

Lexical indicators of profit and loss in Spanish shareholder letters

Blanca Carbajo Coronado
Autonomous University of Madrid / Spain

Abstract – Through a case, frequency, and collocational study, this work aims to detect linguistic differences between the letters to shareholders of profitable and loss-making companies. In a sample of 50 letters from each group, two lexical categories were analysed, verbs and eventive nouns, using the corpus manager *Sketch Engine*. The results indicated that verbs and eventive nouns with an absolute frequency of at least five occurrences overlap in both corpora by 95.7 per cent and 95.8 per cent, respectively. The frequency analysis showed that those with significantly higher frequency in one corpus denoted events or activities that were to be expected for companies in their group, such as *aumentar* ('increase'), *crecimiento* ('growth'), or *cumplimiento* ('compliance'), in the profitable companies; but terms such as *relanzar* ('relaunch'), *revalorización* ('revaluation'), or *pérdida* ('loss') in the loss-making companies. The analysis of the combinatorial properties of these verbs and nouns revealed subtle but significant differences between the two groups. In the case of verbs, the choice of the direct object is key, and in the case of nouns, qualifying and adverbial adjectives are crucial, as well as the Theme complements.

Keywords – finance; lexicon; collocations; verbs; eventive nouns

1. INTRODUCTION¹

Companies publish information periodically on their initiatives or following regulations to enable potential investors and shareholders to assess their results, financial situation, and business plans. A shareholder letter is usually prepared once a year and included at the beginning of the company's annual report. The company's senior executives send this to their shareholders to provide a general summary of its operations for the entire year. The letter discusses the company's key financial results, market position, objectives, and

¹ This publication is part of the project *Computational linguistic methods for the readability and simplification of financial narratives*. CLARA-FINT (PID2020-116001RB-C31), funded by the Spanish Ministry of Science and Innovation and the State Research Agency. The author acknowledges the financial support provided by the FPU grant (FPU20/04007) which has been awarded by the Spanish Ministry of Science, Innovation and Universities.



business approaches. Additionally, specific events of the past year and share price fluctuations may also be stated.

The starting point of this paper is the assumption made by experts in the field of financial communication, who point out that loss-making companies mimic the themes and causality patterns of profitable companies (see Aerts 2005). Other works, such as the computational test by El-Haj *et al.* (2021), can lead to assuming a particular affinity or closeness between both groups. Based on machine learning, that study aimed at developing an automatic classifier of letters in profit or loss. Between 85 per cent and 89 per cent of the letters were correctly classified, but this is a low figure for a binary classification task in artificial intelligence.

This paper seeks to contribute to the study of shareholder letters with the help of the tools and methods of corpus linguistics in order to shed light on the usage and collocations of some lexical items, and to provide an insightful discussion on the language of shareholders' communications. Results can be valuable to linguists, especially working in the field of Spanish for specific purposes, as well as corporate communicators or business communication researchers.

A selection of 100 letters to shareholders are analysed, which are part of a corpus of Spanish annual reports called *FinT-esp* (Moreno-Sandoval *et al.* 2020). The analysis is based on the distinction between two groups of companies: those with profits and those with losses in the year under review. Little research has been carried out with the use of this corpus with the exception of García Toro (2020) on discourse markers. The distinction between companies is essential because, through the letters, shareholders may be able to detect any subtle changes in the company's strategy and learn where the company's senior executives stand on interpreting the results. Some companies may speak openly and transparently about their accounting results, but that is not true for all. In the letters, the results are either truthfully shown or hidden, depending on whether the company has made a profit or registered losses in the previous accounting year (Patelli and Pedrini 2014).

The focus of this paper is on indicators of profit or loss in companies. An 'indicator' is a loose term that refers to lexical items that are distinctive of either profit-making or loss-making companies, as they appear more frequently in one of both corpora. Ideally, lexical items will account for unique or almost unique themes to that group of companies. For example, a clear indicator would be *reparto (de dividendos)* ('(dividend) payout').

Only profitable companies own the required capital to pay out dividends to shareholders. By identifying such indicators, this study aims to provide insights into the linguistic features that distinguish profitable companies from their loss-making counterparts.

The categories chosen for the analysis are verbs and eventive nouns, that is, nouns designating events or something that happens within a period. The reason for this choice is that they account for the letters' most relevant themes. Action verbs, that is, activities, achievements, and accomplishments in Vendler's (1957) terminology (e.g. *competir* 'compete', *delegar* 'delegate', *implementar* 'implement') can shed light on the differences between the profit and loss groups in terms of the initiatives undertaken by the company during the respective financial year. Causative verbs such as *permitir* ('allow') or *provocar* ('provoke') can help understand the argumentation and justification of the company's decisions. They give an account of the point of view taken by companies to explain the events that occurred during the financial year. For example, the reduction in revenue may *have been caused* by a fiscal adjustment beyond the company's control. Another example is that the hard work of employees could be what *has enabled* record highs in revenue to be achieved. Verbs of change of state such as *reducir* ('reduce') and *crecer* ('grow') can give clues as to how the business is doing. On the other hand, copulative verbs, usually *ser* ('be') introduce information about how the year is evaluated in static terms by the company or the senior executive. Like verbs, nouns identify the critical events of the year and those that are expected to occur in the future. For example, one can talk about the *rise* in steel prices, in which the eventive noun identifies the main topic.

With the abovementioned information in mind, this paper aims to compare 50 letters from loss-making companies with 50 letters from profit-making companies to identify the differences between the texts of both groups regarding lexical aspects. More specifically, the research questions addressed are the following:

Q1) Are there any noteworthy lexical differences that set them apart?

Q2) In cases where the letters from both corpora employ the same lexical choices, what methods can be used to disambiguate them?

The paper is structured as follows. Section 2 provides an overview of the work done in financial communication studies and linguistics on letters to shareholders. It also introduces the characteristics of the letters, their communicative purposes, and the topics

they address. Section 3 describes the corpus of letters and the methodology. Section 4 presents the data analysis structured around the two lexical categories chosen. First, the differences between profit and loss letters in the frequency of use are discussed (Section 4.1. for verbs and 4.3. for eventive nouns). Then, the collocations that can help to disambiguate between the two groups of letters are studied (Section 4.2. for verbs and 4.4. for eventive nouns). The conclusions and final remarks are included in Section 5.

2. LETTERS TO SHAREHOLDERS

Letters to shareholders present the factual information contained in the annual report and assess the company's current situation, past performance, and growth potential. They constitute, therefore, a textual genre with an apparent persuasive character, not only because of the topics they cover but also because they explicitly identify their readers; their authors subscribe to what is said and sometimes state their position on the issues (Vogel 2020). These characteristics make letters to shareholders unique compared to other genres in the business world, which is one of the reasons why they have attracted a great deal of attention from experts in corporate investor relations.

Notable works have focused on the letters' move analysis, that is, the small semantic units with communicative purpose into which a text is divided (see Swales 1990; Garzone 2005; Bhatia 2008, among others). In Bhatia's (2008) proposal, a letter to shareholders contains the following movements: 1) reviewing the financial year, 2) identifying important issues, 3) developing those issues, 4) describing expectations and making promises, 5) expressing gratitude to employees and shareholders (optional), 6) talking about the future, and 7) closing positively and politely. Other authors propose variants, but these always include at least three movements: contextualisation/assessment of the year under discussion, description and justification of the financial year's performance, and discussion of future performance and strategy.

Some of those works focusing on moves address the topic from a comparative perspective. These include Garzone (2004), Nickerson and Groot (2005), Ruiz-Garrido *et al.* (2012), and Skorczynska and Giménez-Moreno (2016), among others. The work of Skorczynska and Giménez-Moreno (2016) reveals interesting characteristics of Spanish letters to shareholders, as opposed to British and Polish ones. For example, the fact that they share with the British ones the highest number of moves (a total of 15 moves

compared to 17 for the British and nine for the Polish ones), or that they reveal a preference for providing a great deal of information on the objectives, performance, strategy, investment, and work of the board of directors.

A number of studies have been particularly interested in the mechanisms of evaluation, which is understood as the expression of the speaker's or writer's attitude, or point of view towards the entities or propositions they are addressing (see Thompson and Hunston 2000), as well as the linguistic elements that make them explicit: epistemic modality markers (Kranich 2011; Kranich and Bicsar 2012), and adjectives or collocations (Poole 2016, 2017; Wang 2020; Skorczynska Sznajde 2021).

