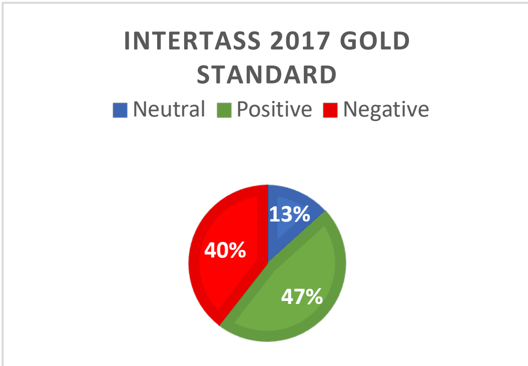

Figure 3: Polarity distribution in Twitter

IBM Watson NLU is a collection of APIs that offer text analysis through NLP[4]. We haven't created a custom model to get specific results that are tailored to the financial domain. In this way, we have maintained the same level of domain adaptation in all three systems. In the case of our lexicon-based system, we have not developed a specific one for financial terms.

[1] http://www.sepln.org/workshops/tass/

[2] All the systems participating in TASS 2017 "are based on the use of deep learning techniques as the state-of-the-art in SA in Twitter" (Martínez-Cámara et al. 2017).

[3] https://azure.microsoft.com/es-es/services/cognitive-services/text-analytics/

[4] https://natural-language-understanding-demo.ng.bluemix.net/

	Prec N	Prec. +	Prec. -	**Macro Prec.**	**Acc.**	Recall N	Recall +	Recall -	**Macro Recall**	F1 N	F1 +	F1 -	**M F1**
Watson	0.26	**0.89**	**1.00**	**0.72**	0.43	**0.88**	0.32	0.08	0.43	0.41	0.47	0.15	**0.54**
Azure	**0.50**	0.73	0.33	0.52	**0.70**	0.06	**0.94**	**0.33**	0.45	0.11	**0.83**	**0.33**	0.48
FinT-esp	0.35	0.83	0.23	0.47	0.57	0.58	0.59	**0.33**	**0.50**	**0,44**	0,69	0,27	0.48

Table 4: Results against the FinT-esp Gold Standard

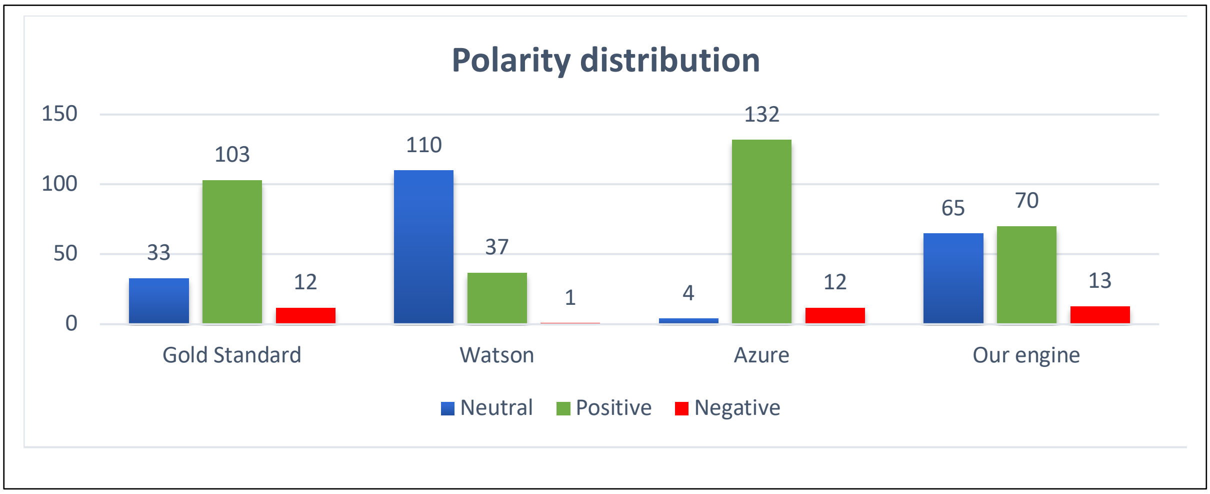

Figure 4: Polarity distribution in the FinT-esp GS and the three systems

Table 4 gives the detailed results of the evaluation, disaggregating the scores by polarity value. Figure 4 shows a wide variety in the performance of the systems, an indicator that each method uses very different technologies.