Adjectives and other evaluative categories, such as adverbs, can help determine how a company or senior executive perceives a particular event, action, or state. As a result, researchers in both linguistics and corporate communication studies have paid considerable attention to them. However, other lexical categories (primarily nouns or verbs) have received more attention in financial research, where there is a line of works focused on form-oriented content analysis, that is, the type and frequency of lexical units used (see Laskin 2018). This paper continues this line of research and provides new linguistic insights into these categories.

3. METHODOLOGY

A total of 100 letters of 24 companies from the period 2014–2018 were extracted from the *FinT-esp* corpus. All these letters belong to reports of companies listed on the Madrid Stock Exchange General Index and are distributed across eight different sectors following NACE,² the industry standard classification system of the European Union (see Table 1). The classification into profit and loss was carried out with the help of an expert in the field of finance, although the results of the year are available in the annual accounts contained in each company's report.

² <https://nacev2.com/>

Business activities	Profit corpus companies	Token count	%	Loss corpus companies	Token count	%
Mining and quarrying	Repsol	2,000	3.13	Repsol	1,786	2.85
Manufacturing	Gamesa, Acerinox, Ebro Foods, Almirall, Tubacex, Pharma Mar, Ercros, PRISA, Duro Felguera, Adveo, Adolfo Domínguez	40,808	63.79	Gamesa, Acerinox, Almirall, Deoleo, Tubacex, Pharma Mar, Ercros, PRISA, Duro Felguera, Adveo, Adolfo Domínguez	37,292	59.53
Electricity, gas, steam, and air conditioning supply	Gas Natural Fenosa	1,002	1.57	Naturgy Gas Natural	1,226	1.96
Water supply, sewerage, waste management, and remediation activities	Fluidra	2,392	3.74	Fluidra	2,606	4.16
Construction	Ferrovial, FCC, Sacyr	8,649	13.52	Ferrovial, FCC, Sacyr	12,823	20.49
Accommodation and food service activities	NH	1,756	2.75	NH	1,503	2.40
Information and communication	Telefónica, Indra	5,136	8.03	Indra	3,021	4.82
Real estate activities	Realia, Urbas	2,229	3.48	Realia, Urbas	2,390	3.82
Total		63,972			62,647	

Table 1: Companies in each business activity with their token count and percentage of words in the corpus

The underlying idea of the selection was to avoid biased results due to imbalances in the distribution of sectors in the corpora. The choice was conditioned from the outset by the number of loss letters available (60 vs. 332 for earnings), so the distribution of companies in the loss corpus was emulated in the profit corpus, with preference to the letters that belonged to the same company. Although the original *FinT-esp* corpus contained letters from 15 different business sectors, only eight were chosen for this study since only eight of them had companies incurring losses.

As the corpus includes letters from different companies, the results may be affected by differences between industries. Each sector displays a specific lexicon and tone in their corporate communications. For example, a mining company may use more technical language related to exploration, excavation, and extraction. In contrast, a real estate company may use more terms related to development, leasing, and sales. Additionally, the tone may also differ depending on the sector. For example, letters from companies in

the information and communication sector may be more upbeat and optimistic, while those in the mining and quarrying sector may be more cautious and focused on risk management. Luo and Zhou (2019) found that in industries related to oil and gas, tobacco, alcohol, or firearms less optimistic tones are used in order to avoid regulatory and public scrutiny. Despite focusing on detecting differences between company letters, this study recognises a marked similarity at the baseline between both corpora. Focusing on the differences between companies may overlook the similarities—in this case, the lexical ones—which are much more pronounced, as the literature reveals. Furthermore, the fact that the corpus tool chosen for this study is designed to look for differences may lead to bias when analysing data (Taylor 2013, 2018).

Sketch Engine (Kilgarrif *et al.* 2014) was the corpus manager tool chosen to conduct the lexical analysis. The texts were uploaded and subsequently divided into two sub-corpora: profit letters, which contained 63,972 tokens distributed across 1,823 sentences, and loss letters, with 62,647 tokens distributed across 2,069 sentences.

The lists of noun and verb lemmas from both sub-corpora were drawn using the *Word List* tool of *Sketch Engine*. The analysis was limited to representative lexical units in terms of frequency. Only the lemmas that appeared in the corresponding sub-corpus with an absolute frequency equal to or higher than five were selected. Admittedly, below this frequency, a large number of verbs and nouns can be found which show the greatest variation between the sub-corpora, but their low frequency prevent their consideration as prototypical units, which is the focus of the present work.

Subsequently, the lists from both sub-corpora were compared to obtain data for the same lemmas in both groups. Finally, in the case of eventive nouns, a manual selection was conducted based on the criteria for this lexical category described for Spanish in Bosque (1999). Eventive nouns share two key distributional properties: 1) they designate entities with temporal limits, which is why they are usually combined with verbs such as *empezar* ('start'), *comenzar* ('begin'), or *concluir* ('conclude'), and 2) they take the preposition *durante* ('during') as their object.

Cases of ambiguity were found in those nouns where there is an eventive and an objectual interpretation, such as *gobierno* ('government'), *consejo* ('council'), or *negocio* ('business'). Another type of ambiguity arose when there was an eventive and resultative interpretation in nouns, such as *beneficio* ('profit'), *pérdida* ('loss'), or *ganancia* ('gain'). In the first interpretation, the referent is an event, as illustrated in (1). However, in the

second interpretation, the referent is the effect of the event designated by the verbal base of the noun, *perder* ('lose'). Disambiguation occurs within the predicates where the noun phrase is located, as in (2). Due to their thematic relevance and the increased frequency of the eventive interpretation, these ambiguous cases were not discarded.

(1) Se produjo la **pérdida** del mercado asiático.

'The loss of the Asian market occurred.'

(2) Tuvo unas **pérdidas** de 61,9 millones de euros.

'It made a loss of 61.9 million euros.'

Following this process, a total of 345 verbs and 165 eventive nouns were retrieved, which are analysed in Section 4 with regard to their frequency and collocations. For the frequency comparison between the sub-corpora, the normalised frequency per thousand tokens has been adopted. Moreover, in cases where a particular unit occurs in both sub-corpora, the log-likelihood ratio has been calculated with Paul Rayson's UCREL log-likelihood calculator³ to determine whether the difference in frequencies is statistically significant. All those units that do not exceed 3.84 ($p < 0.05$) were discarded (McEnery and Xiao 2005).

4. RESULTS AND ANALYSIS

The analysis of the data is divided into two parts. First, the differences between the two corpora regarding the frequency of verbs and eventive nouns are examined. The focus is primarily on those verbs and nouns that may be key to differentiate between the two groups of letters according to their frequency of use: firstly, those that do not appear in one of the two corpora and, secondly, those in both groups of letters, but in different proportions. On the other hand, the second part of the analysis deals with collocations. The purpose is to determine whether the more immediate context of verbs and nouns is relevant in those cases where the frequency of use does not allow significant differences to be drawn.

³ <https://ucrel.lancs.ac.uk/llwizard.html>

4.1. Frequency analysis of verbs

A total 345 verbs were extracted, all used with a minimum absolute frequency of five in at least one of the two corpora. The data reveal that both corpora of letters use practically the same verbs. Only seven verbs which appear in the gain corpus are not present in the loss corpus: *colaborar* ('collaborate'), *comportar* ('involve'), *estabilizar* ('stabilise'), *moderar* ('reduce'), *optar* ('opt for'), *propiciar* ('favour'), and *sustituir* ('replace'). In contrast, eight verbs from the loss corpus do not appear in the gain corpus: *centralizar* ('centralise'), *coincidir* ('coincide'), *combinar* ('combine'), *detallar* ('detail'), *indicar* ('show'), *pertenecer* ('belong'), *poseer* ('own'), and *relanzar* ('relaunch'). Thus, there is an overlap of 95.7 per cent between the sets of verbs in both corpora. This lexical similarity is one of the main reasons why it is challenging to classify the letters into two categories of companies, especially automatically.

In the following, there is a discussion on the use of verbs relevant for the differentiation between the two corpora. The term 'relevant' here implies lexical units used exclusively or almost exclusively by one of the two groups of companies, because of their semantic content and collocations. Therefore, if these lexical units appeared in a particular letter and we did not know what group they belonged to, they would be decisive in determining whether it is one group or the other. As can be seen, although the first selection is statistical—the focus is on units that do not occur in the other sub-corpus or with a difference in use statistically significant—there is a second selection of a subjective nature carried out after an in-depth examination of the letters. Therefore, all those verbs for which there is no further discussion are not to be considered relevant for the differentiation of both corpora.

4.1.1. Verbs exclusive to profit and loss-making companies

Three verbs that are only attested in one of the two groups of letters are analysed below. Their frequencies are shown in Table 2. In the following subsections, they are sorted alphabetically.

Verbs	Normalised frequency ⁴ in profit corpus	Normalised frequency in loss corpus
<i>Optar</i> ‘opt for’	0.11	0
<i>Propiciar</i> ‘favour’	0.13	0
<i>Relanzar</i> ‘relaunch’	0	0.08

Table 2: Frequency of selected verbs exclusive to one corpus

4.1.1.1. *Optar* (‘opt for’)

Optar is frequently used by profit-making companies because of its prepositional regime complements introduced by *por* (‘by’), such as *pago* (‘payment’), *modalidad* (‘modality’), or *efectivo* (‘cash’), as in example (3), which are related to shareholders’ decisions about the dividend they will receive. It is worth noting that it is not common for loss-making companies to pay out dividends.

- (3) Nos permitió hacer inversiones por 156 millones de euros, atender el pago del dividendo de aquellos accionistas que **optaron** por efectivo en el “scrip dividend” del ejercicio (...). (Acerinox, Chairman, 2016)

‘It allowed us to make investments of 156 million euros to pay the dividend to those shareholders who opted for cash in the scrip dividend for the year (...).’

4.1.1.2. *Propiciar* (‘favour’)

Propiciar is a causative verb that usually implies a positive connotation. Profit-making companies sometimes use it to disguise their direct involvement in the events described (because its subject usually denotes events and not entities) while, at the same time, making it clear that these events have occurred thanks to the company’s good policies, as in example (4). At other times, *propiciar* describes situations in which the behaviour of some aspect of the market has led to another event considered positive. The former is encoded by the syntactic subject and other sentence elements that usually refer to the agent or cause, such as *evolución* (‘evolution’) or *materias primas* (‘raw materials’). The latter is encoded by direct objects in active sentences and passive subjects, such as *meta* (‘goal’), *desarrollo* (‘development’), *aumento* (‘increase’), *subida* (‘rise’), or *mejora* (‘improvement’).

⁴ Normalised frequency per thousand tokens.

- (4) La consecución de nuestro Plan Estratégico 2018 con año y medio de antelación, las inversiones realizadas a lo largo del año y el crecimiento de nuestra actividad en los mercados maduros **han propiciado** un crecimiento del valor. (Fluidra, Chairman, 2017)

‘The achievement of our 2018 Strategic Plan a year and a half in advance, the investments made throughout the year, and the growth of our activity in mature markets have led to a growth in value.’

4.1.1.3. *Relanzar* (‘relaunch’)

Loss-making firms use *relanzar* when a previously promoted action, initiative, or project could not be carried out or was not as successful as expected. Direct objects include *negocio* (‘business’), as in example (5), *proyecto* (‘project’), *crecimiento* (‘growth’), or even *compañía* (‘company’).

- (5) A partir de ahora trabajaremos para **relanzar** el negocio y lograr una compañía cercana a la producción y orientada al consumidor (...). (Deoleo, CEO, 2016)

‘From now on, we will work to relaunch the business and achieve a company close to production and consumer-oriented [...].’

4.1.2. Verbs with a higher frequency in profit and loss-making companies

Three verbs with a statistically significant higher frequency in one of the two groups of letters are analysed below. Table 3 displays their frequencies. In the following subsections, they are arranged in alphabetical order.

Verbs	Normalised frequency in profit corpus	Normalised frequency in loss corpus	Log-likelihood ratio
<i>Aumentar</i> ‘increase’	0.70	0.37	6.79
<i>Lograr</i> ‘achieve’	0.23	0.73	17.17
<i>Registrar</i> ‘record’	0.23	0.45	4.27

Table 3: Frequency and log-likelihood ratio of selected verbs appearing in both corpora

4.1.2.1. *Aumentar* (‘increase’)

Aumentar is an interesting verb in terms of its distributional properties since, by itself, it cannot disambiguate the two groups of letters. In the profit letters, it occurs typically in past tenses and with direct objects having to do with quantitative aspects of the business, such as *producción* (‘production’), *exportación* (‘export’), or *consumo* (‘consumption’), as well as with direct objects related to qualitative aspects of the business such as

valoración ('valuation'), as seen in example (6), *confianza* ('trust'), *esfuerzo* ('effort'), and *valor* ('value'). The subjects of this verb in the profit letters are typically *beneficio* ('profit'), *ebit*, or *ebitda* (two different financial profit indicators). As shown, *aumentar* is generally associated with indicators that have improved in the fiscal year.

- (6) La propuesta de valor dirigida a cumplir y superar las expectativas de los consumidores está permitiendo que **aumente** la valoración sobre los hoteles. (NH, CEO, 2015)

'The value proposition aimed at meeting and exceeding consumer expectations is allowing the valuation of hotels to increase.'

4.1.2.2. *Lograr* ('achieve')

The essential complements for this verb are the direct objects. Loss-making companies may aim at *equilibrio* ('break-even') or *éxito* ('success') but not at *ahorro* ('savings') or *contratación* ('recruitment'), as they typically prioritise more pressing objectives, such as *saneamiento* ('company restructuring'), as in example (7), *renovación* ('renewal'), *recuperación* ('turnaround'), *estabilidad* ('stability'), *eficiencia* ('efficiency'), *reducción de costes* ('cost reduction'), *endeudamiento equilibrado* ('balanced indebtedness'), or *mejora* ('improvement'), among other essential aspects. These concerns would explain the frequency of *lograr* in the loss corpus, as these companies engage with their shareholders for more purposes than the profit-making companies because they need to convince them to make a future improvement in their performance.

- (7) La contención del gasto debe seguir siendo un pilar de gestión ineludible en el Grupo FCC para **lograr** el pleno saneamiento de la compañía, a través de iniciativas alineadas con el ajuste y la reducción presupuestaria. (FCC, CEO, 2016)

'Cost containment must remain an unavoidable management pillar in the FCC Group in order to achieve the full restructuring of the company through initiatives aligned with budget adjustment and reduction.'

The direct objects of *lograr* in the profit letters are most commonly *ahorro* ('savings'), as in example (8), *contratación* ('recruitment'), *equilibrio* ('balance'), and *éxito* ('success'), among others.

- (8) Hemos conseguido el 49 % de nuestro objetivo de **lograr** un ahorro recurrente anual de 50 millones de euros [...]. (Acerinox, CEO, 2017)

'We have achieved 49 % of our target to achieve annual recurring savings of 50 million euros [...].'

4.1.2.3. *Registrar* ('record')

Registrar usually occurs in the loss corpus combined with *pérdida* ('loss'), as in example (9), *caída* ('fall'), or *reducción* ('reduction') as its direct objects. However, in the profit corpus, it also appears with nouns of positive connotation as direct objects, such as *crecimiento* ('growth'), *evolución* ('evolution'), or *beneficio* ('profit').

- (9) El resultado atribuible en 2016 **registró** una pérdida neta de -162 millones de euros (...). (FCC, CEO, 2016)

'The attributable result in 2016 recorded a net loss of -162 million euros [...].'

4.1.3. Further discussion

Even though there is an overwhelming similarity in lexical choice, some minor differences have been found as regards the verbs used in the two corpora of letters. The most relevant verbs in profit letters reflect typical activities of profitable companies. For example, there is talk of how *las inversiones propician un crecimiento del valor* ('investments lead to value growth'). Additionally, shareholders can *optar por el pago* ('opt for payment'), meaning they can receive dividends. The letters also mention improvements in performance indicators and business margins. For example, the following events are discussed: *aumentar la producción / la exportación / el consumo / la valoración / la confianza* ('increase production / export / consumption / valuation / confidence'), as well as *aumentar el beneficio / ebit / ebitda* ('increase profit / ebit / ebitda'), or *registrar un crecimiento / evolución / beneficio* ('register growth / evolution / profit').

On the other hand, in letters from loss-making companies, the most frequent verbs refer to the closing of the fiscal year without profits or refer to the objectives they promise to fulfil in the future. They use *relanzar el proyecto / el negocio / el crecimiento / la compañía* ('relaunch the project / business / growth / company') and talk about *lograr equilibrio / éxito / saneamiento / estabilidad / eficiencia / mejora* ('achieve (balance / success / sanitation / stability / efficiency / improvement)'), or *registrar pérdidas* ('register losses'). Specific indicators such as *ebit* or *ebitda* are not referenced.