Azure approximates the distribution of the FinT-esp GS and obtains better results in Accuracy than the other systems (see Table 5). However, MacroF1 scores show that none of the 3 systems meets the objective of classifying the polarity of the sentences acceptably.

In the following section, we examine the peculiarities of the systems and the analyzed sentences.

5 Discussion

A general feature of the systems is that they provide very good precision results with positive sentences (from Azure 0.73 to Watson 0.89). However, in recall only Azure stands out (0.94). Bearing in mind that positive messages account for 70% of the GS, this largely explains why Azure wins in the benchmark (0.83 F1 score for +).

Conversely, Azure is the worst at detecting neutrals (F1 score of 0.11). None of the three systems works acceptably with negative messages either (F1 between 0.15 and 0.33).

Systems	M-F1	Acc.
ELiRF-UPV-run1	0.49	0.61
Watson	**0.54**	0.43
Azure	0.48	**0.70**
FinT-esp	0.48	0.57

Table 5: Macro F1 and Accuracy

None of the three systems has been specifically trained for the financial reporting domain. Therefore, it is striking, that each has a very different analysis strategy. Watson is decidedly inclined towards neutral messages, while Azure bets almost exclusively on interpreting sentences as positive. The contingency table displays the distribution of the variables (Table 6).

Watson and Azure are based on ML technology, whereas FinT-esp is based on a general polarity lexicon and phrase rules. In the test results, the distribution of positives and neutrals is similar. In all three systems, the focus is on the subjective part of the texts and not on the description of the facts. Something that escapes all the tools evaluated here.

FinT-esp			
GS	negative	neutral	positive
negative	3	4	5
neutral	3	23	7
positive	7	38	58

Azure			
GS	negative	neutral	positive
negative	4	0	8
neutral	4	2	27
positive	4	2	97

Watson			
GS	negative	neutral	positive
negative	1	11	0
neutral	0	29	4
positive	0	70	33

Table 6: Contingency tables

Next, we will show an example of each polarity, where none of the three systems has been able to analyze correctly.

- **Negative**: "*Seguimos siendo líderes, pero nuestro mercado ha quedado reducido al 20% del total*." (trans. 'We are still leaders, but our market has been reduced to 20% of the total'). From the investor's point of view, the strong reduction in the market share is considered as bad news. Azure and the Fin-tesp system classified the sentence as positive, probably because of the presence of "leaders."
- **Positive**: "*La deuda a diciembre de 2016 se redujo en los últimos doce meses de 305 millones a 188 millones de euros, es decir, hemos bajado 117 millones de euros en un año*." (Trans. 'Debt at December 2016 was reduced in the last twelve months from 305 million euros to 188 million euros, in other words, we have reduced 117 million euros in one year'). Although "debt" is an inherently negative word, the message is positive for investors, as the debt has been drastically reduced. Watson and the FinT-esp tool classified the sentence as neutral; Azure as negative.
- **Neutral**: "*A pesar de todo ello, la eficiencia de la actividad en una sola planta se fue poniendo de manifiesto a lo largo del año*." (Trans. 'In spite of all this, the efficiency of the activity in a single plant became evident throughout the year'). In this sentence two opposite movements are neutralized, expressed by "in spite of" and by "efficiency". Azure and the FinT-esp tool classified the sentence as positive. Watson as neutral.

These examples reflect the argumentative complexity of financial narratives. It is common for two opposing ideas to appear in the same sentence. In some cases, they are neutralized but in others one is stronger than the other.

6 Conclusions and future work

Financial narratives have boosted across the last decade, offering a unique setting to test different computational linguistics methodologies for sentiment analysis across a specific language domain: financial reporting texts. Additionally, whereas most of the current studies have been centered in English financial narratives, the access to non-English financial and non-financial qualitative disclosures offers a great opportunity for sentiment analysis in other languages.