4.2. Verb disambiguation through collocations

As seen in Section 4.1., the gain and loss corpora share 95.7 per cent of the verbs with a minimum absolute frequency of five. Therefore, one must resort to syntactic collocations to find clear differences between both corpora. To this purpose, the 50 most frequent verbs were selected in the two corpora (see Appendix 1) and their distributional properties were studied. In this section, three types of collocations need to be distinguished:

1. Collocations which are exclusive to one corpus and not semantically equivalent to the collocations in the other corpus. In these cases, the verb or noun in question is combined in each corpus with different lexical units whose meanings are also markedly different. For example, in the case of direct objects, one may note the difference between *mantener la exigencia* ('keep (the) demand') in the profit letters and *mantener la rentabilidad* ('keep (the) profitability') in the loss letters.
2. Those which are exclusive to one corpus, but semantically equivalent or close to the collocations in the other corpus. Different lexical units are used in each corpus, but their meaning is similar; that is, verbs and eventive nouns appear in similar contexts. For example, the collocation of *mejorar* ('improve') with *considerablemente* ('considerably') in the profit letters, and with *sustancialmente* ('substantially') in the loss letters, both adverbial modifiers of gradation or intensity with similar meanings.
3. Those which are identical in both corpora. For example, the adverb *intensamente* ('intensely') modifies *trabajar* ('work') in both corpora.

For our purpose, type (1) collocations are the only ones suitable to discern between the two groups of companies. The following is a description of their combinations, grouped according to their syntactic functions. The verbs are listed in alphabetical order.

4.2.1. *Alcanzar* ('reach')

Alcanzar combines with direct objects referring to scalar attributes and their values, which are vital to differentiate between the two groups of letters. In the case of the profitable companies, *récord* ('record'), *competitividad* ('competitiveness'), *beneficio* ('profit'), or *tasa* ('rate') are used, alluding to the significant advances (high or maximum values of

the respective dimensions) that have been achieved. By contrast, these values are usually lower in loss letters: *alcanzar el estándar* ('reaching the standard') or *alcanzar la expectativa* ('reaching the expectation').

4.2.2. *Conseguir* ('achieve')

This verb is usually attested in different syntactic structures in both types of letters. In the profit letters, *conseguir* is used with direct nominal objects referring to different profit indicators, such as *cota de excelencia* ('excellence benchmark'), *ebitda*, *eficiencia* ('efficiency'), or *crecimiento* ('growth'). Although loss-making companies can point to their *ebitda*, this term will never be the direct object of *conseguir* in loss letters, as it is a verb with positive connotations and this indicator will be a negative figure. In the letters of loss-making companies, *conseguir* is usually followed by infinitive objects (in the terminology of *Sketch Engine*), such as *batir la tendencia general del mercado* ('to beat the market general trend') and *volver a la rentabilidad* ('to return to profitability').

4.2.3. *Marcar* ('mark')

Marcar is used in the profit corpus with complements referring to the positive events that have taken place during the financial year, although it is used with two different meanings. Taking the *Diccionario de la Lengua Española* (DLE 2022) as a reference, *marcar* is used in the sense 'prescribe, determine, or fix' with direct objects such as *cumplimiento* ('compliance') or with subjects in passive sentences such as *previsión* ('forecast'). On the other hand, the ninth sense found in the DLE ('divide spaces, with milestones or signs of any kind, or divide them mentally') reflects the use of this verb when combined with nouns such as *hito* ('milestone'), where this lexical unit acts as a subject in a passive structure.

Likewise, loss-making companies use *marcar* with this last meaning. It combines with direct objects such as *camino* ('path'), *línea* ('line'), or *diferencia* ('difference'). Having failed to meet targets, companies argue that the year has been a turning point.

4.2.4. *Seguir* ('continue')

This verb is the head of the durative periphrasis *seguir* ('keep') + gerund. In the profit letters, the main verb refers to processes (often incremental) with a positive connotation, as in *cosechando* ('earning'), *amortizando* ('amortising'), *evolucionando* ('evolving'), *enriqueciendo* ('enriching'), *prosperando* ('prospering'), *progresando* ('progressing'), *perfeccionando* ('perfecting'), *compitiendo* ('competing'), *consolidando* ('consolidating'), *afianzando* ('strengthening'), or *incrementando* ('increasing'). In the letters from loss-making companies, the main verb of the periphrasis alludes to difficulties that are being overcome or must be overcome, as in *fortaleciéndonos* ('strengthening (ourselves)') and *viviendo* ('living'), to indicate that challenges have marked the year, or *ajustando* ('adjusting'), which refers to costs.

4.2.5. *Superar* ('exceed')

In the sense 'exceed a limit', *superar* resembles *alcanzar* ('reach') in that it often combines with direct objects referring to scalar attributes and their values. In the profit letters, it appears with direct objects such as *récord* ('record'), *índice* ('index'), *cifra* ('figure'), or *previsión* ('forecast'), which denote specific limits. In the loss letters, it often serves a different meaning, that is, to 'overcome obstacles or difficulties.' In these letters, it is combined with objects such as *año* ('year'), *crisis* ('crisis'), and *dificultad* ('difficulty'), referring to the intricacies of the year.

4.2.6. Further discussion

Some conclusions can be made drawn from what has been discussed in the previous subsections as regards the collocations of verbs in both corpora. The most relevant syntactic function for distinguishing between letters from profitable companies and those with losses is the direct object, which is used differently in both groups of letters. On the one hand, profitable companies use direct objects related to qualitative milestones (*cota de excelencia* 'excellence threshold', *hito* 'milestone', *liderazgo* 'leadership') and quantitative business issues (*records* 'records', *índices* 'indices'), suggesting a concern for excellence and growth in different business indicators.

On the other hand, loss-making companies tend to mention qualitative milestones, as the quantitative ones are unfavourable — they have not achieved profitability. For example, they use *alcanzar el estándar / las expectativas* ‘reach the standard / expectations’, *conseguir la vuelta a la rentabilidad* ‘achieve the return to profitability’, *marcar la diferencia* ‘make a difference’ (in their sector), *seguir fortaleciéndonos* ‘continue to strengthen (ourselves)’, *superar una dificultad* ‘overcome a difficulty’. They focus on the possibility of change and improvement, while profitable companies talk about the opportunity for growth and expansion.

4.3. Frequency analysis of eventive nouns

A total of 165 eventive nouns were detected in the list of noun lemmas with a minimum absolute frequency of five. The data reveal that practically the same nouns are used in both corpora. There are only three nouns exclusive to the profit letters (*globalización* ‘globalisation’, *protección* ‘protection’, and *laminación* ‘lamination’) and four nouns exclusive to the loss letters (*fraude* ‘fraud’, *negociación* ‘negotiation’, *recogida*, ‘collection’, and *revalorización* ‘revaluation’), representing an overlap of 95.8 per cent between the nouns of both corpora. In the following, the distinctive nouns of each corpus are included, as well as their relevance in differentiating the two types of companies.

As explained in Section 4.1, the criterion for considering a unit as distinctive is primarily statistical: they are only attested in one the corpora or are more frequent in one of them, and this difference in frequency is statistically significant. In addition, there is an eminently subjective criterion; after reading and studying the letters, only those which, due to their semantic content, tip the balance in favour of one group or the other are selected. For example, *recogida* (‘collection’) only appears in the loss letters, but it refers to *waste collection*, so that it cannot be a relevant unit in the loss corpus since, in the profit group, there are also manufacturing companies that may refer to waste recycling. This consideration is not absolute; instead, there are units that may attested in both corpora but combine with other items that can be used to disambiguate between letters.

4.3.1. Nouns exclusive to profit and loss-making companies

Nouns that are only found in one of the two groups of letters are discussed below. Their frequencies are shown in Table 4. In the following subsections, they are sorted in alphabetical order.

Nouns	Normalised frequency in profit corpus	Normalised frequency in loss corpus
<i>Globalización</i> ‘globalisation’	0.08	0
<i>Negociación</i> ‘negotiation’	0	0.13
<i>Protección</i> ‘protection’	0.09	0
<i>Revalorización</i> ‘revaluation’	0	0.08

Table 4: Frequency of selected nouns exclusive to one corpus

4.3.1.1. *Globalización* (‘globalisation’)

This noun is used with a negative connotation in the profit letters. Companies are concerned about the globalisation of markets, which is perceived as a threat that companies must face, as illustrated in example (10).

- (10) Ninguna otra empresa productora de acero inoxidable en el mundo goza de tan buena posición para afrontar el difícil proceso de **globalización** de la economía, en su nueva versión de “**globalización** con barreras” (...). (Acerinox, CEO 2016)
 ‘No other stainless steel producing company in the world enjoys such a good position to face the difficult process of globalisation of the economy, in its new version of “globalisation with barriers” [...].’

4.3.1.2. *Negociación* (‘negotiation’)

The chairmen or chairwomen and CEOs of loss-making companies report the negotiations with banks to define debt interest rates. Therefore, *negociación* can be combined with *banca* (‘banking’) as the object of the preposition *con* (‘with’), or with *refinanciación* (‘refinancing’), the direct object (Theme) of the base verb *negociar* (‘negotiate’), as in example (11).

- (11) Este documento (...) es la base sobre la que se está apoyando la **negociación** con la banca cuyo primer objetivo es cerrar un “Acuerdo de Espera” o (“Standstill”) para alcanzar posteriormente una reestructuración de la deuda. (Duro Felguera, Chairman, 2016)

‘This document [...] constitutes the basis for the negotiations with the banks. Its first objective is to conclude a ‘standstill’ in order to reach a debt restructuring subsequently.’

4.3.1.3. *Protección* (‘protection’)

Profit-making companies refer to the protection of three areas: investment (*protección de los intereses de los accionistas* ‘protection of shareholders’ interests’), environment (*protección del entorno ambiental* ‘protection of the environment’), and labour (*protección y seguridad de su personal expatriado* ‘protection and safety of their expatriate staff’).

4.3.1.4. *Revalorización* (‘revaluation’)

Revalorización implies that an asset or security is at low levels at the time of speech. Letters from loss-making companies consider it relevant to comment on this (possible) increase in the value of their company’s stock, as illustrated in example (12). This rise is good news, implying that investors are confident in the business performance. Profit-making companies with good results do not need to appeal to investors’ confidence in this manner.

- (12) Estamos convencidos de que los mercados no están reflejando adecuadamente el valor intrínseco de nuestro valor y que existe un importante potencial de **revalorización** de nuestra acción. (Acerinox, Chairman, 2018)

‘We are convinced that the markets are not adequately reflecting the intrinsic value of our stock and that there is significant upside potential for our share.’

4.3.2. Nouns with a higher frequency in profit and loss-making companies

The following is an analysis of eventive nouns that occur with a statistically significant higher frequency in one of the two corpora of letters. Table 5 displays their frequencies. In the subsections below, they are sorted alphabetically.

Nouns	Normalised frequencies in profit corpus	Normalised frequencies in loss corpus	Log-likelihood ratio
<i>Ahorro</i> ‘saving’	0.17	0.37	4.58
<i>Ampliación</i> ‘expansion’	0.22	0.61	12.02
<i>Cumplimiento</i> ‘compliance’	0.42	0.19	5.61
<i>Organización</i> ‘organisation’	0.17	0.35	3.97
<i>Pérdida</i> ‘loss’	0.22	0.49	6.95
<i>Reestructuración</i> ‘restructuring’	0.22	0.49	6.95

Table 5: Frequency and log-likelihood ratio of selected nouns appearing in both corpora

4.3.2.1. *Ahorro* (‘saving’)

Saving costs and funds are a critical factor for loss-making companies, which encourage *medidas de ahorro* (‘savings measures’), *acciones de ahorro* (‘savings actions’), or *planes de ahorro* (‘savings plans’). For these companies, it is an obligation while for profitable companies, it is a supplement or an additional merit of their management. Therefore, instead of acting as a complement of nouns referring to different types of projects, *ahorro* usually appears as a noun head modified by the adjective in profit letters, as, for example, *ahorro adicional / recurrente* (‘additional / recurrent savings’). It also functions as a direct object of verbs, such as *destacar* (‘highlight’) and *lograr* (‘achieve’), with positive connotations.

4.3.2.2. *Ampliación* (‘expansion’)

Ampliaciones de capital (‘capital increases’) are frequently referred to in the loss letters. They are operations aimed at increasing the company’s resources, which can be done by increasing the number of shares or their nominal value, that is, the value assigned by the owner. Although a profit-making company may also carry out capital increases to make investments, capital increases are typical of loss-making companies, as they allow them to meet their debts. Its occurrence should therefore tip the balance in favour of loss-making. When *ampliación* is modified by adjectives such as *segunda* (‘second’) or *última* (‘last’), it becomes more evident that it belongs to the loss corpus. In the case of gains, texts mention the *ampliación de la capacidad* (‘expansion of capacity’), referring to the increase in plant output in the case of manufacturing companies.

4.3.2.3. *Cumplimiento* ('compliance')

Profit-making companies refer to all the commitments and obligations they have managed to meet, be their environmental or governance policies, measures, laws, or business plans, as illustrated in example (13). There is even discussion of a 'culture of compliance' (*cultura del cumplimiento*), of which, for obvious reasons, loss-making companies cannot be a part of.

- (13) Dará como resultado un Grupo más fuerte, más innovador, más competitivo y más comprometido con los valores tradicionales de nuestra compañía: la prudencia, la austeridad, la calidad y el **cumplimiento** de los compromisos asumidos. (Sacyr, Chairman, 2016)

'It will result in a stronger, more innovative, more competitive Group that is more committed to the traditional values of our company: prudence, austerity, quality and compliance with our commitments.'

4.3.2.4. *Organización* ('organisation')

Loss letters refer to the different organisations they support and provide them with the necessary social prestige. Note, for example, (14), where there is an ambiguity between the objectual and eventive interpretations. Loss letters can also allude to the need for efficiency, as illustrated in (15). By contrast, profit letters praise their *cultura de organización* ('organisational culture').

- (14) Fluidra ha reforzado significativamente su **organización** mundial de I+D. (Fluidra, Chairman, 2018)

'Fluidra has significantly strengthened its worldwide R&D organisation.'

- (15) En paralelo a estas medidas se ha definido una **organización** más eficiente (Deoleo, Chairman, 2016)

'In parallel to these measures a more efficient organisation has been defined'.

4.3.2.5. *Pérdida* ('loss')

Unsurprisingly, *pérdidas* ('losses') are mentioned much more frequently in the loss letters than in the profit letters. These include *pérdidas contables* ('accounting losses'), *pérdidas operativas* ('operating losses'), *pérdidas de margen* ('margin losses'), and *pérdidas de volumen* ('volume losses'), among others. The profit letters, on the other hand, point out

losses *en las ventas* ('on sales'), and highlight gains compared to losses in previous years or the value of the shares in falling markets, as illustrated in example (16).

(16) Se revalorizó en un 33,9% durante el ejercicio frente a la **pérdida** del 2,0% que sufrió el IBEX-35. (Acerinox, CEO, 2016)

'It was revalued by 33.9% during the year compared to the 2.0% loss suffered by the IBEX-35.'

4.3.2.6. *Reestructuración* ('restructuring')

In the loss letters, *reestructuración* is a complement of nouns referring to a process, such as *actuación* ('performance') or a part of it, *culminación* ('completion'), and to other nouns, such as *esfuerzo* ('effort'), implying that it is a change forced by the situation. In loss-making companies, in turn, it also combines with *industrial* ('industrial'), *logística* ('logistical'), a classificatory relational adjective, *organizativa* ('organisational', where *organizativa* refers to the subject of the base verb) or *de plantilla* ('workforce'), among others. Profit-making companies use these as well. However, unlike the loss-making companies, they refer to a 'shareholding restructuring' (*reestructuración accionarial*) or the restructuring of a 'business division' (*división del negocio*), in specific company areas.

4.3.3. Further discussion

The previous analysis revealed that the distinctive eventive nouns of each corpus depict activities typical of their group. Specific nouns, including *negociación* ('negotiation') and *revalorización* ('revaluation'), exhibit significant differences between the two groups. Loss-making companies use them to refer to their debt management and the possibility of increasing the value of their shares in the market. These companies also focus on the actions they must take to achieve profits in the future (*planes de ahorro* 'savings plans', *ampliación de capital* 'capital expansion', *reestructuración logística / organizativa* '(logistics / organisational restructuring)'). On the other hand, profitable companies focus on achieving previously established growth and objectives (*cumplimiento* 'compliance'). Additionally, corporate social responsibility is an essential topic for both, as it can affect their social prestige and, therefore, their image before shareholders and the general public. In the case of loss-making companies, their letters reference the *organizaciones* ('organisations') they support.

4.4. Noun disambiguation through collocations

As discussed in Section 4.3, the profit and loss corpora share 95.8 per cent of the eventive nouns with a minimum absolute frequency of five in at least one corpus. This overlap means that, in most cases, to distinguish between the gain and the loss letters, one must resort to collocations. To this end, as in the case of verbs, the 50 most frequent eventive nouns in each corpus were selected (see Appendix 2).