This paper confirms that the Spanish financial narrative suffers the Pollyanna effect (Rutherford 2005). That is, a communicative bias towards positive terms. This bias is consistent with the managers' aversion to communicate bad news that may affect the company's capital market value or the company's reputation. The positive bias in the narratives affects the accuracy of the different semantic analysis tools. Compared to the use of SA techniques in other narrative contexts as the social media, differences in the performance of the three tested systems suggests that the specific language complexity of these texts requires more domain-specific methods for tone analysis. Particularly, across bad-performing companies where sentences including words such as "debt" or "restructuring" can be misclassified as "negative" whereas the overall context of the message is positive. Or bad news related to decreases in performance can be masked with the use of positive expressions (i.e. *we are still leaders, although our market has been reduced to 20% of the total*).

Additionally, the different distribution of polarity in two gold standards is consistent with

Chen, Huang and Chen (2018) who showed that there is a clear difference in the language of "market sentiment of social media data" compared to other "formal reports". In the financial reporting domain, the tone is directly linked to measurable facts (financial performance, sales and gross margin increases, debt reductions, EBITDA), past performance and investors' future expectations about the company. Therefore, while expectations can be measured with classic sentiment analysis, measuring financial facts needs the participation of financial experts to create specialized lexicons, train models and offer a neat assessment of those sentences that present discrepancies. This explains why the three tools evaluated had poor results.

Žnidaršič et al. (2018) have studied the importance of the expressions of "trust" and "doubt" in financial communications, and the correlations with the financial activity of companies. An extension of this article will address lexical and terminological issues based on the FinT-esp corpus.

We contribute to language resources with the compilation of the first corpus of "letters to shareholders" in Spanish. Additionally, we create a gold standard (and the corresponding annotation guidelines) to evaluate opinion systems and we have carried out a first sentiment analysis comparison. Both the corpus and the GS will be freely available to researchers on the project site.

Future work aims to develop a financial polarity lexicon, including verbs and adverbs expressing epistemic modality (probability and certainty) as suggested in Malo et al. (2014). We will also explore Machine Learning methods trained on dataset to classify the tone of financial narratives.

Acknowledgments

This research is supported by Spanish Ministry of Economy and Industry, via the FinT-esp project (TIN2017-89351-R).

References

Kristian D. Allee and Matthew D. Deangelis. 2015. The structure of voluntary disclosure narratives: evidence from tone dispersion. *Journal of Accounting Research*. 53 (2): 241-274. https://doi.org/10.1111/1475-679X.12072

Özgür Arslan-Ayaydin, Kris Boudt and James Thewissen. 2016. Managers set the tone: Equity incentives and the tone of earnings press releases. *Journal of Banking and Finance*, 72 (Supplement): S132-S147. https://doi.org/10.1016/j.jbankfin.2015.10.007

Jane E. Baird and Robert C. Zelin II. 2000. The effects of information ordering on investor perceptions: An experiment utilizing presidents' letters. *Journal of Financial and Strategic Decisions*, 13(3): 71-81.

Gaétan Breton and Richard J. Taffler 2001. Accounting information and analyst stock recommendation decisions: a content analysis approach. *Accounting and Business Research*, 31 (2): 91-101. https://doi.org/10.1080/00014788.2001.9729604

Vivien Beattie, 2014. Accounting narratives and the narrative turn in accounting research. *The British Accounting Review*, 46(2): 111-134. https://dx.doi.org/10.1016/j.bar.2014.05.001

Kris Boudt, James Thewissen and Wouter Torsin. 2018. When does the tone of earnings press releases matter? *International Review of Financial Analysis*, 57: 231-245. https://doi.org/10.1016/j.irfa.2018.02.002.

Chung-Chi Chen, Hen-Hsen Huang and Hsin-Hsi Chen. 2018. NTUSD-Fin: A Market Sentiment Dictionary for Financial Social Media Data Applications. In *Proceedings of the LREC 2018 Workshop "The First Financial Narrative Processing Workshop (FNP 2018)"*. Miyazaki, Japan, pages 59–65

Russell J. Craig and Niamh M. Brennan. 2012. An exploration of the relationship between language choice in CEO letters to shareholders and corporate reputation. *Accounting Forum*. 36 (3): 166-177. https://doi.org/10.1016/j.accfor.2012.02.004.