As in the case of verbs (see Section 4.2), the focus will be placed on the collocations of type (1), the most suitable to differentiate between both groups of letters. It is worth recalling that type (1) collocations are exclusive to one corpus and bear no semantic resemblance to those of the other corpus. The nouns are presented below in alphabetical order.

4.4.1. *Acuerdo* ('agreement')

The profit-making companies use adverbial adjectives such as *definitivo* ('definitive') and evaluative adjectives such as *oportuno* ('timely') to define the agreements they sign. They also close *acuerdos de fusión* ('merger deals'), which can be perceived as a business growth operation. Loss-making companies use evaluative qualifying adjectives with a negative connotation to define the agreements they sign, such as *malo* ('bad').

4.4.2. *Aumento* ('increase')

Aumento ('increase') is a noun that introduces gradual events. Profit letters speak of *aumento significativo del volumen de inversión* ('significant increase in the company's investment volume') or *aumento de la productividad de la compañía* ('increase in the company's productivity') and, in other cases, allude to *aumento salarial* ('wage increase'). Another indication of earnings is *aumento de dividendos* ('increase in dividends'). It is an action that can be considered unequivocal of earnings because their distribution is not mandatory and their rise less so. To a lesser extent, *aumento de las ventas o de la demanda* ('increase in sales or demand') can also be distinctive of profit. In all these cases, the preposition object is the Theme argument of *aumento*.

4.4.3. *Comportamiento* ('behaviour')

The adjective modifiers in the profit letters make it clear that the behaviour has been good in its different aspects. *Comportamiento ético* ('ethical behaviour') is defined as *impecable* ('impeccable'), *estrategia de comunicación* ('communication strategy') as *intachable* ('faultless'), *nuevas gamas* ('new ranges') as *extraordinarias* ('outstanding'), and the markets in which the company participates as *excelentes* ('excellent'). The implementation of a *guía de comportamiento íntegro* ('code of integrity') is also mentioned. By contrast, loss letters prefer to use the comparative adjective *peor* ('worse') to assess the company's performance, as in (17).

- (17) Acerinox, que había evolucionado en línea con el Ibex 35 durante el primer semestre, tuvo un peor **comportamiento** que este índice en el cuarto trimestre (Acerinox, Chairman, 2018).

'Acerinox, which had evolved in line with the Ibex 35 during the year's first half, performed worse than this index in the fourth quarter.'

4.4.4. *Compromiso* ('commitment')

To define their commitment, profitable companies use evaluative adjectives that denote the highest degree of a property, such as *ejemplar* ('model'), *altísimo* ('very high'), and *mayor* ('higher'), the latter used to compare the current year with past years. They also employ adjectives such as *constante* ('constant'). On the other hand, the loss-making companies avoid qualifying adjectives and use relational adjectives such as *público* ('public') or adverbial adjectives such as *mutuo* ('mutual'), referring to the commitment between the companies and the employees of the group.

4.4.5. *Crecimiento* ('growth')

In the profit letters, the modifiers of *crecimiento* are adjectives of degree and intensity, used as evaluative and adverbial modifiers, such as *fuerte* ('strong'), *relevante* ('relevant'), *mayor* ('major'), *potente* ('powerful'), *importante* ('important'), *significativo* ('significant'), *mejor* ('better'), *notable* ('remarkable'), *continuo* ('continuous'), *sostenible* ('sustainable'), *rentable* ('profitable'), *sólido* ('solid'), *firme* ('steady'), or *exponencial* ('exponential'), among others.

Loss letters use specific adjective modifiers. They judge their growth as *desigual* ('uneven') and promise to meet more *ambiciosos* ('ambitious') growth rates. Also, *crecimiento* ('growth') appears as a direct object of verbs such as *relanzar* ('relaunch'), implying that growth has not occurred in past years.

4.4.6. *Ejercicio* ('year')

The adjectival modifiers of *ejercicio* are different in each corpus. In the profit letters, it is combined with evaluative qualifying adjectives such as *bueno* ('good'). By contrast, in the loss letters, it occurs together with the adjectives *arriesgado* ('risky'), *decisivo* ('decisive'), or *complejo* ('complex'), among others.

4.4.7. *Evolución* ('evolution')

The profitable companies describe *evolución de las ventas* ('sales evolution') with qualifying adjectives such as *alcista* ('bullish'), with a descriptive interpretation, and *espectacular* ('spectacular'), with an evaluative interpretation. On the other hand, the loss letters use evaluative modifiers to define the *evolución* of different aspects of their business as *desfavorable* ('unfavourable') and *peor* ('worse').

4.4.8. *Inversión* ('investment')

The profit letters refer to 'investments' *arranque* ('beginning') or *fomento* ('promotion'), with investment playing the semantic role of Theme (effected in the first case and affected in the second). The quantitative dimension of investments is introduced by the phrases *grado de inversión* ('investment degree'), *importe de inversión* ('investment amount'), and *volumen de inversión* ('investment volume'). Companies with losses mention their *reducción* ('reduction'), *contracción* ('contraction'), or *parálisis* ('stoppage') as part of nominal groups in which *inversión* is assigned the role of affected Theme.

4.4.9. *Mejora* ('improvement')

The profitable firms use evaluative adjectives such as *grandes* ('great') to refer to the technological improvements (*mejoras tecnológicas*) they have made and use the adverbial

adjective *constante* ('constant') to define them. By contrast, the loss-making companies prefer to use adjectives such as *paulatinas* ('gradual') or *inminentes* ('imminent').

4.4.10. *Proceso* ('process')

The profitable companies mention *procesos de producción* ('production processes'). In the loss letters, *proceso* is combined with prepositional complements with nouns such as *refinanciación* ('refinancing') and *desinversión* ('disinvestment'), which are symptoms of losses in a business.

4.4.11. *Reducción* ('reduction')

Like *aumento*, the noun *reducción* introduces gradual events. In the profit letters, it is accompanied by argument Themes such as *accidentes* ('accidents') or *accidentabilidad* ('accidentability'), referring to production plants, and *emisiones de carbono* ('carbon emissions') or *emisiones de gases de efecto invernadero* ('greenhouse gas emissions'). In the letters of loss-making companies, its arguments include *producción* ('production'), *capital* ('capital'), or *inversiones* ('investments'). In addition, loss-making companies refer to *reducción presupuestaria* ('budget reduction'). Also, *reducción* is part of the prepositional complements of nouns such as *esfuerzo* ('effort'), *programa* ('programme'), *política* ('policy'), or *cultura* ('culture'), which are expressions that imply that the company is committed to reducing costs and debt.

4.4.12. *Result* ('resultado')

Adjective modifiers are vital for the disambiguation of *resultado*, especially those that are evaluative and qualifying and denote a high or maximum degree of a property. Profit letters assess the last fiscal year's results as *positivos* ('positive'), *buenos* ('good'), *excelentes* ('excellent'), *sólidos* ('solid'), *magníficos* ('great'), *extraordinarios* ('extraordinary'), *significativos* ('significant'), *grandes* ('big'), or as *los mejores* 'the best'. By contrast, in the loss letters, they are *inferiores* ('lower').

4.4.13. Further discussion

The analysis of collocations showed that adjectival modifiers, prepositional complements introduced by *de* and head of noun groups are the most important syntactic functions to distinguish between both corpora of letters. Moreover, profitable companies use evaluative adjectives with positive connotations to describe their performance (*crecimiento significativo* ‘significant growth’, *mejoras grandes* ‘great improvements’, *resultados excelentes* ‘excellent results’). In contrast, loss-making companies use evaluative adjectives with negative connotations (*malos acuerdos* ‘bad deals’, *peor comportamiento* ‘worse performance’, *evolución desfavorable* ‘unfavourable development’, *resultados inferiores* ‘inferior results’). Concerning other types of collocations, profitable companies discuss increases in various positive areas of business performance (*aumento del volumen de la inversión / de la productividad / de dividendos / salarial*, ‘increase in investment volume / productivity / dividends / wages’), while loss-making companies talk about downsizing and *procesos de desinversión / refinanciación* (‘disinvestment/ financing processes’).

5. CONCLUSION

This paper has analysed a selection of one hundred letters to shareholders written by companies listed on the Madrid Stock Exchange General Index between 2014–2018. The aim was to find lexical clues that would allow to differentiate between letters from profitable and letters from loss-making companies.

However, it is worth noticing that the present study has some limitations and contextualise the results accurately. First, the corpus size may not represent all letters to shareholders listed on the Madrid Stock Exchange General Index. Secondly, the study may have mainly focused on the differences between the corpora, neglecting the similarities that were clear from the outset. Additionally, the data selection and discussion methodology, which included a mixture of quantitative and subjective analysis, could have played a role on the results. Finally, focusing on two specific grammatical categories may not fully capture the complexity of the language used in these letters.

The initial hypotheses pointed to a lexical similarity in both corpora because the letters from the loss-making companies tended to mimic the profit-making companies’ themes and causality patterns. These hypotheses have been confirmed, as the lexical

choice in both corpora is similar. In the grammatical categories analysed in this paper, that is, verbs and eventive nouns, there is an overlap of 95.7 per cent and 95.8 per cent, respectively.

The characteristic verbs that profitable companies make use of are those that report the increase in business margins or the distribution of dividends (*optar* ‘opt for’, *propiciar* ‘propitiate’, *aumentar* ‘increase’). On the other hand, those related to closing the fiscal year without profits, meeting their objectives, or promising to fulfil them in the future (*relanzar* ‘relaunch’, *registrar* ‘register’) are characteristic of the loss-making companies.

As far as the nouns are concerned, those with a higher frequency in profitable companies refer to the achievement of the objectives proposed in previous years (*cumplimiento* ‘compliance’). Similarly, the nouns with significantly more frequency in the loss-making companies refer to the losses themselves or the actions the company must take concerning products and business management to overcome those poor results: *ahorro* (‘savings’), *ampliación* (‘expansion’), *reestructuración* (‘restructuring’), *organización* (‘organisation’), etc.

Furthermore, in the analysis of the distributional properties of verbs and eventive nouns, the study revealed that the syntactic element that showed the most prominent differences between the two groups is the direct object, rather than the subject or other arguments. The most noticeable difference in the collocations is that profitable companies can talk about qualitative achievements (the company is a leader in the sector, has proved to be a company that delivers, has the confidence of shareholders, and has managed well) and quantitative achievements (business margins are positive). By contrast, the loss-making companies can only mention qualitative achievements (e.g., the company’s leading position) and negative quantitative data.

In the case of eventive nouns, the critical syntactic functions in the distinction between the letters are the adjectival modifiers, especially qualifying and adverbial, prepositional complements introduced by *de* (usually argumental with the role of Theme), and head in those cases where the noun in question acts as a complement to other nouns.

In short, despite the similarity between the two corpora, there are lexical signs that allow us to differentiate between them. However, such differences are very subtle because the linguistic reality of the letters is very complex. In fact, an examination carried out solely at the lexical or even morphosyntactic level will likely not reveal any real

significant differences. Therefore, it would be more productive to analyse the structure and argumentation patterns of the letters in depth, as well as the data that the companies choose to omit. These are central aspects for our future work, which will help us to understand the genre of letters to shareholders in Spanish better.

REFERENCES

- Aerts, Walter. 2005. Picking up the pieces: Impression management in the retrospective attributional framing of accounting outcomes. *Accounting, Organizations and Society* 30: 493–517.
- Bhatia, Vijay K. 2008. Towards critical genre analysis. In Vijay K. Bhatia, John Flowerdew and Rodney H. Jones eds. *Advances in Discourse Studies*. London: Routledge, 166–177.
- Bosque, Ignacio. 1999. El nombre común. In Violeta Demonte and Ignacio Bosque eds. *Gramática Descriptiva de la Lengua Española*. Madrid: Espasa Calpe, 3–75.
- El-Haj, Mahmoud, Antonio Moreno-Sandoval and José Antonio Jiménez Millán. 2021. Machine learning models for classifying Spanish beaters and non-beaters financial reports. In Antonio Moreno-Sandoval ed. *Financial Narrative Processing in Spanish*. Valencia: Tirant, 179–198.
- García Toro, Ana. 2022. Los marcadores del discurso en la narrativa financiera: Análisis de las cartas a los accionistas. *ELUA. Estudios de Lingüística* 38: 187–214.
- Garzone, Giuliana. 2004. Annual company reports and CEO's letters: Discoursal features and cultural markedness. In Christopher Candlin and Maurizio Gotti eds. *Intercultural Aspects of Specialized Discourse*. Bern: Peter Lang, 311–341.
- Garzone, Giuliana. 2005. Letters to shareholders and chairman's statements: Textual variability and generic integrity. In Paul Gillaerts and Maurizio Gotti eds. *Genre Variation in Business Letters*. Bern: Peter Lang, 179–204.
- Kranich, Svenja. 2011. The use of epistemic expressions in letters to shareholders in France, the United States and Germany. *Langage et Société* 3/3: 115–134.
- Kranich, Svenja and Andrea Bicsar. 2012. "These forecasts may be substantially different from actual results". The use of epistemic modal markers in English and German original letters to shareholders and in English-German translations. *Linguistik Online* 55/5: 41–55.
- Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý and Vít Suchomel. 2014. The Sketch Engine: Ten years on. *Lexicography* 1/1: 7–36.
- Laskin, Alexander V. 2018. The narrative strategies of winners and losers: Analyzing annual reports of publicly traded corporations. *International Journal of Business Communication* 55/3: 338–356.
- Luo, Yan and Linying Zhou. 2019. Tone of earnings announcements in sin industries. *Asian Review of Accounting* 27/2: 228–246.
- McEnery, Anthony and Zhonghua Xiao. 2005. *Help or Help to: What do corpora have to say?* *English Studies* 86/2: 161–187.
- Moreno-Sandoval, Antonio, Ana Gisbert and Helena Montoro. 2020. FinT-esp: A corpus of financial reports in Spanish. In Miguel Fuster-Márquez, Carmen Gregori-Signes and José Santaemilia Ruiz eds. *Multiperspectives in Analysis and Corpus Design*. Granada: Comares, 89–102.

- Nickerson, Catherine and Elizabeth de Groot. 2005. Dear shareholder, dear stockholder, dear stakeholder: The business letter genre in the annual general report. In Paul Gillaerts and Maurizio Gotti eds. *Genre Variation in Business Letters*. Bern: Peter Lang, 325–246.
- Patelli, Lorena and Matteo Pedrini. 2014. Is the optimism in CEO's letters to shareholders sincere? Impression management versus communicative action during the economic crisis. *Journal of Business Ethics* 124/1: 19–34.
- Poole, Robert. 2016. Good times, bad times: A keyword analysis of letters to shareholders of two fortune 500 banking institutions. *International Journal of Business Communication* 53/1: 55–73.
- Poole, Robert. 2017. "New opportunities" and "strong performance": Evaluative adjectives in letters to shareholders and potential for pedagogically-downsized specialized corpora. *English for Specific Purposes* 47: 40–51.
- Real Academia Española. n.d. *Diccionario de la Lengua Española* (23rd ed., 23.5 online version). <https://dle.rae.es/>
- Ruiz-Garrido, Miguel, Inmaculada Fortanet-Gómez and Juan Carlos Palmer-Silveira. 2012. Introducing British and Spanish companies to investors: Building the corporate image through the chairman's statement. In Jolanta Aritz and Robyn C. Walker eds. *Discourse Perspectives on Organizational Communication*. Plymouth: Fairleigh Dickinson University Press, 159–178.
- Skorzynska Sznajder, Hanna. 2021. The analysis of adjectives in letters to shareholders of British and Spanish companies from the Valuation Theory. *Pragmalinguistics* 29: 421–439.
- Skorzynska Sznajder, Hanna and Rosa Giménez-Moreno. 2016. Variation in letters to shareholders from British, Polish and Spanish companies: A comparative study. *Journal of Intercultural Communication* 40: 1–21.
- Swales, John. 1990. *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.
- Taylor, Charlotte. 2013. Searching for similarity using corpus-assisted discourse studies. *Corpora* 8/1: 81–113.
- Taylor, Charlotte. 2018. Similarity. In Charlotte Taylor and Anna Marchi eds. *Corpus Approaches to Discourse: A Critical Review*. London: Routledge, 19–37.
- Thompson, Geoff and Susan Hunston. 2000. Evaluation: An introduction. In Susan Hunston and Geoff Thompson eds. *Evaluation in Text: Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press, 1–27.
- Vendler, Zeno. 1957. Verbs and times. *The Philosophical Review* 66/2: 143–160.
- Vogel, Radek. 2020. Persuasion in business discourse: Strategic use of evaluative lexical means in corporate annual reports. In Olga Dontcheva-Navratilova, Martin Adam, Renata Povolná and Radek Vogel eds. *Persuasion in Specialised Discourses*. Cham: Palgrave MacMillan, 159–195.
- Wang, Jinxiao. 2020. A corpus-based analysis of collocations in Chinese and American letters to shareholders. *International Journal Advances in Social Science and Humanities* 8/6: 10–19.