Angela K. Davis and Isho Tama-Sweet. 2012. Managers' use of language across alternative disclosure outlets: earnings press releases versus MD&A. *Contemporary Accounting Research*, 29 (3): 804-837. https://doi.org/10.1111/j.1911-3846.2011.01125.x

Ronen Feldman, Suresh Govindaraj, Joshua Livnat and Benjamin Segal. 2010. Management's tone change, post earnings announcement drift and accruals. *Review of Accounting Studies* 15(4):915-953. https://doi.org/10.1007/s11142-009-9111-x

John R. Graham, Campbell R. Harvey and Shiva Rajgopal. 2005. "The economic implications of corporate financial reporting", *Journal of Accounting and Economics*, 40(1-3): 3-73. https://doi.org/10.1016/jacceco.2005.01.002

Reggy Hooghiemstra, 2010. Letter to the shareholders: A content analysis comparison of letters written by

CEOs in the United States and Japan, *The International Journal of Accounting*, 45(3): 275-300. https://doi.org/10.1013/j.intacc.2010.06.006

Ken Hylan. 1998. Exploring corporate rhetoric: Metadiscourse in the CEO's letter. *Journal of Business Communication*, 35(2): 224-245. https://doi.org/10.1177/002194369803500203

George E. Iatridis. 2016. Financial reporting language in financial statements: Does pessimism restrict the potential for managerial opportunism? *International Review of Financial Analysis*, 45(1): 1-17. https://doi.org/10.1016/j.irfa.2016.02.004.

Feng Li. 2008. Annual Report Readability, Current Earnings, and Earnings Persistence. *Journal of Accounting and Economics*, 45(2-3): 221-247. https://doi.org/10.1016/j.acceco.2008.02.003.

Tim Loughran and Bill McDonald. 2016. Textual Analysis in Accounting and Finance: A survey. *Journal of Accounting Research*, 54 (4): 1187-1230. https://doi.org/10.1111/1475-679X.12123.

Pekka Malo, Ankur Sinha, Pekka Korhonen, Jyrki Wallenius and Pyry Takala. 2014. Good Debt or Bad Debt: Detecting Semantic Orientation in Economic Texts. *Journal of the Association For Information Science And Technology*, 65(4): 782-796. https://doi.org/10.1002/asi.23062

Eugenio Martínez-Cámara, Manuel C. Díaz-Galiano, M. Ángel García-Lumbreras, Manuel García-Vega and Julio Villena-Román. 2017. Overview of TASS 2017. In *Proceedings of TASS 2017: Workshop on Semantic Analysis at SEPLN*, September 2017, pages 13-21. https://www.researchgate.net/publication/320410270

Doris M. Merkl-Davis and Niamh M. Brennan. 2007. Discretionary disclosure strategies in corporate narratives: incremental information or impression management? *Journal of Accounting Literature*, 27: 116-196.

Antonio Moreno-Sandoval, Ana Gisbert and Helena Montoro. 2019. FinT-esp: a corpus of financial reports in Spanish. Presented at the *XI International Conference on Corpus Linguistics*, Valencia, Spain, 15-17 May 2019 (to be published as a chapter).

Lluis Padró and Evgeny Stanilovsky. 2012. FreeLing 3.0: Towards Wider Multilinguality. In Proceedings of LREC 2012 2473-2479. http://hdl.handle.net/2117/15986

Marlene Plumlee, Darrell Brown, Rachel M. Hayes and R. Scott Marshall. 2015. Voluntary environmental disclosure quality and firm value: Further evidence. *Journal of Accounting and Public Policy.* 34(4): 336-361. https://doi.org/10.1016/j.accpubpol.2015.04.0004

Brian A. Rutherford. 2005. Genre Analysis of Corporate Annual Report Narratives: A Corpus Linguistics-Based Approach. *Journal of Business and Communication,* 42(4): 349-378. https://doi.org/10.1177/0021943605279244

Liafisu S. Yekini, Tomasz P. Wisniewski, Yuval Millo. 2016. Market reaction to the positiveness of annual report narratives. *The British Accounting Review*. 48 (4): 415-430. https://doi.org/10.1016/j.bar.2015.12.001

Martin Žnidaršič, Jasmina Smailović, Jan Gorše, Miha Grčar, Igor Mozetič and Senja Pollak. 2018. Trust and Doubt Terms in Financial Tweets and Periodic Reports. In *Proceedings of the LREC 2018 Workshop "The First Financial Narrative Processing Workshop (FNP 2018)"*. Miyazaki, Japan, pages 59–65.