Corresponding author

Blanca Carbajo Coronado
Facultad de Filosofía y Letras
Autonomous University of Madrid
Francisco Tomás y Valiente, 1
28049 Madrid
Spain
E-mail: blanca.carbajo@uam.es

received: December 2022
accepted: March 2023

APPENDICES

Appendix 1: The 50 most frequent verbs in the profit and loss corpora

	Profit corpus	Normalised frequency	Loss corpus	Normalised frequency
1	<i>Ser</i>	9.43	<i>Ser</i>	10.21
2	<i>Estar</i>	3.19	<i>Estar</i>	3.85
3	<i>Tener</i>	2.58	<i>Tener</i>	3.18
4	<i>Permitir</i>	2.17	<i>Hacer</i>	2.20
5	<i>Seguir</i>	2.14	<i>Poder</i>	1.72
6	<i>Hacer</i>	1.89	<i>Permitir</i>	1.52
7	<i>Alcanzar</i>	1.69	<i>Seguir</i>	1.48
8	<i>Poder</i>	1.30	<i>Alcanzar</i>	1.47
9	<i>Mejorar</i>	1.27	<i>Pasar</i>	1.47
10	<i>Mantener</i>	1.14	<i>Querer</i>	1.13
11	<i>Crecer</i>	1.09	<i>Deber</i>	1.13
12	<i>Querer</i>	1.09	<i>Mantener</i>	1.04
13	<i>Realizar</i>	1.06	<i>Dar</i>	1.02
14	<i>Destacar</i>	1.06	<i>Continuar</i>	0.94
15	<i>Pasar</i>	1.03	<i>Llevar</i>	0.93
16	<i>Desarrollar</i>	1.02	<i>Trabajar</i>	0.88
17	<i>Unir</i>	1.00	<i>Suponer</i>	0.86
18	<i>Conseguir</i>	0.98	<i>Unir</i>	0.81
19	<i>Ir</i>	0.94	<i>Mejorar</i>	0.80
20	<i>Dar</i>	0.92	<i>Consolidar</i>	0.77
21	<i>Suponer</i>	0.89	<i>Reducir</i>	0.75
22	<i>Generar</i>	0.80	<i>Ver</i>	0.73
23	<i>Reducir</i>	0.77	<i>Poner</i>	0.73
24	<i>Deber</i>	0.70	<i>Lograr</i>	0.73
25	<i>Aumentar</i>	0.70	<i>Estimar</i>	0.73
26	<i>Trabajar</i>	0.70	<i>Conseguir</i>	0.73
27	<i>Avanzar</i>	0.69	<i>Crecer</i>	0.72
28	<i>Continuar</i>	0.69	<i>Desarrollar</i>	0.70
29	<i>Impulsar</i>	0.69	<i>Ir</i>	0.67
30	<i>Situvar</i>	0.67	<i>Contar</i>	0.67
31	<i>Contar</i>	0.67	<i>Destacar</i>	0.65
32	<i>Poner</i>	0.66	<i>Realizar</i>	0.65
33	<i>Cumplir</i>	0.64	<i>Esperar</i>	0.64
34	<i>Llevar</i>	0.64	<i>Avanzar</i>	0.62
35	<i>Consolidar</i>	0.59	<i>Presentar</i>	0.61
36	<i>Superar</i>	0.58	<i>Decir</i>	0.59

37	<i>Esperar</i>	0.58	<i>Producir</i>	0.54
38	<i>Marcar</i>	0.56	<i>Obtener</i>	0.54
39	<i>Demostrar</i>	0.56	<i>Reforzar</i>	0.51
40	<i>Afrontar</i>	0.55	<i>Mostrar</i>	0.51
41	<i>Estimar</i>	0.53	<i>Afrontar</i>	0.48
42	<i>Contribuir</i>	0.50	<i>Centrar</i>	0.48
43	<i>Obtener</i>	0.50	<i>Iniciar</i>	0.48
44	<i>Ver</i>	0.48	<i>Cerrar</i>	0.48
45	<i>Prever</i>	0.47	<i>Generar</i>	0.48
46	<i>Presentar</i>	0.47	<i>Situar</i>	0.45
47	<i>Ascender</i>	0.47	<i>Registrar</i>	0.45
48	<i>Producir</i>	0.47	<i>Reflejar</i>	0.41
49	<i>Operar</i>	0.45	<i>Marcar</i>	0.41
50	<i>Agradecer</i>	0.44	<i>Llegar</i>	0.41

Appendix 2: The 50 most frequent eventive nouns in the profit and loss corpora

	Profit corpus	Normalised frequency	Loss corpus	Normalised frequency
1	<i>Ejercicio</i>	3.17	<i>Negocio</i>	2.98
2	<i>Resultado</i>	2.91	<i>Ejercicio</i>	2.63
3	<i>Crecimiento</i>	2.77	<i>Resultado</i>	2.20
4	<i>Negocio</i>	2.42	<i>Crecimiento</i>	1.61
5	<i>Compromiso</i>	1.36	<i>Venta</i>	1.53
6	<i>Mejora</i>	1.31	<i>Proceso</i>	1.31
7	<i>Venta</i>	1.27	<i>Gestión</i>	1.28
8	<i>Desarrollo</i>	1.17	<i>Cambio</i>	1.13
9	<i>Esfuerzo</i>	1.17	<i>Deuda</i>	1.13
10	<i>Producción</i>	1.14	<i>Coste</i>	0.97
11	<i>Inversión</i>	1.09	<i>Desarrollo</i>	0.96
12	<i>Gestión</i>	1.02	<i>Reducción</i>	0.88
13	<i>Administración</i>	0.84	<i>Compromiso</i>	0.85
14	<i>Evolución</i>	0.83	<i>Esfuerzo</i>	0.85
15	<i>Proceso</i>	0.83	<i>Administración</i>	0.83
16	<i>Acuerdo</i>	0.77	<i>Mejora</i>	0.83
17	<i>Innovación</i>	0.72	<i>Construcción</i>	0.75
18	<i>Reducción</i>	0.70	<i>Onversión</i>	0.73
19	<i>Coste</i>	0.69	<i>Cierre</i>	0.72
20	<i>Trabajo</i>	0.67	<i>Operación</i>	0.70
21	<i>Consumo</i>	0.64	<i>Trabajo</i>	0.70
22	<i>Cambio</i>	0.63	<i>Acuerdo</i>	0.67
23	<i>Gobierno</i>	0.63	<i>Generación</i>	0.64
24	<i>Ingreso</i>	0.56	<i>Transformación</i>	0.64
25	<i>Beneficio</i>	0.54	<i>Ampliación</i>	0.61
26	<i>Deuda</i>	0.54	<i>Evolución</i>	0.61
27	<i>Comportamiento</i>	0.52	<i>Ingreso</i>	0.57
28	<i>Construcción</i>	0.52	<i>Gasto</i>	0.51
29	<i>Emisión</i>	0.50	<i>Pérdida</i>	0.49
30	<i>Generación</i>	0.48	<i>Producción</i>	0.49
31	<i>Operación</i>	0.48	<i>Reestructuración</i>	0.49
32	<i>Recuperación</i>	0.48	<i>Integración</i>	0.41
33	<i>Cierre</i>	0.47	<i>Fobierno</i>	0.40
34	<i>Gasto</i>	0.43	<i>Aumento</i>	0.38
35	<i>Aumento</i>	0.42	<i>Consumo</i>	0.38
36	<i>Cumplimiento</i>	0.42	<i>Ahorro</i>	0.37
37	<i>Iniciativa</i>	0.42	<i>Caída</i>	0.37
38	<i>Transformación</i>	0.39	<i>Incremento</i>	0.37

39	<i>Solución</i>	0.38	<i>Innovación</i>	0.35
40	<i>Adquisición</i>	0.36	<i>Organización</i>	0.35
41	<i>Creación</i>	0.36	<i>Actuación</i>	0.34
42	<i>Demanda</i>	0.36	<i>Beneficio</i>	0.34
43	<i>Incremento</i>	0.36	<i>Comportamiento</i>	0.34
44	<i>Actuación</i>	0.30	<i>Compra</i>	0.34
45	<i>Instalación</i>	0.30	<i>Control</i>	0.34
46	<i>Aprobación</i>	0.28	<i>Oferta</i>	0.34
47	<i>Control</i>	0.28	<i>Iniciativa</i>	0.32
48	<i>División</i>	0.28	<i>Solución</i>	0.32
49	<i>Pago</i>	0.28	<i>Comercialización</i>	0.30
50	<i>Caída</i>	0.27	<i>Demanda</i>	0.30
